

IEEE Standard for Ethernet

SECTION SIX

This section includes Clause 78 through Clause 90 and Annex 83A through Annex 86A.

Contents

78. Energy-Efficient Ethernet (EEE)	21
78.1 Overview	21
78.1.1 LPI Signaling	21
78.1.1.1 Interlayer service interfaces	22
78.1.1.2 Responsibilities of LPI Client	22
78.1.2 LPI Client service interface	22
78.1.2.1 LP_IDLE.request	22
78.1.2.1.1 Function	22
78.1.2.1.2 Semantics of the service primitive	23
78.1.2.1.3 When generated	23
78.1.2.1.4 Effect of receipt	23
78.1.2.2 LP_IDLE.indication	23
78.1.2.2.1 Function	23
78.1.2.2.2 Semantics of the service primitive	23
78.1.2.2.3 When generated	23
78.1.2.2.4 Effect of receipt	23
78.1.3 Reconciliation sublayer operation	24
78.1.3.1 RS LPI assert function	24
78.1.3.2 LPI detect function	25
78.1.3.3 PHY LPI operation	25
78.1.3.3.1 PHY LPI transmit operation	25

78.1.3.3.2	PHY LPI receive operation	26
78.1.4	EEE Supported PHY types	26
78.2	LPI mode timing parameters description	26
78.3	Capabilities Negotiation	27
78.4	Data Link Layer Capabilities	28
78.4.1	Data Link Layer capabilities timing requirements	29
78.4.2	Control state diagrams	29
78.4.2.1	Conventions	29
78.4.2.2	Constants	29
78.4.2.3	Variables	29
78.4.2.4	Functions	31
78.4.2.5	State diagrams	32
78.4.3	State change procedure across a link	34
78.4.3.1	Transmitting link partner's state change procedure across a link	35
78.4.3.2	Receiving link partner's state change procedure across a link	35
78.5	Communication link access latency	36
78.5.1	10 Gb/s PHY extension using XGXS	37
78.6	Protocol implementation conformance statement (PICS) proforma for EEE Data Link Layer Capabilities	37
78.6.1	Introduction	37
78.6.2	Identification	38
78.6.2.1	Implementation identification	38
78.6.2.2	Protocol summary	38
78.6.3	Major capabilities/options	38
78.6.4	DLL requirements	39
79.	IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements	41
79.1	Overview	41
79.1.1	IEEE 802.3 LLDP frame format	41
79.1.1.1	Destination Address field	42
79.1.1.2	Source Address field	42
79.1.1.3	Length/Type field	42
79.1.1.4	LLDPDU field	42
79.1.1.5	Pad field	42
79.1.1.6	Frame Check Sequence field	42
79.2	Requirements of the IEEE 802.3 Organizationally Specific TLV set	42
79.3	IEEE 802.3 Organizationally Specific TLVs	42
79.3.1	MAC/PHY Configuration/Status TLV	43
79.3.1.1	Auto-negotiation support/status	43
79.3.1.2	PMD auto-negotiation advertised capability field	43
79.3.1.3	Operational MAU type	43
79.3.1.4	MAC/PHY Configuration/Status TLV usage rules	44
79.3.2	Power Via MDI TLV	44
79.3.2.1	MDI power support	44
79.3.2.2	PSE power pair	45
79.3.2.3	Power class	45
79.3.2.4	Requested power type/source/priority	45
79.3.2.4.1	Power type	46
79.3.2.4.2	Power source	46
79.3.2.4.3	Power priority	46
79.3.2.5	PD requested power value	46
79.3.2.6	PSE allocated power value	47

79.3.2.7	Power Via MDI TLV usage rules	47
79.3.3	Link Aggregation TLV (deprecated)	47
79.3.3.1	Aggregation status	47
79.3.3.2	Aggregated port ID	48
79.3.3.3	Link Aggregation TLV usage rules	48
79.3.4	Maximum Frame Size TLV	48
79.3.4.1	Maximum frame size	48
79.3.4.2	Maximum Frame Size TLV usage rules	48
79.3.5	EEE TLV	49
79.3.5.1	Transmit Tw	49
79.3.5.2	Receive Tw	49
79.3.5.3	Fallback Tw	49
79.3.5.4	Echo Transmit and Receive Tw	50
79.3.5.5	EEE TLV usage rules	50
79.4	IEEE 802.3 Organizationally Specific TLV selection management	50
79.4.1	IEEE 802.3 Organizationally Specific TLV selection variable/LLDP Configuration managed object class cross reference	50
79.4.2	IEEE 802.3 Organizationally Specific TLV/LLDP Local and Remote System group managed object class cross references	50
79.5	Protocol implementation conformance statement (PICS) proforma for IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements	53
79.5.1	Introduction	53
79.5.2	Identification	53
79.5.2.1	Implementation identification	53
79.5.2.2	Protocol summary	53
79.5.3	Major capabilities/options	54
79.5.4	IEEE 802.3 Organizationally Specific TLV	54
79.5.5	MAC/PHY Configuration/Status TLV	54
79.5.6	EEE TLV	55
79.5.7	Power Via MDI TLV	56
79.5.8	Link Aggregation TLV	57
79.5.9	Maximum Frame Size TLV	57
80.	Introduction to 40 Gb/s and 100 Gb/s networks	59
80.1	Overview	59
80.1.1	Scope	59
80.1.2	Objectives	59
80.1.3	Relationship of 40 Gigabit and 100 Gigabit Ethernet to the ISO OSI reference model	59
80.1.4	Nomenclature	61
80.1.5	Physical Layer signaling systems	61
80.2	Summary of 40 Gigabit and 100 Gigabit Ethernet sublayers	62
80.2.1	Reconciliation Sublayer (RS) and Media Independent Interface	62
80.2.2	Physical Coding Sublayer (PCS)	62
80.2.3	Forward Error Correction (FEC) sublayer	63
80.2.4	Physical Medium Attachment (PMA) sublayer	63
80.2.5	Physical Medium Dependent (PMD) sublayer	63
80.2.6	Auto-Negotiation	63
80.2.7	Management interface (MDIO/MDC)	63
80.2.8	Management	63
80.3	Service interface specification method and notation	64
80.3.1	Inter-sublayer service interface	64

80.3.2	Instances of the Inter-sublayer service interface.....	64
80.3.3	Semantics of inter-sublayer service interface primitives.....	67
80.3.3.1	IS_UNITDATA_i.request.....	67
80.3.3.1.1	Semantics of the service primitive.....	67
80.3.3.1.2	When generated.....	67
80.3.3.1.3	Effect of receipt.....	67
80.3.3.2	IS_UNITDATA_i.indication.....	67
80.3.3.2.1	Semantics of the service primitive.....	67
80.3.3.2.2	When generated.....	68
80.3.3.2.3	Effect of receipt.....	68
80.3.3.3	IS_SIGNAL.indication.....	68
80.3.3.3.1	Semantics of the service primitive.....	68
80.3.3.3.2	When generated.....	68
80.3.3.3.3	Effect of receipt.....	68
80.4	Delay constraints.....	68
80.5	Skew constraints.....	69
80.6	State diagrams.....	73
80.7	Protocol implementation conformance statement (PICS) proforma.....	73
81.	Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation (XLGMII and CGMII).....	75
81.1	Overview.....	75
81.1.1	Summary of major concepts.....	76
81.1.2	Application.....	76
81.1.3	Rate of operation.....	76
81.1.4	Delay constraints.....	76
81.1.5	Allocation of functions.....	77
81.1.6	XLGMII/CGMII structure.....	77
81.1.7	Mapping of XLGMII/CGMII signals to PLS service primitives.....	78
81.1.7.1	Mapping of PLS_DATA.request.....	78
81.1.7.1.1	Function.....	78
81.1.7.1.2	Semantics of the service primitive.....	78
81.1.7.1.3	When generated.....	78
81.1.7.1.4	Effect of receipt.....	78
81.1.7.2	Mapping of PLS_DATA.indication.....	79
81.1.7.2.1	Function.....	79
81.1.7.2.2	Semantics of the service primitive.....	79
81.1.7.2.3	When generated.....	79
81.1.7.2.4	Effect of receipt.....	79
81.1.7.3	Mapping of PLS_CARRIER.indication.....	79
81.1.7.4	Mapping of PLS_SIGNAL.indication.....	79
81.1.7.5	Mapping of PLS_DATA_VALID.indication.....	79
81.1.7.5.1	Function.....	79
81.1.7.5.2	Semantics of the service primitive.....	79
81.1.7.5.3	When generated.....	80
81.1.7.5.4	Effect of receipt.....	80
81.2	XLGMII/CGMII data stream.....	80
81.2.1	Inter-frame <inter-frame>.....	81
81.2.2	Preamble <preamble> and start of frame delimiter <sfd>.....	81
81.2.3	Data <data>.....	82
81.2.4	End of frame delimiter <efd>.....	82
81.2.5	Definition of Start of Packet and End of Packet Delimiters.....	82
81.3	XLGMII/CGMII functional specifications.....	82

81.3.1	Transmit	82
81.3.1.1	TX_CLK	82
81.3.1.2	TXC<7:0> (transmit control)	83
81.3.1.3	TXD<63:0> (transmit data)	83
81.3.1.4	Start control character alignment	85
81.3.2	Receive	86
81.3.2.1	RX_CLK (receive clock)	86
81.3.2.2	RXC<7:0> (receive control)	86
81.3.2.3	RXD (receive data)	88
81.3.3	Error and fault handling	89
81.3.3.1	Response to error indications by the XLGMII/CGMII	89
81.3.3.2	Conditions for generation of transmit Error control characters	90
81.3.3.3	Response to received invalid frame sequences	90
81.3.4	Link fault signaling	90
81.3.4.1	Variables and counters	91
81.3.4.2	State Diagram	91
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	93
81.4.1	Introduction	93
81.4.2	Identification	93
81.4.2.1	Implementation identification	93
81.4.2.2	Protocol summary	93
81.4.2.3	Major capabilities/options	94
81.4.3	PICS proforma tables for Reconciliation Sublayer and Media Independent Interface for 40 Gb/s and 100 Gb/s operation	94
81.4.3.1	General	94
81.4.3.2	Mapping of PLS service primitives	94
81.4.3.3	Data stream structure	95
81.4.3.4	XLGMII/CGMII signal functional specifications	96
81.4.3.5	Link fault signaling state diagram	97
82.	Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R	99
82.1	Overview	99
82.1.1	Scope	99
82.1.2	Relationship of 40GBASE-R and 100GBASE-R to other standards	99
82.1.3	Summary of 40GBASE-R and 100GBASE-R sublayers	100
82.1.3.1	Physical Coding Sublayer (PCS)	100
82.1.4	Inter-sublayer interfaces	101
82.1.4.1	PCS service interface (XLGMII/CGMII)	101
82.1.4.2	Physical Medium Attachment (PMA) or Forward Error Correction (FEC) service interface	101
82.1.5	Functional block diagram	101
82.2	Physical Coding Sublayer (PCS)	102
82.2.1	Functions within the PCS	102
82.2.2	Use of blocks	103
82.2.3	64B/66B transmission code	104
82.2.3.1	Notation conventions	104
82.2.3.2	Transmission order	104
82.2.3.3	Block structure	107
82.2.3.4	Control codes	108
82.2.3.5	Valid and invalid blocks	109
82.2.3.6	Idle (I)	109

82.2.3.7	Start (/S).....	109
82.2.3.8	Terminate (/T).....	109
82.2.3.9	ordered_set (/O).....	110
82.2.3.10	Error (/E).....	110
82.2.4	Transmit process.....	110
82.2.5	Scrambler.....	110
82.2.6	Block distribution.....	110
82.2.7	Alignment marker insertion.....	111
82.2.8	BIP calculations.....	113
82.2.9	PMA or FEC Interface.....	114
82.2.10	Test-pattern generators.....	115
82.2.11	Block synchronization.....	115
82.2.12	PCS lane deskew.....	115
82.2.13	PCS lane reorder.....	115
82.2.14	Alignment marker removal.....	116
82.2.15	Descrambler.....	116
82.2.16	Receive process.....	116
82.2.17	Test-pattern checker.....	116
82.2.18	Detailed functions and state diagrams.....	117
82.2.18.1	State diagram conventions.....	117
82.2.18.2	State variables.....	117
82.2.18.2.1	Constants.....	117
82.2.18.2.2	Variables.....	117
82.2.18.2.3	Functions.....	119
82.2.18.2.4	Counters.....	120
82.2.18.2.5	Timers.....	121
82.2.18.3	State diagrams.....	121
82.3	PCS Management.....	122
82.3.1	PMD MDIO function mapping.....	122
82.4	Loopback.....	123
82.5	Delay constraints.....	123
82.6	Auto-Negotiation.....	124
82.7	Protocol implementation conformance statement (PICS) proforma for Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R.....	130
82.7.1	Introduction.....	130
82.7.2	Identification.....	130
82.7.2.1	Implementation identification.....	130
82.7.2.2	Protocol summary.....	130
82.7.3	Major capabilities/options.....	131
82.7.4	PICS Proforma Tables for PCS, type 40GBASE-R and 100GBASE-R.....	132
82.7.4.1	Coding rules.....	132
82.7.4.2	Scrambler and Descrambler.....	132
82.7.4.3	Deskew and Reordering.....	133
82.7.4.4	Alignment Markers.....	133
82.7.5	Test-pattern modes.....	133
82.7.5.1	Bit order.....	134
82.7.6	Management.....	134
82.7.6.1	State diagrams.....	135
82.7.6.2	Loopback.....	136
82.7.6.3	Delay constraints.....	136
82.7.6.5	Auto-Negotiation for Backplane Ethernet functions.....	136
83.	Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R.....	137

83.1	Overview.....	137
83.1.1	Scope.....	137
83.1.2	Position of the PMA in the 40GBASE-R or 100GBASE-R sublayers.....	137
83.1.3	Summary of functions.....	137
83.1.4	PMA sublayer positioning.....	137
83.2	PMA interfaces.....	140
83.3	PMA service interface.....	140
83.4	Service interface below PMA.....	142
83.5	Functions within the PMA.....	143
83.5.1	Per input-lane clock and data recovery.....	143
83.5.2	Bit-level multiplexing.....	143
83.5.3	Skew and Skew Variation.....	146
83.5.3.1	Skew generation toward SP1.....	146
83.5.3.2	Skew tolerance at SP1.....	146
83.5.3.3	Skew generation toward SP2.....	146
83.5.3.4	Skew tolerance at SP5.....	146
83.5.3.5	Skew generation at SP6.....	146
83.5.3.6	Skew tolerance at SP6.....	147
83.5.4	Delay constraints.....	147
83.5.5	Clocking architecture.....	147
83.5.6	Signal drivers.....	147
83.5.7	Link status.....	148
83.5.8	PMA local loopback mode.....	148
83.5.9	PMA remote loopback mode (optional).....	148
83.5.10	PMA test patterns (optional).....	149
83.6	PMA MDIO function mapping.....	150
83.7	Protocol implementation conformance statement (PICS) proforma for Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R.....	155
83.7.1	Introduction.....	155
83.7.2	Identification.....	155
83.7.2.1	Implementation identification.....	155
83.7.2.2	Protocol summary.....	155
83.7.3	Major capabilities/options.....	156
83.7.4	Skew generation and tolerance.....	158
83.7.5	Test patterns.....	158
83.7.6	Loopback modes.....	159
84.	Physical Medium Dependent sublayer and baseband medium, type 40GBASE-KR4.....	161
84.1	Overview.....	161
84.2	Physical Medium Dependent (PMD) service interface.....	161
84.3	PCS requirements for Auto-Negotiation (AN) service interface.....	162
84.4	Delay constraints.....	163
84.5	Skew constraints.....	163
84.6	PMD MDIO function mapping.....	163
84.7	PMD functional specifications.....	163
84.7.1	Link block diagram.....	163
84.7.2	PMD transmit function.....	163
84.7.3	PMD receive function.....	165
84.7.4	Global PMD signal detect function.....	165
84.7.5	PMD lane-by-lane signal detect function.....	165
84.7.6	Global PMD transmit disable function.....	165
84.7.7	PMD lane-by-lane transmit disable function.....	166
84.7.8	Loopback mode.....	166

84.7.9	PMD_fault function	166
84.7.10	PMD transmit fault function	166
84.7.11	PMD receive fault function.....	166
84.7.12	PMD control function	167
84.8	40GBASE-KR4 electrical characteristics	167
84.8.1	Transmitter characteristics	167
84.8.1.1	Test fixture	167
84.8.2	Receiver characteristics	167
84.8.2.1	Receiver interference tolerance	167
84.9	Interconnect characteristics.....	167
84.10	Environmental specifications.....	167
84.10.1	General safety	167
84.10.2	Network safety	168
84.10.3	Installation and maintenance guidelines	168
84.10.4	Electromagnetic compatibility.....	168
84.10.5	Temperature and humidity	168
84.11	Protocol implementation conformance statement (PICS) proforma for Clause 84, Physical Medium Dependent sublayer and baseband medium, type 40GBASE-KR4	169
84.11.1	Introduction.....	169
84.11.2	Identification	169
84.11.2.1	Implementation identification	169
84.11.2.2	Protocol summary	169
84.11.3	Major capabilities/options.....	170
84.11.4	PICS proforma tables for Clause 84, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-KR4.....	171
84.11.4.1	PMD functional specifications.....	171
84.11.4.2	Management functions.....	172
84.11.4.3	Transmitter electrical characteristics	172
84.11.4.4	Receiver electrical characteristics.....	172
84.11.4.5	Environmental specifications.....	173
85.	Physical Medium Dependent sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10.....	175
85.1	Overview.....	175
85.2	Physical Medium Dependent (PMD) service interface	176
85.3	PCS requirements for Auto-Negotiation (AN) service interface.....	177
85.4	Delay constraints.....	177
85.5	Skew constraints	177
85.6	PMD MDIO function mapping.....	178
85.7	PMD functional specifications.....	180
85.7.1	Link block diagram	180
85.7.2	PMD Transmit function	181
85.7.3	PMD Receive function.....	181
85.7.4	Global PMD signal detect function	181
85.7.5	PMD lane-by-lane signal detect function	182
85.7.6	Global PMD transmit disable function	182
85.7.7	PMD lane-by-lane transmit disable function	182
85.7.8	Loopback mode.....	182
85.7.9	PMD_fault function	183
85.7.10	PMD transmit fault function	183
85.7.11	PMD receive fault function.....	183
85.7.12	PMD control function	183
85.8	MDI Electrical specifications for 40GBASE-CR4 and 100GBASE-CR10	183

85.8.1	Signal levels.....	183
85.8.2	Signal paths.....	184
85.8.3	Transmitter characteristics.....	184
85.8.3.1	Transmitter differential output return loss.....	185
85.8.3.2	Transmitter noise parameter measurements.....	185
85.8.3.3	Transmitter output waveform.....	186
85.8.3.3.1	Coefficient initialization.....	188
85.8.3.3.2	Coefficient step size.....	188
85.8.3.3.3	Coefficient range.....	188
85.8.3.3.4	Waveform acquisition.....	188
85.8.3.3.5	Linear fit to the waveform measurement at TP2.....	188
85.8.3.3.6	Transfer function between the transmit function and TP2.....	189
85.8.3.4	Insertion loss TP0 to TP2 or TP3 to TP5.....	190
85.8.3.5	Test fixture.....	191
85.8.3.6	Test fixture impedance.....	192
85.8.3.7	Test fixture insertion loss.....	192
85.8.3.8	Data dependent jitter (DDJ).....	192
85.8.3.9	Signaling rate range.....	192
85.8.4	Receiver characteristics at TP3 summary.....	193
85.8.4.1	Receiver differential input return loss.....	193
85.8.4.2	Receiver interference tolerance test.....	194
85.8.4.2.1	Test setup.....	194
85.8.4.2.2	Test channel.....	195
85.8.4.2.3	Test channel calibration.....	195
85.8.4.2.4	Pattern generator.....	196
85.8.4.2.5	Test procedure.....	196
85.8.4.3	Bit error ratio.....	196
85.8.4.4	Signaling rate range.....	196
85.8.4.5	AC coupling.....	196
85.9	Channel characteristics.....	197
85.10	Cable assembly characteristics.....	197
85.10.1	Characteristic impedance and reference impedance.....	197
85.10.2	Cable assembly insertion loss.....	197
85.10.3	Cable assembly insertion loss deviation (ILD).....	199
85.10.4	Cable assembly return loss.....	200
85.10.5	Cable assembly multiple disturber near-end crosstalk (MDNEXT) loss.....	201
85.10.6	Cable assembly multiple disturber far-end crosstalk (MDFEXT) loss.....	201
85.10.7	Cable assembly integrated crosstalk noise (ICN).....	202
85.10.8	Cable assembly test fixture.....	204
85.10.9	Mated test fixtures.....	204
85.10.9.1	Mated test fixtures insertion loss.....	205
85.10.9.2	Mated test fixtures return loss.....	205
85.10.9.3	Mated test fixtures common-mode conversion loss.....	206
85.10.9.4	Mated test fixtures integrated crosstalk noise.....	207
85.10.10	Shielding.....	207
85.10.11	Crossover function.....	207
85.11	MDI specification.....	208
85.11.1	40GBASE-CR4 MDI connectors.....	208
85.11.1.1	Style-1 40GBASE-CR4 MDI connectors.....	208
85.11.1.1.1	Style-1 AC coupling.....	209
85.11.1.2	Style-2 40GBASE-CR4 MDI connectors.....	210
85.11.1.2.1	Style-2 40GBASE-CR4 Connector pin assignments.....	210
85.11.2	100GBASE-CR10 MDI connectors.....	211
85.11.2.1	100GBASE-CR10 MDI AC coupling.....	213

85.11.3	Electronic keying	213
85.12	Environmental specifications	213
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	214
85.13.1	Introduction	214
85.13.2	Identification	214
85.13.2.1	Implementation identification	214
85.13.2.2	Protocol summary	214
85.13.3	Major capabilities/options	215
85.13.4	PICS proforma tables for Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10	216
85.13.4.1	PMD functional specifications	216
85.13.4.2	Management functions	217
85.13.4.3	Transmitter specifications	217
85.13.4.4	Receiver specifications	218
85.13.4.5	Cable assembly specifications	219
85.13.4.6	MDI connector specifications	220
85.13.4.7	Environmental specifications	220
86.	Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-SR4 and 100GBASE-SR10	221
86.1	Overview	221
86.2	Physical Medium Dependent (PMD) service interface	223
86.3	Delay and Skew	224
86.3.1	Delay constraints	224
86.3.2	Skew and Skew Variation constraints	224
86.4	PMD MDIO function mapping	224
86.5	PMD functional specifications	225
86.5.1	PMD block diagram	225
86.5.2	PMD transmit function	226
86.5.3	PMD receive function	226
86.5.4	PMD global signal detect function	226
86.5.5	PMD lane-by-lane signal detect function	227
86.5.6	PMD reset function	227
86.5.7	PMD global transmit disable function (optional)	227
86.5.8	PMD lane-by-lane transmit disable function (optional)	228
86.5.9	PMD fault function (optional)	228
86.5.10	PMD transmit fault function (optional)	228
86.5.11	PMD receive fault function (optional)	228
86.6	Lane assignments	228
86.7	PMD to MDI specifications for 40GBASE-SR4 or 100GBASE-SR10	228
86.7.1	Transmitter optical specifications	229
86.7.2	Characteristics of signal within, and at the receiving end of, a compliant optical channel	230
86.7.3	40GBASE-SR4 or 100GBASE-SR10 receiver optical specifications	231
86.7.4	40GBASE-SR4 or 100GBASE-SR10 illustrative link power budget	232
86.8	Definitions of optical and dual-use parameters and measurement methods	232
86.8.1	Test points and compliance boards	232
86.8.2	Test patterns and related subclauses	232
86.8.2.1	Multi-lane testing considerations	235
86.8.3	Parameters applicable to both electrical and optical signals	235
86.8.3.1	Skew and Skew Variation	235

86.8.3.2	Eye diagrams.....	236
86.8.3.2.1	Eye mask acceptable hit count examples.....	236
86.8.3.3	Jitter	236
86.8.3.3.1	J2 Jitter.....	236
86.8.3.3.2	J9 Jitter.....	236
86.8.4	Optical parameter definitions.....	237
86.8.4.1	Wavelength and spectral width.....	237
86.8.4.2	Average optical power.....	237
86.8.4.3	Optical Modulation Amplitude (OMA).....	237
86.8.4.4	Transmitter and dispersion penalty (TDP).....	237
86.8.4.5	Extinction ratio	237
86.8.4.6	Transmitter optical waveform (transmit eye)	237
86.8.4.6.1	Optical transmitter eye mask	238
86.8.4.7	Stressed receiver sensitivity.....	239
86.8.4.8	Receiver jitter tolerance	239
86.9	Safety, installation, environment, and labeling.....	240
86.9.1	General safety	240
86.9.2	Laser safety	240
86.9.3	Installation	240
86.9.4	Environment.....	240
86.9.5	PMD labeling.....	240
86.10	Optical channel	240
86.10.1	Fiber optic cabling model	240
86.10.2	Characteristics of the fiber optic cabling (channel).....	241
86.10.2.1	Optical fiber cable.....	241
86.10.2.2	Optical fiber connection.....	241
86.10.2.2.1	Connection insertion loss.....	242
86.10.2.2.2	Maximum discrete reflectance.....	242
86.10.3	Medium Dependent Interface (MDI).....	242
86.10.3.1	Optical lane assignments for 40GBASE-SR4	242
86.10.3.2	Optical lane assignments for 100GBASE-SR10	243
86.10.3.3	Medium Dependent Interface (MDI) requirements	243
86.11	Protocol implementation conformance statement (PICS) proforma for Clause 86, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-SR4 and 100GBASE-SR10.....	245
86.11.1	Introduction.....	245
86.11.2	Identification	245
86.11.2.1	Implementation identification.....	245
86.11.2.2	Protocol summary	245
86.11.3	Major capabilities/options.....	246
86.11.4	PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, types 40GBASE-SR4 and 100GBASE-SR10	247
86.11.4.1	PMD functional specifications.....	247
86.11.4.2	Management functions.....	248
86.11.4.3	Optical specifications for 40GBASE-SR4 or 100GBASE-SR10.....	248
86.11.4.4	Definitions of parameters and measurement methods	249
86.11.4.5	Environmental and safety specifications	249
86.11.4.6	Optical channel and MDI.....	250
87.	Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4	251
87.1	Overview.....	251
87.2	Physical Medium Dependent (PMD) service interface	251
87.3	Delay and Skew	253

87.3.1	Delay constraints.....	253
87.3.2	Skew constraints	253
87.4	PMD MDIO function mapping.....	253
87.5	PMD functional specifications.....	253
87.5.1	PMD block diagram.....	253
87.5.2	PMD transmit function	254
87.5.3	PMD receive function	255
87.5.4	PMD global signal detect function	255
87.5.5	PMD lane-by-lane signal detect function	256
87.5.6	PMD reset function.....	256
87.5.7	PMD global transmit disable function (optional)	256
87.5.8	PMD lane-by-lane transmit disable function	256
87.5.9	PMD fault function (optional)	257
87.5.10	PMD transmit fault function (optional)	257
87.5.11	PMD receive fault function (optional).....	257
87.6	Wavelength-division-multiplexed lane assignments	257
87.7	PMD to MDI optical specifications for 40GBASE-LR4	257
87.7.1	40GBASE-LR4 transmitter optical specifications	258
87.7.2	40GBASE-LR4 receive optical specifications.....	259
87.7.3	40GBASE-LR4 illustrative link power budget	259
87.8	Definition of optical parameters and measurement methods.....	260
87.8.1	Test patterns for optical parameters.....	260
87.8.2	Skew and Skew Variation	260
87.8.3	Wavelength	261
87.8.4	Average optical power	261
87.8.5	Optical Modulation Amplitude (OMA).....	261
87.8.6	Transmitter and dispersion penalty.....	262
87.8.6.1	Reference transmitter requirements	262
87.8.6.2	Channel requirements	262
87.8.6.3	Reference receiver requirements	263
87.8.6.4	Test procedure.....	263
87.8.7	Extinction ratio	263
87.8.8	Relative Intensity Noise (RIN _{20OMA}).....	263
87.8.9	Transmitter optical waveform (transmit eye)	263
87.8.10	Receiver sensitivity.....	264
87.8.11	Stressed receiver sensitivity.....	264
87.8.11.1	Stressed receiver conformance test block diagram.....	264
87.8.11.2	Stressed receiver conformance test signal characteristics and calibration	265
87.8.11.3	Stressed receiver conformance test signal verification.....	267
87.8.11.4	Sinusoidal jitter for receiver conformance test.....	267
87.8.11.5	Stressed receiver conformance test procedure for WDM conformance testing	268
87.8.12	Receiver 3 dB electrical upper cutoff frequency	269
87.9	Safety, installation, environment, and labeling.....	269
87.9.1	General safety	269
87.9.2	Laser safety	269
87.9.3	Installation	270
87.9.4	Environment.....	270
87.9.4.1	Electromagnetic emission	270
87.9.4.2	Temperature, humidity, and handling.....	270
87.9.5	PMD labeling requirements	270
87.10	Fiber optic cabling model	271
87.11	Characteristics of the fiber optic cabling (channel).....	271
87.11.1	Optical fiber cable.....	272
87.11.2	Optical fiber connection.....	272

87.11.2.1	Connection insertion loss.....	272
87.11.2.2	Maximum discrete reflectance.....	272
87.11.3	Medium Dependent Interface (MDI) requirements.....	272
87.12	Protocol implementation conformance statement (PICS) proforma for Clause 87, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4.....	273
87.12.1	Introduction.....	273
87.12.2	Identification.....	273
87.12.2.1	Implementation identification.....	273
87.12.2.2	Protocol summary.....	273
87.12.3	Major capabilities/options.....	274
87.12.4	PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4.....	275
87.12.4.1	PMD functional specifications.....	275
87.12.4.2	Management functions.....	276
87.12.4.3	PMD to MDI optical specifications for 40GBASE-LR4.....	276
87.12.4.4	Optical measurement methods.....	277
87.12.4.5	Environmental specifications.....	277
87.12.4.6	Characteristics of the fiber optic cabling and MDI.....	277
88.	Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE-LR4 and 100GBASE-ER4.....	279
88.1	Overview.....	279
88.2	Physical Medium Dependent (PMD) service interface.....	279
88.3	Delay and Skew.....	281
88.3.1	Delay constraints.....	281
88.3.2	Skew constraints.....	281
88.4	PMD MDIO function mapping.....	281
88.5	PMD functional specifications.....	281
88.5.1	PMD block diagram.....	281
88.5.2	PMD transmit function.....	283
88.5.3	PMD receive function.....	283
88.5.4	PMD global signal detect function.....	283
88.5.5	PMD lane-by-lane signal detect function.....	284
88.5.6	PMD reset function.....	284
88.5.7	PMD global transmit disable function (optional).....	284
88.5.8	PMD lane-by-lane transmit disable function.....	284
88.5.9	PMD fault function (optional).....	285
88.5.10	PMD transmit fault function (optional).....	285
88.5.11	PMD receive fault function (optional).....	285
88.6	Wavelength-division-multiplexed lane assignments.....	285
88.7	PMD to MDI optical specifications for 100GBASE-LR4 and 100GBASE-ER4.....	285
88.7.1	100GBASE-LR4 and 100GBASE-ER4 transmitter optical specifications.....	286
88.7.2	100GBASE-LR4 and 100GBASE-ER4 receive optical specifications.....	288
88.7.3	100GBASE-LR4 and 100GBASE-ER4 illustrative link power budgets.....	289
88.8	Definition of optical parameters and measurement methods.....	289
88.8.1	Test patterns for optical parameters.....	289
88.8.2	Wavelength.....	289
88.8.3	Average optical power.....	290
88.8.4	Optical Modulation Amplitude (OMA).....	290
88.8.5	Transmitter and dispersion penalty (TDP).....	291
88.8.5.1	Reference transmitter requirements.....	291
88.8.5.2	Channel requirements.....	291
88.8.5.3	Reference receiver requirements.....	292

88.8.5.4	Test procedure.....	292
88.8.6	Extinction ratio	292
88.8.7	Relative Intensity Noise (RIN200MA).....	292
88.8.8	Transmitter optical waveform (transmit eye)	292
88.8.9	Receiver sensitivity.....	293
88.8.10	Stressed receiver sensitivity.....	293
88.8.11	Receiver 3 dB electrical upper cutoff frequency	293
88.9	Safety, installation, environment, and labeling.....	293
88.9.1	General safety	293
88.9.2	Laser safety	293
88.9.3	Installation	294
88.9.4	Environment.....	294
88.9.5	Electromagnetic emission	294
88.9.6	Temperature, humidity, and handling	294
88.9.7	PMD labeling requirements	294
88.10	Fiber optic cabling model	295
88.11	Characteristics of the fiber optic cabling (channel).....	295
88.11.1	Optical fiber cable.....	296
88.11.2	Optical fiber connection.....	296
88.11.2.1	Connection insertion loss.....	296
88.11.2.2	Maximum discrete reflectance.....	296
88.11.3	Medium Dependent Interface (MDI) requirements	296
88.12	Protocol implementation conformance statement (PICS) proforma for Clause 88, Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE–LR4 and 100GBASE–ER4	297
88.12.1	Introduction.....	297
88.12.2	Identification	297
88.12.2.1	Implementation identification.....	297
88.12.2.2	Protocol summary	297
88.12.3	Major capabilities/options.....	298
88.12.4	PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, types 100GBASE–LR4 and 100GBASE–ER4.....	299
88.12.4.1	PMD functional specifications.....	299
88.12.4.2	Management functions.....	300
88.12.4.3	PMD to MDI optical specifications for 100GBASE–LR4	300
88.12.4.4	PMD to MDI optical specifications for 100GBASE–ER4	300
88.12.4.5	Optical measurement methods.....	301
88.12.4.6	Environmental specifications.....	301
88.12.4.7	Characteristics of the fiber optic cabling and MDI.....	301
89.	Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR	303
89.1	Overview.....	303
89.2	Physical Medium Dependent (PMD) service interface	304
89.3	Delay and skew	305
89.3.1	Delay constraints.....	305
89.3.2	Skew constraints	305
89.4	PMD MDIO function mapping.....	305
89.5	PMD functional specifications.....	306
89.5.1	PMD block diagram.....	306
89.5.2	PMD transmit function	306
89.5.3	PMD receive function	306
89.5.4	PMD global signal detect function	307
89.5.5	PMD reset function.....	308

89.5.6	PMD global transmit disable function (optional)	308
89.5.7	PMD fault function (optional)	308
89.5.8	PMD transmit fault function (optional)	308
89.5.9	PMD receive fault function (optional)	308
89.6	PMD to MDI optical specifications for 40GBASE-FR	308
89.6.1	40GBASE-FR transmitter optical specifications	309
89.6.2	40GBASE-FR receive optical specifications	309
89.6.3	40GBASE-FR illustrative link power budget	310
89.6.4	Comparison of power budget methodology	310
89.7	Definition of optical parameters and measurement methods	311
89.7.1	Test patterns for optical parameters	311
89.7.2	Skew and Skew Variation	312
89.7.3	Wavelength	312
89.7.4	Average optical power	312
89.7.5	Dispersion penalty	312
89.7.5.1	Channel requirements	312
89.7.5.2	Reference receiver requirements	313
89.7.5.3	Test procedure	313
89.7.6	Extinction ratio	313
89.7.7	Relative Intensity Noise (RIN _{200MA})	314
89.7.8	Transmitter optical waveform (transmit eye)	314
89.7.9	Receiver sensitivity	314
89.7.10	Receiver jitter tolerance	314
89.7.11	Receiver 3 dB electrical upper cutoff frequency	314
89.8	Safety, installation, environment, and labeling	315
89.8.1	General safety	315
89.8.2	Laser safety	315
89.8.3	Installation	315
89.8.4	Environment	315
89.8.4.1	Electromagnetic emission	316
89.8.4.2	Temperature, humidity, and handling	316
89.8.5	PMD labeling requirements	316
89.9	Fiber optic cabling model	316
89.10	Characteristics of the fiber optic cabling (channel)	316
89.10.1	Optical fiber cable	316
89.10.2	Optical fiber connection	317
89.10.2.1	Connection insertion loss	317
89.10.2.2	Maximum discrete reflectance	317
89.10.3	Medium Dependent Interface (MDI) requirements	318
89.11	Protocol implementation conformance statement (PICS) proforma for Clause 89, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR	319
89.11.1	Introduction	319
89.11.2	Identification	319
89.11.2.1	Implementation identification	319
89.11.2.2	Protocol summary	319
89.11.3	Major capabilities/options	320
89.11.4	PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR	320
89.11.4.1	PMD functional specifications	320
89.11.4.2	Management functions	321
89.11.4.3	PMD to MDI optical specifications for 40GBASE-FR	321
89.11.4.4	Optical measurement methods	322
89.11.4.5	Environmental specifications	322
89.11.4.6	Characteristics of the fiber optic cabling and MDI	322

90. Ethernet support for time synchronization protocols.....	323
90.1 Introduction.....	323
90.2 Overview.....	323
90.3 Relationship with other IEEE standards.....	323
90.4 Time Synchronization Service Interface (TSSI).....	323
90.4.1 Introduction.....	323
90.4.1.1 Interlayer service interfaces.....	324
90.4.1.2 Responsibilities of TimeSync Client.....	324
90.4.2 TSSI.....	325
90.4.3 Detailed service specification.....	325
90.4.3.1 TS_TX.indication primitive.....	325
90.4.3.1.1 Semantics.....	325
90.4.3.1.2 Condition for generation.....	325
90.4.3.1.3 Effect of receipt.....	325
90.4.3.2 TS_RX.indication primitive.....	325
90.4.3.2.1 Semantics.....	325
90.4.3.2.2 Condition for generation.....	326
90.4.3.2.3 Effect of receipt.....	326
90.5 generic Reconciliation Sublayer (gRS).....	326
90.5.1 TS_SFD_Detect_TX function.....	326
90.5.2 TS_SFD_Detect_RX function.....	326
90.6 Overview of management features.....	327
90.7 Data delay measurement.....	328
90.8 Protocol implementation conformance statement (PICS) proforma for Clause 90, Ethernet support for time synchronization protocols.....	330
90.8.1 Introduction.....	330
90.8.2 Identification.....	330
90.8.2.1 Implementation identification.....	330
90.8.2.2 Protocol summary.....	330
90.8.3 TSSI indication.....	331
90.8.4 Data delay reporting.....	331
Annex 83A (normative) 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI).....	333
83A.1 Overview.....	333
83A.1.1 Summary of major concepts.....	334
83A.1.2 Rate of operation.....	334
83A.2 XLAUI/CAUI link block diagram.....	334
83A.2.1 Transmitter compliance points.....	335
83A.2.2 Receiver compliance points.....	336
83A.3 XLAUI/CAUI electrical characteristics.....	336
83A.3.1 Signal levels.....	336
83A.3.2 Signal paths.....	337
83A.3.3 Transmitter characteristics.....	337
83A.3.3.1 Output amplitude.....	337
83A.3.3.2 Rise/fall time.....	338
83A.3.3.3 Differential output return loss.....	339
83A.3.3.4 Common-mode output return loss.....	339
83A.3.3.5 Transmitter eye mask and transmitter jitter definition.....	340
83A.3.4 Receiver characteristics.....	341
83A.3.4.1 Bit error ratio.....	342
83A.3.4.2 Input signal definition.....	342

83A.3.4.3	Differential input return loss	342
83A.3.4.4	Differential to common-mode input return loss	343
83A.3.4.5	AC coupling	344
83A.3.4.6	Jitter tolerance	344
83A.4	Interconnect characteristics	345
83A.4.1	Characteristic impedance	346
83A.5	Electrical parameter measurement methods	347
83A.5.1	Transmit jitter	347
83A.5.2	Receiver tolerance	347
83A.6	Environmental specifications	348
83A.6.1	General safety	348
83A.6.2	Network safety	348
83A.6.3	Installation and maintenance guidelines	348
83A.6.4	Electromagnetic compatibility	348
83A.6.5	Temperature and humidity	348
83A.7	Protocol implementation conformance statement (PICS) proforma for Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)	349
83A.7.1	Introduction	349
83A.7.2	Identification	349
83A.7.2.1	Implementation identification	349
83A.7.2.2	Protocol summary	349
83A.7.3	Major capabilities/options	350
83A.7.4	XLAUI/CAUI transmitter requirements	350
83A.7.5	XLAUI/CAUI receiver requirements	351
83A.7.6	Electrical measurement methods	351
83A.7.7	Environmental specifications	351
Annex 83B (normative)	Chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)	353
83B.1	Overview	353
83B.2	Compliance point specifications for chip-module XLAUI/CAUI	355
83B.2.1	Module specifications	357
83B.2.2	Host specifications	360
83B.2.3	Host input signal tolerance	361
83B.3	Environmental specifications	362
83B.3.1	General safety	362
83B.3.2	Network safety	362
83B.3.3	Installation and maintenance guidelines	362
83B.3.4	Electromagnetic compatibility	362
83B.3.5	Temperature and humidity	363
83B.4	Protocol implementation conformance statement (PICS) proforma for Annex 83B, Chip- module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)	364
83B.4.1	Introduction	364
83B.4.2	Identification	364
83B.4.2.1	Implementation identification	364
83B.4.2.2	Protocol summary	364
83B.4.3	Major capabilities/options	365
83B.4.4	Module requirements	365
83B.4.5	Host requirements	365
83B.4.6	Environmental specifications	366
Annex 83C (normative)	PMA sublayer partitioning examples	367

83C.1	Partitioning examples with FEC	367
83C.1.1	FEC implemented with PCS	367
83C.1.2	FEC implemented with PMD	368
83C.2	Partitioning examples without FEC	368
83C.2.1	Single PMA sublayer without FEC	368
83C.2.2	Single XLAUI/CAUI without FEC	369
83C.2.3	Separate SERDES for optical module interface	369
Annex 85A (informative) 40GBASE-CR4 and 100GBASE-CR10 TP0 and TP5 test point parameters ...		371
85A.1	Overview	371
85A.2	Transmitter characteristics at TP0	371
85A.3	Receiver characteristics at TP5	372
85A.4	Transmitter and receiver differential printed circuit board trace loss	372
85A.5	Channel insertion loss	373
85A.6	Channel return loss	374
85A.7	Channel insertion loss deviation (ILD)	374
85A.8	Channel integrated crosstalk noise (ICN)	375
Annex 86A (normative) Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPP) and 100GBASE-SR10 (CPPI)		377
86A.1	Overview	377
86A.2	Block diagram and test points	377
86A.3	Lane assignments	377
86A.4	Electrical specifications for nPPI	378
86A.4.1	nPPI host to module electrical specifications	378
86A.4.1.1	Differential return losses at TP1 and TP1a	379
86A.4.2	nPPI module to host electrical specifications	379
86A.4.2.1	Differential return losses at TP4 and TP4a	380
86A.5	Definitions of electrical parameters and measurement methods	381
86A.5.1	Test points and compliance boards	381
86A.5.1.1	Compliance board parameters	382
86A.5.1.1.1	Reference insertion losses of HCB and MCB	383
86A.5.1.1.2	Electrical specifications of mated HCB and MCB	384
86A.5.2	Test patterns and related subclauses	386
86A.5.3	Parameter definitions	386
86A.5.3.1	AC common-mode voltage	387
86A.5.3.2	Termination mismatch	387
86A.5.3.3	Transition time	388
86A.5.3.4	Data Dependent Pulse Width Shrinkage (DDPWS)	388
86A.5.3.5	Signal to noise ratio Qsq	389
86A.5.3.6	Eye mask for TP1a and TP4	389
86A.5.3.7	Reference impedances for electrical measurements	390
86A.5.3.8	Host input signal tolerance	390
86A.5.3.8.1	Introduction	390
86A.5.3.8.2	Test equipment and setup	390
86A.5.3.8.3	Stressed eye jitter characteristics	390
86A.5.3.8.4	Calibration	392
86A.5.3.8.5	Calibration procedure	392
86A.5.3.8.6	Test procedure	393
86A.6	Recommended electrical channel	394
86A.7	Safety, installation, environment, and labeling	395
86A.7.1	General safety	395

86A.7.2	Installation	395
86A.7.3	Environment.....	395
86A.7.4	PMD labeling.....	395
86A.8	Protocol implementation conformance statement (PICS) proforma for Annex 86A, Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPPI) and 100GBASE-SR10 (CPPI)	396
86A.8.1	Introduction.....	396
86A.8.2	Identification.....	396
86A.8.2.1	Implementation identification.....	396
86A.8.2.2	Protocol summary.....	396
86A.8.3	Major capabilities/options.....	397
86A.8.4	PICS proforma tables for Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPPI) and 100GBASE-SR10 (CPPI).....	397
86A.8.4.1	PMD functional specifications.....	397
86A.8.4.2	Electrical specifications for nPPI.....	398
86A.8.4.3	Definitions of parameters and measurement methods.....	398
86A.8.4.4	Environmental and safety specifications	399

78. Energy-Efficient Ethernet (EEE)

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 the 100BASE-TX PHY, the 1000BASE-T PHY, and the 10GBASE-T PHY. For operation over electrical backplanes, EEE supports the 1000BASE-KX PHY, the 10GBASE-KX4 PHY, and the 10GBASE-KR PHY. EEE also supports XGMII extension using the XGXS for 10 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.

78.1.1 LPI Signaling

LPI signaling allows the LPI Client to indicate to the PHY, and to the link partner, that a break in the data stream is expected, and the LPI Client can use this information to enter power-saving modes that require additional time to resume normal operation. LPI signaling also informs the LPI Client when the link partner has sent such an indication.

The definition of LPI signaling assumes the use of the MAC 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 the LPI mode.

78.1.1.1 Interlayer service interfaces

Figure 78–1 depicts the LPI Client and the RS interlayer service interfaces.

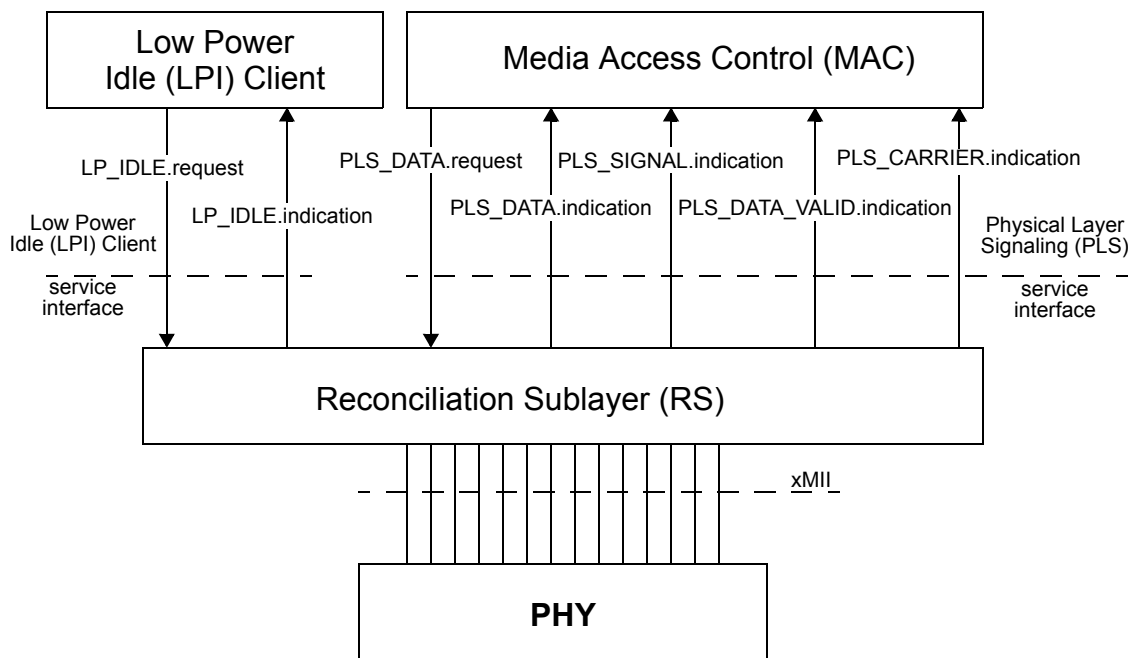


Figure 78–1—LPI Client and RS interlayer service interfaces

78.1.1.2 Responsibilities of LPI Client

The decision on when to signal LPI to the link partner is made by the LPI Client and communicated to the PHY through the RS. The LPI Client is also informed when the link partner is signaling LPI by the RS.

The conditions under which the LPI Client decides to send LPI, and what action are taken by the LPI Client when it receives LPI from the link partner, are implementation specific and beyond the scope of this standard.

78.1.2 LPI Client service interface

The following specifies the service interface provided by the RS to the LPI Client. These services are described in an abstract manner and do not imply any particular implementation.

The following primitives are defined:

- LP_IDLE.request
- LP_IDLE.indication

78.1.2.1 LP_IDLE.request

78.1.2.1.1 Function

A primitive used by the LPI Client to start or stop the signaling of LPI to the link partner.

78.1.2.1.2 Semantics of the service primitive

The semantics of the service primitive are as follows:

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 in any of the following cases:

- a) link_status is not OK (see 28.2.6.1.1)
- b) LPI_REQUEST=ASSERT within 1 s of the change of link_status to OK
- c) The PHY is indicating LOCAL FAULT
- d) The PHY is indicating REMOTE FAULT

78.1.2.1.3 When generated

Specification of the time when this primitive is generated by the LPI client is out of the scope of this standard.

78.1.2.1.4 Effect of receipt

The receipt of this primitive will cause the RS to start or stop signaling LPI to the link partner.

78.1.2.2 LP_IDLE.indication

78.1.2.2.1 Function

A primitive that is used to indicate to the LPI Client that the link partner has started or stopped signaling LPI.

78.1.2.2.2 Semantics of the service primitive

The semantics of the service primitive are as follows:

LP_IDLE.indication (LPI_INDICATION)

The LPI_INDICATION parameter can take one of two values: ASSERT or DE-ASSERT. ASSERT indicates that the link partner has started signaling LPI. DE-ASSERT indicates that the link partner has stopped signaling LPI.

78.1.2.2.3 When generated

This primitive is generated by the PHY when it receives an LPI signal or a wake signal from its link partner.

78.1.2.2.4 Effect of receipt

The effect of receipt of this primitive by the LPI client is unspecified.

78.1.3 Reconciliation sublayer operation

LPI assert and detect functions are contained in the Reconciliation Sublayer as shown in Figure 78–2. The xMII in this diagram represents any of the family of medium independent interfaces supported by IEEE.

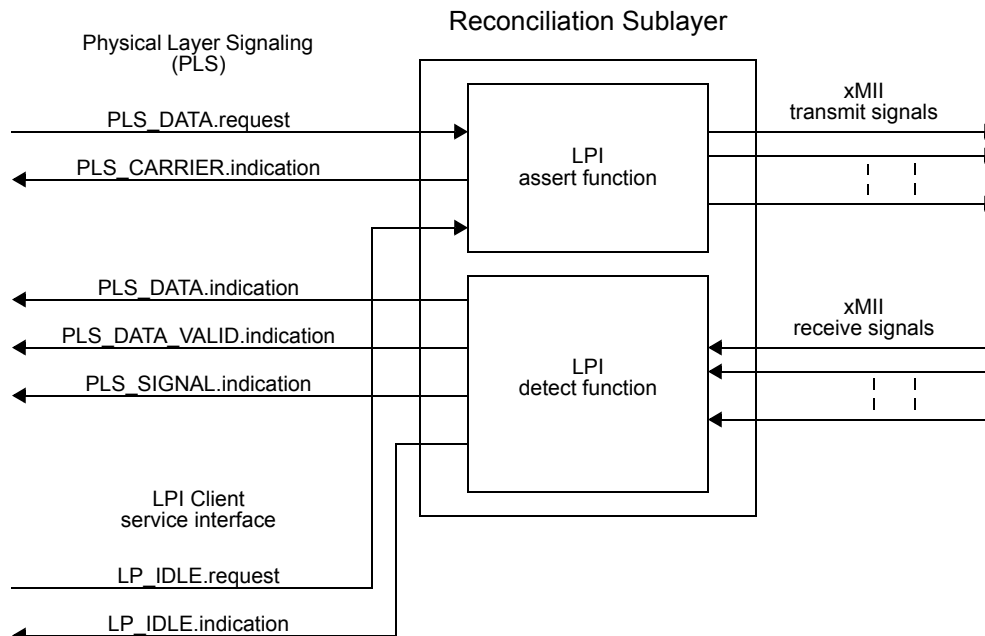


Figure 78–2—RS LPI assert and detect functions

The following provides an overview of RS LPI operation. The actual specification of RS LPI operation can be found in the respective RS clauses.

78.1.3.1 RS LPI assert function

In the absence of an LPI request, indicated by the LPI_REQUEST parameter set to DE-ASSERT in the LP_IDLE.request primitive of the LPI Client interface, the LPI assert function maps the PLS service interface to the transmit xMII signals as under normal conditions.

When an LPI request is asserted, indicated by the LPI_REQUEST parameter set to ASSERT in the LP_IDLE.request primitive of the LPI Client interface, the LPI assert function starts to transmit the “Assert LPI” encoding on the xMII. The LPI assert function also sets the CARRIER_STATUS parameter to CARRIER_ON in the PLS_CARRIER.indication primitive of the PLS service interface. This will prevent the MAC from transmitting.

When the LPI request is de-asserted, indicated by the LPI_REQUEST parameter set to DE-ASSERT in the LP_IDLE.request primitive of the LPI Client interface, the LPI assert function starts to transmit the normal interframe encoding on the xMII. After a delay, the LPI assert function sets the CARRIER_STATUS parameter to CARRIER_OFF in the PLS_CARRIER.indication primitive of the PLS service interface, allowing the MAC to start transmitting again. This delay is provided to allow the link partner to prepare for normal operation. This delay has a PHY dependent default value but this value can be adjusted using the Data Link Layer capabilities defined in 78.4.

78.1.3.2 LPI detect function

In the absence of LPI, indicated by an encoding other than “Assert LPI” on the receive xMII, the LPI detect function maps the receive xMII signals to the PLS service interface as under normal conditions.

At the start of LPI, indicated by the transition from normal interframe encoding to the “Assert LPI” encoding on the receive xMII, the LPI detect function continues to indicate idle on the PLS service interface, but sets LP_IDLE.indication(LPI_INDICATION) to ASSERT.

At the end of LPI, indicated by the transition from the “Assert LPI” encoding to any other encoding on the receive xMII, LP_IDLE.indication(LPI_INDICATION) is set to DE-ASSERT and the RS receive function resumes normal decode operation.

78.1.3.3 PHY LPI operation

The following provides an overview of PHY LPI operation. The specification of PHY LPI operation can be found in the respective PHY clauses (see Table 78–1).

78.1.3.3.1 PHY LPI transmit operation

When the start of “Assert LPI” encoding on the xMII is detected, the PHY signals sleep to its link partner to indicate that the local transmitter is entering LPI mode.

The EEE capability in most PHYs (for example, 100BASE-TX, 10GBASE-T, 1000BASE-KX, 10GBASE-KR, and 10GBASE-KX4) requires the local PHY transmitter to go quiet after sleep is signalled.

In the 1000BASE-T LPI mode, the local PHY transmitter goes quiet only after the local PHY signals sleep and receives a sleep signal from the remote PHY. If the remote PHY chooses not to signal LPI, then neither PHY can go into a low power mode; however, LPI requests are passed from one end of the link to the other regardless and system energy savings can be achieved even if the PHY link does not go into a low power mode.

The transmit function of the local PHY is enabled periodically to transmit refresh signals that are used by the link partner to update adaptive filters and timing circuits in order to maintain link integrity.

This quiet-refresh cycle continues until the reception of the normal interframe encoding on the xMII. The transmit function in the PHY communicates this to the link partner by sending a wake signal for a predefined period of time. The PHY then enters the normal operating state.

Figure 78–3 illustrates general principles of the EEE-capable transmitter operation.

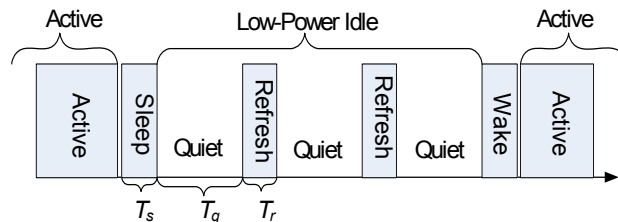


Figure 78–3—Overview of EEE LPI operation

No data frames are lost or corrupted during the transition to or from the LPI mode.

78.1.3.3.2 PHY LPI receive operation

In the receive direction, entering the LPI mode is triggered by the reception of a sleep signal from the link partner, which indicates that the link partner is about to enter the LPI mode. After sending the sleep signal, the link partner ceases transmission. When the receiver detects the sleep signal, the local PHY indicates “Assert LPI” on the xMII and the local receiver can disable some functionality to reduce power consumption.

The link partner periodically transmits refresh signals that are used by the local PHY to update adaptive coefficients and timing circuits. This quiet-refresh cycle continues until the link partner initiates transition back to normal mode by transmitting the wake signal for a predetermined period of time controlled by the LPI assert function in the RS. This allows the local receiver to prepare for normal operation and transition from the “Assert LPI” encoding to the normal interframe encoding on the xMII. After a system specified recovery time, the link supports the nominal operational data rate.

78.1.4 EEE Supported PHY types

EEE defines a low power mode of operation for the IEEE 802.3 PHYs and the XGXS listed in Table 78–1. The table also lists the clauses associated with each PHY or sublayer. Normative requirements for the EEE capability for each PHY type and for XGXS are in the associated clauses.

Table 78–1—Clauses associated with each interface 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

78.2 LPI mode timing parameters description

- T_s The period of time that the PHY transmits the sleep signal before turning all transmitters off
- T_q The period of time that the PHY remains quiet before sending the refresh signal
- T_r Duration of the refresh signal
- $T_{phy_prop_tx}$ The propagation delay of a given unit of data from the xMII to the MDI
- $T_{phy_prop_rx}$ The propagation delay of a given unit of data from the MDI to the xMII
- $T_{phy_shrink_tx}$ Transmitter shrinkage time, defined as the absolute time difference between the following two timing parameters:
 - Delay between a transition from the “Assert LPI” to “Normal Idle” at the xMII and the corresponding start of the wake signal at the MDI
 - $T_{phy_prop_tx}$
- $T_{phy_shrink_rx}$ Receiver shrinkage time, defined as the absolute time difference between the following two timing parameters:

—Delay between start of the wake signal at the MDI and the corresponding transition from “Assert LPI” to “Normal Idle” at the xMII

— $T_{\text{phy_prop_rx}}$

$T_{\text{w_phy}}$ Parameter employed by the system that corresponds to the behavior of the PHY. It is the period of time between reception of an IDLE signal on the xMII and when the first data code-words are permitted on the xMII. The wake time of a compliant PHY does not exceed $T_{\text{w_phy}}$ (min).

$T_{\text{w_sys_tx}}$ Parameter employed by the system that corresponds to its requirements. It is the longest period of time the system has to wait between a request to transmit and its readiness to transmit.

$T_{\text{w_sys_rx}}$ Parameter employed by the system that corresponds to its requirements. It is the minimum time required by the system between a request to wake and its readiness to receive data.

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

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

Protocol	T_s (μs)		T_q (μs)		T_r (μs)	
	Min	Max	Min	Max	Min	Max
100BASE-TX	200	220	20 000	22 000	200	220
1000BASE-T	182.0	202.0	20 000	24 000	198.0	218.2
1000BASE-KX	19.9	20.1	2 500	2 600	19.9	20.1
XGXS (XAUI)	19.9	20.1	2 500	2 600	19.9	20.1
10GBASE-KX4	19.9	20.1	2 500	2 600	19.9	20.1
10GBASE-KR	4.9	5.1	1 700	1 800	16.9	17.5
10GBASE-T	2.88	3.2	39.68	39.68	1.28	1.28

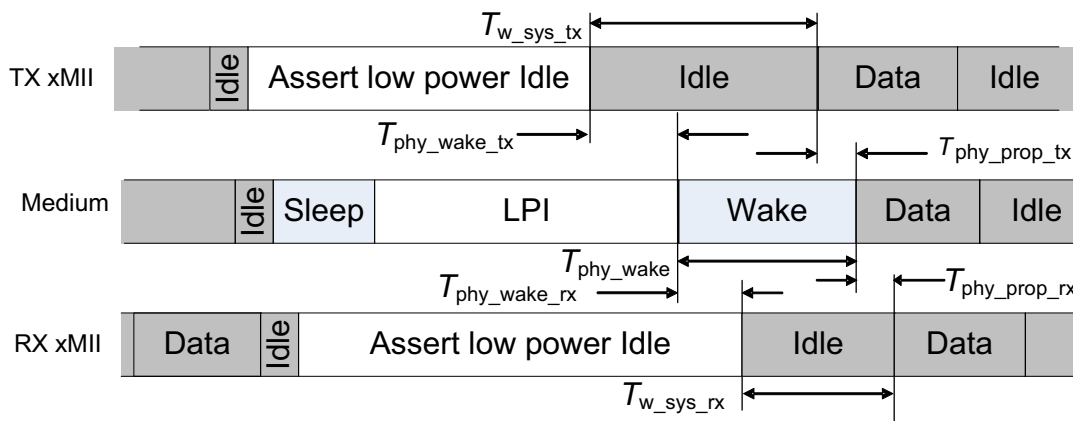
Figure 78–4 illustrates the relationship between the LPI mode timing parameters and the minimum system wake time.

78.3 Capabilities Negotiation

The EEE capability shall be advertised during the Auto-Negotiation stage. Auto-Negotiation provides a linked device with the capability to detect the abilities (modes of operation) supported by the device at the other end of the link, determine common abilities, and configure for joint operation. Auto-Negotiation is performed at power up, on command from management, due to link failure, or due to user intervention.

During Auto-Negotiation, both link partners indicate their EEE capabilities. EEE is supported only if during Auto-Negotiation both the local device and link partner advertise the EEE capability for the resolved PHY type. If EEE is not supported, all EEE functionality is disabled and the LPI client does not assert LPI. If EEE is supported by both link partners for the negotiated PHY type, then the EEE function can be used independently in either direction.

Additional capabilities and settings using L2 protocol frames, including the adjustment of the $T_{\text{w_sys_tx}}$ parameter, are described in 78.4.



$$\begin{aligned}
 T_{w_sys_tx}(\text{min}) &= T_{w_sys_rx}(\text{min}) + T_{phy_shrink_tx}(\text{max}) + T_{phy_shrink_rx}(\text{max}) \\
 T_{w_phy}(\text{min}) &= T_{phy_wake}(\text{min}) + T_{phy_shrink_tx} \\
 T_{w_sys_res}(\text{min}) &\text{ is greater of } T_{w_sys_tx}(\text{min}) \text{ and } T_{w_phy}(\text{min}) \\
 T_{phy_shrink_tx}(\text{max}) &= (T_{phy_wake_tx}(\text{max}) - T_{phy_prop_tx}(\text{min})) \\
 T_{phy_shrink_rx}(\text{max}) &= (T_{phy_wake_rx}(\text{max}) - T_{phy_prop_rx}(\text{min}))
 \end{aligned}$$

where

- $T_{phy_wake_tx}$ = xMII start of wake to MDI start of wake delay
- $T_{phy_prop_tx}$ = xMII to MDI data propagation delay
- $T_{phy_wake_rx}$ = MDI start of wake to xMII start of wake delay
- $T_{phy_prop_rx}$ = MDI to xMII data propagation delay
- T_{phy_wake} = Minimum wake duration required by PHY

Figure 78-4—LPI mode timing parameters and their relationship to minimum system wake time

78.4 Data Link Layer Capabilities

Additional capabilities and settings are supported using frames based on the IEEE 802.3 Organizationally Specific TLVs are defined in Annex F of IEEE Std 802.1AB™-2009. Devices that require longer wake-up times prior to being able to accept data on their receive paths may use the Data Link Layer capabilities defined in this subclause to negotiate for extended system wake-up times from the transmitting link partner. This mechanism may allow for more or less aggressive energy saving modes.

The Data Link Layer capabilities shall be implemented for devices with an operating speed equal to or greater than 10 Gb/s and may be implemented for all other devices.

Implementations that use the Data Link Layer capabilities shall comply with all mandatory parts of IEEE Std 802.1AB-2009; shall support the EEE Type, Length, Value (TLV) defined in 79.3.5; timing requirement in 78.4.1; and shall support the control state diagrams defined in 78.4.2.

The Data Link Layer capabilities are described from a unidirectional perspective on the link between transmitting and receiving link partners. For duplex EEE links that implement the Data Link Layer capabilities, each link partner shall implement the TLV, control and state diagrams for a transmitter as well as a receiver.

For purposes of Data Link Layer capabilities, all values that are negotiated and/or exchanged that have a fractional value shall be rounded up to the nearest integer number in microseconds.

78.4.1 Data Link Layer capabilities timing requirements

An EEE link partner shall send an LLDPDU containing an EEE TLV within 10 s of the Link Layer capability exchange being enabled when both the variables `dll_enabled` and `dll_ready` are asserted.

An LLDPDU containing an EEE TLV with an updated value for the “Echo Transmit $T_{w_sys_tx}$ ” field shall be sent within 10 s of receipt of an LLDPDU containing an EEE TLV where the value of “Transmit $T_{w_sys_tx}$ ” field is different from the previously communicated value.

An LLDPDU containing an EEE TLV with an updated value for the “Echo Receive $T_{w_sys_tx}$ ” field shall be sent within 10 s of receipt of an LLDPDU containing an EEE TLV where the value of “Receive $T_{w_sys_tx}$ ” field is different from the previously communicated value.

78.4.2 Control state diagrams

The control state diagrams for an EEE transmitting link partner and an EEE receiving link partner specify the externally observable behavior of an EEE transmitting link partner and an EEE receiving link partner implementing Data Link Layer capabilities respectively. EEE transmitting link partners implementing Data Link Layer capabilities shall provide the behavior of the state diagram as shown in Figure 78–5. EEE receiving link partners implementing Data Link Layer capabilities shall provide the behavior of the state diagram as shown in Figure 78–6.

78.4.2.1 Conventions

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

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

78.4.2.2 Constants

PHY WAKE VALUE

Integer (2 octets wide) representing the $T_{w_sys_tx}$ (min) defined for the PHY that is in use for the link. This parameter should be rounded up to the nearest integer number when it is calculated and examined according to 78.2 and Table 78–4.

78.4.2.3 Variables

Unless otherwise specified, all integers are assumed to be 2 octets wide.

LocTxSystemValue

Integer that indicates the value of $T_{w_sys_tx}$ that the local system can support. This value is updated by the EEE DLL Transmitter state diagram. This variable maps into the `aLldpXdot3LocTxTwSys` attribute.

RemTxSystemValueEcho

Integer that indicates the value Transmit $T_{w_sys_tx}$ echoed back by the remote system. This value maps from the `aLldpXdot3RemTxTwSysEcho` attribute.

LocRxSystemValue

Integer that indicates the value of $T_{w_sys_tx}$ that the local system requests from the remote system. This value is updated by the EEE Receiver L2 state diagram. This variable maps into the aLldpXdot3LocRxTwSys attribute.

RemRxSystemValueEcho

Integer that indicates the value of Receive $T_{w_sys_tx}$ echoed back by the remote system. This value maps from the aLldpXdot3RemRxTwSysEcho attribute.

LocFbSystemValue

Integer that indicates the value of fallback $T_{w_sys_tx}$ that the local system requests from the remote system. This value is updated by the local system software.

RemTxSystemValue

Integer that indicates the value of $T_{w_sys_tx}$ that the remote system can support. This value maps from the aLldpXdot3RemTxTwSys attribute.

LocTxSystemValueEcho

Integer that indicates the remote system's Transmit $T_{w_sys_tx}$ that was used by the local system to compute the $T_{w_sys_tx}$ that it wants to request from the remote system. This value maps into the aLldpXdot3LocTxTwSysEcho attribute.

RemRxSystemValue

Integer that indicates the value of $T_{w_sys_tx}$ that the remote system requests from the local system. This value maps from the aLldpXdot3RemRxTwSys attribute.

LocRxSystemValueEcho

Integer that indicates the remote systems Receive $T_{w_sys_tx}$ that was used by the local system to compute the $T_{w_sys_tx}$ that it can support. This value maps into the aLldpXdot3LocRxTwSysEcho attribute.

LocResolvedTxSystemValue

Integer that indicates the current $T_{w_sys_tx}$ supported by the local system.

LocResolvedRxSystemValue

Integer that indicates the current $T_{w_sys_tx}$ supported by the remote system.

TempTxVar

Integer used to store the value of $T_{w_sys_tx}$.

TempRxVar

Integer used to store the value of $T_{w_sys_tx}$.

local_system_change

An implementation specific control variable that indicates that the local system wants to change either the Transmit $T_{w_sys_tx}$ or the Receive $T_{w_sys_tx}$.

tx_dll_ready

Data Link Layer ready: This variable indicates that the tx system initialization is complete and is ready to update/receive LLDPDU containing EEE TLV. This variable is updated by the local system software.

rx_dll_ready

Data Link Layer ready: This variable indicates that the rx system initialization is complete and is ready to update/receive LLDPDU containing EEE TLV. This variable is updated by the local system software.

NEW_TX_VALUE

Integer that indicates the value of $T_{w_sys_tx}$ that the local system can support.

NEW_RX_VALUE

Integer that indicates the value of $T_{w_sys_tx}$ that the local system wants the remote system to support.

A summary of cross-references between the EEE object class attributes and the transmit and receive control state diagrams, including the direction of the mapping, is provided in Table 78–3.

Table 78–3—Attribute to state diagram variable cross-reference

Entity	Object class	Attribute	Mapping	State diagram variable
TX	oLldpXdot3Loc-SystemsGroup	aLldpXdot3LocTxTwSys	←	LocTxSystemValue
		aLldpXdot3LocRxTwSysEcho	←	LocRxSystemValueEcho
		aLldpXdot3LocDllEnabled	⇒	tx_dll_enabled
		aLldpXdot3LocTxDllReady	←	tx_dll_ready
	oLldpXdot3Rem-SystemsGroup	aLldpXdot3RemRxTwSys	⇒	RemRxSystemValue
		aLldpXdot3RemTxTwSysEcho	⇒	RemTxSystemValueEcho
RX	oLldpXdot3Loc-SystemsGroup	aLldpXdot3LocRxTwSys	←	LocRxSystemValue
		aLldpXdot3LocTxTwSysEcho	←	LocTxSystemValueEcho
		aLldpXdot3LocFbTwSys	←	LocFbSystemValue
		aLldpXdot3LocDllEnabled	⇒	rx_dll_enabled
		aLldpXdot3LocRxDllReady	←	rx_dll_ready
	oLldpXdot3Rem-SystemsGroup	aLldpXdot3RemTxTwSys	⇒	RemTxSystemValue
		aLldpXdot3RemRxTwSysEcho	←	RemRxSystemValueEcho

78.4.2.4 Functions

examine_Tx_change

This function computes the new value of $T_{w_sys_tx}$ that the local system can support when there is an updated request from the remote system or if local system conditions require a change in the value of the presently supported $T_{w_sys_tx}$.

examine_Rx_change

This function computes the new value of $T_{w_sys_tx}$ that the local system wants the remote system to support. This function is called when the remote system wants to change its presently allocated $T_{w_sys_tx}$ or if local system conditions require a change in the value of $T_{w_sys_tx}$ presently supported by the remote system.

78.4.2.5 State diagrams

Control for placing data on the medium rests with the transmitting side, hence $T_{w_sys_tx}$ is enforced by the transmitter. For a given path between link partners (i.e., a transmitter and its associated receiver), the transmitting link partner shall wait for the time indicated by the Transmit $T_{w_sys_tx}$ after de-asserting LPI (at the xMII) before sending data frames. The receiving link partner shall be ready to accept data based on its echoed value of Transmit link partner's $T_{w_sys_tx}$. This ensures that the link partners transition out of LPI mode and receive frames without loss or corruption.

The general state change procedure for transmitter is shown in Figure 78–5.

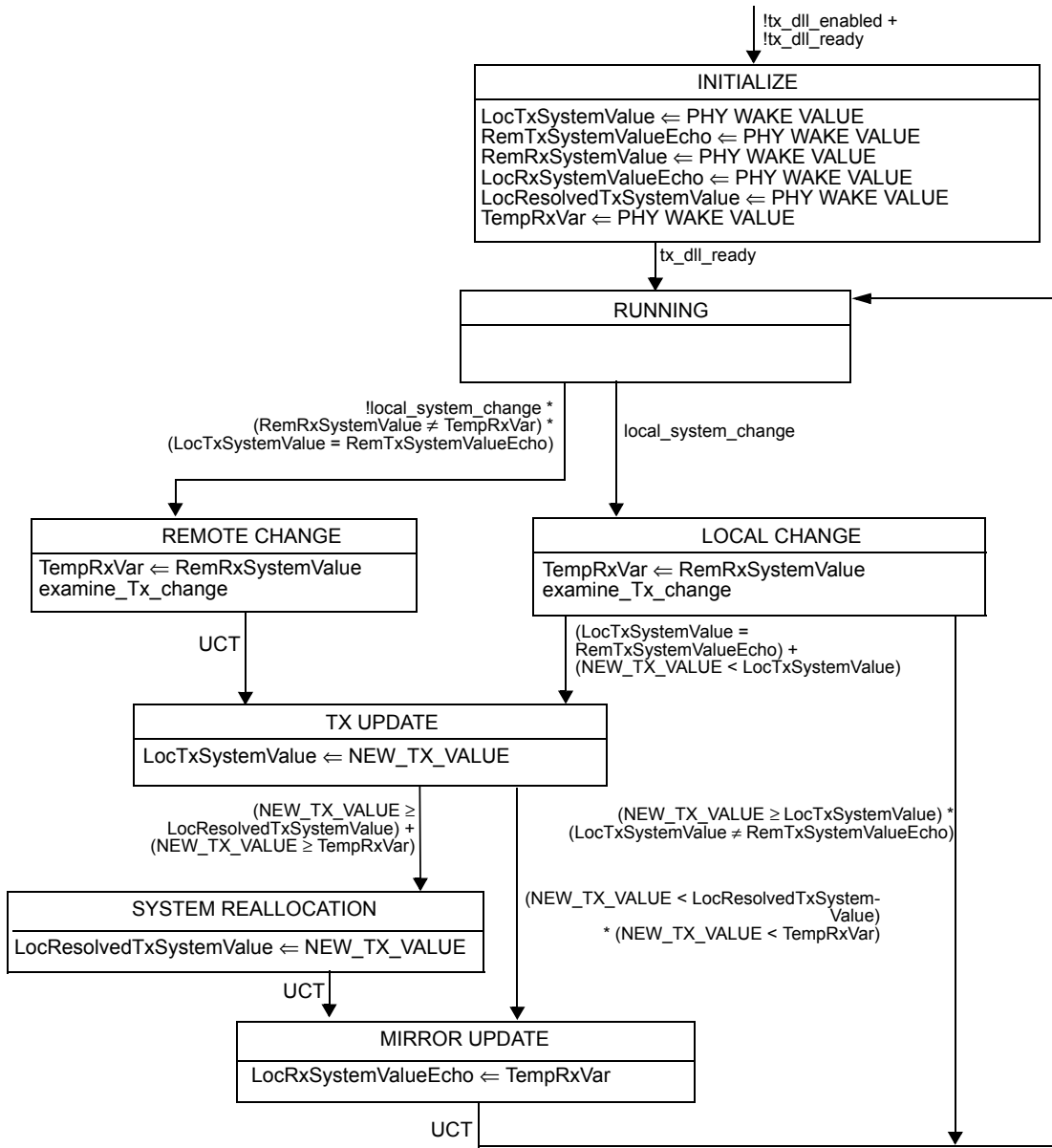


Figure 78-5—EEE DLL Transmitter state diagram

The general state change procedure for receiver is shown in Figure 78–6.

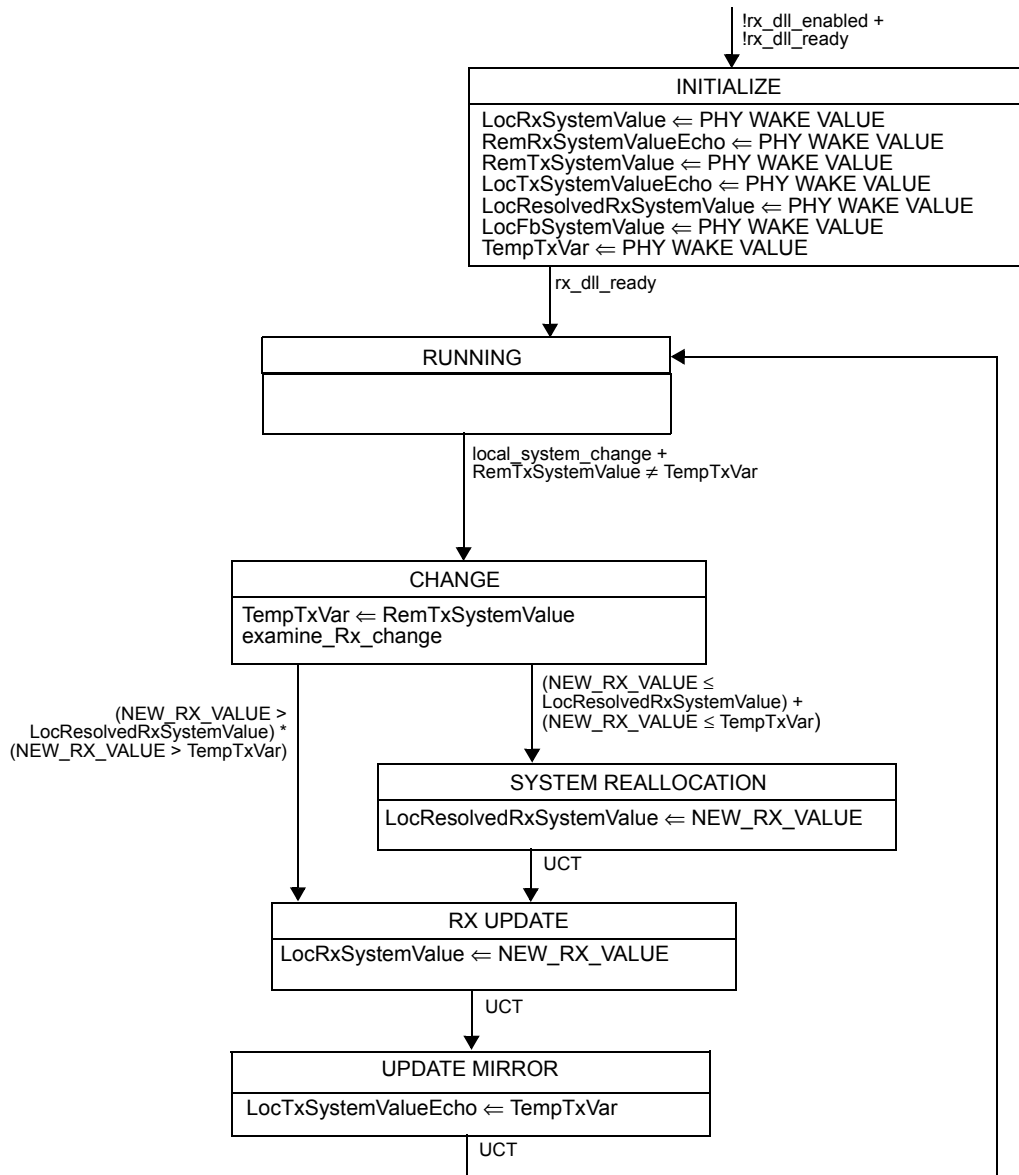


Figure 78–6—EEE DLL Receiver state diagram

78.4.3 State change procedure across a link

The transmitting and receiving link partners utilize the LLDP mechanism to advertise their various attributes to the other entity.

The initial $T_{w_sys_tx}$ defaults governing the EEE operation of the link default to the wake values required by the PHYs. This provides for EEE operation and functionality on initialization and prior to the exchange and processing of the TLVs.

The receiving link partner may request a new $T_{w_sys_tx}$ value through the `aLldpXdot3LocRxTwSys` (30.12.2.1.24) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2). The request appears to the transmitting link partner as a change to the `aLldpXdot3RemRxTwSys` (30.12.3.1.21) attribute in the `LldpXdot3RemSystemsGroup` managed (30.12.3) object class. The transmitting link partner responds to its receiving partner's request through the `aLldpXdot3LocTxTwSys` (30.12.2.1.22) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2). The transmitting link partner also copies the value of the `aLldpXdot3RemRxTwSys` (30.12.3.1.21) attribute in the `LldpXdot3RemSystemsGroup` managed (30.12.3) object class to the `aLldpXdot3LocRxTwSysEcho` (30.12.2.1.25) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2).

The transmitting link partner may advertise new value of $T_{w_sys_tx}$ through the `aLldpXdot3LocTxTwSys` (30.12.2.1.22) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2). This appears to the receiving link partner as a change to the `aLldpXdot3RemTxTwSys` (30.12.3.1.19) attribute in the `LldpXdot3RemSystemsGroup` managed (30.12.3) object class. The receiving link partner responds to a transmitter's request through the `aLldpXdot3LocRxTwSys` (30.12.2.1.24) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2). The receiving link partner also copies the value of the `aLldpXdot3RemTxTwSys` (30.12.3.1.19) attribute in the `LldpXdot3RemSystemsGroup` managed (30.12.3) object class to the `aLldpXdot3LocTxTwSysEcho` (30.12.2.1.23) attribute in the `LldpXdot3LocSystemsGroup` managed object class (30.12.2). This appears to the transmitting link partner as a change to the `aLldpXdot3RemTxTwSysEcho` (30.12.3.1.20) attribute in the `LldpXdot3RemSystemsGroup` managed (30.12.3).

The state diagrams in Figure 78–5 and Figure 78–6 describe the preceding behavior.

78.4.3.1 Transmitting link partner's state change procedure across a link

A transmitting link partner is said to be in sync with the receiving link partner if the presently advertised value of Transmit $T_{w_sys_tx}$ and the corresponding echoed value are equal.

During normal operation, the transmitting link partner is in the RUNNING state. If the transmitting link partner wants to initiate a change to the presently resolved value of $T_{w_sys_tx}$, the `local_system_change` is asserted and the transmitting link partner enters the LOCAL CHANGE state where `NEW_TX_VALUE` is computed. If the new value is smaller than the presently advertised value of $T_{w_sys_tx}$ or if the transmitting link partner is in sync with the receiving link partner, then it enters TX UPDATE state. Otherwise, it returns to the RUNNING state.

If the transmitting link partner sees a change in the $T_{w_sys_tx}$ requested by the receiving link partner, it recognizes the request only if it is in sync with the transmitting link partner. The transmitting link partner examines the request by entering the REMOTE CHANGE state where a `NEW_TX_VALUE` is computed and it then enters the TX UPDATE state.

Upon entering the TX UPDATE state, the transmitter updates the advertised value of Transmit $T_{w_sys_tx}$ with `NEW_TX_VALUE`. If the `NEW_TX_VALUE` is equal to or greater than either the resolved $T_{w_sys_tx}$ value or the value requested by the receiving link partner then it enters the SYSTEM REALLOCATION state where it updates the value of resolved $T_{w_sys_tx}$ with `NEW_TX_VALUE`. The transmitting link partner enters the MIRROR UPDATE state either from the SYSTEM REALLOCATION state or directly from the TX UPDATE state. The UPDATE MIRROR state then updates the echo for the Receive $T_{w_sys_tx}$ and returns to the RUNNING state.

78.4.3.2 Receiving link partner's state change procedure across a link

A receiving link partner is said to be in sync with the transmitting link partner if the presently requested value of Receive $T_{w_sys_tx}$ and the corresponding echoed value are equal.

During normal operation, the receiving link partner is in the RUNNING state. If the receiving link partner wants to request a change to the presently resolved value of $T_{w_sys_tx}$, the `local_system_change` is asserted. When `local_system_change` is asserted or when the receiving link partner sees a change in the $T_{w_sys_tx}$ advertised by the transmitting link partner, it enters the CHANGE state where `NEW_RX_VALUE` is computed. If `NEW_RX_VALUE` is less than either the presently resolved value of $T_{w_sys_tx}$ or the presently advertised value by the transmitting link partner, it enters the SYSTEM REALLOCATION state where it updates the resolved value of $T_{w_sys_tx}$ to `NEW_RX_VALUE`. The receiving link partner ultimately enters the RX UPDATE state, either from the SYSTEM REALLOCATION state or directly from the CHANGE state.

In the RX UPDATE state, it updates the presently requested value to `NEW_RX_VALUE`, then it updates the echo for the Transmit $T_{w_sys_tx}$ in the UPDATE MIRROR state and finally goes back to the RUNNING state.

78.5 Communication link access latency

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

The EEE capability adds latency that has to be considered by the network designer. When in the LPI mode, the PHY link is not available immediately for transmission of data. The system has to wake it up by sending the normal IDLE code on the MAC interface. Following the reception of an IDLE code on the MAC interface, the PHY starts the wake-up process. The maximal PHY recovery time, T_{w_phy} is defined for each PHY.

Transmit and/or Receive wait time shrinkage can happen when $T_{phy_shrink_rx}$ or $T_{phy_shrink_tx}$ (as defined in 78.1.3) are not zero. This has to be taken into consideration in designing or configuring the network.

Table 78–4 summarizes critical timing parameters for supported PHYs. These are listed here to assist the system designer in assessing the impact of EEE on the operation of the link.

Case-1 of the 1000BASE-T PHY applies to PHYs in Master mode. Case-2 of the 1000BASE-T PHY applies to PHYs in Slave mode.

Case-1 of the 10GBASE-KR PHY applies to PHYs without FEC. Case-2 of the 10GBASE-KR PHY applies to PHYs with FEC.

Case-1 of the 10GBASE-T PHY applies when the PHY is requested to transmit the Wake signal before transmission of the Sleep signal to the Link Partner is complete. Case-2 of the 10GBASE-T PHY applies when the PHY is requested to transmit the Wake signal after transmission of the Sleep signal to the Link Partner is complete and if the PHY has not indicated LOCAL FAULT at any time during the previous 10 ms.

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

PHY Type	Case	$T_{w_sys_tx}$ (min) (μ s)	T_{w_phy} (min) (μ s)	$T_{phy_shrink_tx}$ (max) (μ s)	$T_{phy_shrink_rx}$ (max) (μ s)	$T_{w_sys_rx}$ (min) (μ s)
100BASE-TX		30	20.5	5.0	15	10
1000BASE-T	Case-1	16.5	16.5	5.0	2.5	1.76
	Case-2	16.5	16.5	12.24	9.74	1.76
1000BASE-KX		13.26	11.25	5.0	6.5	1.76
XGXS (XAUI)		12.38	9.25	5.0	4.5	2.88
10GBASE-T	Case-1	7.36	7.36	4.48	0	2.88
	Case-2	4.48	4.48	1.6	0	2.88
10GBASE-KX4		12.38	9.25	5.0	4.5	2.88
10GBASE-KR	Case-1	15.38	12.25	5.0	7.5	2.88
	Case-2	17.38	14.25	5.0	9.5	2.88

78.5.1 10 Gb/s PHY extension using XGXS

The XGXS can be inserted between the RS and a 10 Gb/s PHY to transparently extend the physical reach of the XGMII. The LPI signaling can operate through the XGXS 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 DTE XS XAUI stop enable bit (5.0.9) is asserted, the DTE XS may stop signaling on the XAUI in the transmit direction to conserve energy. If the DTE XS XAUI stop enable bit is asserted, the RS defers sending data following deassertion of LPI by an additional time equal to $T_{w_sys_tx} - T_{w_sys_rx}$ for the XGXS as shown in Table 78–4 (see 46.4.2.1).

If the PHY XS XAUI stop enable bit (4.0.9) is asserted, the PHY XS may stop signaling on the XAUI in the receive direction to conserve energy. The receiver negotiates an additional time for the remote T_{w_sys} equal to $T_{w_sys_tx} - T_{w_sys_rx}$ for the XGXS as shown in Table 78–4 before setting the PHY XS XAUI stop enable bit.

78.6 Protocol implementation conformance statement (PICS) proforma for EEE Data Link Layer Capabilities¹

78.6.1 Introduction

The supplier of a protocol implementation that is claimed to conform to 78.4 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.

¹Copyright 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.

78.6.2 Identification

78.6.2.1 Implementation identification

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

78.6.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, 78.4, EEE Data Link Layer Capabilities
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 <input type="checkbox"/> Yes <input type="checkbox"/> (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3-2012.	
Date of Statement	

78.6.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
10G	Support 10G or higher operation	78.4	Support for 10 Gb/s or higher operation	O	Yes <input type="checkbox"/> No <input type="checkbox"/>
DLL1	DLL	78.4	DLL	10G:M	Yes <input type="checkbox"/> N/A <input type="checkbox"/>
DLL2	DLL	78.4	DLL	!10G:O	Yes <input type="checkbox"/> No <input type="checkbox"/>

In addition, the following predicate name is defined for use when different implementations from the preceding set have common parameters:

$$\text{DLL} = \text{DLL1 OR DLL 2}$$

78.6.4 DLL requirements

Item	Feature	Subclause	Value/Comment	Status	Support
DLR1	DLL Timing	78.4.1	Timing requirements	DLL:M	Yes [] N/A []
DLR2	DLL Control state diagrams	78.4.2	State machines for TX and RX	DLL:M	Yes [] N/A []

79. IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements

79.1 Overview

The Link Layer Discovery Protocol (LLDP) specified in IEEE Std 802.1AB-2009 is a MAC Client protocol that allows stations attached to an IEEE 802 LAN to advertise to all other stations attached to the same IEEE 802 LAN: the major capabilities provided by the system incorporating that station, the management address or addresses of the entity or entities that provide management of those capabilities, and the identification of the station’s point of attachment to the IEEE 802 LAN.

The information fields in each LLDP frame are contained in a Link Layer Discovery Protocol Data Unit (LLDPDU) as a sequence of short, variable length, information elements known as TLVs that each include a type, length, and value field.

Organizationally Specific TLVs can be defined by either the professional organizations or the individual vendors that are involved with the particular functionality being implemented within a system. The basic format and procedures for defining Organizationally Specific TLVs are provided in subclause 9.6 of IEEE Std 802.1AB-2009.

79.1.1 IEEE 802.3 LLDP frame format

The IEEE 802.3 LLDP frame format is illustrated in Figure 79–1.

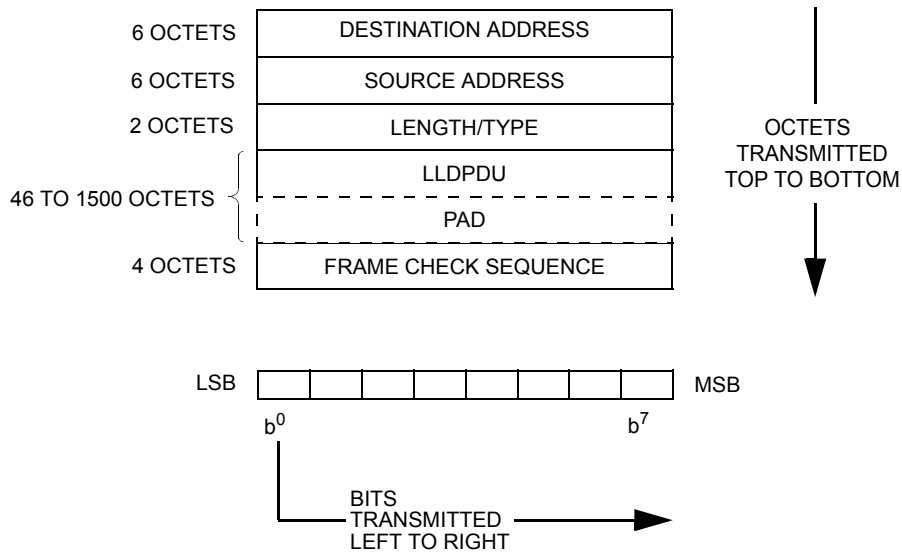


Figure 79–1—IEEE 802.3 LLDP frame format

NOTE—The illustration shows the simplest form of an IEEE 802.3 LLDP frame; i.e., where the frame has no IEEE Std 802.1Q™ tag header, or IEEE Std 802.1AE™ security tag, or any other form of encapsulation applied to it.²

²NOTES in text, tables, and figures are given for information only and do not contain requirements needed to implement the standard.

79.1.1.1 Destination Address field

The Destination Address field of an IEEE 802.3 LLDP frame contains a MAC address specified by 7.1 of IEEE Std 802.1AB-2009 (see 79.2).

79.1.1.2 Source Address field

The Source Address field of an IEEE 802.3 LLDP frame contains the 48-bit individual address of the station sending the frame.

79.1.1.3 Length/Type field

The Length/Type field of an IEEE 802.3 LLDP frame is a 2-octet field that contains the hexadecimal value: 88-CC. This value carries the Type interpretation (see 3.2.6), and has been universally assigned for LLDP.

79.1.1.4 LLDPDU field

The LLDPDU field of an IEEE 802.3 LLDP frame contains the LLDPDU which is a sequence of short, variable length, information elements known as TLVs that each include type, length, and value fields.

79.1.1.5 Pad field

A minimum MAC frame size is required for correct operation and, if necessary, a Pad field is appended after the LLDPDU field as defined in 3.2.8.

79.1.1.6 Frame Check Sequence field

The Frame Check Sequence (FCS) field contains the Frame Check Sequence, as defined in 3.2.9.

79.2 Requirements of the IEEE 802.3 Organizationally Specific TLV set

All IEEE 802.3 Organizationally Specific TLVs shall conform to the LLDPDU bit and octet ordering conventions of 8.1 of IEEE Std 802.1AB-2009.

79.3 IEEE 802.3 Organizationally Specific TLVs

The currently defined IEEE 802.3 Organizationally Specific TLVs are listed in Table 79–1. Any additions or changes to these TLVs will be included in this clause.

Table 79–1—IEEE 802.3 Organizationally Specific TLVs

IEEE 802.3 subtype	TLV name	Subclause reference
1	MAC/PHY Configuration/Status	79.3.1
2	Power Via Medium Dependent Interface (MDI)	79.3.2
3	Link Aggregation (deprecated)	79.3.3
4	Maximum Frame Size	79.3.4
5	Energy-Efficient Ethernet	79.3.5
6–255	<i>Reserved</i>	—

79.3.1 MAC/PHY Configuration/Status TLV

The MAC/PHY Configuration/Status TLV is an optional TLV that identifies the following:

- a) The duplex and bit-rate capability of the sending IEEE 802.3 LAN node that is connected to the physical medium.
- b) The current duplex and bit-rate settings of the sending IEEE 802.3 LAN node.
- c) Whether the current duplex and bit-rate settings are the result of auto-negotiation during link initiation or of manual set override action.

Figure 79–2 shows the format of this TLV.

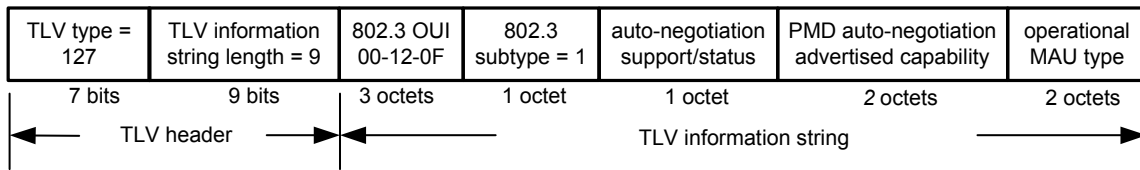


Figure 79–2—MAC/PHY configuration/status TLV format

79.3.1.1 Auto-negotiation support/status

The auto-negotiation support/status field shall contain a bitmap that identifies the auto-negotiation support and current status of the local IEEE 802.3 LAN station as defined in Table 79–2. If the auto-negotiation support bit (bit 0) is one and the auto-negotiation status bit (bit 1) is zero, the IEEE 802.3 physical media dependent sublayer (PMD) operating mode is determined by the operational Medium Attachment Unit (MAU) type field value rather than by auto-negotiation.

Table 79–2—IEEE 802.3 auto-negotiation support/status

Bit	Function	Value/meaning	IETF RFC 4836 reference
0	Auto-negotiation support	1 = supported 0 = not supported	ifMauAutoNegSupported
1	Auto-negotiation status	1 = enabled 0 = not enabled	ifMauAutoNegAdminStatus
2–7	—	Reserved for future standardization	—

79.3.1.2 PMD auto-negotiation advertised capability field

The ‘PMD auto-negotiation capability’ field shall contain a 2-octet value that provides a bitmap of the ifMauAutoNegCapAdvertisedBits object, defined in IETF RFC 4836, of the sending device. Bit zero is the high order (left-most) bit in an octet string.

79.3.1.3 Operational MAU type

The operational MAU type field contains an integer value indicating the MAU type of the sending device. This value shall be derived from the list position of the corresponding dot3MauType as listed in IETF RFC 4836 (or subsequent revisions) and is equal to the last number in the respective dot3MauType Object

Identifier (OID). For example, if the ifMauType object is dot3MauType1000BaseTHD which corresponds to {dot3MauType 29}, the numerical value of this field is 29. For MAU types not listed in IETF RFC 4836 (or subsequent revisions), the value of this field shall be set to zero.

79.3.1.4 MAC/PHY Configuration/Status TLV usage rules

An LLDPDU should contain no more than one MAC/PHY Configuration/Status TLV.

79.3.2 Power Via MDI TLV

Clause 33 defines two option power entities: a Powered Device (PD) and Power Sourcing Equipment (PSE). These entities allow devices to draw/supply power over the sample generic cabling as used for data transmission. The Power Via MDI TLV allows network management to advertise and discover the MDI power support capabilities of the sending IEEE 802.3 LAN station. This TLV is also required to perform Data Link Layer classification as defined in 33.6. Figure 79–3 shows the format of this TLV.

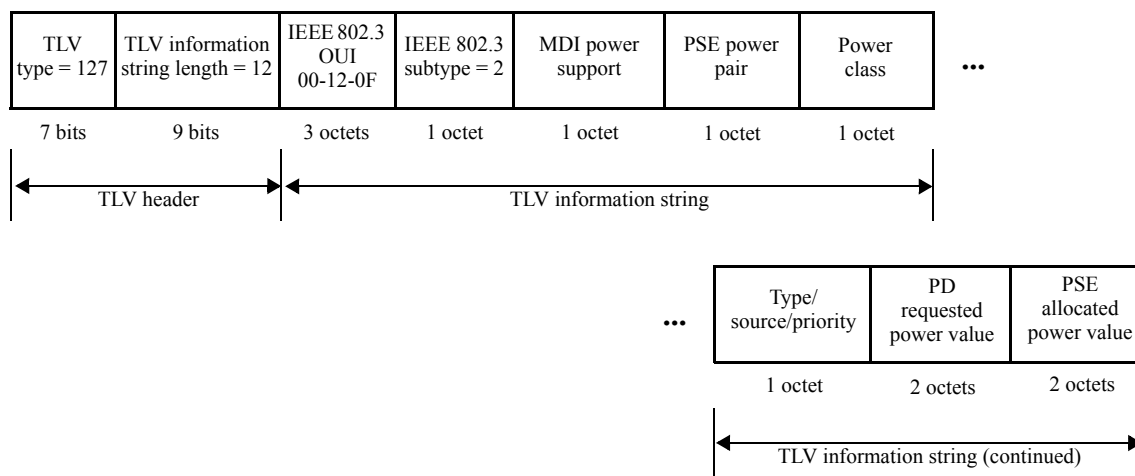


Figure 79–3—Power Via MDI TLV format

The TLV shown in Figure 79–3 is a revision of the legacy Power via MDI TLV originally defined in IEEE Std 802.1AB-2009 Annex F.3. The legacy TLV had only the first three fields of the TLV shown in the figure. These three fields enable discovery and advertisement of MDI power support capabilities. The newly added fields provide Data Link Layer classification capabilities. The revised TLV can be used by the PSE only when it is supplying power to a PI encompassed within an MDI and by the PD only when it is drawing power from the PI. Power entities may continue to use the legacy TLV prior to supplying/drawing power to/from the PI. If the power entity implements Data Link Layer classification, it shall use the Power via MDI TLV shown in Figure 79–3 after the PI has been powered.

79.3.2.1 MDI power support

The MDI power support field shall contain a bitmap of the MDI power capabilities and status as defined in Table 79–3.

NOTE 1—Port class information is implied by the support of the PSE or PD groups.

NOTE 2—MDI power support information is implied by support of IETF RFC 3621.

NOTE 3—If bit 1 is zero, bit 2 has no meaning.

Table 79–3—MDI power capabilities/status

Bit	Function	Value/meaning	IETF RFC 3621 object reference
0	Port class	1 = PSE 0 = PD	See Note 1
1	Power Sourcing Equipment (PSE) MDI power support	1 = supported 0 = not supported	See Note 2 and Note 3
2	PSE MDI power state	1 = enabled 0 = disabled	pethPsePortAdminEnable
3	PSE pairs control ability	1 = pair selection can be controlled 0 = pair selection can not be controlled	pethPsePortPowerPairContolAbility
4–7	Reserved for future standardization	—	—

79.3.2.2 PSE power pair

The PSE power pair field shall contain an integer value as defined by the pethPsePortPowerPairs object in IETF RFC 3621.

79.3.2.3 Power class

The power class field shall contain an integer value as defined by the pethPsePortPowerClassifications object in IETF RFC 3621.

79.3.2.4 Requested power type/source/priority

The power type/source/priority field shall contain a bit-map of the power type, source and priority defined in Table 79–4 and is reported for the device generating the TLV.

Table 79–4—Power type/source/priority field

Bit	Function	Value/meaning
7:6	power type	$\begin{array}{cc} 7 & 6 \\ 1 & 1 = \text{Type 1 PD} \\ 1 & 0 = \text{Type 1 PSE} \\ 0 & 1 = \text{Type 2 PD} \\ 0 & 0 = \text{Type 2 PSE} \end{array}$
5:4	power source	Where power type = PD $\begin{array}{cc} 5 & 4 \\ 1 & 1 = \text{PSE and local} \\ 1 & 0 = \text{Reserved} \\ 0 & 1 = \text{PSE} \\ 0 & 0 = \text{Unknown} \end{array}$ Where power type = PSE $\begin{array}{cc} 5 & 4 \\ 1 & 1 = \text{Reserved} \\ 1 & 0 = \text{Backup source} \\ 0 & 1 = \text{Primary power source} \\ 0 & 0 = \text{Unknown} \end{array}$

Table 79–4—Power type/source/priority field (continued)

Bit	Function	Value/meaning															
3:2	Reserved	Transmit as zero, ignore on receive															
1:0	power priority	<table border="0"> <tr> <td><u>1</u></td> <td><u>0</u></td> <td></td> </tr> <tr> <td>1</td> <td>1</td> <td>= low</td> </tr> <tr> <td>1</td> <td>0</td> <td>= high</td> </tr> <tr> <td>0</td> <td>1</td> <td>= critical</td> </tr> <tr> <td>0</td> <td>0</td> <td>= unknown (default)</td> </tr> </table>	<u>1</u>	<u>0</u>		1	1	= low	1	0	= high	0	1	= critical	0	0	= unknown (default)
<u>1</u>	<u>0</u>																
1	1	= low															
1	0	= high															
0	1	= critical															
0	0	= unknown (default)															

79.3.2.4.1 Power type

This field shall be set according to Table 79–4.

79.3.2.4.2 Power source

When the power type is PD, this field shall be set to 01 when the PD is being powered only through the PI; to 11 when the PD is being powered from both; and to 00 when this information is not available.

When the power type is PSE, this field shall be set to 01 when the PSE is sourcing its power through the PI from its primary supply; to 10 when the PSE is sourcing its power through the PI from a backup source; and to 00 when this information is not available.

79.3.2.4.3 Power priority

When the power type is PD, this field shall be set to the power priority configured for the device. If a PD is unable to determine its power priority or it has not been configured, then this field shall be set to 00.

When the power type is PSE, this field reflects the PD priority that the PSE advertises to assign to the PD.

79.3.2.5 PD requested power value

The PD requested power value field shall contain the PD’s requested power value defined in Table 79–5.

Table 79–5—PD requested power value field

Bit	Function	Value/meaning
15:0	PD requested power value	Power = 0.1 × (decimal value of bits) Watts. Valid values for these bits are decimal 1 through 255.

The PD requested power value is encoded according to Equation (79–1).

$$Power = \{0.1 \times X\}_W \quad (79-1)$$

where

Power is the effective requested PD power value

X is the decimal value of the power value field, bits 15:0

“PD requested power value” is the maximum input average power (see 33.3.7.2) the PD wants to draw. “PD requested power value” is the power value at the input to the PD’s PI.

79.3.2.6 PSE allocated power value

The PSE allocated power value field shall contain the PSE’s allocated power value defined in Table 79–6.

Table 79–6—PSE allocated power value field

Bit	Function	Value/meaning
15:0	PSE allocated power value	Power = 0.1 × (decimal value of bits) Watts. Valid values for these bits are decimal 1 through 255.

The PSE allocated power value is encoded according to Equation (79–2).

$$Power = \{0.1 \times X\}_W \tag{79-2}$$

where

Power is the effective allocated PSE power value

X is the decimal value of the power value field, bits 15:0

“PSE allocated power value” is the maximum input average power (see 33.3.7.2) the PSE expects the PD to draw. “PSE allocated power value” is the power at the input to the PD’s PI. The PSE uses this value to compute P_{Class} as defined in 33.2.6.

79.3.2.7 Power Via MDI TLV usage rules

An LLDPDU should contain no more than one Power Via MDI TLV.

79.3.3 Link Aggregation TLV (deprecated)

The Link Aggregation TLV is an optional TLV that indicates whether the link is capable of being aggregated, whether the link is currently in an aggregation, and if in an aggregation, the port identification of the aggregation. Figure 79–4 shows the format for this TLV.

NOTE—As the Link Aggregation specification has now been removed from IEEE Std 802.3 and is now standardized as IEEE Std 802.1AX, new implementations of this standard are encouraged to make use of the Link Aggregation TLV that is now part of the IEEE 802.1 extension MIB specified in Annex E of IEEE Std 802.1AB-2009.

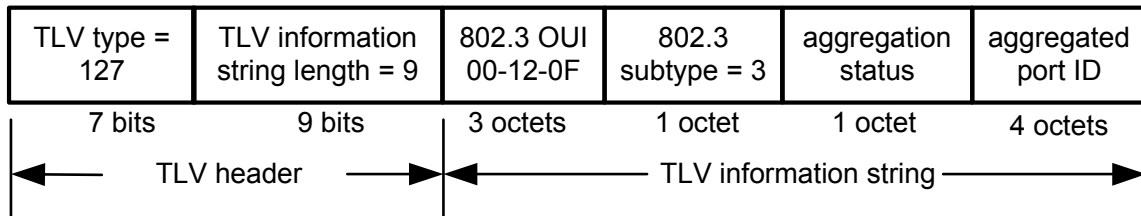


Figure 79–4—Link Aggregation TLV format

79.3.3.1 Aggregation status

The link aggregation status field shall contain a bitmap of the link aggregation capabilities and the current aggregation status of the link as defined in Table 79–7.

Table 79–7—Link aggregation capability/status

Bit	Function	Value/meaning
0	Aggregation capability	0 = not capable of being aggregated 1 = capable of being aggregated
1	Aggregation status	0 = not currently in aggregation 1 = currently in aggregation
2–7	reserved for future standardization	—

79.3.3.2 Aggregated port ID

The aggregated port ID field shall contain the IEEE 802.3 aggregated port identifier, aAggPortID, derived from the ifNumber in the ifIndex for the interface (see 30.7.2.1.1).

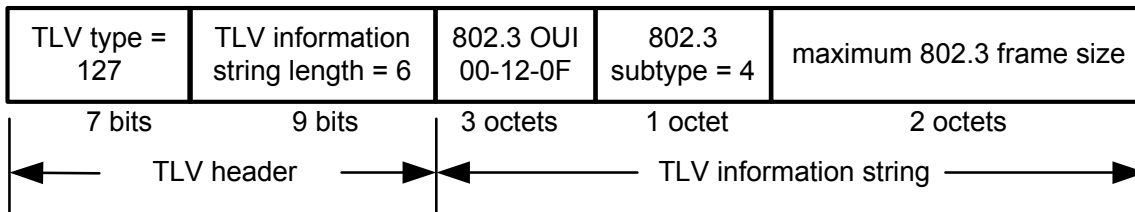
79.3.3.3 Link Aggregation TLV usage rules

An LLDPDU should contain no more than one Link Aggregation TLV.

79.3.4 Maximum Frame Size TLV

The Maximum Frame Size TLV is an optional TLV that indicates the maximum frame size capability of the implemented MAC and PHY. Figure 79–5 shows the format of this TLV.

NOTE—MAC and PHY support for a given frame size doesn't necessarily mean that the upper layers support that frame size.

**Figure 79–5—Maximum Frame Size TLV format****79.3.4.1 Maximum frame size**

The maximum frame size field shall contain an integer value indicating the maximum supported frame size in octets as determined by the following:

- If the MAC/PHY supports only basic frames (see 3.2.7) the maximum frame size field shall be set to 1518.
- If the MAC/PHY supports Q-tagged frames (see 3.2.7) the maximum frame size field shall be set to 1522.
- If the MAC/PHY supports envelope frames (see 3.2.7) the maximum frame size field shall be set to 2000.

79.3.4.2 Maximum Frame Size TLV usage rules

An LLDPDU should contain no more than one Maximum Frame Size TLV.

79.3.5 EEE TLV

The EEE TLV is used to exchange information about the EEE Data Link Layer capabilities. Figure 79–6 shows the format of this TLV.

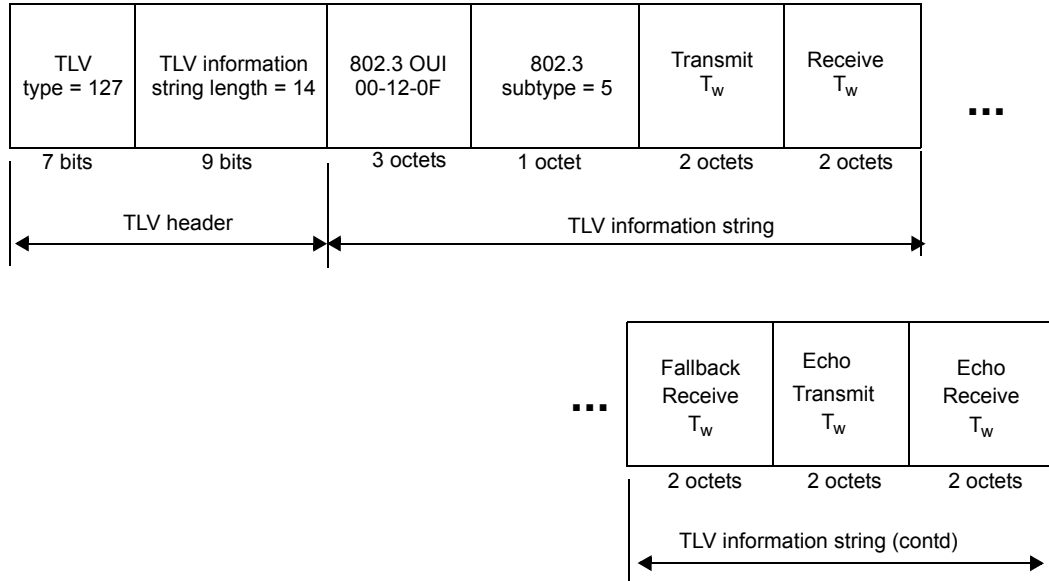


Figure 79–6—EEE TLV format

79.3.5.1 Transmit T_w

Transmit $T_{w_sys_tx}$ (2 octets wide) shall be defined as the time (expressed in microseconds) that the transmitting link partner will wait before it starts transmitting data after leaving the Low Power Idle (LPI) mode. This is a function of the transmit system design and may be constrained, for example, by the transmit path buffering. The default value for Transmit $T_{w_sys_tx}$ is the T_{w_phy} defined for the PHY that is in use for the link. The Transmitting link partner expects that the Receiving link partner will be able to accept data after the time delay Transmit $T_{w_sys_tx}$ (expressed in microseconds).

79.3.5.2 Receive T_w

Receive $T_{w_sys_tx}$ (2 octets wide) shall be defined as the time (expressed in microseconds) that the receiving link partner is requesting the transmitting link partner to wait before starting the transmission data following the LPI. The default value for Receive $T_{w_sys_tx}$ is the T_{w_phy} defined for the PHY that is in use for the link. The Receive $T_{w_sys_tx}$ value can be larger but not smaller than the default. The extra wait time may be used by the receive link partner for power-saving mechanisms that require a longer wake-up time than the PHY-layer definitions.

79.3.5.3 Fallback T_w

A receiving link partner may inform the transmitter of an alternate desired $T_{w_sys_tx}$. Since a receiving link partner is likely to have discrete levels for savings, this provides the transmitter with additional information that it may use for a more efficient allocation. As with the Receive $T_{w_sys_tx}$, this is 2 octets wide. Systems that do not implement this option default the value to be the same as that of the Receive $T_{w_sys_tx}$.

79.3.5.4 Echo Transmit and Receive T_w

The respective echo values shall be defined as the local link partner's reflection (echo) of the remote link partner's respective values. When a local link partner receives its echoed values from the remote link partner it can determine whether or not the remote link partner has received, registered, and processed its most recent values. For example, if the local link partner receives echoed parameters that do not match the values in its local MIB, then the local link partner infers that the remote link partner's request was based on stale information.

79.3.5.5 EEE TLV usage rules

An LLDPDU should contain no more than one EEE TLV.

79.4 IEEE 802.3 Organizationally Specific TLV selection management

TLV selection management consists of providing the network manager with the means to select which specific IEEE 802.3 Organizationally Specific TLVs are enabled for inclusion in an LLDPDU. The following LLDP variable cross references the LLDP local systems configuration MIB tables (see Clause 11 of IEEE Std 802.1AB-2009) to indicate which specific TLVs are enabled for the particular port(s) on the system. The specific port(s) through which each TLV is enabled for transmission may be set (or reset) by the network manager:

- a) **mibXdot3TLVsTxEnable:** This variable lists the single-instance use IEEE 802.3 Organizationally Specific TLVs, each with a bitmap indicating the system ports through which the referenced TLV is enabled for transmission.

79.4.1 IEEE 802.3 Organizationally Specific TLV selection variable/LLDP Configuration managed object class cross reference

Table 79–8 lists the relationship both between IEEE 802.3 TLV selection variable and the corresponding LLDP Configuration managed object class (see 30.12.1) attribute.

Table 79–8—IEEE 802.3 Organizationally Specific TLV selection variable/LLDP MIB object cross reference

IEEE 802.3 TLV selection variable	LLDP Configuration managed object class attribute
mibXdot3TLVsTxEnable	aLldpXdot3PortConfigTLVsTxEnable

79.4.2 IEEE 802.3 Organizationally Specific TLV/LLDP Local and Remote System group managed object class cross references

The cross references between the IEEE 802.3 TLVs and the LLDP Local System Group managed object class (see 30.12.2) attributes are listed in Table 79–9. The cross references between the IEEE 802.3 TLVs and the LLDP Remote System Group managed object class (see 30.12.3) attributes are listed in Table 79–10.

The cross-references between the EEE TLV and the EEE local (30.12.2) and remote (30.12.3) object class attributes are listed in Table 79–9 and Table 79–10.

Table 79–9—IEEE 802.3 Organizationally Specific TLV/LLDP Local System Group managed object class cross references

TLV name	TLV variable	LLDP Local System Group managed object class attribute
MAC/PHY Configuration/Status	Auto-negotiation support	aLldpXdot3LocPortAutoNegSupported
	Auto-negotiation status	aLldpXdot3LocPortAutoNegEnabled
	PMD auto-negotiation advertised capability	aLldpXdot3LocPortAutoNegAdvertisedCap
	Operational MAU type	aLldpXdot3LocPortOperMauType
Power via MDI	Port class	aLldpXdot3LocPowerPortClass
	PSE MDI power support	aLldpXdot3LocPowerMDISupported
	PSE MDI power state	aLldpXdot3LocPowerMDIEnabled
	PSE pairs control ability	aLldpXdot3LocPowerPairControlable
	PSE power pair	aLldpXdot3LocPowerPairs
	Power class	aLldpXdot3LocPowerClass
	Power type	aLldpXdot3LocPowerType
	Power source	aLldpXdot3LocPowerSource
	Power priority	aLldpXdot3LocPowerPriority
	PD requested power value	aLldpXdot3LocPDRequestedPowerValue
	PSE allocated power value	aLldpXdot3LocPSEAllocatedPowerValue
Link Aggregation (deprecated)	aggregation status	aLldpXdot3LocLinkAggStatus
	aggregated port ID	aLldpXdot3LocLinkAggPortId
Maximum Frame Size	maximum frame size	aLldpXdot3LocMaxFrameSize
EEE	Transmit $T_{w_sys_tx}$	aLldpXdot3LocTxTwSys
	Receive $T_{w_sys_tx}$	aLldpXdot3LocRxTwSys
	Echo Transmit $T_{w_sys_tx}$	aLldpXdot3LocTxTwSysEcho
	Echo Receive $T_{w_sys_tx}$	aLldpXdot3LocRxTwSysEcho
	Fallback $T_{w_sys_tx}$	aLldpXdot3LocFbTwSys

Table 79–10—IEEE 802.3 Organizationally Specific TLV/LLDP Remote System Group managed object class cross references

TLV name	TLV variable	LLDP Remote System Group managed object class attribute
MAC/PHY Configuration/Status	Auto-negotiation support	aLldpXdot3RemPortAutoNegSupported
	Auto-negotiation status	aLldpXdot3RemPortAutoNegEnabled
	PMD auto-negotiation advertised capability	aLldpXdot3RemPortAutoNegAdvertisedCap
	Operational MAU type	aLldpXdot3RemPortOperMauType
Power via MDI	Port class	aLldpXdot3RemPowerPortClass
	PSE MDI power support	aLldpXdot3RemPowerMDISupported
	PSE MDI power state	aLldpXdot3RemPowerMDIEnabled
	PSE pairs control ability	aLldpXdot3RemPowerPairControlable
	PSE power pair	aLldpXdot3RemPowerPairs
	Power class	aLldpXdot3RemPowerClass
	Power type	aLldpXdot3RemPowerType
	Power source	aLldpXdot3RemPowerSource
	Power priority	aLldpXdot3RemPowerPriority
	PD requested power value	aLldpXdot3RemPDRequestedPowerValue
	PSE allocated power value	aLldpXdot3RemPSEAllocatedPowerValue
Link Aggregation (deprecated)	aggregation status	aLldpXdot3RemLinkAggStatus
	aggregated port ID	aLldpXdot3RemLinkAggPortId
Maximum Frame Size	maximum frame size	aLldpXdot3RemMaxFrameSize
EEE	Transmit $T_{w_sys_tx}$	aLldpXdot3RemTxTwSys
	Receive $T_{w_sys_tx}$	aLldpXdot3RemRxTwSys
	Echo Transmit $T_{w_sys_tx}$	aLldpXdot3RemTxTwSysEcho
	Echo Receive $T_{w_sys_tx}$	aLldpXdot3RemRxTwSysEcho
	Fallback $T_{w_sys_tx}$	aLldpXdot3RemFbTwSys

79.5 Protocol implementation conformance statement (PICS) proforma for IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements³

79.5.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 79, IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements, 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.

79.5.2 Identification

79.5.2.1 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	
NOTES	
1—Required for all implementations.	
2—May be completed as appropriate in meeting the requirements for the identification.	
3—The terms Name and Version should be interpreted appropriately to correspond with a supplier's terminology (e.g., Type, Series, Model).	

79.5.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 79, IEEE 802.3 Organizationally Specific Link Layer Discovery Protocol (LLDP) type, length, and value (TLV) information elements
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 802.3-2012.)	
Date of Statement	

³Copyright 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.

79.5.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*MP	MAC/PHY Configuration/Status TLV	79.3.1		O	Yes [] No []
*PV	Power Via MDI TLV	79.3.2		O	Yes [] No []
*LA	Link Aggregation TLV	79.3.3	TLV deprecated	O	Yes [] No []
*FS	Maximum Frame Size TLV	79.3.4		O	Yes [] No []
*EE	EEE TLV	79.3.5		O	Yes [] No []

79.5.4 IEEE 802.3 Organizationally Specific TLV

Item	Feature	Subclause	Value/Comment	Status	Support
TLV1	Group MAC addresses	79.2	<i>Nearest device</i> group MAC addresses listed in Table 7-1 of IEEE Std 802.1AB-2009	M	Yes []
TLV2	LLDPDU bit and octet ordering	79.2	Defined in subclause 8.1 of IEEE Std 802.1AB-2009	M	Yes []

79.5.5 MAC/PHY Configuration/Status TLV

Item	Feature	Subclause	Value/Comment	Status	Support
MPT1	auto-negotiation support/status field	79.3.1.1	Identifies support and current status as defined in Table 79-2	MP:M	Yes [] N/A []
MPT2	PMD auto-negotiation capability field	79.3.1.2	Bitmap of the ifMauAutoNeg-CapAdvertisedBits object defined in IETF RFC 4836	MP:M	Yes [] N/A []
MPT3	operational MAU type field	79.3.1.3	Derived from the list position of the corresponding dot3MauType as listed in IETF RFC 4836 (or subsequent revisions)	MP:M	Yes [] N/A []
MPT4	operational MAU type field	79.3.1.3	Set to zero for MAU types not listed in IETF RFC 4836 (or subsequent revisions)	MP:M	Yes [] N/A []
MPT5	Usage rules	79.3.1.4	LLDPDU contains no more than one MAC/PHY Configuration/Status TLV	MP:O	Yes [] No [] N/A []

79.5.6 EEE TLV

Item	Feature	Subclause	Value/Comment	Status	Support
EET1	Transmit T_w field	79.3.5.1	2 octets representing time (expressed in microseconds) that the transmitting link partner will wait before it starts transmitting data after leaving the LPI mode	EE:M	Yes [] N/A []
EET2	Receive T_w field	79.3.5.2	2 octets representing time (expressed in microseconds) that the receiving link partner is requesting the transmitting link partner to wait before it starts transmitting data following the LPI	EE:M	Yes [] N/A []
EET3	Fallback field	79.3.5.3	2 octets representing time (expressed in microseconds)	EE:O	Yes [] N/A []
EET4	Echo Transmit and Receive T_w fields	79.3.5.4	2 octets representing time (expressed in microseconds)	EE:M	Yes [] N/A []
EET5	Usage rules	79.3.5.5	LLDPDU contains no more than one EEE TLV	EE:O	Yes [] No [] N/A []

79.5.7 Power Via MDI TLV

Item	Feature	Subclause	Value/Comment	Status	Support
PVT1	MDI power support field	79.3.2.1	Bit map of the MDI power capabilities and status as defined in Table 79–3	PV:M	Yes [] N/A []
PVT2	PSE power pair field	79.3.2.2	Integer value as defined by the pethPsePortPowerPairs object in IETF RFC 3621	PV:M	Yes [] N/A []
PVT3	power class field	79.3.2.3	Integer value as defined by the pethPsePortPowerClassifications object in IETF RFC 3621	PV:M	Yes [] N/A []
PVT4	Power type/source/priority field	79.3.2.4	Contains a bit-map of the power type, source, and priority defined in Table 79–4	PV:M	Yes [] N/A []
PVT5	Power type field	79.3.2.4.1	Set according to Table 79–4	PV:M	Yes [] N/A []
PVT6	Power source field when power type is PD	79.3.2.4.2	Set to ‘01’ when powered only through the PI; set to ‘11’ when powered from both; set to ‘00’ when information is not available	PV:M	Yes [] N/A []
PVT7	Power source field when power type is PSE	79.3.2.4.2	When sourcing power through the PI, set to ‘01’ when using primary supply; set to ‘10’ when using backup source; set to ‘00’ when information is not available	PV:M	Yes [] N/A []
PVT8	Power priority field when power type is PD	79.3.2.4.3	Set to the power priority configured for the device; set to ‘00’ if power priority is undetermined	PV:M	Yes [] N/A []
PVT9	PD requested power value field	79.3.2.5	Contains the PD’s requested power value defined in Table 79–5	PV:M	Yes [] N/A []
PVT10	PSE allocated power value field	79.3.2.6	Contains the PSE’s allocated power value defined in Table 79–6	PV:M	Yes [] N/A []
PVT11	Usage rules	79.3.2.7	LLDPDU contains no more than one Power Via MDI TLV	PV:O	Yes [] No [] N/A []

79.5.8 Link Aggregation TLV

Item	Feature	Subclause	Value/Comment	Status	Support
LAT1	link aggregation status field	79.3.3.1	Bitmap of the link aggregation capabilities and the current aggregation status as defined in Table 79–7	LA:M	Yes [] N/A []
LAT2	aggregated port ID	79.3.3.2	IEEE 802.3 aggregated port identifier, aAggPortID	LA:M	Yes [] N/A []
LAT3	Usage rules	79.3.3.3	LLDPDU contains no more than one Link Aggregation TLV	LA:O	Yes [] No [] N/A []

79.5.9 Maximum Frame Size TLV

Item	Feature	Subclause	Value/Comment	Status	Support
FST1	maximum frame size field	79.3.4.1	Integer value indicating the maximum supported frame size	FS:M	Yes [] N/A []
FST2	maximum frame size field	79.3.4.1	1518 for basic frames	FS:O/1	Yes [] No [] N/A []
FST3	maximum frame size field	79.3.4.1	1522 for Q-tagged frames	FS:O/1	Yes [] No [] N/A []
FST4	maximum frame size field	79.3.4.1	2000 for envelope frames	FS:O/1	Yes [] No [] N/A []
FST5	Usage rules	79.3.4.2	LLDPDU contains no more than one Maximum Frame Size TLV	FS:O	Yes [] No [] N/A []

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

80.1 Overview

80.1.1 Scope

40 Gigabit and 100 Gigabit Ethernet uses the IEEE 802.3 MAC sublayer, connected through a Media Independent Interface to 40 Gb/s and 100 Gb/s Physical Layer entities such as those specified in Table 80–1.

40 Gigabit and 100 Gigabit Ethernet is defined for full duplex operation only.

80.1.2 Objectives

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

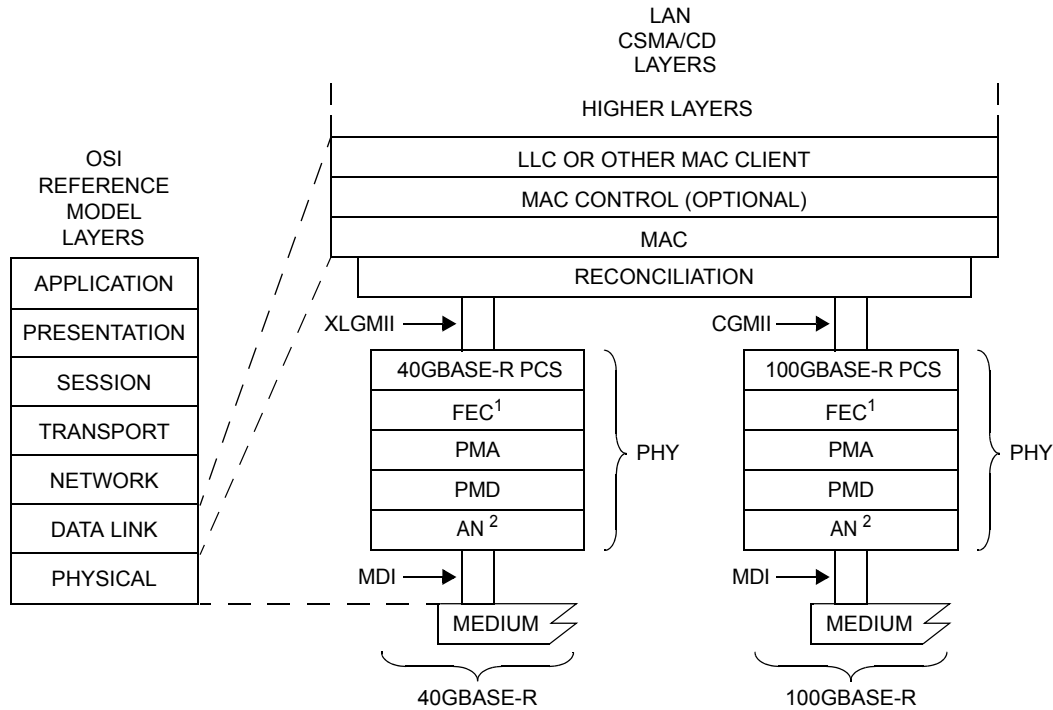
- a) Support the full duplex Ethernet MAC.
- b) Preserve the IEEE 802.3 Ethernet frame format utilizing the IEEE 802.3 MAC.
- c) Preserve minimum and maximum frame size of IEEE Std 802.3.
- d) Support a BER better than or equal to 10^{-12} at the MAC/PLS service interface.
- e) Provide appropriate support 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

40 Gigabit and 100 Gigabit Ethernet couples the IEEE 802.3 MAC to a family of 40 Gb/s and 100 Gb/s Physical Layers. The relationships among 40 Gigabit and 100 Gigabit Ethernet, the IEEE 802.3 MAC, and the ISO Open System Interconnection (OSI) reference model are shown in Figure 80–1.

While this specification defines interfaces in terms of bits, octets, and frames, implementations may choose other data-path widths for implementation convenience. The only exceptions are as follows:

- a) The XLGMII and CGMII, which, when implemented as a logical interconnection port between the MAC sublayer and the Physical Layer (PHY), uses a 64-bit wide data path as specified in Clause 81.
- b) The management interface, which, when physically implemented as the MDIO/MDC (Management Data Input/Output and Management Data Clock) at an observable interconnection port, uses a bit-wide data path as specified in Clause 45.
- c) The PMA service interface, which, when physically implemented as XLAUI (40 Gigabit Attachment Unit Interface) at an observable interconnection port, uses a 4 lane data path as specified in Annex 83A or Annex 83B.



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

- d) The PMA service interface, which, when physically implemented as CAUI (100 Gigabit Attachment Unit Interface) at an observable interconnection port, uses a 10 lane data path as specified in Annex 83A or Annex 83B.
- e) The PMD service interface, which, when physically implemented as XLPPi (40 Gigabit Parallel Physical Interface) at an observable interconnection port, uses a 4 lane data path as specified in Annex 86A.
- f) The PMD service interface, which, when physically implemented as CPPI (100 Gigabit Parallel Physical Interface) at an observable interconnection port, uses a 10 lane data path as specified in Annex 86A.
- g) The MDI as specified in Clause 89 for 40GBASE-FR uses a single lane data path.
- h) The MDIs as specified in Clause 84 for 40GBASE-KR4, in Clause 85 for 40GBASE-CR4, in Clause 86 for 40GBASE-SR4, Clause 87 for 40GBASE-LR4 and in Clause 88 for 100GBASE-LR4 and 100GBASE-ER4 all use a 4 lane data path.
- i) The MDIs as specified in Clause 85 for 100GBASE-CR10, and in Clause 86 for 100GBASE-SR10 use a 10 lane data path.

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 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.

Table 80–1—40 Gb/s and 100 Gb/s PHYs

Name	Description
40GBASE-KR4	40 Gb/s PHY using 40GBASE-R encoding over four lanes of an electrical backplane, with reach up to at least 1 m (see Clause 84)
40GBASE-CR4	40 Gb/s PHY using 40GBASE-R encoding over four lanes of shielded balanced copper cabling, with reach up to at least 7 m (see Clause 85)
40GBASE-SR4	40 Gb/s PHY using 40GBASE-R encoding over four lanes of multimode fiber, with reach up to at least 100 m (see Clause 86)
40GBASE-FR	40 Gb/s PHY using 40GBASE-R encoding over one lane on single-mode fiber, with reach up to at least 2 km (see Clause 89)
40GBASE-LR4	40 Gb/s PHY using 40GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 10 km (see Clause 87)
100GBASE-CR10	100 Gb/s PHY using 100GBASE-R encoding over ten lanes of shielded balanced copper cabling, with reach up to at least 7 m (see Clause 85)
100GBASE-SR10	100 Gb/s PHY using 100GBASE-R encoding over ten lanes of multimode fiber, with reach up to at least 100 m (see Clause 86)
100GBASE-LR4	100 Gb/s PHY using 100GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 10 km (see Clause 88)
100GBASE-ER4	100 Gb/s PHY using 100GBASE-R encoding over four WDM lanes on single-mode fiber, with reach up to at least 40 km (see Clause 88)

80.1.5 Physical Layer signaling systems

This standard specifies a family of Physical Layer implementations. The generic term 40 Gigabit and 100 Gigabit Ethernet refers to any use of the 40 Gb/s and 100 Gb/s IEEE 802.3 MAC (the 40 Gigabit and 100 Gigabit Ethernet MAC) coupled with any IEEE 802.3 40GBASE or 100GBASE Physical Layer implementations. Table 80–2 specifies the correlation between nomenclature and clauses. Implementations conforming to one or more nomenclatures must meet the requirements of the corresponding clauses.

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	XLAI	CAUI	XLAI	CAUI	40GBASE-KR4 PMD	40GBASE-CR4 PMD	100GBASE-CR10 PMD	40GBASE-SR4 PMD	100GBASE-SR10 PMD	XLPI	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	

^aO = Optional, M = Mandatory.

80.2 Summary of 40 Gigabit and 100 Gigabit Ethernet sublayers

80.2.1 Reconciliation Sublayer (RS) and Media Independent Interface

The Media Independent Interface (Clause 81) provides a logical interconnection between the MAC sublayer and Physical Layer entities (PHY). The Media Independent Interface is not intended to be physically instantiated, rather it can logically connect layers within a device.

The XLGMII supports 40 Gb/s and CGMII supports 100 Gb/s operation through its 64-bit-wide transmit and receive data paths. The Reconciliation Sublayer (RS) provides a mapping between the signals provided at the Media Independent Interface (XLGMII and CGMII) and the MAC/PLS service definition.

While XLGMII and CGMII are optional interfaces, they are used extensively in this standard as a basis for functional specification and provides a common service interface for the physical coding sublayers defined in Clause 82.

80.2.2 Physical Coding Sublayer (PCS)

The terms 40GBASE-R and 100GBASE-R refer to a specific family of Physical Layer implementations based upon the 64B/66B data coding method specified in Clause 82 and the PMA specification defined in Clause 83. The 40GBASE-R and 100GBASE-R PCSs perform encoding (decoding) of data from (to) the

XLGMII/CGMII to 64B/66B code blocks, distribute the data to multiple lanes, and transfer the encoded data to the PMA.

80.2.3 Forward Error Correction (FEC) sublayer

The Forward Error Correction sublayer is an optional sublayer for 40GBASE-R and 100GBASE-R copper and backplane 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 FEC sublayer is specified in Clause 74.

80.2.4 Physical Medium Attachment (PMA) sublayer

The PMA provides a medium-independent means for the PCS to support the use of a range of physical media. The 40GBASE-R and 100GBASE-R PMAs perform the mapping of transmit and receive data streams between the PCS and PMA via the PMA service interface, and the mapping and multiplexing of transmit and receive data streams between the PMA and PMD via the PMD service interface. In addition, the PMAs perform retiming of the received data stream when appropriate, optionally provide data loopback at the PMA or PMD service interface, and optionally provide test pattern generation and checking.

The 40GBASE-R and 100GBASE-R PMAs are specified in Clause 83.

80.2.5 Physical Medium Dependent (PMD) sublayer

The Physical Medium Dependent sublayer is responsible for interfacing to the transmission medium. The PMD is located just above the Medium Dependent Interface (MDI). The MDI, logically subsumed within each PMD subclause, is the actual medium attachment for the various supported media.

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

80.2.6 Auto-Negotiation

Auto-Negotiation provides a linked device with the capability to detect the abilities (modes of operation) supported by the device at the other end of the link, determine common abilities, and configure for joint operation.

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

80.2.7 Management interface (MDIO/MDC)

The optional MDIO/MDC management interface (Clause 45) provides an interconnection between MDIO Manageable Devices (MMDs) and Station Management (STA) entities.

80.2.8 Management

Managed objects, attributes, and actions are defined for all 40 Gigabit and 100 Gigabit Ethernet components. These items are defined in Clause 30.

80.3 Service interface specification method and notation

The service interface specification for 40GBASE-R and 100GBASE-R Physical Layers is as per the definition in 1.2.2. Note that the 40GBASE-R and 100GBASE-R inter-sublayer service interfaces use multiple scalar REQUEST and INDICATION primitives, to indicate the transfer of multiple independent streams of data units, as explained in 80.3.1 through 80.3.3.

80.3.1 Inter-sublayer service interface

The inter-sublayer service interface is described in an abstract manner and does not imply any particular implementation. The inter-sublayer service interface primitives are defined as follows:

```
IS_UNITDATA_i.request  
IS_UNITDATA_i.indication  
IS_SIGNAL.indication
```

The IS_UNITDATA_i.request (where $i = 0$ to $n - 1$, and n is the number of streams of data units) primitive is used to define the transfer of multiple streams of data units from a sublayer N to the next lower sublayer $N - 1$. The IS_UNITDATA_i.indication (where $i = 0$ to $n - 1$, and n is the number of streams of data units) primitive is used to define the transfer of multiple streams of data units from a sublayer $N - 1$ to the next higher sublayer N . The IS_SIGNAL.indication primitive is used to define the transfer of signal status from a sublayer $N - 1$ to the next higher sublayer N .

80.3.2 Instances of the Inter-sublayer service interface

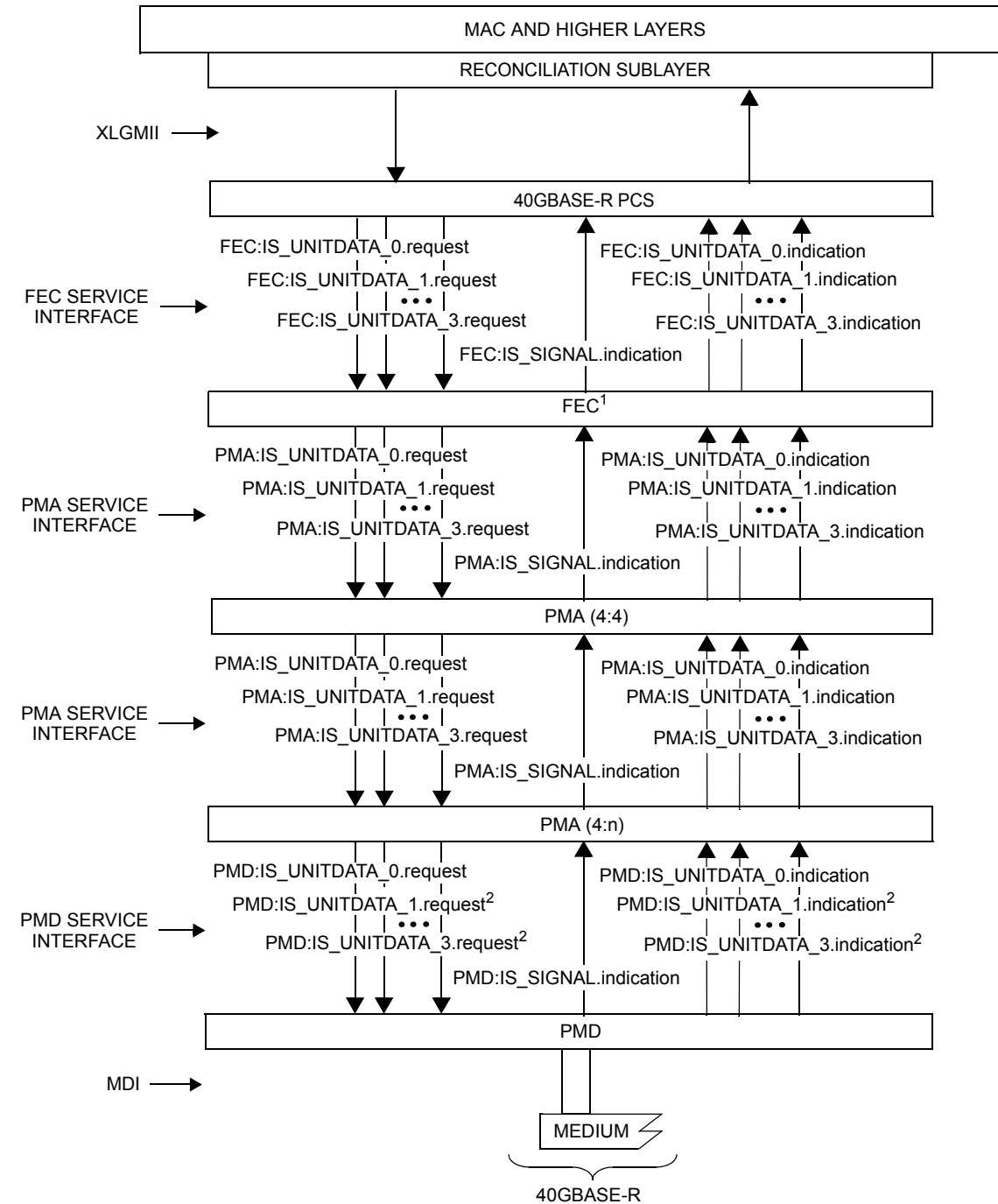
The inter-sublayer interface can be instantiated between different sublayers, hence a prefix notation is defined to identify a specific instance of an inter-sublayer service interface. The following prefixes are defined:

- a) PMD:—for primitives issued on the interface between the PMD sublayer and the PMA sublayer called the PMD service interface.
- b) PMA:—for primitives issued on the interface between the PMA sublayer and the PCS (or the FEC) sublayer called the PMA service interface.
- c) FEC:—for primitives issued on the interface between the FEC sublayer and the PCS (or the PMA) sublayer called the FEC service interface.

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. For example, the primitives for one instance of the inter-sublayer service interface, named the PMD service interface, are identified as follows:

```
PMD:IS_UNITDATA_i.request  
PMD:IS_UNITDATA_i.indication  
PMD:IS_SIGNAL.indication.
```

Primitives for other instances, of inter-sublayer interfaces, are represented in a similar manner as described above.

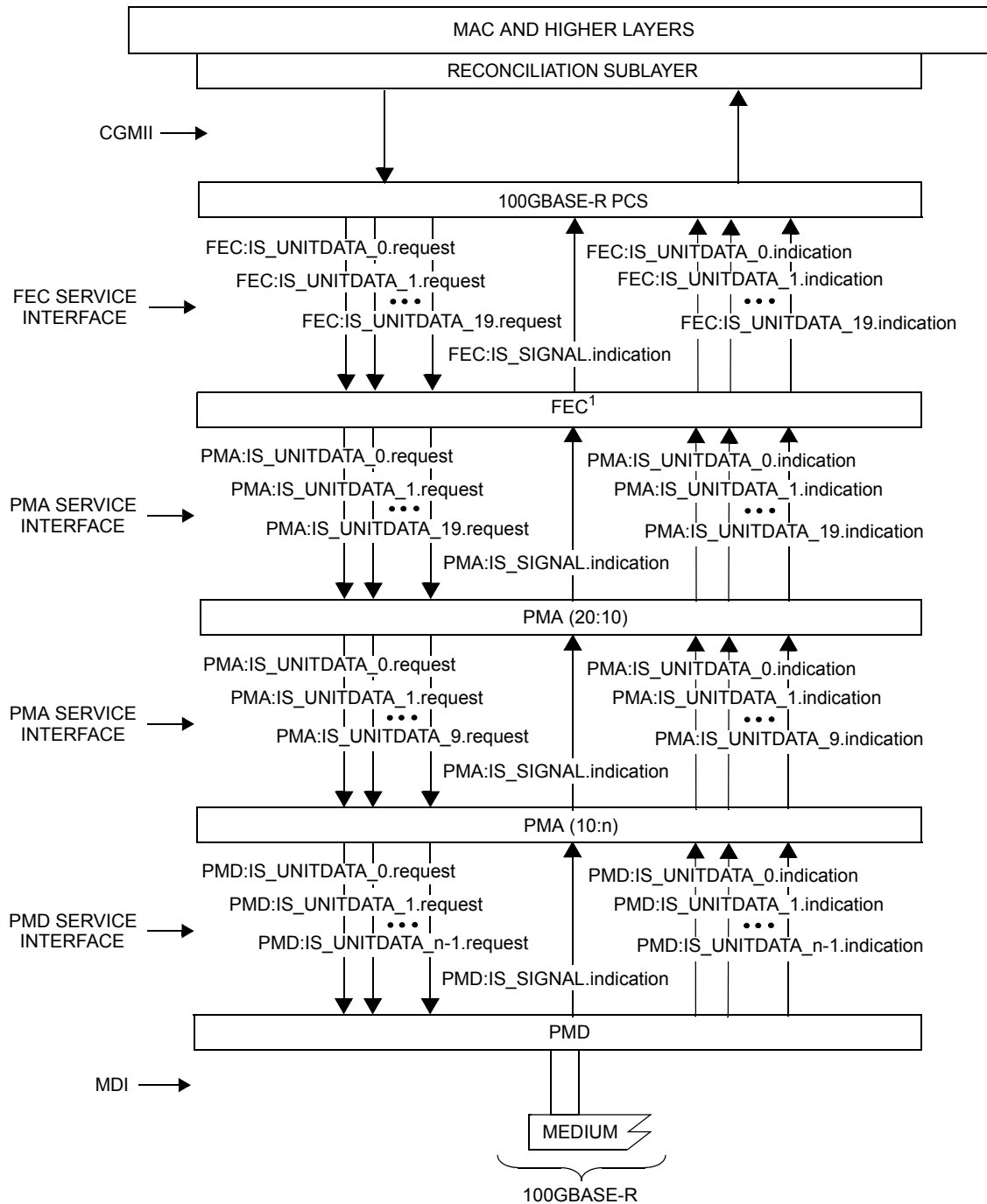


FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT

XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE
n = NUMBER OF PARALLEL STREAMS OF DATA UNITS

NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE
NOTE 2—DOES NOT EXIST FOR 40GBASE-FR PMD

Figure 80-2—40GBASE-R inter-sublayer service interfaces



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

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

80.3.3 Semantics of inter-sublayer service interface primitives

The semantics of the inter-sublayer service interface primitives for the 40GBASE-R and 100GBASE-R sublayers are described in 80.3.3.1 through 80.3.3.3.

80.3.3.1 IS_UNITDATA_*i*.request

The IS_UNITDATA_*i*.request (where $i = 0$ to $n - 1$) primitive is used to define the transfer of multiple streams of data units from a sublayer N to the next lower sublayer $N - 1$, where n is the number of parallel streams of data units.

80.3.3.1.1 Semantics of the service primitive

```
IS_UNITDATA_0.request(tx_bit)
IS_UNITDATA_1.request(tx_bit)
...
IS_UNITDATA_n-1.request(tx_bit)
```

The data conveyed by IS_UNITDATA_0.request to IS_UNITDATA_n-1.request consists of n 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.

80.3.3.1.2 When generated

The sublayer N continuously sends n parallel bit streams IS_UNITDATA_*i*.request(tx_bit) to the next lower sublayer $N - 1$, each at a nominal signaling rate defined by a specific instance of the inter-sublayer service interface.

80.3.3.1.3 Effect of receipt

The effect of receipt of this primitive is defined by the sublayer that receives this primitive.

80.3.3.2 IS_UNITDATA_*i*.indication

The IS_UNITDATA_*i*.indication (where $i = 0$ to $n - 1$) primitive is used to define the transfer of multiple streams of data units from the sublayer $N - 1$ to the next higher sublayer N , where n is the number of parallel streams of data units.

80.3.3.2.1 Semantics of the service primitive

```
IS_UNITDATA_0.indication(rx_bit)
IS_UNITDATA_1.indication(rx_bit)
...
IS_UNITDATA_n-1.indication(rx_bit)
```

The data conveyed by IS_UNITDATA_0.indication to IS_UNITDATA_n-1.indication consists of n 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.

80.3.3.2.2 When generated

The sublayer N-1 continuously sends n parallel bit streams IS_UNITDATA_i.indication(rx_bit) to the next higher sublayer N, each at a nominal signaling rate defined by a specific instance of the inter-sublayer service interface.

80.3.3.2.3 Effect of receipt

The effect of receipt of this primitive is defined by the sublayer that receives this primitive.

80.3.3.3 IS_SIGNAL.indication

The IS_SIGNAL.indication primitive is generated by the sublayer N - 1 to the next higher sublayer N 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., no valid signal being received by the sublayer that generates this primitive) to the next higher sublayer N.

80.3.3.3.1 Semantics of the service primitive

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 sublayer N - 1 to the next higher sublayer N. A value of OK does not guarantee valid data is being presented by the sublayer N - 1 to the next higher sublayer N.

80.3.3.3.2 When generated

The sublayer N - 1 generates the IS_SIGNAL.indication primitive to the next higher sublayer N whenever there is change in the value of the SIGNAL_OK parameter.

80.3.3.3.3 Effect of receipt

The effect of receipt of this primitive is defined by the sublayer that receives this primitive.

80.4 Delay constraints

Predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementors must conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices. Table 80-3 contains the values of maximum sublayer delay (sum of transmit and receive delays at one end of the link) in bit times as specified in 1.4 and pause_quanta as specified in 31B.2. If a PHY contains an Auto-Negotiation sublayer, the delay of the Auto-Negotiation sublayer is included within the delay of the PMD and medium.

See 44.3 for the calculation of bit time per meter of fiber or electrical cable.

See 31B.3.7 for PAUSE reaction timing constraints for stations at operating speeds of 40 Gb/s and 100 Gb/s.

Table 80–3—Sublayer delay constraints

Sublayer	Maximum (bit time) ^a	Maximum (pause_quanta) ^b	Maximum (ns)	Notes ^c
40G MAC, RS, and MAC Control	16384	32	409.6	See 81.1.4.
40GBASE-R PCS	11264	22	281.6	See 82.5.
40GBASE-R FEC	24576	48	614.4	See 74.6.
40GBASE-R PMA	4096	8	102.4	See 83.5.4.
40GBASE-KR4 PMD	2048	4	51.2	Includes delay of one direction through backplane medium. See 84.4.
40GBASE-CR4 PMD	4096	8	102.4	Does not include delay of cable medium. See 85.4.
40GBASE-SR4 PMD	1024	2	25.6	Includes 2 m of fiber. See 86.3.1.
40GBASE-FR PMD	1024	2	25.6	Includes 2 m of fiber. See 89.3.1.
40GBASE-LR4 PMD	1024	2	25.6	Includes 2 m of fiber. See 87.3.1.
100G MAC, RS, and MAC Control	24576	48	245.76	See 81.1.4.
100GBASE-R PCS	35328	69	353.28	See 82.5.
100GBASE-R FEC	122880	240	1228.8	See 74.6.
100GBASE-R PMA	9216	18	92.16	See 83.5.4.
100GBASE-CR10 PMD	9728	19	97.28	Does not include delay of cable medium. See 85.4.
100GBASE-SR10 PMD	2048	4	20.48	Includes 2 m of fiber. See 86.3.1.
100GBASE-LR4 PMD	2048	4	20.48	Includes 2 m of fiber. See 88.3.1.
100GBASE-ER4 PMD	2048	4	20.48	Includes 2 m of fiber. See 88.3.1.

^a 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 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.)

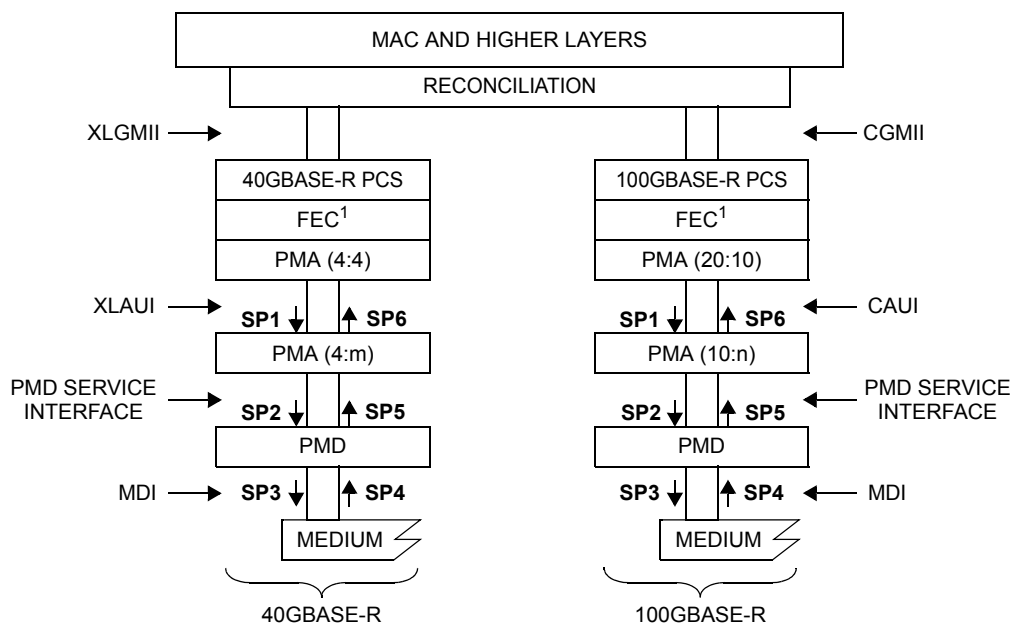
^c Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.

80.5 Skew constraints

Skew (or relative delay) can be introduced between lanes by both active and passive elements of a 40GBASE-R or 100GBASE-R link. Skew is defined as the difference between the times of the earliest PCS lane and latest PCS lane for the one to zero transition of the alignment marker sync bits. The PCS deskew function (see 82.2.12) compensates for all lane-to-lane Skew observed at the receiver. The Skew between

the lanes must be kept within limits as shown in Table 80–4 so that the transmitted information on the lanes can be reassembled by the receive PCS.

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



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
 m = 1 or 4
 n = 4 or 10

NOTE1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE

Figure 80–4—40GBASE-R and 100GBASE-R Skew points for single XLAUI or CAUI

The maximum Skew and Skew Variation at physically instantiated interfaces is specified at Skew points SP1, SP2, and SP3 for the transmit direction and SP4, SP5, and SP6 for the receive direction as illustrated in Figure 80–4 (single XLAUI or CAUI interface) and Figure 80–5 (multiple XLAUI or CAUI interfaces).

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

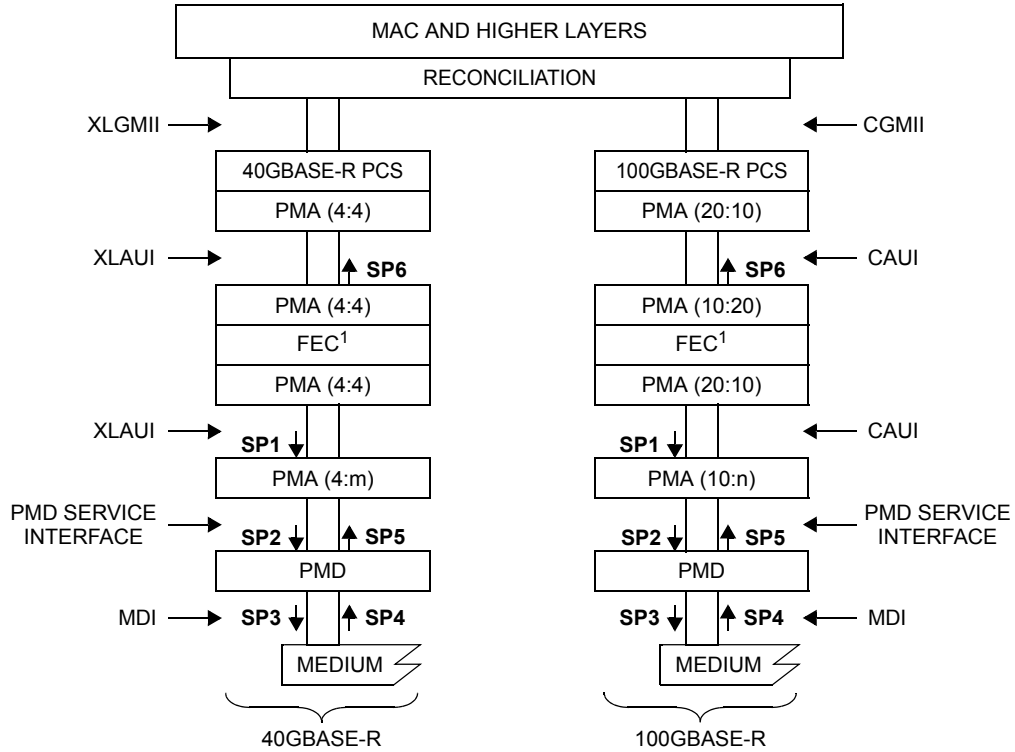
- SP1 on the XLAUI/CAUI interface, at the input of the PMA closest to the PMD;
- SP2 on the PMD service interface, at the input of the PMD;
- SP3 at the output of the PMD, at the MDI.

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

- SP4 at the MDI, at the input of the PMD;

- SP5 on the PMD service interface, at the output of the PMD;
- SP6 on the XLAUI/CAUI interface, at the output of the PMA closest to the PCS.

The allowable limits for Skew are shown in Table 80–4 and the allowable limits for Skew Variation are shown in Table 80–5.



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

PMD = PHYSICAL MEDIUM DEPENDENT
 XLAUI = 40 Gb/s ATTACHMENT UNIT INTERFACE
 XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE
 m = 1 or 4
 n = 4 or 10

NOTE1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE

Figure 80–5—40GBASE-R and 100GBASE-R Skew points for multiple XLAUI or CAUI

The Skew requirements for the PCS, PMA and PMD sublayers are specified in the respective clauses as noted in Table 80–4 and Table 80–5.

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
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
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
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
SP6	160	≈ 1649	≈ 824	See 83.5.3.5
At PCS receive	180	≈ 1856	≈ 928	See 82.2.12

^aThe Skew limit includes 1 ns allowance for PCB traces that are associated with the Skew points.

^bThe 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.

^cThe 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.

^dShould there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sub-layer 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
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
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
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
SP6	3.8	≈ 39	N/A	See 83.5.3.5
At PCS receive	4	≈ 41	N/A	See 82.2.12

^aThe 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.

^bThe 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.

^cShould there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sub-layer clause prevails.

80.6 State diagrams

State diagrams take precedence over text.

The conventions of 1.2 are adopted, along with the extensions listed in 21.5.

Multiple states of a function that have a transition to a common state utilizing different qualifiers (for example, multiple exit conditions to an IDLE or WAIT state) may be indicated by a shared arrow. An exit transition arrow must connect to the shared arrow, and the qualifier must be met prior to termination of the transition arrow on the shared arrow. The shared arrow has no qualifier.

80.7 Protocol implementation conformance statement (PICS) proforma

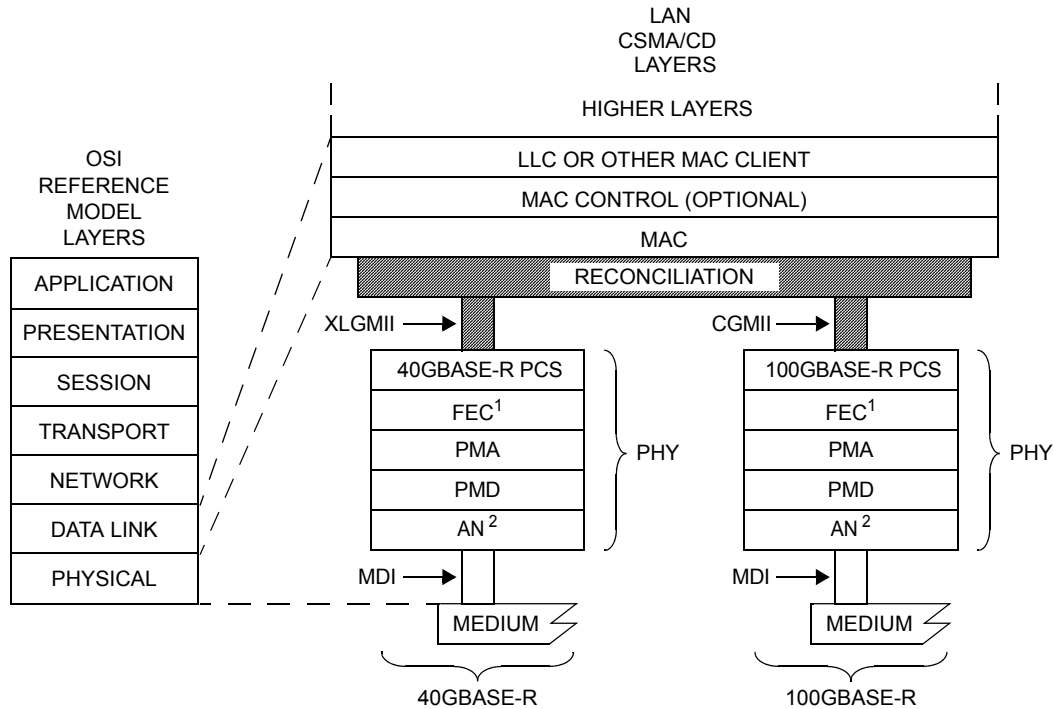
The supplier of a protocol implementation that is claimed to conform to any part of IEEE Std 802.3, Clause 45, Clause 73, Clause 74, Clause 81 through Clause 89, and related annexes demonstrates compliance by completing a protocol implementation conformance statement (PICS) proforma.

A completed PICS proforma is the PICS for the implementation in question. The PICS is a statement of which capabilities and options of the protocol have been implemented. A PICS is included at the end of each clause as appropriate. Each of the 40 Gigabit and 100 Gigabit Ethernet PICS conforms to the same notation and conventions used in 21.6.

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

81.1 Overview

This clause defines the characteristics of the Reconciliation Sublayer (RS) and the Media Independent Interface between CSMA/CD media access controllers and various PHYs. Figure 81–1 shows the relationship of the RS and Media Independent Interface to the ISO/IEC OSI reference model. Note that there are two variants of the Media Independent Interface in this clause, the 40 Gb/s Media Independent Interface (XLGMII) and the 100 Gb/s Media Independent Interface (CGMII).



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 81–1—RS and MII relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

The XLGMII and the CGMII are optional logical interfaces between the MAC sublayer and the Physical Layer (PHY).

The RS adapts the bit serial protocols of the MAC to the parallel format of the PCS service interface. Though the XLGMII/CGMII is an optional interface, it is used in this standard as a basis for specification. The Physical Coding Sublayer (PCS) is specified to the XLGMII/CGMII, so if not implemented, a conforming implementation shall behave functionally as if the RS and XLGMII/CGMII were implemented.

The XLGMII/CGMII has the following characteristics:

- a) The XLGMII supports a speed of 40 Gb/s.
- b) The CGMII supports a speed of 100 Gb/s.
- c) Data and delimiters are synchronous to a clock reference.
- d) It provides independent 64-bit wide transmit and receive data paths.
- e) It supports full duplex operation only.

81.1.1 Summary of major concepts

The following are the major concepts of the XLGMII/CGMII:

- a) The XLGMII/CGMII is functionally similar to other media independent interfaces that have been defined for lower speeds, as they all define an interface allowing independent development of MAC and PHY logic.
- b) The RS converts between the MAC serial data stream and the parallel data paths of the XLGMII/CGMII.
- c) The RS maps the signal set provided at the XLGMII/CGMII to the PLS service primitives provided at the MAC.
- d) Each direction of data transfer is independent and serviced by data, control, and clock signals.
- e) The RS generates continuous data or control characters on the transmit path and expects continuous data or control characters on the receive path.
- f) The RS participates in link fault detection and reporting by monitoring the receive path for status reports that indicate an unreliable link, and generating status reports on the transmit path to report detected link faults to the DTE on the remote end of the connecting link.

81.1.2 Application

This clause applies to the interface between the MAC and PHY. This logical interface is used to provide media independence so that an identical media access controller may be used with supported PHY types.

81.1.3 Rate of operation

The XLGMII is specified to support 40 Gb/s operation and the CGMII is specified to support 100 Gb/s operation.

81.1.4 Delay constraints

The maximum cumulative MAC Control, MAC, and RS delay (sum of transmit and receive delays at one end of the link) shall meet the values specified in Table 81–1. 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.

Table 81–1—Delay constraints

Sublayer	Maximum (bit time)	Maximum (pause_quanta)	Maximum (ns)
40 Gb/s MAC, RS, and MAC Control	16384	32	409.6
100 Gb/s MAC, RS, and MAC Control	24576	48	245.76

81.1.5 Allocation of functions

The allocation of functions at the XLGMII/CGMII balances the need for media independence with interface simplicity. The XLGMII and CGMII maximize media independence by separating the Data Link and Physical Layers of the OSI seven-layer reference model.

81.1.6 XLGMII/CGMII structure

The XLGMII/CGMII is composed of independent transmit and receive paths. Each direction uses 64 data signals (TXD<63:0> and RXD<63:0>), 8 control signals (TXC<7:0> and RXC<7:0>), and a clock (TX_CLK and RX_CLK). Figure 81–2 depicts a schematic view of the RS inputs and outputs.

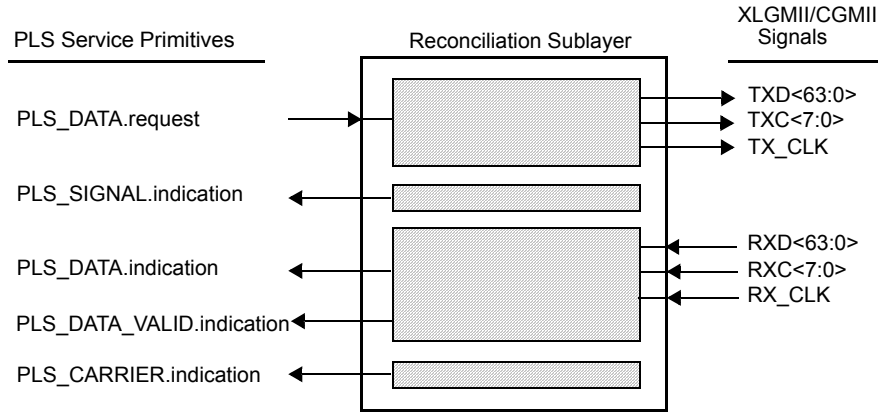


Figure 81–2—Reconciliation Sublayer (RS) inputs and outputs

The 64 TXD and 8 TXC signals shall be organized into eight data lanes, as shall the 64 RXD and 8 RXC signals (see Table 81–2). The eight lanes in each direction share a common clock, TX_CLK for transmit and RX_CLK for receive. The eight lanes are used in round-robin sequence to carry an octet stream. On transmit, each eight PLS_DATA.request transactions represent an octet transmitted by the MAC. The first octet is aligned to lane 0, the second to lane 1, the third to lane 2, the fourth to lane 3, the fifth to lane 4, the sixth to lane 5, the seventh to lane 6 and the eighth to lane 7, then repeating with the ninth to lane 0, etc. Delimiters and interframe idle characters are encoded on the TXD and RXD signals with the control code indicated by assertion of TXC and RXC, respectively.

Table 81–2—Transmit and receive lane associations

TXD, RXD	TXC, RXC	Lane
<7:0>	<0>	0
<15:8>	<1>	1
<23:16>	<2>	2
<31:24>	<3>	3
<39:32>	<4>	4
<47:40>	<5>	5
<55:48>	<6>	6
<63:56>	<7>	7

81.1.7 Mapping of XLGMII/CGMII signals to PLS service primitives

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.

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

81.1.7.1 Mapping of PLS_DATA.request

81.1.7.1.1 Function

The RS maps the primitive PLS_DATA.request to the XLGMII/CGMII signals TXD<63:0>, TXC<7:0>, and TX_CLK.

81.1.7.1.2 Semantics of the service primitive

PLS_DATA.request(OUTPUT_UNIT)

The OUTPUT_UNIT parameter can take one of three values: one, zero, or DATA_COMPLETE. One or zero represents a single data bit. The DATA_COMPLETE value signifies that the MAC sublayer has no more data to output.

81.1.7.1.3 When generated

This primitive is generated by the MAC sublayer to request the transmission of a single data bit on the physical medium or to stop transmission.

81.1.7.1.4 Effect of receipt

The OUTPUT_UNIT values are conveyed to the PHY by the signals TXD<63:0> and TXC<7:0> on each TX_CLK rising edge. Each PLS_DATA.request transaction shall be mapped to a TXD signal in sequence (TXD<0>, TXD<1>, ... TXD<63>, TXD<0>) as described in 81.2, and for every eight transactions, a TXC signal in sequence (TXC<0>, TXC<1>, ... TXC<7>, TXC<0>) is generated, as described in 81.3.1.2. After 64 PLS_DATA.request transactions from the MAC sublayer (eight octets of eight PLS_DATA.request transactions each), the RS requests transmission of 64 data bits by the PHY. The first octet of preamble shall be converted to a Start control character and aligned to lane 0. The TXD<63:0> and TXC<7:0> shall be generated by the RS every 64 bit-times of the MAC sublayer.

The DATA_COMPLETE value shall be mapped to a Terminate control character encoded on the next eight TXD signals in sequence after the last data octet; and is transferred to the PHY at the next TX_CLK rising edge. This may be on the same TX_CLK rising edge as the last data octet or the subsequent TX_CLK rising edge. When the Terminate control character is in lane 0, 1, 2, 3, 4, 5, or 6, the lanes following in sequence are encoded with an Idle control character.

81.1.7.2 Mapping of PLS_DATA.indication**81.1.7.2.1 Function**

The RS maps the XLGMII/CGMII signals RXD<63:0>, RXC<7:0>, and RX_CLK to the primitive PLS_DATA.indication.

81.1.7.2.2 Semantics of the service primitive

PLS_DATA.indication (INPUT_UNIT)

The INPUT_UNIT parameter can take one of two values: one or zero. It represents a single data bit.

81.1.7.2.3 When generated

The INPUT_UNIT values are derived from the signals RXC<7:0> and RXD<63:0> received from the PHY on each rising edge of RX_CLK. Each primitive generated to the MAC sublayer entity corresponds to a PLS_DATA.request issued by the MAC at the remote end of the link connecting two DTEs. For each RXD<63:0> during frame reception, the RS shall generate 64 PLS_DATA.indication transactions until the end of frame (Terminate control character), where 0, 8, 16, 24, 32, 40, 48, or 56 PLS_DATA.indication transactions will be generated from the RXD<63:0> containing the Terminate. During frame reception, each RXD signal shall be mapped in sequence into a PLS_DATA.indication transaction (RXD<0>, RXD<1>, ... RXD<63>, RXD<0>) as described in 81.2.

The RS shall convert a valid Start control character to a preamble octet prior to generation of the associated PLS_DATA.indication transactions. The RS shall not generate any PLS_DATA.indication primitives for a Terminate control character. To assure robust operation, the value of the data transferred to the MAC may be changed by the RS as required by XLGMII/CGMII error indications (see 81.3.3). Sequence ordered_sets are not indicated to the MAC (see 81.3.4).

81.1.7.2.4 Effect of receipt

The effect of receipt of this primitive by the MAC sublayer is unspecified.

81.1.7.3 Mapping of PLS_CARRIER.indication

40 Gb/s and 100 Gb/s operation supports full duplex operation only. The RS never generates this primitive.

81.1.7.4 Mapping of PLS_SIGNAL.indication

40 Gb/s and 100 Gb/s operation supports full duplex operation only. The RS never generates this primitive.

81.1.7.5 Mapping of PLS_DATA_VALID.indication**81.1.7.5.1 Function**

The RS maps the XLGMII/CGMII signals RXC<7:0> and RXD<63:0> to the primitive PLS_DATA_VALID.indication.

81.1.7.5.2 Semantics of the service primitive

PLS_DATA_VALID.indication (DATA_VALID_STATUS)

The `DATA_VALID_STATUS` parameter can take one of two values: `DATA_VALID` or `DATA_NOT_VALID`. The `DATA_VALID` value indicates that the `INPUT_UNIT` parameter of the `PLS_DATA.indication` primitive contains valid data of an incoming frame. The `DATA_NOT_VALID` value indicates that the `INPUT_UNIT` parameter of the `PLS_DATA.indication` primitive does not contain valid data of an incoming frame.

81.1.7.5.3 When generated

The `PLS_DATA_VALID.indication` service primitive shall be generated by the RS whenever the `DATA_VALID_STATUS` parameter changes from `DATA_VALID` to `DATA_NOT_VALID` or vice versa.

`DATA_VALID_STATUS` shall assume the value `DATA_VALID` when a `PLS_DATA.indication` transaction is generated in response to reception of a Start control character on lane 0 if the prior `RXC<7:0>` and `RXD<63:0>` contained eight Idle characters, a Sequence ordered set, or a Terminate character. `DATA_VALID_STATUS` shall assume the value `DATA_NOT_VALID` when `RXC` of the current lane in sequence is asserted for anything except an Error control character. In the absence of errors, `DATA_NOT_VALID` is caused by a Terminate control character. When `DATA_VALID_STATUS` changes from `DATA_VALID` to `DATA_NOT_VALID` because of a control character other than Terminate, the RS shall ensure that the MAC will detect a `FrameCheckError` prior to indicating `DATA_NOT_VALID` to the MAC (see 81.3.3.1).

81.1.7.5.4 Effect of receipt

The effect of receipt of this primitive by the MAC sublayer is unspecified.

81.2 XLGMII/CGMII data stream

Packets transmitted through the XLGMII/CGMII shall be transferred within the XLGMII/CGMII data stream. A data stream is a sequence of bytes, where each byte conveys either a data octet or control character. The parts of the data stream are shown in Figure 81–3.

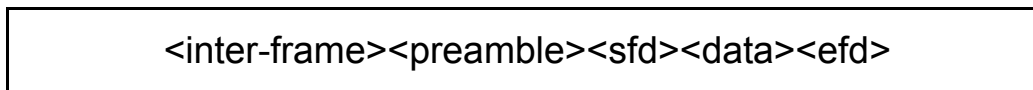


Figure 81–3—XLGMII/CGMII data stream

For the XLGMII/CGMII, transmission and reception of each bit and mapping of data octets to lanes shall be as shown in Figure 81–4.

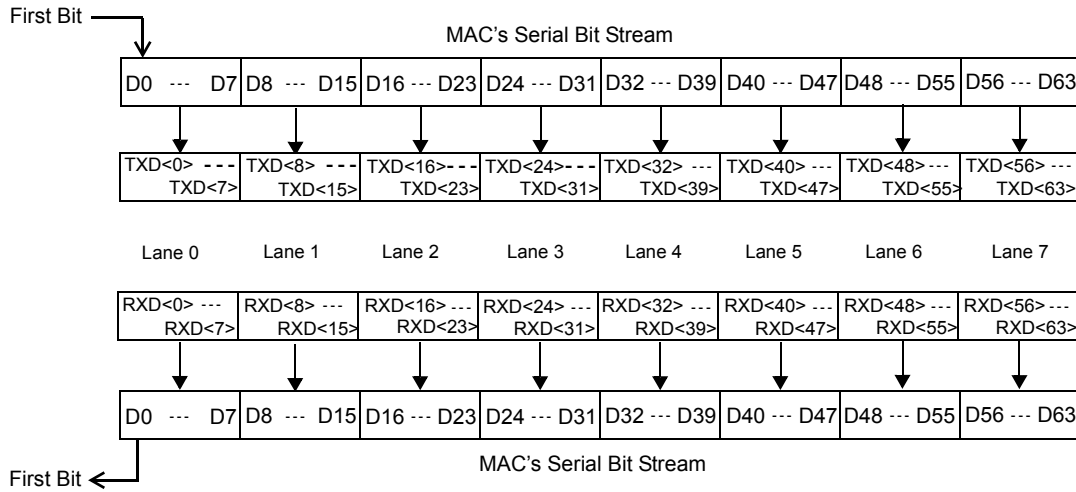


Figure 81–4—Relationship of data lanes to MAC serial bit stream

81.2.1 Inter-frame <inter-frame>

The inter-frame <inter-frame> period on an XLGMII/CGMII transmit or receive path is an interval during which no frame data activity occurs. The <inter-frame> corresponding to the MAC interpacket gap begins with the Terminate control character, continues with Idle control characters and ends with the Idle control character prior to a Start control character. The length of the interpacket gap may be changed between the transmitting MAC and receiving MAC by one or more functions (e.g., RS lane alignment or PHY clock rate compensation). The minimum IPG at the XLGMII/CGMII of the receiving RS is one octet.

The signaling of link status information logically occurs in the <inter-frame> period (see 81.3.4). Frame processing when signaling of link status information is initiated or terminated is described in 81.3.3.

81.2.2 Preamble <preamble> and start of frame delimiter <sfd>

The preamble <preamble> begins a frame transmission by a MAC as specified in 4.2.5 and when generated by a MAC consists of 7 octets with the following bit values:

10101010 10101010 10101010 10101010 10101010 10101010 10101010

The Start control character indicates the beginning of MAC data on the XLGMII/CGMII. On transmit, the RS converts the first data octet of preamble transferred from the MAC into a Start control character. On receive, the RS converts the Start control character into a preamble data octet. The Start control character is aligned to lane 0 of the XLGMII/CGMII by the RS on transmit and by the PHY on receive.

The start of frame delimiter <sfd> indicates the start of a frame and immediately follows the preamble. The bit value of <sfd> at the XLGMII/CGMII is the same as the Start Frame Delimiter (SFD) specified in 4.2.6 and is equal to the following:

10101011

81.3.1.2 TXC<7:0> (transmit control)

TXC<7:0> indicate that the RS is presenting either data or control characters on the XLGMII/CGMII for transmission. The TXC signal for a lane shall be de-asserted when a data octet is being sent on the corresponding lane and asserted when a control character is being sent. In the absence of errors, the TXC signals are de-asserted by the RS for each octet of the preamble (except the first octet that is replaced with a Start control character) and remain de-asserted while all octets to be transmitted are presented on the lanes of the XLGMII/CGMII. TXC<7:0> are driven by the RS and shall transition synchronously with respect to the rising edge of TX_CLK. Table 81–3 specifies the permissible encodings of TXD and TXC for an XLGMII/CGMII transmit lane. Additional requirements apply for proper code sequences and in which lanes particular codes are valid (e.g., Start control character is to be aligned to lane 0).

Table 81–3—Permissible encodings of TXC and TXD

TXC	TXD	Description	PLS_DATA.request parameter
0	0x00 through 0xFF	Normal data transmission	Zero, one (eight bits)
1	0x00 through 0x06	Reserved	—
1	0x07	Idle	No applicable parameter (normal inter-frame)
1	0x08 through 0x9B	Reserved	—
1	0x9C	Sequence (only valid in lane 0)	No applicable parameter (inter-frame status signal)
1	0x9D through 0xFA	Reserved	—
1	0xFB	Start (only valid in lane 0)	No applicable parameter, replaces first eight zero, one of a frame (preamble octet)
1	0xFC	Reserved	—
1	0xFD	Terminate	DATA_COMPLETE
1	0xFE	Error	No applicable parameter
1	0xFF	Reserved	—

81.3.1.3 TXD<63:0> (transmit data)

TXD is a bundle of 64 data signals organized into eight lanes of eight signals each (TXD<7:0>, TXD<15:8>, TXD<23:16>, TXD<31:24>, TXD<39:32>, TXD<47:40>, TXD<55:48>, and TXD<63:56>) that are driven by the RS. Each lane is associated with a TXC signal as shown in Table 81–2 and shall be encoded as shown in Table 81–3. TXD<63:0> shall transition synchronously with respect to the rising edge of TX_CLK. For each high TX_CLK transition, data and/or control are presented on TXD<63:0> to the PHY for transmission. TXD<0> is the least significant bit of lane 0, TXD<8> the least significant bit of lane 1, TXD<16> the least significant bit of lane 2, TXD<24> the least significant bit of lane 3, TXD<32> the least significant bit of lane 4, TXD<40> the least significant bit of lane 5, TXD<48> the least significant bit of lane 6, and TXD<56> the least significant bit of lane 7.

Assertion on a lane of appropriate TXD values when TXC is asserted will cause the PHY to generate code-groups associated with either Idle, Start, Terminate, Sequence, or Error control characters. While the TXC of a lane is de-asserted, TXD of the lane is used to request the PHY to generate code-groups corresponding to the data octet value of TXD. An example of normal frame transmission is illustrated in Figure 81–5.

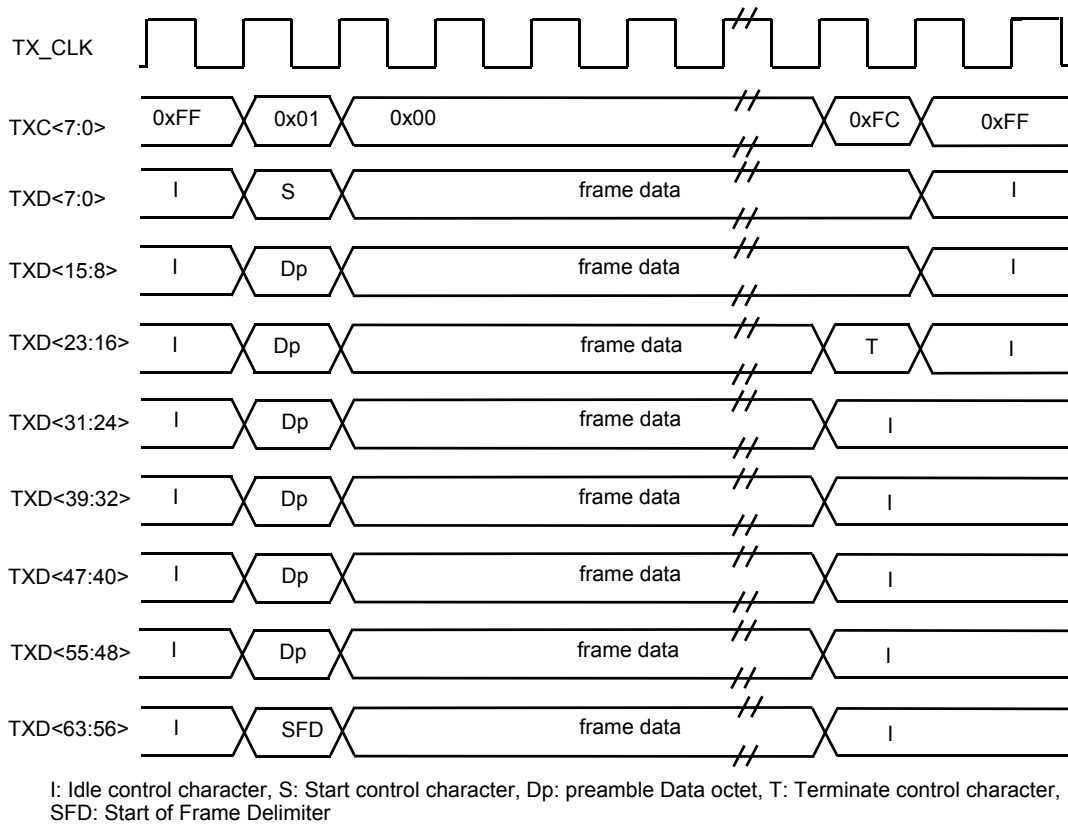
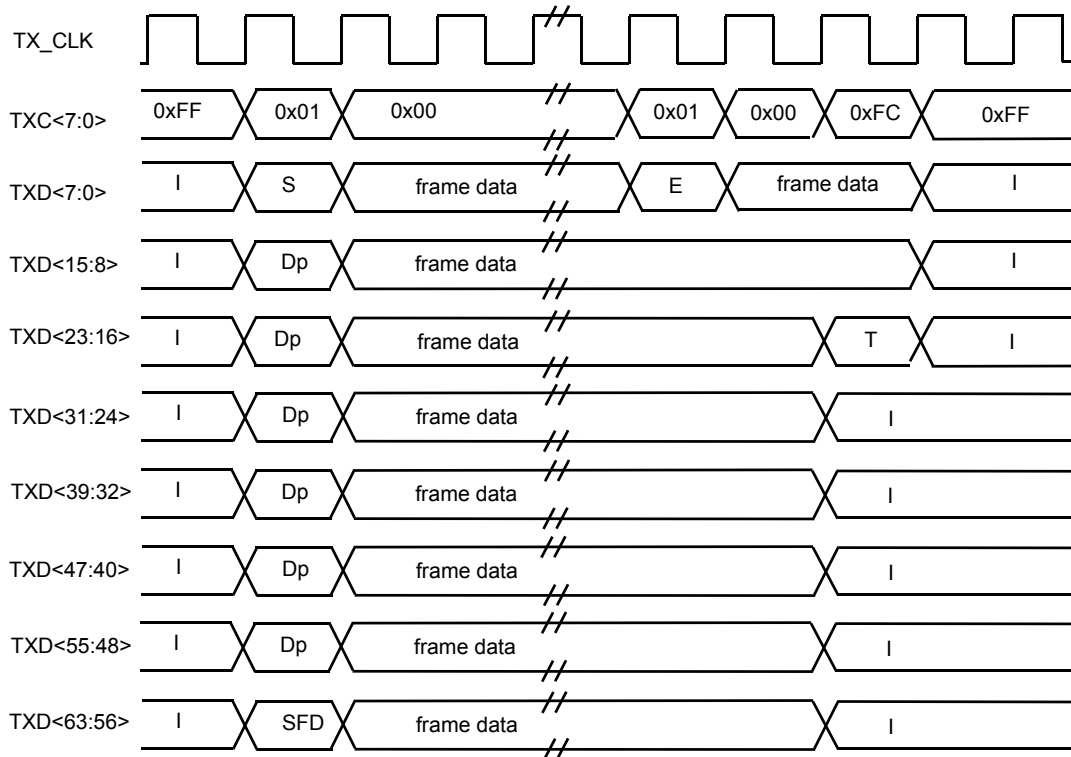


Figure 81-5—Normal frame transmission

Figure 81–6 shows the behavior of TXD and TXC during an example transmission of a frame propagating an error.



I: Idle control character, S: Start control character, Dp: preamble Data octet, T: Terminate control character, E: Error control character, SFD: Start of Frame Delimiter

Figure 81–6—Transmit Error Propagation

81.3.1.4 Start control character alignment

On transmit, it may be necessary for the RS to modify the length of the <inter-frame> in order to align the Start control character (first octet of preamble) on lane 0. This shall be accomplished in one of the following two ways:

- A MAC implementation may incorporate this RS function into its design and always insert additional idle characters to align the start of preamble on an eight byte boundary. Note that this will reduce the effective data rate for certain packet sizes separated with minimum inter-frame spacing.
- Alternatively, the RS may maintain the effective data rate by sometimes inserting and sometimes deleting idle characters to align the Start control character. When using this method the RS must maintain a Deficit Idle Count (DIC) that represents the cumulative count of idle characters deleted or inserted. The DIC is incremented for each idle character deleted, decremented for each idle character inserted, and the decision of whether to insert or delete idle characters is constrained by bounding the DIC to a minimum value of zero and maximum value of seven. Note that this may result in inter-frame spacing observed on the transmit XLGMII/CGMII that is up to seven octets shorter than the minimum transmitted inter-frame spacing specified in Clause 4; however, the frequency of shortened inter-frame spacing is constrained by the DIC rules. The DIC is only reset at

initialization and is applied regardless of the size of the IPG transmitted by the MAC sublayer. An equivalent technique may be employed to control RS alignment of the Start control character provided that the result is the same as if the RS implemented the DIC as described.

81.3.2 Receive

81.3.2.1 RX_CLK (receive clock)

RX_CLK is a continuous clock that provides the timing reference for the transfer of the RXC<7:0> and RXD<63:0> signals from the PHY to the RS. RXC<7:0> and RXD<63:0> shall be sampled by the RS on the rising edge of RX_CLK. RX_CLK is sourced by the PHY.

The frequency of RX_CLK may be derived from the received data or it may be that of a nominal clock (e.g., TX_CLK). When the received data rate at the PHY is within tolerance, the RX_CLK frequency shall be one-sixty-fourth of the MAC receive data rate.

There is no need to transition between the recovered clock reference and a nominal clock reference on a frame-by-frame basis. If loss of received signal from the medium causes a PHY to lose the recovered RX_CLK reference, the PHY shall source the RX_CLK from a nominal clock reference.

NOTE—This standard neither requires nor assumes a guaranteed phase relationship between the RX_CLK and TX_CLK signals.

81.3.2.2 RXC<7:0> (receive control)

RXC<7:0> indicate that the PHY is presenting either recovered and decoded data or control characters on the XLGMII/CGMII. The RXC signal for a lane shall be de-asserted when a data octet is being received on the corresponding lane and asserted when a control character is being received. In the absence of errors, the RXC signals are de-asserted by the PHY for each octet of the preamble (except the first octet that is replaced with a Start control character) and remain de-asserted while all octets to be received are presented on the lanes of the XLGMII/CGMII. RXC<7:0> are driven by the PHY and shall transition synchronously with respect to the rising edge of RX_CLK. Table 81–4 specifies the permissible encodings of RXD and RXC for an XLGMII/CGMII receive lane. Additional requirements apply for proper code sequences and in which lanes particular codes are valid (e.g., Start control character is to be aligned to lane 0).

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

RXC	RXD	Description	PLS_DATA.indication parameter
0	0x00 through 0xFF	Normal data reception	Zero, one (eight bits)
1	0x00 through 0x06	Reserved	—
1	0x07	Idle	No applicable parameter (Normal inter-frame)
1	0x08 through 0x9B	Reserved	—
1	0x9C	Sequence (only valid in lane 0)	No applicable parameter (Inter-frame status signal)
1	0x9D through 0xFA	Reserved	—
1	0xFB	Start (only valid in lane 0)	No applicable parameter, first eight zero, one of a frame (a preamble octet)
1	0xFC	Reserved	—
1	0xFD	Terminate	No applicable parameter (start of inter-frame)
1	0xFE	Error	No applicable parameter
1	0xFF	Reserved	—

Figure 81–7 shows the behavior of RXC<7:0> during an example frame reception with no errors.

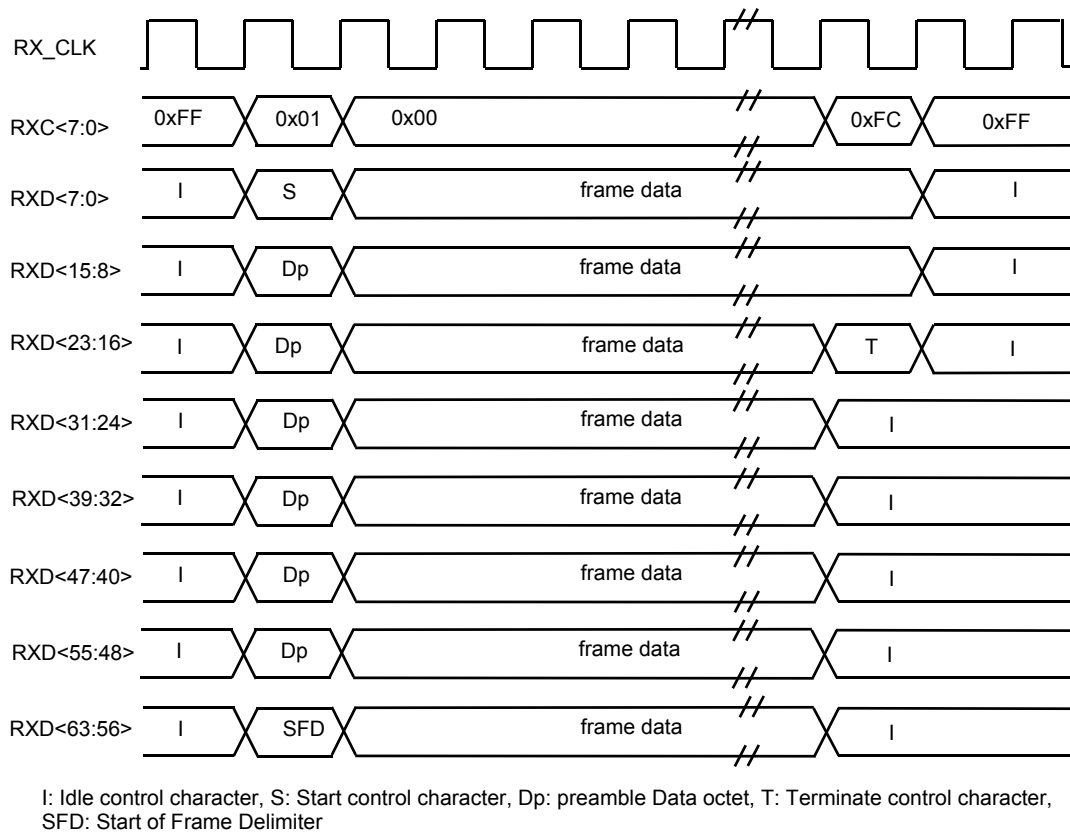


Figure 81–7—Frame reception without error

81.3.2.3 RXD (receive data)

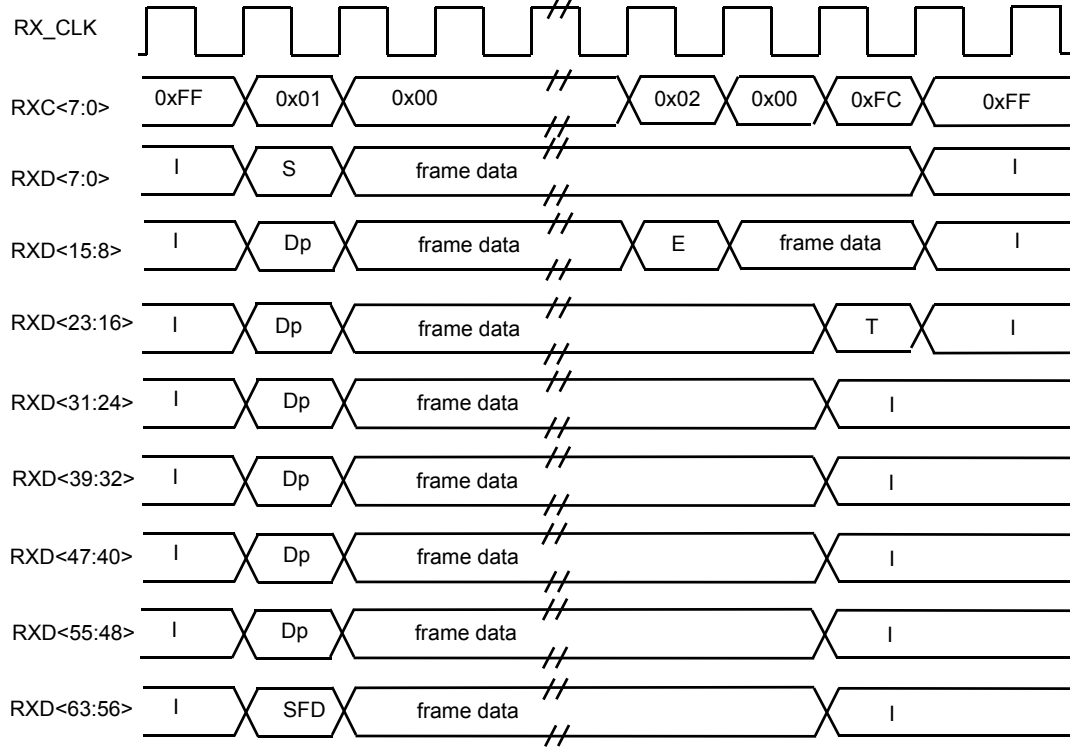
RXD is a bundle of 64 data signals (RXD<63:0>) organized into eight lanes of eight signals each (RXD<7:0>, RXD<15:8>, RXD<23:16>, RXD<31:24>, RXD<39:32>, RXD<47:40>, RXD<55:48>, and RXD<63:56>) that are driven by the PHY. Each lane is associated with an RXC signal as shown in Table 81–2 and shall be decoded by the RS as shown in Table 81–4. RXD<63:0> shall transition synchronously with respect to the rising edge of RX_CLK. For each rising RX_CLK transition, received data and/or control are presented on RXD<63:0> for mapping by the RS. RXD<0> is the least significant bit of lane 0, RXD<8> the least significant bit of lane 1, RXD<16> the least significant bit of lane 2, RXD<24> the least significant bit of lane 3, RXD<32> the least significant bit of lane 3, RXD<32> the least significant bit of lane 4, RXD<40> the least significant bit of lane 5, RXD<48> the least significant bit of lane 6, and RXD<56> the least significant bit of lane 7. Figure 81–7 shows the behavior of RXD<63:0> during frame reception.

While the RXC of a lane is de-asserted, RXD of the lane is used by the RS to generate PLS_DATA indications. Assertion on a lane of appropriate RXD values when RXC is asserted indicates to the RS the Start control character, Terminate control character, Sequence control character, or Error control character that drive its mapping functions.

RXC of a lane is asserted with the appropriate Error control character encoding on RXD of the lane to indicate that an error was detected somewhere in the frame presently being transferred from the PHY to the

RS (e.g., a coding error, or any error that the PHY is capable of detecting, and that may otherwise be undetectable at the MAC sublayer).

The effect of an Error control character on the RS is defined in 81.3.3.1. Figure 81–8 shows the behavior of RXC and RXD during the reception of an example frame with an error.



I: Idle control character, S: Start control character, Dp: preamble Data octet, T: Terminate control character, E: Error control character, SFD: Start of Frame Delimiter

Figure 81–8—Reception with error

81.3.3 Error and fault handling

81.3.3.1 Response to error indications by the XLGMII/CGMII

If, during frame reception (i.e., when DATA_VALID_STATUS = DATA_VALID), a control character other than a Terminate control character is signaled on a received lane, the RS shall ensure that the MAC will detect a FrameCheckError in that frame. This requirement may be met by incorporating a function in the RS that produces a received frame data sequence delivered to the MAC sublayer that is guaranteed to not yield a valid CRC result, as specified by the frame check sequence algorithm (see 3.2.8). This data sequence may be produced by substituting data delivered to the MAC. The RS generates eight PLS_DATA.indication primitives for each Error control character received within a frame, and may generate eight PLS_DATA.indication primitives to ensure FrameCheckError when a control character other than Terminate causes the end of the frame.

Other techniques may be employed to respond to a received Error control character provided that the result is that the MAC sublayer behaves as though a FrameCheckError occurred in the received frame.

81.3.3.2 Conditions for generation of transmit Error control characters

If, during the process of transmitting a frame, it is necessary to request that the PHY deliberately corrupt the contents of the frame in such a manner that a receiver will detect the corruption with the highest degree of probability, then an Error control character may be asserted on a transmit lane by the appropriate encoding of the lane's TXD and TXC signals.

81.3.3.3 Response to received invalid frame sequences

The RS shall not indicate DATA_VALID to the MAC for a Start control character received on any lane other than lane 0. Error free operation will not change the SFD alignment in lane 7. A MAC/RS implementation is not required to process a packet that has an SFD in a position other than lane 7 in the XLGMII/CGMII transfer containing the Start control character.

81.3.4 Link fault signaling

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, 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, and continuously generates Idle control characters. When the RS no longer receives fault status messages, it returns to normal operation, sending MAC data. 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.

Table 81–5—Sequence ordered_sets

Lane 0	Lane 1	Lane 2	Lane 3	Lane 4	Lane 5	Lane 6	Lane 7	Description
Sequence	0x00	0x00	0x00	0x00	0x00	0x00	0x00	Reserved
Sequence	0x00	0x00	0x01	0x00	0x00	0x00	0x00	Local Fault
Sequence	0x00	0x00	0x02	0x00	0x00	0x00	0x00	Remote Fault

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).

81.3.4.1 Variables and counters

The Link Fault Signaling state diagram uses the following variables and counters:

col_cnt

A count of the number of columns received not containing a fault_sequence. This counter increments at RX_CLK rate (on the rising clock transitions) unless reset.

fault_sequence

A new column received on RXC<7:0> and RXD<63:0> comprising a Sequence ordered_set of eight bytes and consisting of a Sequence control character in lane 0 and a seq_type in lanes 1, 2, 3, 4, 5, 6, and 7 indicating either Local Fault or Remote Fault.

last_seq_type

The seq_type of the previous Sequence ordered_set received

Values:Local Fault; 0x00 in lane 1, 0x00 in lane 2, 0x01 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

link_fault

An indicator of the fault status.

Values:OK; No fault.

Local Fault; fault detected by the PHY.

Remote Fault; fault detection signaled by the remote RS.

reset

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 and has not been reset (default).

TRUE: The device has not been completely powered or has been reset.

seq_cnt

A count of the number of received Sequence ordered_sets of the same type.

seq_type

The value received in the current Sequence ordered_set

Values:Local Fault; 0x00 in lane 1, 0x00 in lane 2, 0x01 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3, 0x00 in lane 4, 0x00 in lane 5, 0x00 in lane 6, 0x00 in lane 7.

81.3.4.2 State Diagram

The Link Fault Signaling state diagram specifies the RS monitoring of RXC<7:0> and RXD<63:0> for Sequence ordered_sets. The variable link_fault is set to indicate the value of a received Sequence ordered_set when four fault_sequences containing the same fault value have been received with fault sequence separated by less than 128 columns and no intervening fault_sequences of a different fault value. The state diagram is shown in Figure 81–9.

The variable link_fault is set to OK following any interval of 128 columns not containing a Remote Fault or Local Fault Sequence ordered_set.

The RS output onto TXC<7:0> and TXD<63:0> is controlled by the variable link_fault.

- a) link_fault = OK
The RS shall send MAC frames as requested through the PLS service interface. In the absence of MAC frames, the RS shall generate Idle control characters.
- b) link_fault = Local Fault
The RS shall continuously generate Remote Fault Sequence ordered_sets.
- c) link_fault = Remote Fault
The RS shall continuously generate Idle control characters.

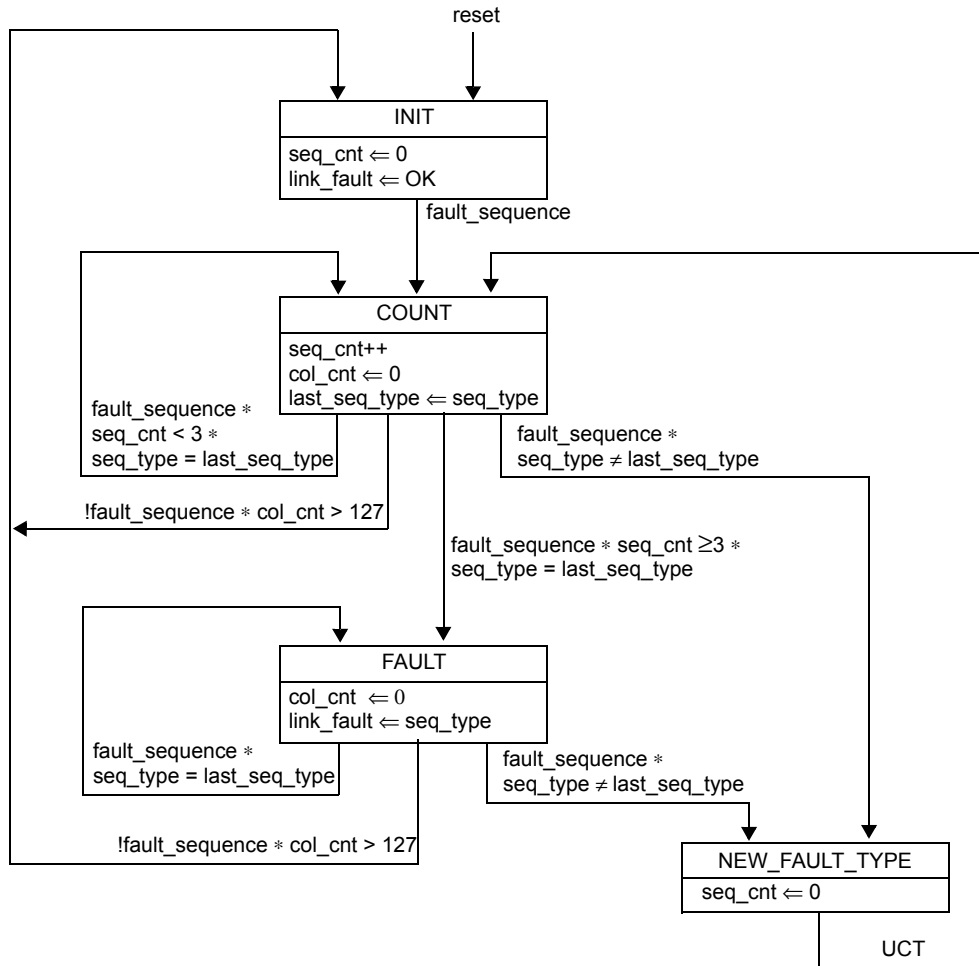


Figure 81-9—Link Fault Signaling state diagram

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 operation⁴

81.4.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 81, Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation, 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.

81.4.2 Identification

81.4.2.1 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).	

81.4.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 81, Reconciliation Sublayer (RS) and Media Independent Interface for 40 Gb/s and 100 Gb/s operation
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 802.3-2012.)	
Date of Statement	

⁴Copyright 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.

81.4.2.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*PHY40	PHY support of XLGMII	81.2, 81.3		O	Yes [] No []
*PHY100	PHY support of CGMII	81.2, 81.3		O	Yes [] No []
*RS40	Reconciliation Sublayer support of XLGMII	81.2, 81.3		O	Yes [] No []
*RS100	Reconciliation Sublayer support of CGMII	81.2, 81.3		O	Yes [] No []

81.4.3 PICS proforma tables for Reconciliation Sublayer and Media Independent Interface for 40 Gb/s and 100 Gb/s operation

81.4.3.1 General

Item	Feature	Subclause	Value/Comment	Status	Support
G1	PHY support of MAC data rate	81.1.3	Support MAC data rate of 40 Gb/s	PHY40:M	Yes [] N/A []
G2	PHY support of MAC data rate	81.1.3	Support MAC data rate of 100 Gb/s	PHY100:M	Yes [] N/A []
G3	Cumulative MAC Control, MAC and RS delay	81.1.4	Per Table 81–1 for 40 Gb/s	RS40:M	Yes [] N/A []
G4	Cumulative MAC Control, MAC and RS delay	81.1.4	Per Table 81–1 for 100 Gb/s	RS100:M	Yes [] N/A []
G5	Lane structure	81.1.6	Per Table 81–2	M	Yes []

81.4.3.2 Mapping of PLS service primitives

Item	Feature	Subclause	Value/Comment	Status	Support
PL1	Mapping to Clause 6	81.1.7	RS implements mapping to Clause 6 PLS service primitives	RS:M	Yes [] N/A []
PL2	Mapping of PLS_DATA.requests	81.1.7.1.4	In sequence TXD<0> to TXD<63>	RS:M	Yes [] N/A []
PL3	Start control character creation	81.1.7.1.4	First octet of preamble converted to Start control character	RS:M	Yes [] N/A []
PL4	TXD and TXC generation	81.1.7.1.4	For each 64 PLS_DATA.requests	RS:M	Yes [] N/A []

Item	Feature	Subclause	Value/Comment	Status	Support
PL5	Terminate control character creation	81.1.7.1.4	DATA_COMPLETE causes creation of Terminate control character in next lane in sequence	RS:M	Yes [] N/A []
PL6	Mapping RXD to PLS_DATA.indications	81.1.7.2.3	Create PLS_DATA.indications in sequence from RXD<0> to RXD<63>	RS:M	Yes [] N/A []
PL7	PLS_DATA.indication generation	81.1.7.2.3	Generate 64 PLS_DATA.indications for each RXD<0:63> until Terminate then generating 0, 8, 16, 24, 32, 40, 48, or 56	RS:M	Yes [] N/A []
PL8	Start control character conversion	81.1.7.2.3	Convert valid Start control character to preamble before generating PLS_DATA.indications	RS:M	Yes [] N/A []
PL9	Terminate control character	81.1.7.2.3	No PLS_DATA.indications generated	RS:M	Yes [] N/A []
PL10	PLS_DATA_VALID.indication generation	81.1.7.5.3	On change of value of DATA_VALID_STATUS	RS:M	Yes [] N/A []
PL11	DATA_VALID_STATUS	81.1.7.5.3	Value of DATA_VALID on a lane 0 Start control character preceded by eight idles, a Sequence ordered set, or a Terminate character	RS:M	Yes [] N/A []
PL12	DATA_VALID_STATUS	81.1.7.5.3	Value of DATA_NOT_VALID on any control character but Error	RS:M	Yes [] N/A []
PL13	Frame not ending with Terminate control character	81.1.7.5.3	Ensure MAC detects CRC error	RS:M	Yes [] N/A []

81.4.3.3 Data stream structure

Item	Feature	Subclause	Value/Comment	Status	Support
DS1	Frame transfer	81.2	Within XLGMII/CGMII data stream	RS:M	Yes [] N/A []
DS2	Bit mapping	81.2	Per Figure 81–4	RS:M	Yes [] N/A []
DS3	Content of <data>	81.2.3	Consist of data octets	RS:M	Yes [] N/A []
DS4	Recognition of <efd>	81.2.4	Terminate recognized in any lane	RS:M	Yes [] N/A []

81.4.3.4 XLGMII/CGMII signal functional specifications

Item	Feature	Subclause	Value/Comment	Status	Support
FS1	TX_CLK active edge	81.3.1.1	TXD and TXC sampled on the rising edge of TX_CLK	XGE:M	Yes [] N/A []
FS2	TX_CLK frequency	81.3.1.1	One-sixty-fourth of the MAC transmit data rate	XGE:M	Yes [] N/A []
FS3	TXC assertion and de-assertion	81.3.1.2	De-asserted for data, asserted for control character	RS:M	Yes [] N/A []
FS4	TXC clock	81.3.1.2	Synchronous to TX_CLK	XGE:M	Yes [] N/A []
FS5	TXD encoding	81.3.1.3	Per Table 81–3	RS:M	Yes [] N/A []
FS6	TXD clock	81.3.1.3	Synchronous to TX_CLK	XGE:M	Yes [] N/A []
FS7	Start alignment	81.3.1.4	Start control character aligned to lane 0	RS:M	Yes [] N/A []
FS8	RX_CLK active edge	81.3.2.1	RXD and RXC sampled on the rising edge of RX_CLK	XGE:M	Yes [] N/A []
FS9	RX_CLK frequency	81.3.2.1	One-sixty-fourth of the MAC receive data rate	XGE:M	Yes [] N/A []
FS10	Loss of receive signal	81.3.2.1	Source RX_CLK from nominal clock	PHY:M	Yes [] N/A []
FS11	RXC assertion and de-assertion	81.3.2.2	De-asserted for data, asserted for control character	PHY:M	Yes [] N/A []
FS12	RXC clock	81.3.2.2	Synchronous to RX_CLK	XGE:M	Yes [] N/A []
FS13	RXD decoding	81.3.2.3	Per Table 81–4	RS:M	Yes [] N/A []
FS14	RXD clock	81.3.2.3	Synchronous to RX_CLK	XGE:M	Yes [] N/A []
FS15	Received Error control character	81.3.3.1	RS cause MAC FrameCheck-Error	RS:M	Yes [] N/A []
FS16	DATA_VALID assertion	81.3.3.3	RS not assert DATA_VALID unless Start control character in lane 0	RS:M	Yes [] N/A []

81.4.3.5 Link fault signaling state diagram

Item	Feature	Subclause	Value/Comment	Status	Support
LF1	Link fault signaling state diagram	81.3.4	Implement per Figure 81-9	RS:M	Yes [] N/A []
LF2	link_fault = OK and MAC frames	81.3.4.2	RS services MAC frame transmission requests	RS:M	Yes [] N/A []
LF3	link_fault = OK and no MAC frames	81.3.4.2	In absence of MAC frames, RS transmits Idle control characters	RS:M	Yes [] N/A []
LF4	link_fault = Local Fault	81.3.4.2	RS transmits continuous Remote Fault Sequence ordered_sets	RS:M	Yes [] N/A []
LF5	link_fault = Remote Fault	81.3.4.2	RS transmits continuous Idle control characters	RS:M	Yes [] N/A []

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

82.1 Overview

82.1.1 Scope

This clause specifies the Physical Coding Sublayer (PCS) 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 40GBASE-R PCS is a sublayer of the 40 Gb/s PHYs listed in Table 80–1. The 100GBASE-R PCS is a sublayer of the 100 Gb/s PHYs listed in Table 80–1. The terms 40GBASE-R and 100GBASE-R are used when referring generally to Physical Layers using the PCS defined in this clause.

Both 40GBASE-R and 100GBASE-R are based on a 64B/66B code. The 64B/66B code supports transmission of data and control characters, while maintaining robust error detection. Data distribution is introduced to support multiple lanes in the Physical Layer. Part of the distribution includes the periodic insertion of an alignment marker, which allows the receive PCS to align data from multiple lanes.

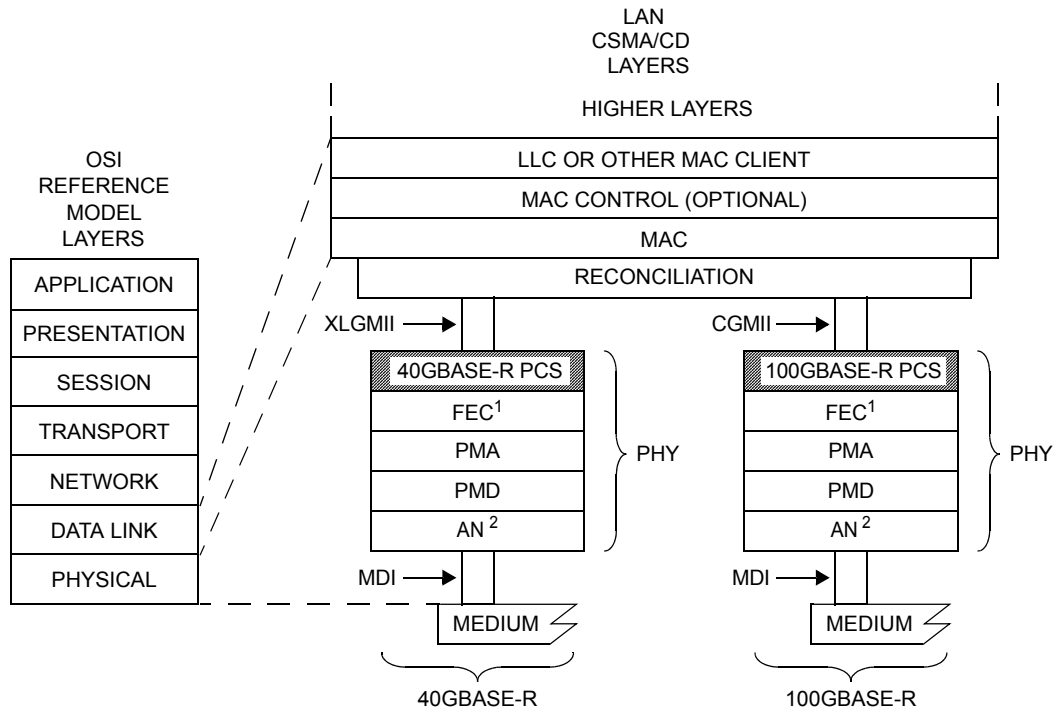
82.1.2 Relationship of 40GBASE-R and 100GBASE-R to other standards

Figure 82–1 depicts the relationships among the 40GBASE-R and 100GBASE-R sublayers (shown shaded), the Ethernet MAC and Reconciliation Sublayers, and the higher layers.

This clause borrows heavily from Clause 49. 64B/66B encoding is reused with appropriate changes made to support 8 byte alignment versus 4 byte alignment in Clause 49. In addition to 64B/66B encoding, there is a methodology to add alignment markers and distribute data to multiple lanes.

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

Figure 82–1 shows the relationship of the 40GBASE-R PCS and 100GBASE-R PCS sublayers (shown shaded) with other sublayers to the ISO Open System Interconnection (OSI) reference model.



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 82–1—40GBASE-R and 100GBASE-R PCS relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

82.1.3.1 Physical Coding Sublayer (PCS)

The PCS service interface is the Media Independent Interface (XLGMII/CGMII), which is defined in Clause 81. The 40 Gb/s variant of this is called the 40 Gb/s Media Independent Interface (XLGMII) and the 100 Gb/s variant of this interface is called the 100 Gb/s Media Independent Interface (CGMII). The XLGMII/CGMII provides a uniform interface to the Reconciliation Sublayer for all 40 Gb/s and 100 Gb/s PHY implementations.

The 40GBASE-R and 100GBASE-R PCSs provide all services required by the XLGMII/CGMII, including the following:

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

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. 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 and is an instance of the inter-sublayer service interface definition in 80.3.

82.1.5 Functional block diagram

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

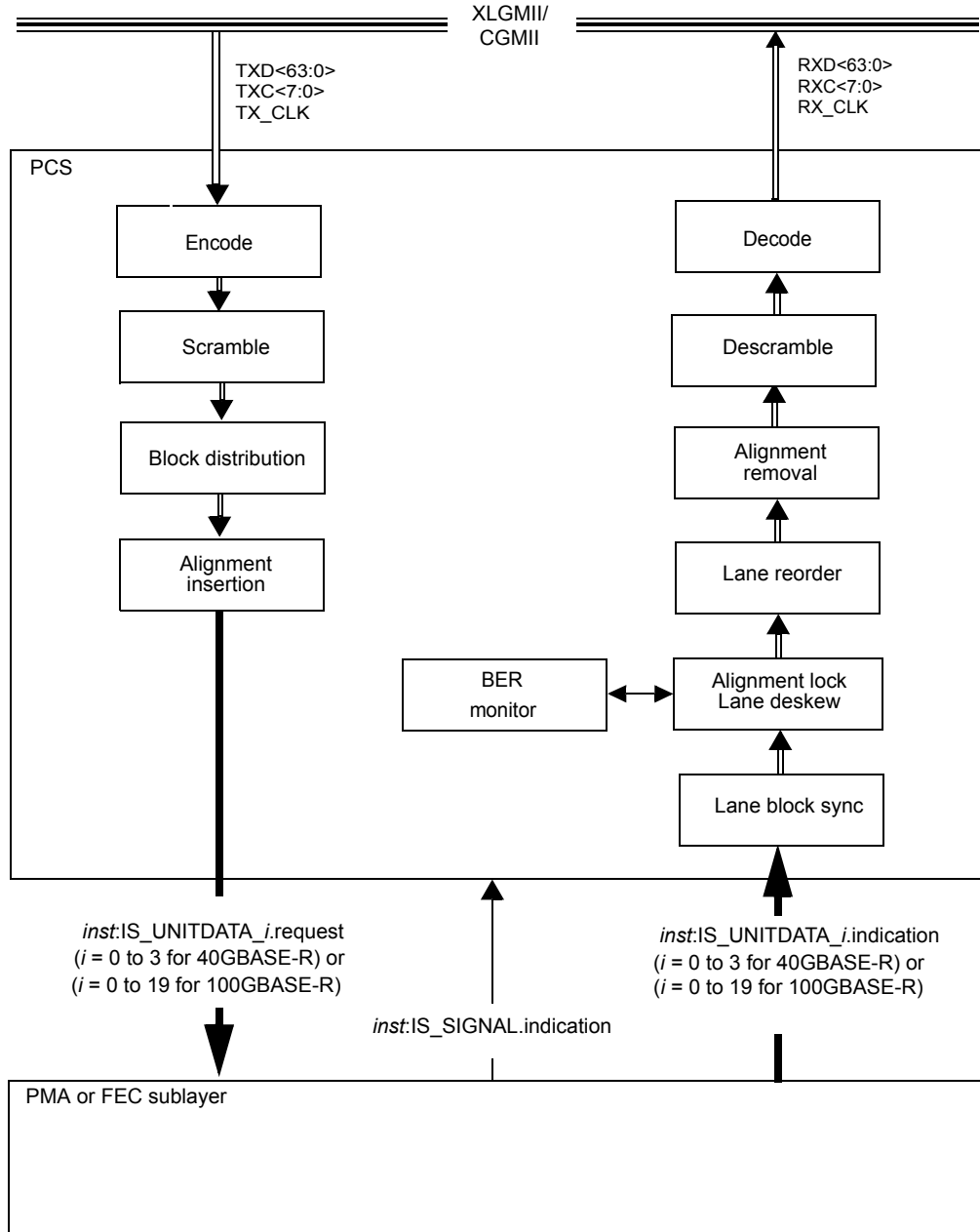


Figure 82-2—Functional block diagram

82.2 Physical Coding Sublayer (PCS)

82.2.1 Functions within the PCS

The 40GBASE-R and 100GBASE-R PCSs comprise the PCS Transmit and PCS Receive processes for each rate of operation. The PCS shields the Reconciliation Sublayer (and MAC) from the specific nature of the underlying channel. The PCS transmit channel and receive channel can each operate in normal mode or test-pattern mode.

When communicating with the XLGMII/CGMII, the PCS uses an eight octet-wide, synchronous data path, with packet delimiting being provided by transmit control signals ($TXC_{<n>} = 1$) and receive control signals ($RXC_{<n>} = 1$). When communicating with the PMA, the PCS uses multiple serial streams, 4 encoded bit streams for a 40GBASE-R PCS, or 20 encoded bit streams for a 100GBASE-R PCS. Alignment to 64B/66B blocks is performed in the PCS. The PMA sublayer operates independently of block and packet boundaries. The PCS provides the functions necessary to map packets between the XLGMII/CGMII format and the PMA/FEC service interface format.

Note that these serial streams originate from a common clock in each direction, but may vary in phase and Skew dynamically.

When the transmit channel is in normal mode, the PCS Transmit process continuously generates blocks based upon the $TXD_{<63:0>}$ and $TXC_{<7:0>}$ signals on the XLGMII/CGMII. The blocks are scrambled and then distributed to individual PCS lanes. After distribution, alignment marker blocks are periodically added to each PCS lane. Transmit data-units are sent to the service interface via the $inst:IS_UNITDATA_i.request$ primitive. In these primitives, the *inst* can be either FEC or PMA depending on which sublayer is adjacent to the PCS sublayer.

When the transmit channel is in test-pattern mode, a test pattern is packed into the transmit data-units that are sent to the PMA/FEC service interface via the $inst:IS_UNITDATA_i.request$ primitive.

When the receive channel is in normal or test-pattern mode, the PCS Synchronization process continuously monitors $inst:IS_SIGNAL.indication(SIGNAL_OK)$. When $SIGNAL_OK$ indicates OK, then the PCS Synchronization process accepts data-units via the $inst:IS_UNITDATA_i.indication$ primitive. It attains block synchronization based on the 2-bit synchronization headers on each one of the PCS lanes. Once block synchronization is found on a PCS lane, then alignment marker lock can be attained by searching for valid alignment markers. After alignment markers are found on all PCS lanes, the PCS lanes can be reordered and deskewed. Note that a particular transmit PCS lane can be received on any receive lane of the service interface due to the Skew and multiplexing that occurs in the path.

The PCS deskew process conveys received blocks to the PCS Receive process. The PCS deskew process deskews and aligns the individual PCS lanes, removes the alignment markers, forms a single stream, and sets the $align_status$ flag to indicate whether the PCS has obtained alignment.

When the PCS deskew process has obtained alignment, the BER monitor process monitors the signal quality asserting hi_ber if excessive errors are detected. When $align_status$ is asserted and hi_ber is de-asserted, the PCS Receive process continuously accepts blocks and generates $RXD_{<63:0>}$ and $RXC_{<7:0>}$ on the XLGMII/CGMII.

When the receive channel is in test-pattern mode, the BER monitor process may be disabled. The Receive process will be held in the RX_INIT state. The received bits will be compared to the test pattern and errors counted.

The PCS shall provide transmit test-pattern mode for the scrambled idle pattern (see 82.2.10), and shall provide receive test-pattern mode for the scrambled idle pattern (see 82.2.10). Test-pattern mode is activated separately for transmit and receive. The PCS shall support transmit test-pattern mode and receive test-pattern mode operating simultaneously so as to support loopback testing.

82.2.2 Use of blocks

The PCS maps XLGMII/CGMII signals into 66-bit blocks, and vice versa, using a 64B/66B coding scheme. The synchronization headers of the blocks allow establishment of block boundaries by the PCS Synchronization process. The PCS functions ENCODE and DECODE generate, manipulate, and interpret blocks as defined in 82.2.3.

82.2.3 64B/66B transmission code

The PCS uses a transmission code to improve the transmission characteristics of information to be transferred across the link and to support transmission of control and data characters. The encodings defined by the transmission code ensure that sufficient transitions are present in the PHY bit stream to make clock recovery possible at the receiver. The encoding also preserves the likelihood of detecting any single or multiple bit errors that may occur during transmission and reception of information. In addition, the synchronization headers of the code enable the receiver to achieve block alignment on the incoming PHY bit stream. The 64B/66B transmission code specified for use in this standard is a run-length-limited code.⁵

The relationship of block bit positions relative to XLGMII/CGMII, PMA, and other PCS functions is illustrated in Figure 82–3 for transmit and Figure 82–4 for receive. These figures illustrate the processing of a block containing 8 data octets. See 82.2.3.3 for information on how blocks containing control characters are mapped into 66-bit blocks. Note that the sync header is generated by the encoder and bypasses the scrambler.

82.2.3.1 Notation conventions

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

64B/66B encodes 8 data octets or control characters into a block. Blocks containing control characters also contain a block type field. Data octets are labeled D_0 to D_7 . The control characters /I/ and /E/ are labeled C_0 to C_7 . The control characters, /Q/ and /Fsig/, for ordered_sets are labeled as O_0 since they are only valid on the first octet of the XLGMII/CGMII. The control character for start is labeled as S_0 for the same reason. The control character for terminate is labeled as T_0 to T_7 . The four trailing zero data octets in ordered_sets are labeled as Z_4 to Z_7 .

One XLGMII/CGMII transfer provides eight characters that are encoded into one 66-bit transmission block. The subscript in the above labels indicates the position of the character in the eight characters from the XLGMII/CGMII transfer.

Contents of block type fields, data octets, and control characters are shown as hexadecimal values. The LSB of the hexadecimal value represents the first transmitted bit. For instance, the block type field $0x1E$ is sent from left to right as 01111000 . The bits of a transmitted or received block are labeled $TxB\langle 65:0 \rangle$ and $RxB\langle 65:0 \rangle$, respectively, where $TxB\langle 0 \rangle$ and $RxB\langle 0 \rangle$ represent the first transmitted bit. The value of the sync header is shown as a binary value. Binary values are shown with the first transmitted bit (the LSB) on the left.

82.2.3.2 Transmission order

Block bit transmission order shall be as illustrated in Figure 82–3 and Figure 82–4. Note that these figures show the mapping from XLGMII/CGMII to 64B/66B block for a block containing eight data characters.

⁵In 10GBASE-R the run length is limited to 66 bits, but in 40GBASE-R and 100GBASE-R, when multiplexing occurs in the PMA, this guaranteed run length limit increases. For example, if two PCS lanes are multiplexed in the PMA, then the possible run length would double.

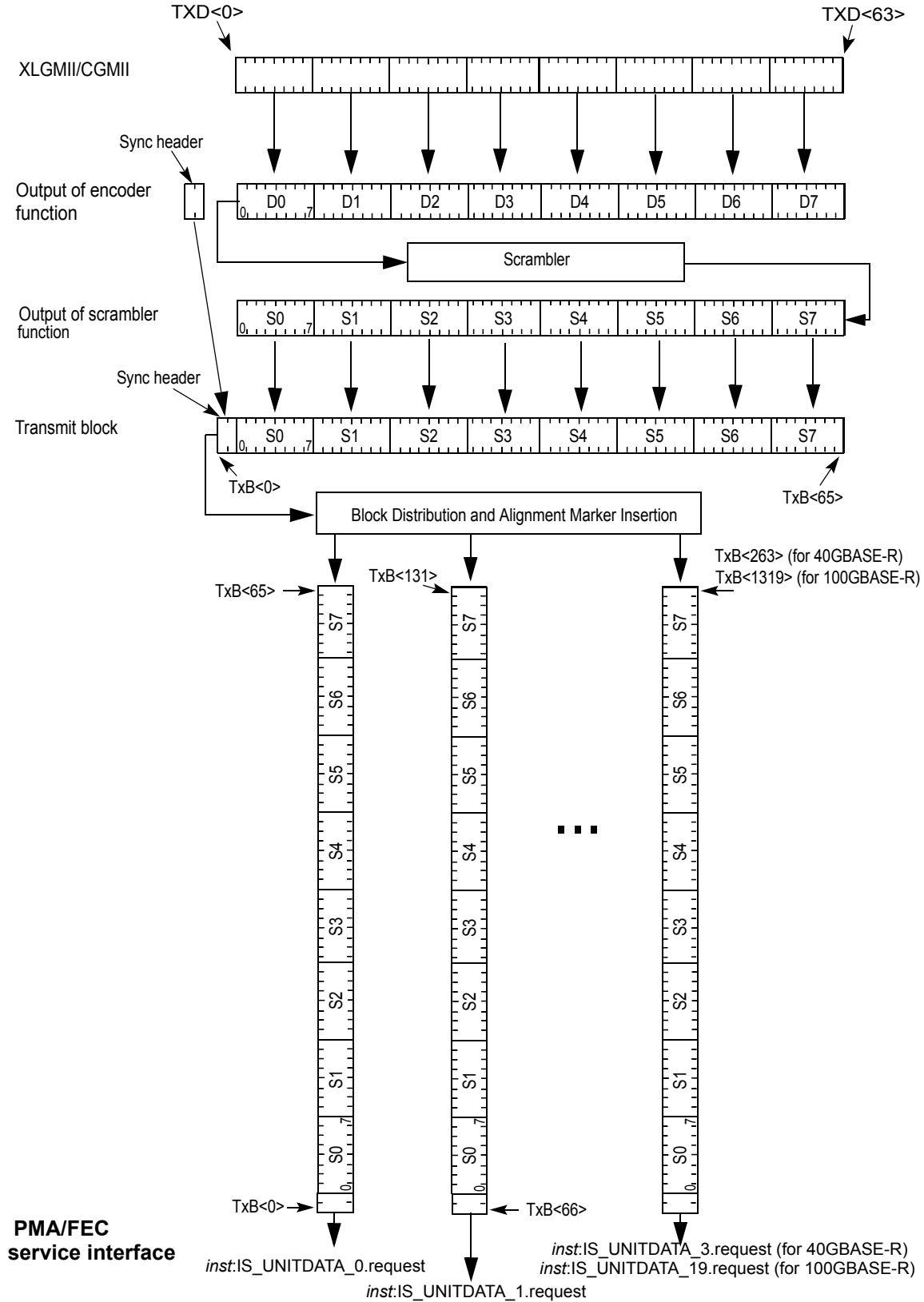


Figure 82-3—PCS Transmit bit ordering

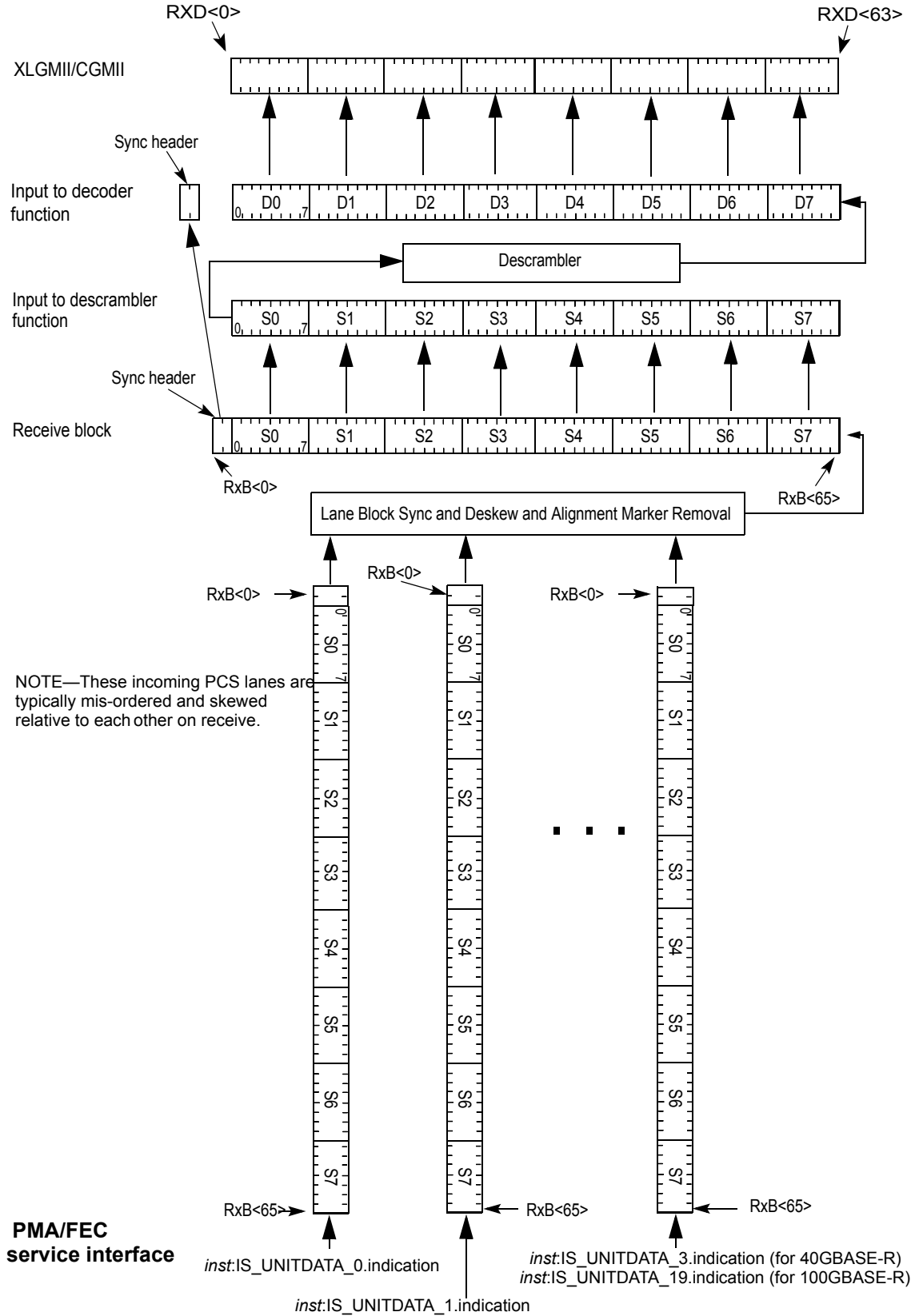


Figure 82-4—PCS Receive bit ordering

82.2.3.3 Block structure

Blocks consist of 66 bits. The first two bits of a block are the synchronization header (sync header). Blocks are either data blocks or control blocks. The sync header is 01 for data blocks and 10 for control blocks. Thus, there is always a transition between the first two bits of a block. The remainder of the block contains the payload. The payload is scrambled and the sync header bypasses the scrambler. Therefore, the sync header is the only position in the block that is always guaranteed to contain a transition. This feature of the code is used to obtain block synchronization.

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

The format of the blocks is as shown in Figure 82–5. In the figure, the column labeled Input Data shows, in abbreviated form, the eight characters used to create the 66-bit block. These characters are either data characters or control characters and, when transferred across the XLGMII/CGMII, the corresponding TXC or RXC bit is set accordingly. Within the Input Data column, D_0 through D_7 are data octets and are transferred with the corresponding TXC or RXC bit set to zero. All other characters are control characters and are transferred with the corresponding TXC or RXC bit set to one. The single bit fields (thin rectangles with no label in the figure) are sent as zero and ignored upon receipt.

All unused values of block type field⁶ are invalid; they shall not be transmitted and shall be considered an error if received.

⁶The block type field values have been chosen to have a 4-bit Hamming distance between them. There are five unused values that maintain this Hamming distance: 0x00, 0x2D, 0x33, 0x55, and 0x66.

Input Data	S y n c	Block Payload									
Bit Position:	0 1 2	65									
Data Block Format:											
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ D ₇	01	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇		
Control Block Formats:		Block Type Field									
C ₀ C ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x1E	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	
S ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ D ₇	10	0x78	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇		
O ₀ D ₁ D ₂ D ₃ Z ₄ Z ₅ Z ₆ Z ₇	10	0x4B	D ₁	D ₂	D ₃	O ₀	0x000_0000				
T ₀ C ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x87			C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
D ₀ T ₁ C ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0x99	D ₀			C ₂	C ₃	C ₄	C ₅	C ₆	C ₇
D ₀ D ₁ T ₂ C ₃ C ₄ C ₅ C ₆ C ₇	10	0xAA	D ₀	D ₁			C ₃	C ₄	C ₅	C ₆	C ₇
D ₀ D ₁ D ₂ T ₃ C ₄ C ₅ C ₆ C ₇	10	0xB4	D ₀	D ₁	D ₂			C ₄	C ₅	C ₆	C ₇
D ₀ D ₁ D ₂ D ₃ T ₄ C ₅ C ₆ C ₇	10	0xCC	D ₀	D ₁	D ₂	D ₃			C ₅	C ₆	C ₇
D ₀ D ₁ D ₂ D ₃ D ₄ T ₅ C ₆ C ₇	10	0xD2	D ₀	D ₁	D ₂	D ₃	D ₄			C ₆	C ₇
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ T ₆ C ₇	10	0xE1	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅			C ₇
D ₀ D ₁ D ₂ D ₃ D ₄ D ₅ D ₆ T ₇	10	0xFF	D ₀	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆		

Figure 82–5—64B/66B block formats

WARNING

The mapping of 40GBASE-R PCS blocks into OPU3 specified in ITU-T G.709 [B51] depends on the set of control block types shown in Figure 82–5. Any deviation from the coding specified in Figure 82–5 will break the mapping and may prevent 40GBASE-R PCS blocks from being mapped into OPU3 (see ITU-T G.709 [B51] for more details).

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.

82.2.3.5 Valid and invalid blocks

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

- a) The sync header has a value of 00 or 11.
- b) The block type field contains an invalid value (one not included in Figure 82–5).
- c) Any control character contains a value not included in Table 82–1.
- d) The set of eight XLGMII/CGMII characters does not have a corresponding block format in Figure 82–5.

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. /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.

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
start	/S/	0xFB		Encoded by block type field
terminate	/T/	0xFD		Encoded by block type field
error	/E/	0xFE		0x1E
Sequence ordered_set	/Q/	0x9C	0x0	Encoded by block type 0x4B plus O code, control codes are set to 0x00
Signal ordered_set ^a	/Fsig/	0x5C	0xF	Encoded by block type 0x4B plus O code, control codes are set to 0x00

^aReserved for INCITS T11 Fibre Channel use.

82.2.3.7 Start (/S/)

The start control character (/S/) indicates the start of a packet. This delimiter is only valid on the first octet of the XLGMII/CGMII (TXD<0:7> and RXD<0:7>). Receipt of an /S/ on any other octet of TXD indicates an error. Block type field values implicitly encode an /S/ as the first character of the block.

82.2.3.8 Terminate (/T/)

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

82.2.3.9 ordered_set (/O/)

Ordered sets are used to extend the ability to send control and status information over the link such as remote fault and local fault status. Ordered sets consist of a control character followed by three data characters followed by four zero data characters on the XLGMII/CGMII. Ordered sets always begin on the first octet of the XLGMII/CGMII. 40 Gigabit and 100 Gigabit Ethernet use one kind of ordered_set: the sequence ordered_set (see 81.3.4). An additional ordered_set, the signal ordered_set, has been reserved and it begins with another control code. The ordered_set control characters (/Q/ and /Fsig/) indicate the start of an ordered_set. The block type field plus the O code encode the specific control character for the ordered_set. See Table 82–1 for the mappings.

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

82.2.3.10 Error (/E/)

In both the 64B/66B encoder and decoder, the /E/ is generated whenever an /E/ is detected. The /E/ is also generated when invalid blocks are detected. The /E/ allows the PCS to propagate detected errors. See R_TYPE and T_TYPE function definitions in 82.2.18.2.3 for further information.

82.2.4 Transmit process

The transmit process generates blocks based upon the TXD<63:0> and TXC<7:0> signals received from the XLGMII/CGMII. One XLGMII/CGMII data transfer is encoded into one 66-bit block. It takes 66 *inst*:IS_UNITDATA_0 transfers to send a 66-bit block of data on each of the PCS lanes. The transmit process must delete idle control characters or sequence ordered sets to accommodate the transmission of alignment markers. If the PCS transmit process spans multiple clock domains, it may also perform clock rate compensation via the deletion of idle control characters or sequence ordered sets or the insertion of idle control characters.

There are sufficient idle control characters to delete in order to make room for alignment markers, in addition to handling clock compensation. Idle control characters or sequence ordered sets are removed, if necessary, to accommodate the insertion of the 66-bit alignment markers. See 82.2.3.6 for more details.

The transmit process generates blocks as specified in the transmit process state diagram. The contents of each block are contained in a vector tx_coded<65:0>, which is passed to the scrambler. tx_coded<1:0> contains the sync header and the remainder of the bits contain the block payload.

82.2.5 Scrambler

The payload, tx_coded<65:2>, is scrambled with a self-synchronizing scrambler. The scrambler is identical to the scrambler used in Clause 49, see 49.2.6 for the definition of the scrambler. The sync bits, tx_coded<1:0>, are not scrambled.

82.2.6 Block distribution

Once the data is encoded and scrambled, it is 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. This allows the PCS to support multiple physical lanes in the PMD and XLAUI or CAUI interfaces (see Annex 83A and Annex 83B). The 40GBASE-R PCS distributes the 66-bit blocks to 4 PCS lanes, and the 100GBASE-R PCS distributes the blocks to 20 PCS lanes. The distribution process is shown in Figure 82–6.

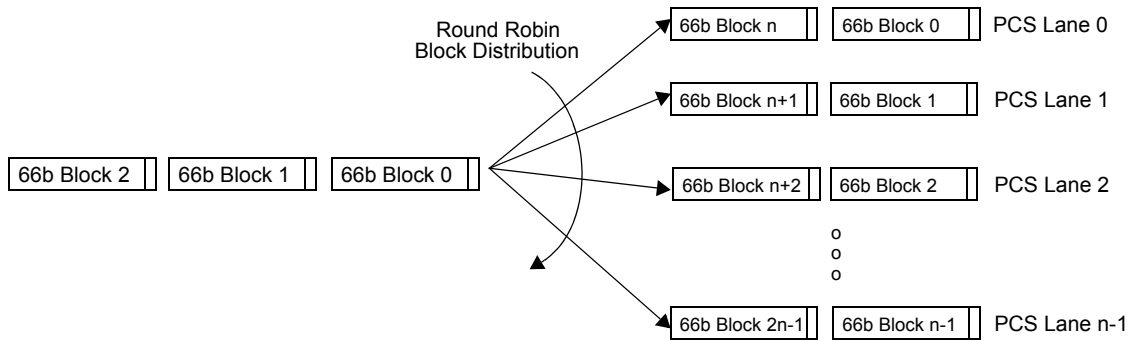


Figure 82–6—PCS Block distribution

82.2.7 Alignment marker insertion

In order to support deskew and reordering of individual PCS lanes at the receive PCS, alignment markers are added periodically to each PCS lane. The alignment marker has the form of a specially defined 66-bit block with a control block sync header. These markers interrupt any data transfer that is already in progress. This allows alignment markers to be inserted into all PCS lanes at the same time. Room for the alignment markers is created by periodically deleting IPG from the XLGMII/CGMII data stream. Other special properties of the alignment markers are that they are not scrambled and do not conform to the encoding rules as outlined in Figure 82–5. This is possible because the alignment markers are added after encoding is performed in the transmit PCS and the alignment markers are removed before 64B/66B decoding is performed in the receive PCS. The alignment markers are not scrambled in order to allow the receiver to find the alignment markers, deskew the PCS lanes, and reassemble the aggregate stream before descrambling is performed. The alignment markers themselves are formed from a known pattern that is defined to be balanced and with many transitions and therefore scrambling is not necessary for the alignment markers. The alignment markers shall be inserted after every 16383 66-bit blocks on each PCS lane. Alignment marker insertion is shown in Figure 82–7 and Figure 82–8.

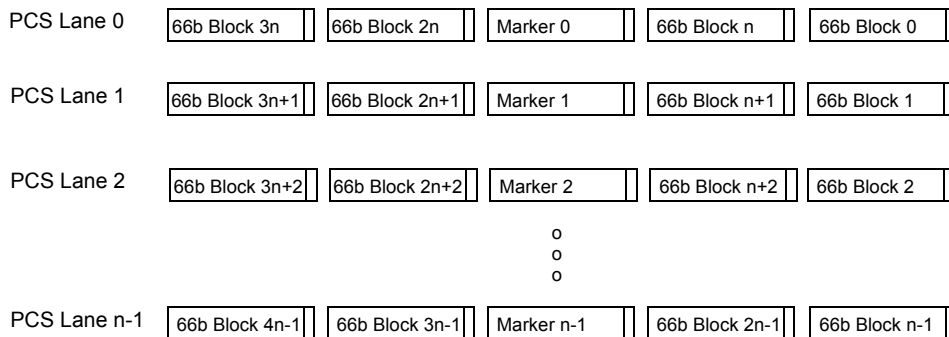


Figure 82–7—Alignment marker insertion

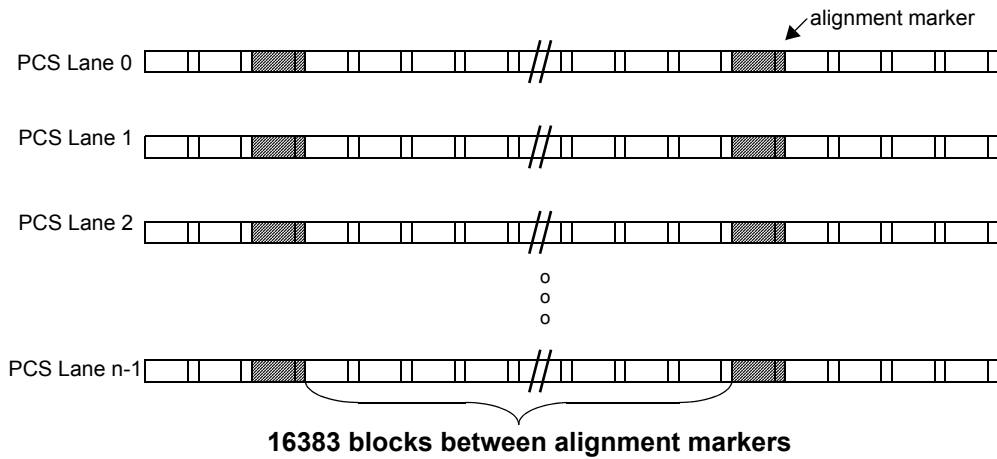


Figure 82–8—Alignment marker insertion period

The format of the alignment markers is shown in Figure 82–9.

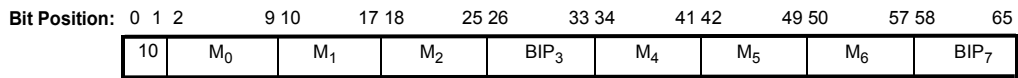


Figure 82–9—Alignment marker format

The content of the alignment markers shall be as shown in Table 82–2 for 100GBASE-R and in Table 82–3 for 40GBASE-R. The contents depend on the PCS lane number and the octet number. Note that M₄ through M₆ are the bit-wise inversion of M₀ through M₂, respectively. Also BIP₇ is the bit-wise inversion of BIP₃. 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 and lane markers 0 to 3 from Table 82–3 are used for the 40GBASE-R PCS. As an example, the lane marker for 100GBASE-R lane number 0 is sent as (left most bit sent first):

10 1000011 00010110 10000100 BIP₃ 01111100 11101001 01111011 BIP₇

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

Table 82–2—100GBASE-R Alignment marker encodings

PCS lane number	Encoding ^a {M ₀ , M ₁ , M ₂ , BIP ₃ , M ₄ , M ₅ , M ₆ , BIP ₇ }	PCS lane number	Encoding ^a {M ₀ , M ₁ , M ₂ , BIP ₃ , M ₄ , M ₅ , M ₆ , BIP ₇ }
0	0xC1, 0x68, 0x21, BIP ₃ , 0x3E, 0x97, 0xDE, BIP ₇	10	0xFD, 0x6C, 0x99, BIP ₃ , 0x02, 0x93, 0x66, BIP ₇
1	0x9D, 0x71, 0x8E, BIP ₃ , 0x62, 0x8E, 0x71, BIP ₇	11	0xB9, 0x91, 0x55, BIP ₃ , 0x46, 0x6E, 0xAA, BIP ₇
2	0x59, 0x4B, 0xE8, BIP ₃ , 0xA6, 0xB4, 0x17, BIP ₇	12	0x5C, 0xB9, 0xB2, BIP ₃ , 0xA3, 0x46, 0x4D, BIP ₇
3	0x4D, 0x95, 0x7B, BIP ₃ , 0xB2, 0x6A, 0x84, BIP ₇	13	0x1A, 0xF8, 0xBD, BIP ₃ , 0xE5, 0x07, 0x42, BIP ₇
4	0xF5, 0x07, 0x09, BIP ₃ , 0x0A, 0xF8, 0xF6, BIP ₇	14	0x83, 0xC7, 0xCA, BIP ₃ , 0x7C, 0x38, 0x35, BIP ₇
5	0xDD, 0x14, 0xC2, BIP ₃ , 0x22, 0xEB, 0x3D, BIP ₇	15	0x35, 0x36, 0xCD, BIP ₃ , 0xCA, 0xC9, 0x32, BIP ₇
6	0x9A, 0x4A, 0x26, BIP ₃ , 0x65, 0xB5, 0xD9, BIP ₇	16	0xC4, 0x31, 0x4C, BIP ₃ , 0x3B, 0xCE, 0xB3, BIP ₇
7	0x7B, 0x45, 0x66, BIP ₃ , 0x84, 0xBA, 0x99, BIP ₇	17	0xAD, 0xD6, 0xB7, BIP ₃ , 0x52, 0x29, 0x48, BIP ₇
8	0xA0, 0x24, 0x76, BIP ₃ , 0x5F, 0xDB, 0x89, BIP ₇	18	0x5F, 0x66, 0x2A, BIP ₃ , 0xA0, 0x99, 0xD5, BIP ₇
9	0x68, 0xC9, 0xFB, BIP ₃ , 0x97, 0x36, 0x04, BIP ₇	19	0xC0, 0xF0, 0xE5, BIP ₃ , 0x3F, 0x0F, 0x1A, BIP ₇

^aEach octet is transmitted LSB to MSB.

Table 82–3—40GBASE-R Alignment marker encodings

PCS lane number	Encoding ^a {M ₀ , M ₁ , M ₂ , BIP ₃ , M ₄ , M ₅ , M ₆ , BIP ₇ }
0	0x90, 0x76, 0x47, BIP ₃ , 0x6F, 0x89, 0xB8, BIP ₇
1	0xF0, 0xC4, 0xE6, BIP ₃ , 0x0F, 0x3B, 0x19, BIP ₇
2	0xC5, 0x65, 0x9B, BIP ₃ , 0x3A, 0x9A, 0x64, BIP ₇
3	0xA2, 0x79, 0x3D, BIP ₃ , 0x5D, 0x86, 0xC2, BIP ₇

^aEach octet is transmitted LSB to MSB.

82.2.8 BIP calculations

A PCS lane BIP field is carried in each PCS Lane alignment marker. This allows an accurate and fast measure of the bit error ratio of a given PCS Lane. This information is used to update error counters; no state machines use this information.

Each alignment marker has two Bit Interleaved Parity fields, BIP₃ and BIP₇. BIP₇ is a bit-wise inversion of BIP₃ in order to keep the alignment marker DC balanced. The BIP₃ field contains the result of a bit

interleaved parity calculation. Each bit in the BIP field is an even parity calculation over all of the previous specified bits of a given PCS Lane, from and including the previous alignment marker, but not including the current alignment marker. Using the bit definitions as shown in Figure 82–9, Table 82–4 has the bit assignments for each BIP₃ bit. As an example, BIP₃ bit 0 contains the result of XORing 131072 bits from 16384 66-bit words. BIP₃ bit 3 and bit 4 also include one sync header bit from each 66-bit word. Bit 3 and bit 4 each contain the result of XORing 147456 bits.

Table 82–4—BIP₃ bit assignments

BIP ₃ bit number	Assigned 66-bit word bits
0	2, 10, 18, 26, 34, 42, 50, 58
1	3, 11, 19, 27, 35, 43, 51, 59
2	4, 12, 20, 28, 36, 44, 52, 60
3	0, 5, 13, 21, 29, 37, 45, 53, 61
4	1, 6, 14, 22, 30, 38, 46, 54, 62
5	7, 15, 23, 31, 39, 47, 55, 63
6	8, 16, 24, 32, 40, 48, 56, 64
7	9, 17, 25, 33, 41, 49, 57, 65

BIP₃ and BIP₇ are transmitted LSB to MSB. As an example, with BIP₃ = 0x0F, the PCS lane marker for 100GBASE-R lane number 0 is sent as (left most bit sent first):

10 10000011 00010110 10000100 11110000 01111100 11101001 01111011 00001111

82.2.9 PMA or FEC Interface

When the transmit channel is operating in normal mode, the 40GBASE-R PCS sends four data streams via *inst:IS_UNITDATA_i.request* primitives and the 100GBASE-R PCS sends twenty data streams via *inst:IS_UNITDATA_i.request* primitives.

The *inst:IS_UNITDATA_i.request* primitives are separate serial streams of bits. Since 66-bit blocks were distributed to each lane, that means for the 40GBASE-R PCS: bits 0 to 65 are sent on PCSL 0, bits 66 to 131 are sent on PCSL 1; bits 132 to 197 are sent on PCSL 2, bits 198 to 263 are sent on PCSL 3, then bits 264 to 329 are sent on PCSL 0 etc.

For 100GBASE-R it is: bits 0 to 65 on PCSL 0, bits 66 to 131 on PCSL 1, bits 132 to 197 on PCSL 2, bits 198 to 263 on PCSL 3, bits 264 to 329 on PCSL 4, bits 330 to 395 on PCSL 5, bits 396 to 461 on PCSL 6, bits 462 to 527 on PCSL 7, bits 528 to 593 on PCSL 8, bits 594 to 659 on PCSL 9, bits 660 to 725 on PCSL 10, bits 726 to 791 on PCSL 11, bits 792 to 857 on PCSL 12, bits 858 to 923 on PCSL 13, bits 924 to 989 on PCSL 14, bits 990 to 1055 on PCSL 15, bits 1056 to 1121 on PCSL 16, bits 1122 to 1187 on PCSL 17, bits 1188 to 1253 on PCSL 18, bits 1254 to 1319 on PCSL 19, then bits 1320 to 1385 on PCSL 0, etc.

82.2.10 Test-pattern generators

The PCS shall have the ability to generate and detect a scrambled idle test pattern. This test-pattern mode is suitable for receiver tests and for certain transmitter tests.

When a scrambled idle pattern is enabled, the test pattern is generated by the scrambler. No seeding of the scrambler is required during test-pattern operation. The input to the scrambler is a control block (block type=0x1E) with all idles as defined in Figure 82–5. Note that the sync headers and alignment markers are added to the stream so that the receive PCS can align and deskew the PCS lanes.

If a Clause 45 MDIO is implemented, then control of the test-pattern generation is from the BASE-R PCS test-pattern control register (bit 3.42.3).

When the transmit channel is operating in test-pattern mode, the encoded bit stream is distributed to the PCS Lanes as in normal operation (see 82.2.4).

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.

82.2.12 PCS lane deskew

Once the receiver achieves block lock on a lane, then it begins obtaining alignment marker lock as specified in 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 lanes (4 or 20 lanes), then all inter-lane Skew is removed as shown in the PCS deskew state diagram in Figure 82–12. The Skew budget that the PCS receiver shall support is shown in Table 82–5.

Note that Skew is defined in 80.5.

Table 82–5—Skew tolerance requirements

PCS	Maximum Skew	Maximum Skew Variation
40GBASE-R	180 ns (~1856 bits)	4ns (~41 bits)
100GBASE-R	180 ns (~928 bits)	4ns (~21 bits)

82.2.13 PCS lane reorder

Transmit 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 receive PCS shall order the received PCS lanes according to the PCS lane number.

82.2.14 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 deleted from the data stream. The difference in rate from the deleted alignment markers is compensated for by inserting idle control characters by a function in the Receive process. Note that an alignment marker is always deleted when a given PCS Lane is in `am_lock=true` 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 delete the block in the alignment marker position.

As part of the alignment marker removal process, the `BIP3` field is compared to the calculated BIP value for each PCS lane. If a Clause 45 MDIO is implemented, then the appropriate BIP error counter register (registers 3.200 through 3.219) is incremented by one each time the calculated BIP value does not equal the value received in the `BIP3` field. The incoming bit error ratio can be estimated by dividing the BIP block error ratio by a factor of 1081344.

82.2.15 Descrambler

The descrambler is identical to that used in Clause 49, see 49.2.10 for the definition.

82.2.16 Receive process

The receive process decodes blocks to produce `RXD<63:0>` and `RXC<7:0>` for transmission to the XLGMII/CGMII. One XLGMII/CGMII data transfer is decoded from each block. The receive process must insert idle control characters to compensate for the removal of alignment markers. If the PCS receive process spans multiple clock domains, it may also perform clock rate compensation via the deletion of idle control characters or sequence ordered sets or the insertion of idle control characters.

The receive process decodes blocks as specified in the receive state diagram shown in Figure 82–15.

82.2.17 Test-pattern checker

When the receive channel is operating in scrambled idle test-pattern mode, the scrambled idle test-pattern checker checks the bits received via `inst:IS_UNITDATA_i.indication` primitives.

The scrambled idle test-pattern checker utilizes the block lock state diagram, the alignment marker state diagram, the PCS deskew state diagram, and the descrambler operating as they do during normal data reception. The BER monitor state diagram is disabled during receive test-pattern mode. When `align_status` is true and the scrambled idle receive test-pattern mode is active, the scrambled idle test-pattern checker observes the sync header and the output from the descrambler. When the sync header and the output of the descrambler is the all idle pattern, a match is detected. When operating in scrambled idle test pattern, the test-pattern error counter counts blocks with a mismatch. Any mismatch indicates an error and shall increment the test-pattern error counter. Because of the error multiplication characteristics of the descrambler, the incoming bit error ratio can be estimated by dividing the 66-bit block error ratio by a factor of 124.

If a Clause 45 MDIO is implemented, then control of the test-pattern reception is from the BASE-R PCS test-pattern control register (bit 3.42.2). In addition errors are counted in the BASE-R PCS test-pattern error counter register (3.43.15:0).

82.2.18 Detailed functions and state diagrams

82.2.18.1 State diagram conventions

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

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

82.2.18.2 State variables

82.2.18.2.1 Constants

EBLOCK_R<71:0>

72 bit vector to be sent to the XLGMII/CGMII containing /E/ in all the eight character locations.

EBLOCK_T<65:0>

66 bit vector to be sent to the PMA containing /E/ in all the eight character locations.

LBLOCK_R<71:0>

72 bit vector to be sent to the XLGMII/CGMII containing one Local Fault ordered_set. The Local Fault ordered_set is defined in 81.3.4.

LBLOCK_T<65:0>

66 bit vector to be sent to the PMA containing one Local Fault ordered_set.

82.2.18.2.2 Variables

align_status

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

alignment_valid

Boolean variable that is set true if all PCS lanes are aligned. It is valid when each lane is in am_lock, with each PCS lane locked to a unique alignment marker from Table 82–2 or Table 82–3, and when all PCS lanes are deskewed. Otherwise, alignment_valid is false.

am_lock<x>

Boolean variable that is set true when receiver acquires alignment marker delineation for a given lane of the service interface, where $x = 0:3$ for 40GBASE-R and $x = 0:19$ for 100GBASE-R.

am_slip_done

Boolean variable that is asserted true when the AM_SLIP requested by the alignment marker lock state diagram has been completed indicating that the next candidate 66-bit block position can be tested.

am_status

A Boolean variable that is true when all lanes are in am_lock and false when at least one lane is not in am_lock.

am_counter_done

Boolean variable that indicates that the alignment marker counter is done.

am_valid

Boolean variable that is set true if received block rx_coded is a valid alignment marker. A valid alignment marker will match one of the encodings in Table 82–2 or Table 82–3, excluding the BIP₃ and BIP₇ fields, and it will be repeated every 16384 blocks.

- ber_test_sh**
Boolean variable that is set true when a new sync header is available for testing and false when BER_TEST_SH state is entered. A new sync header is available for testing when the Block Sync process has accumulated enough bits from the PMA to evaluate the header of the next block.
- block_lock<x>**
Boolean variable that is set true when receiver acquires block delineation for a given lane of the service interface, where $x = 0:3$ for 40GBASE-R and $x = 0:19$ for 100GBASE-R.
- current_am**
This variable holds the value of the current alignment marker. This is compared to the variable first_am to determine if we have alignment marker lock and is always $n \times 16384$ 66-bit blocks away from the first_am.
- enable_deskew**
A Boolean variable that indicates the enabling and disabling of the deskew process. Blocks may be discarded whenever deskew is enabled. True when deskew is enabled, false when deskew is disabled.
- first_am**
A variable that holds the value of the first alignment marker that is recognized on a given lane. This is used later to compare to future alignment markers.
- hi_ber**
Boolean variable which is asserted true when the ber_cnt equals or exceeds 97 indicating a bit error ratio $>10^{-4}$.
- lane_mapping<x>**
This variable indicates which PCS lane is received on lane x of the service interface when am_lock<x> = true, where $x = 0:3$ for 40GBASE-R and $x = 0:19$ for 100GBASE-R.
- PCS_status**
A Boolean variable that is true when align_status is true and hi_ber is false.
- r_block_type**
This variable contains the rx_coded<65:0> vector classification results, returned by the R_TYPE function. It can assume one of the following five values {C,S,T,D,E}, as defined by the R_TYPE function.
- r_block_type_next**
This variable contains the rx_coded<65:0> vector classification results, returned by the R_TYPE_NEXT function. It can assume one of the following five values {C,S,T,D,E}, as defined by the R_TYPE_NEXT function.
- r_test_mode**
Boolean variable that is asserted true when the receiver is in test-pattern mode.
- reset**
Boolean variable that controls the resetting of the PCS. It is true whenever a reset is necessary including when reset is initiated from the MDIO, during power on, and when the MDIO has put the PCS into low-power mode.
- rx_coded<65:0>**
Vector containing the input to the 64B/66B decoder. The format for this vector is shown in Figure 82–5. The leftmost bit in the figure is rx_coded<0> and the rightmost bit is rx_coded<65>.
- rx_raw<71:0>**
Vector containing one XLGMII/CGMII transfer. RXC<0> through RXC<7> are from rx_raw<0> through rx_raw<7>, respectively. RXD<0> through RXD<63> are from rx_raw<8> through rx_raw<71>, respectively.
- sh_valid**
Boolean variable that is set true if received block rx_coded has valid sync header bits. That is, sh_valid is asserted if rx_coded<0> \neq rx_coded<1> and de-asserted otherwise.
- 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 asserted true when the SLIP requested by the Block Lock state diagram has been completed indicating that the next candidate block sync position can be tested.

t_block_type

This variable contains the tx_raw<71:0> vector classification results, returned by the T_TYPE function. It can assume one of the following five values {C,S,T,D,E}, as defined by the T_TYPE function.

test_am

Boolean variable that is set true when a new block is available for testing and false when FIND_1ST state is entered. A new block is available for testing when the Block Sync process has accumulated enough bits from the PMA to evaluate the next block

test_sh

Boolean variable that is set true when a new sync header is available for testing and false when TEST_SH and TEST_SH2 state are entered. A new sync header is available for testing when the Block Sync process has accumulated enough bits from the PMA to evaluate the header of the next block

tx_coded<65:0>

Vector containing the output from the 64B/66B encoder. The format for this vector is shown in Figure 82–5. The leftmost bit in the figure is tx_coded<0> and the rightmost bit is tx_coded<65>.

tx_raw<71:0>

Vector containing one XLGMII/CGMII transfer. TXC<0> through TXC<7> are placed in tx_raw<0> through tx_raw<7>, respectively. TXD<0> through TXD<63> are placed in tx_raw<8> through tx_raw<71>, respectively.

82.2.18.2.3 Functions

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/;
- b) A block type field of 0x4B.

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.

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 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/;
- b) One valid ordered set.

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.

82.2.18.2.4 Counters

am_counter

This counter counts 16383 66-bit blocks that separate two consecutive alignment markers.

am_invld_cnt

Count of the number of invalid alignment markers seen in a row. This counter is always reset to all zeros when a valid marker is detected.

ber_count

A 22-bit counter that counts each time BER_BAD_SH state is entered. This counter is reflected in MDIO register bits 3.33.13:8 and 3.44.15:0.

ber_cnt

Count up to a maximum of 97 of the number of invalid sync headers within the current 1.25 ms (40GBASE-R) or 500 μs (100GBASE-R) period.

errored_block_count:

When the receiver is in normal mode, this 22-bit counter counts once for each time RX_E state is entered. This counter is reflected in MDIO register bits 3.33.7:0 and 3.45.13:0.

sh_cnt

Count of the number of sync headers checked within the current 64 or 1024 block window.

sh_invld_cnt

Count of the number of invalid sync headers within the current 64 or 1024 block window.

test_pattern_error_count:

When the receiver is in test-pattern mode, this 16-bit counter counts errors as described in 82.2.17. This counter is reflected in MDIO register bits 3.43.15:0.

82.2.18.2.5 Timers

State diagram timers follow the conventions of 14.2.3.2.

`xus_timer`

Timer that is triggered every 1.25 ms +1%, –25% for 40GBASE-R or 500 μ s +1%, –25% for 100GBASE-R.

82.2.18.3 State diagrams

The 40GBASE-R PCS shall implement four block lock processes as depicted in Figure 82–10. The 100GBASE-R PCS shall implement twenty block lock processes as depicted in Figure 82–10. A block lock process operates independently on each lane. Each block lock process looks for 64 valid sync headers in a row to declare lock. A valid sync header is either a 01 or a 10. Once in lock, the lock process looks for 65 invalid sync headers within a 1024 sync window to declare out of lock. An invalid sync header is a 11 or 00. Once block lock is achieved on a lane, then the alignment marker process starts.

The 40GBASE-R PCS shall implement four alignment marker lock processes as depicted in Figure 82–11. The 100GBASE-R PCS shall implement twenty alignment marker lock processes as depicted in Figure 82–11. An alignment marker lock process operates independently on each lane. The alignment marker lock state diagram shown in Figure 82–11 determines when the PCS has obtained alignment marker lock to the received bit stream for a given lane of the service interface. Each alignment marker lock process looks for two valid alignment markers 16384 66-bit blocks apart to gain alignment marker lock. On a given lane of the service interface, lane markers must match each other and an entry from Table 82–2 for 100GBASE-R or Table 82–3 for 40GBASE-R. Note that the BIP₃ and BIP₇ fields are excluded from the markers when making a match to each other or the tables. Once in lock, a lane will go out of alignment marker lock if four markers are received in a row that do not match the alignment marker that the lane is currently locked to. When the alignment marker lock process achieves lock for a lane, and if a Clause 45 MDIO is implemented, the PCS shall record the number of the PCS lane received on that lane of the service interface in the appropriate lane mapping register (3.400 to 3.419).

The PCS shall run the deskew process as depicted in Figure 82–12. The PCS deskew process is responsible for determining if the PCS is capable of presenting coherent data to the XLGMII/CGMII. The deskew process ensures that all PCS lanes have alignment marker lock, are locked to different alignment markers, and that the Skew is within the boundaries of what the PCS can deskew.

The BER Monitor state diagram shown in Figure 82–13 monitors the received aggregate signal for high bit error ratio. The high BER state shall be entered if 97 invalid 66-bit sync headers are detected within a 500 μ s window for 100GBASE-R, or a 1.25 ms window for 40GBASE-R. The high BER state is exited once there are less than 97 invalid sync headers in the same window.

The Transmit state diagram shown in Figure 82–14 controls the encoding of transmitted blocks. It makes exactly one transition for each transmit block processed. Though the Transmit state diagram sends Local Fault ordered sets when reset is asserted, the scrambler may not be operational during reset. Thus, the Local Fault ordered sets may not appear on the PMA/FEC service interface.

The Receive state diagram shown in Figure 82–15 controls the decoding of received blocks. It makes exactly one transition for each receive block processed.

The PCS shall perform the functions of block lock, alignment marker lock, PCS deskew, BER Monitor, Transmit, and Receive as specified in the respective state diagrams.

82.3 PCS Management

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

82.3.1 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that may provide control and status information for and about the PCS. Mapping of MDIO control variables to PCS control variables is shown in Table 82–6. Mapping of MDIO status variables to PMD status variables is shown in Table 82–7.

Table 82–6—MDIO/PMD control variable mapping

MDIO control variable	PCS register name	Register/ bit number	PCS control variable
Reset	PCS control 1 register	3.0.15	reset
Loopback	PCS control 1 register	3.0.14	Loopback
Transmit test-pattern enable	BASE-R PCS test-pattern control register	3.42.3	tx_test_mode
Receive test-pattern enable	BASE-R PCS test-pattern control register	3.42.2	rx_test_mode

Table 82–7—MDIO/PMD status variable mapping

MDIO status variable	PCS register name	Register/ bit number	PCS status variable
BASE-R and 10GBASE-T receive link status	BASE-R and 10GBASE-T PCS status 1 register	3.32.12	PCS_status
BASE-R and 10GBASE-T PCS high BER	BASE-R and 10GBASE-T PCS status 1 register	3.32.1	hi_ber
Block <i>x</i> lock	Multi-lane BASE-R PCS alignment status 1 and 2 registers	3.50.7:0 3.51.11:0	block_lock< <i>x</i> >
Lane <i>x</i> aligned	Multi-lane BASE-R PCS alignment status 3 and 4 registers	3.52.7:0 3.53.11:0	am_lock< <i>x</i> >
PCS lane alignment status	Multi-lane BASE-R PCS alignment status 1 register	3.50.12	align_status
BER	BASE-R and 10GBASE-T PCS status 2 register BER high order counter register	3.33.13:8 3.44.15:0	ber_count
Errored blocks	BASE-R and 10GBASE-T PCS status 2 register Errored blocks high order counter register	3.33.7:0 3.45.13:0	errored_block_count
Test-pattern error counter	BASE-R PCS test-pattern error counter register	3.43.15:0	test_pattern_error_count
BIP error counter, lane <i>x</i>	BIP error counter, lane <i>x</i> register	3.200 through 3.219	bip_counter
Lane <i>x</i> mapping	Lane <i>x</i> mapping register	3.400 through 3.419	lane_mapping

82.4 Loopback

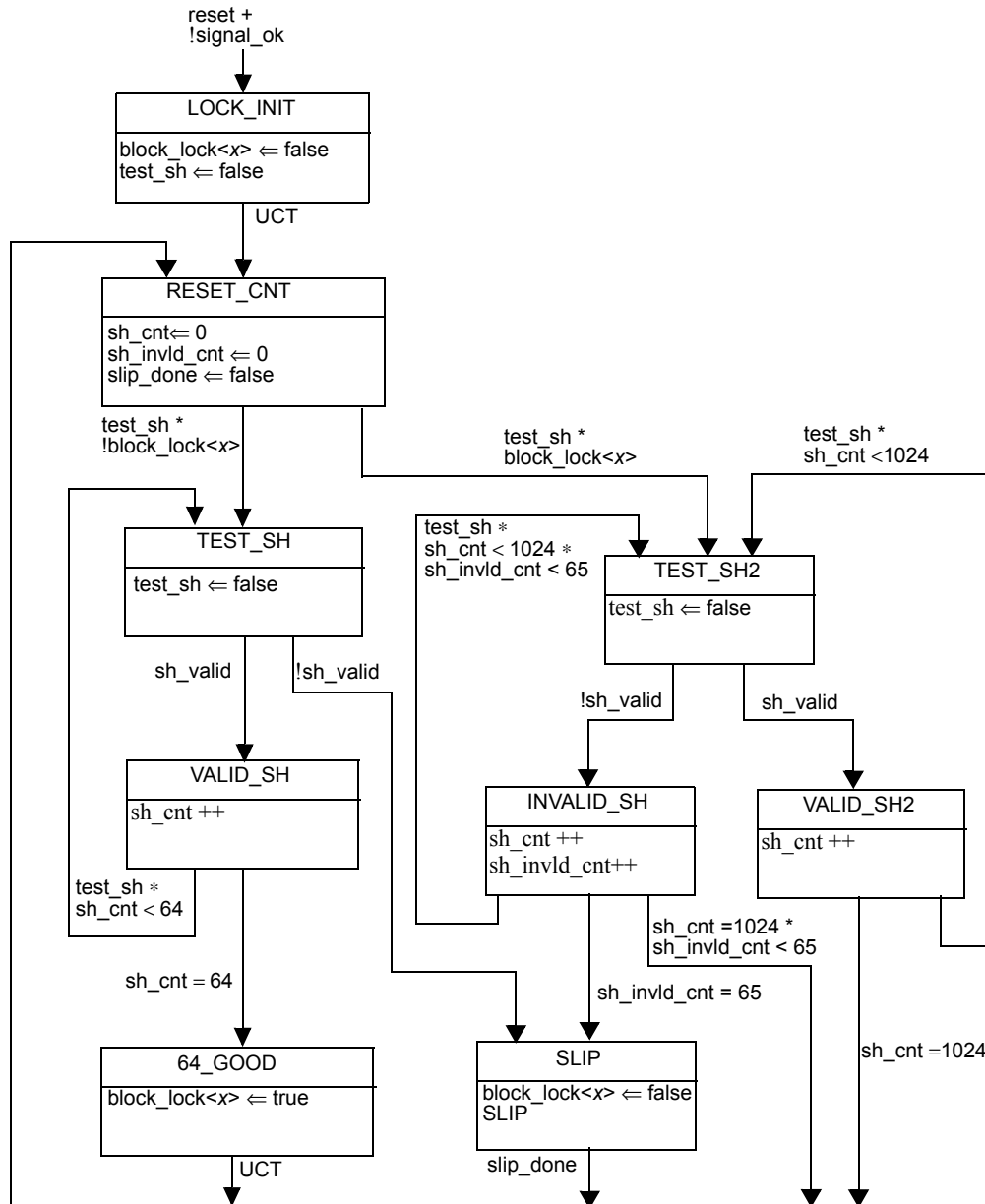
If a Clause 45 MDIO is implemented, then the PCS shall be placed in the loopback mode when the loopback bit from the PCS control 1 register (bit 3.0.14) is set to a one. In this mode, the PCS shall accept data on the transmit path from the XLGMII/CGMII and return it on the receive path to the XLGMII/CGMII. In addition, the PCS shall transmit what it receives from the XLGMII/CGMII to the PMA/FEC sublayer, and shall ignore all data presented to it by the PMA/FEC sublayer.

82.5 Delay constraints

The maximum delay contributed by the 40GBASE-R PCS (sum of transmit and receive delays at one end of the link) shall be no more than 11264 BT (22 pause_quanta or 281.6 ns). The maximum delay contributed by the 100GBASE-R PCS (sum of transmit and receive delays at one end of the link) shall be no more than 35328 BT (69 pause_quanta or 353.28 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.

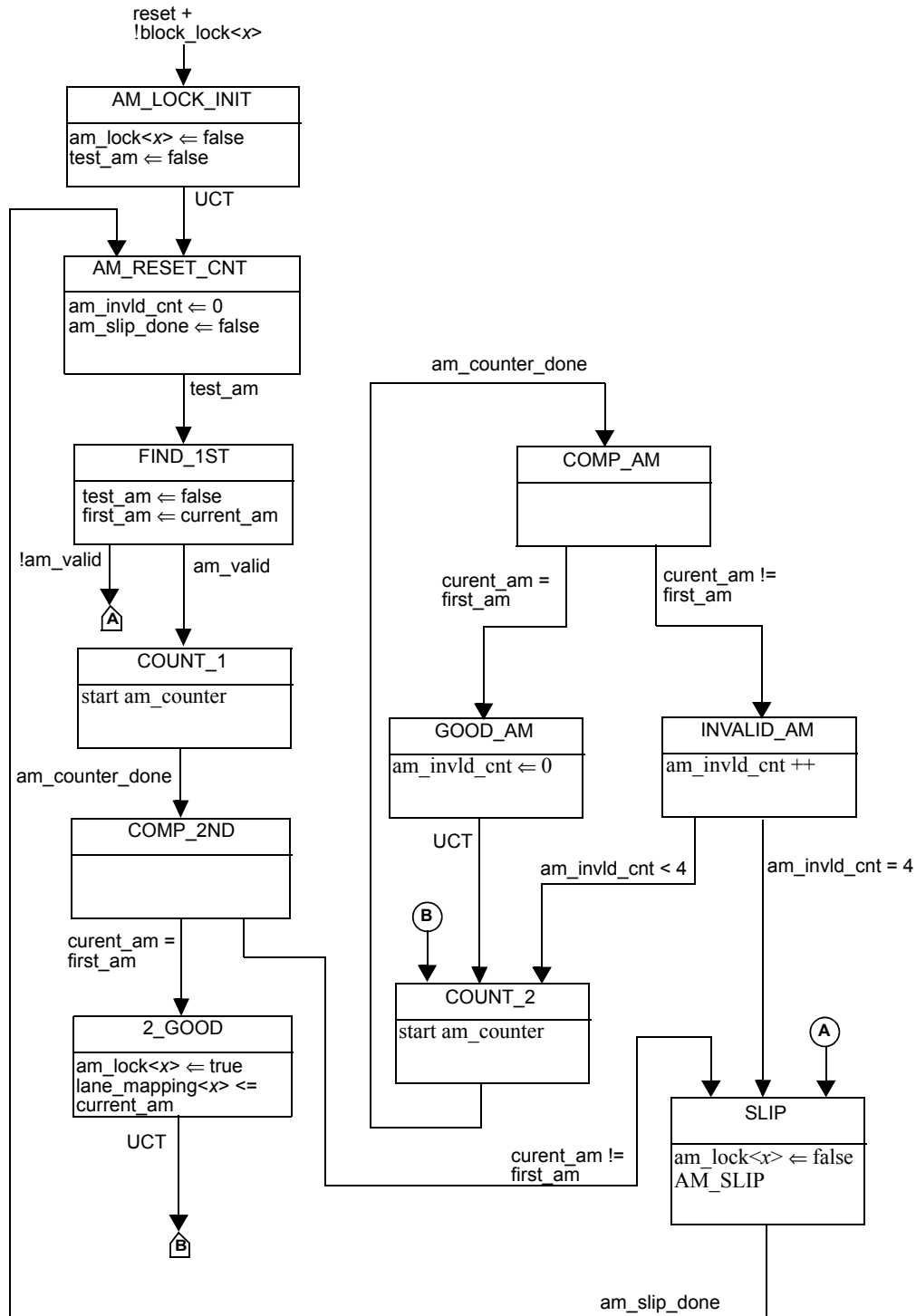
82.6 Auto-Negotiation

The following requirements apply to a PCS used with a 40GBASE-KR4 PMD, 40GBASE-CR4 PMD, or 100GBASE-CR10 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— 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

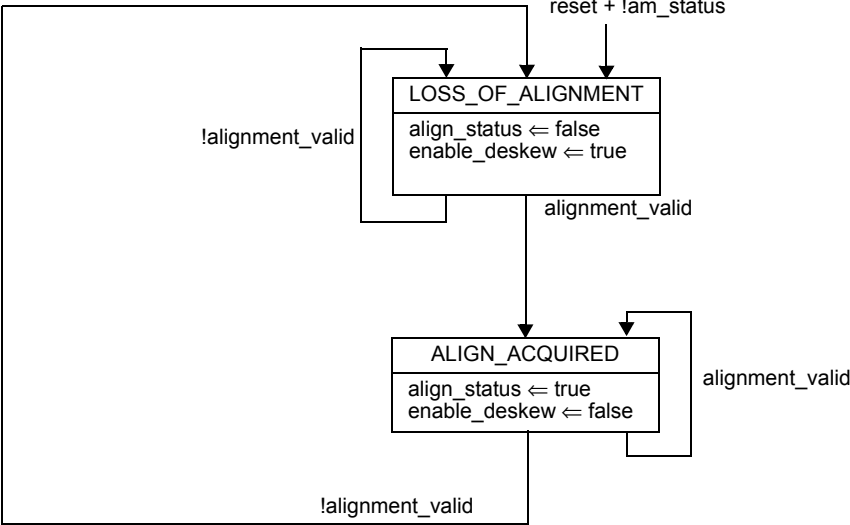


Figure 82-12—PCS deskew state diagram

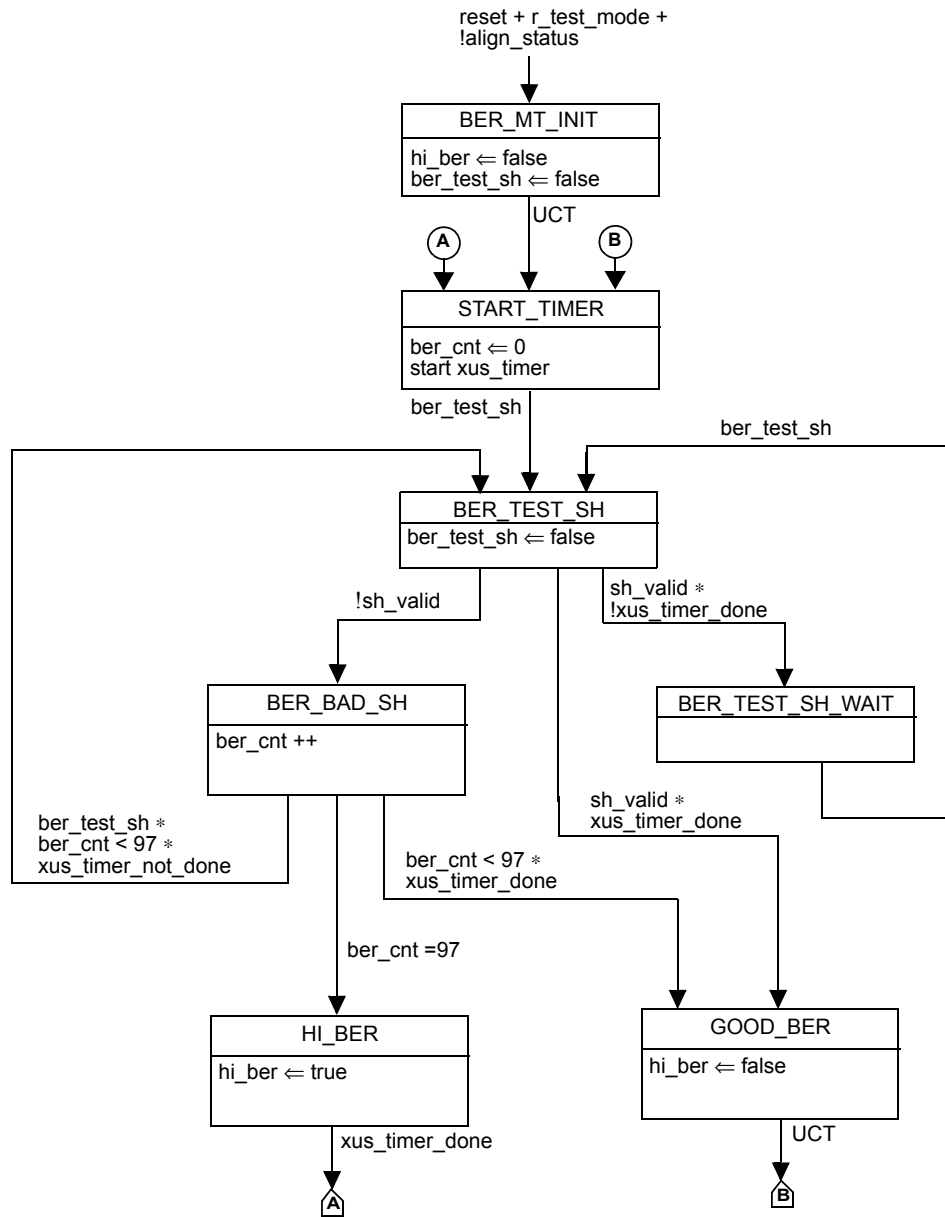


Figure 82-13—BER monitor state diagram

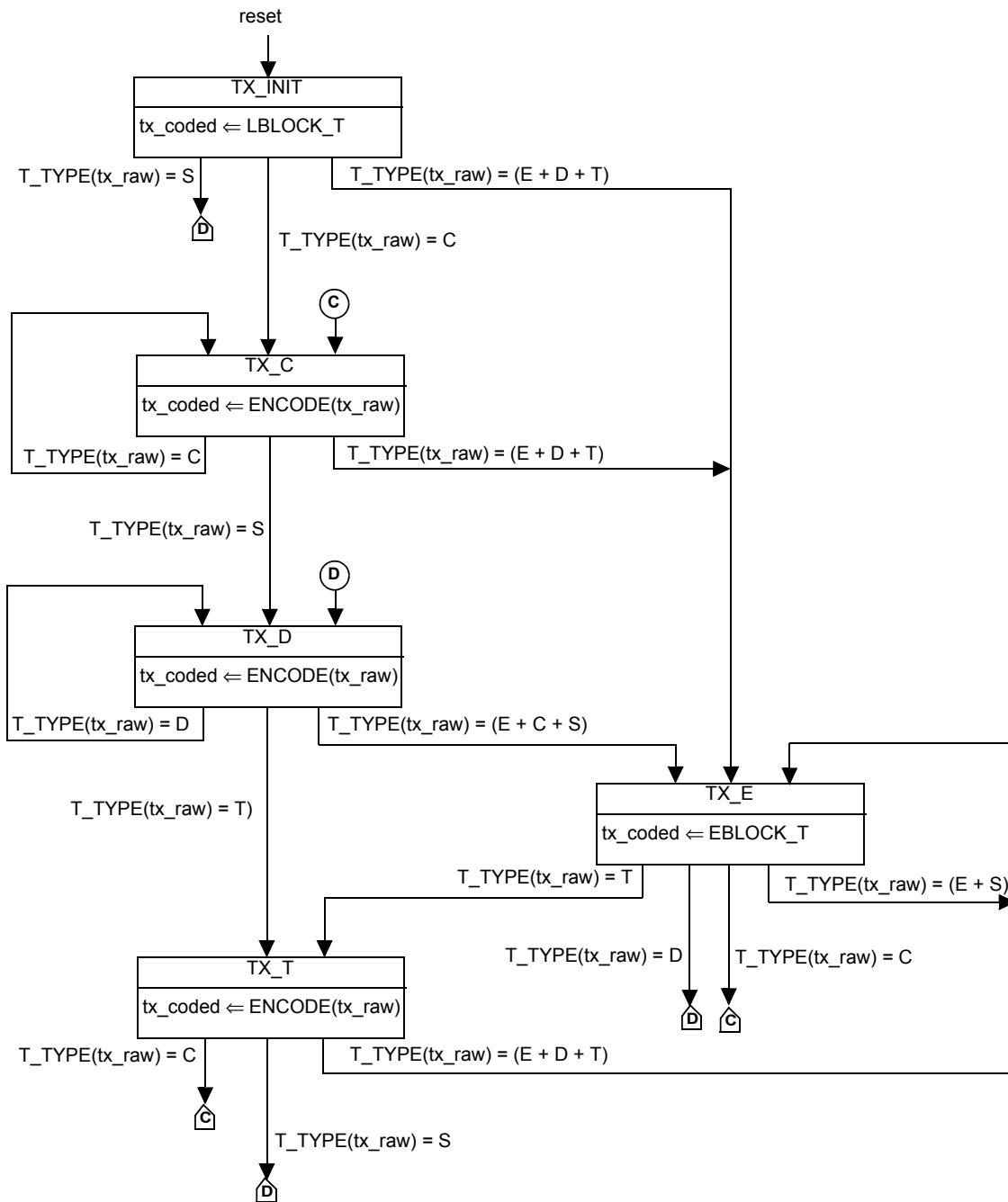


Figure 82-14—Transmit state diagram

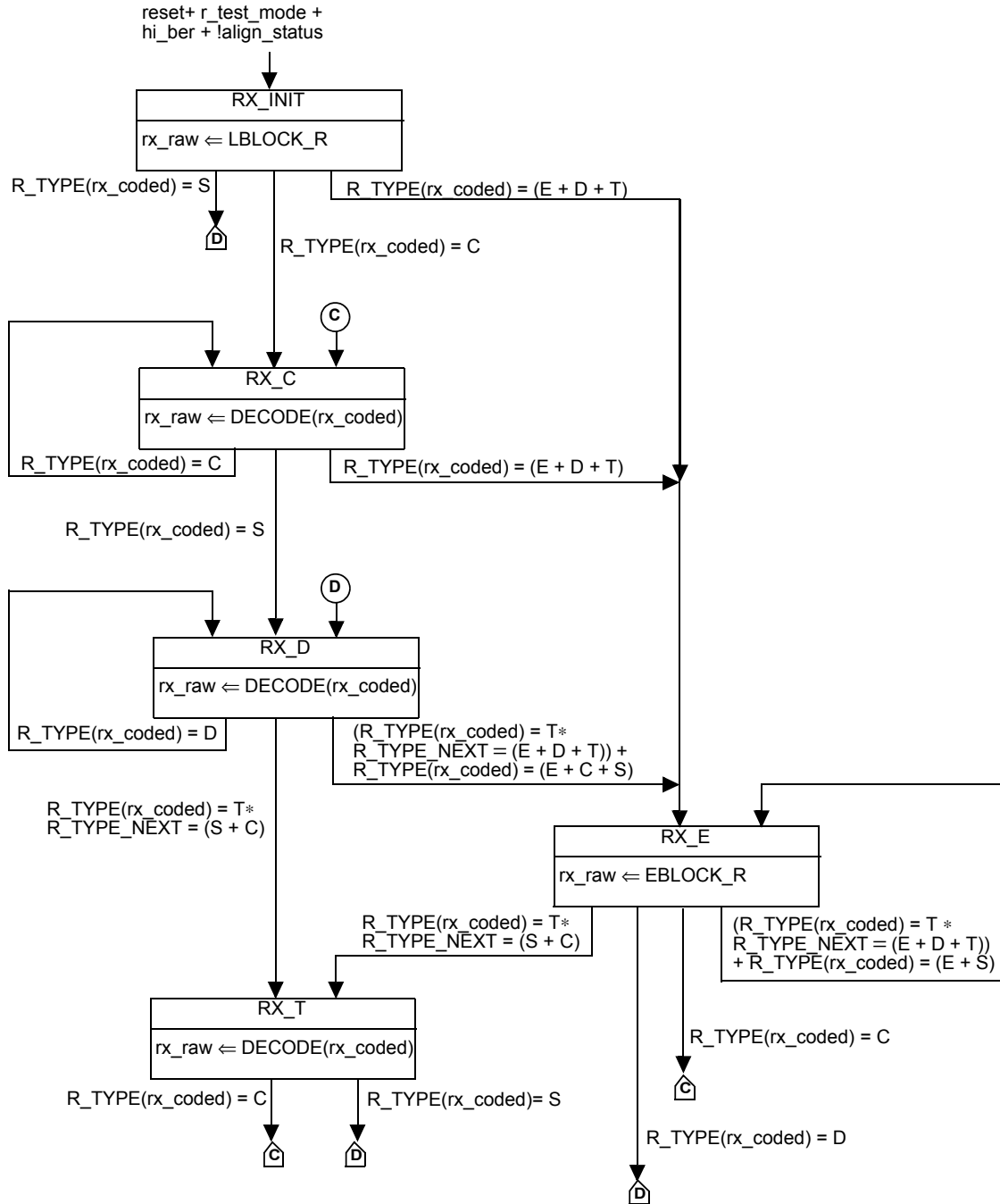


Figure 82–15—Receive state diagram

82.7 Protocol implementation conformance statement (PICS) proforma for Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R⁷

82.7.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R, 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.

82.7.2 Identification

82.7.2.1 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	
<p>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).</p>	

82.7.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 82, Physical Coding Sublayer (PCS) for 64B/66B, type 40GBASE-R and 100GBASE-R
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
<p>Have any Exception items been required? No [] Yes [] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3-2012.)</p>	

Date of Statement	
-------------------	--

⁷Copyright 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.

82.7.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*PCS40	PCS for 40GBASE-R	82.1.1		O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*PCS100	PCS for 100GBASE-R	82.1.1		O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
XGE40	XLGMII logical interface	81, 82.1.4	Logical interface is supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
XGE100	XLGMII logical interface	81, 82.1.4	Logical interface is supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
MD	MDIO	45, 82.3	Registers and interface supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
PMA	Supports operation directly connected to a PMA	82.1.4.2		O.1	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
FEC	Supports operation directly connected to a FEC sublayer	82.1.4.2		O.1	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*JTM	Supports test-pattern mode	82.2.1		PMA:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/> N/A [<input type="checkbox"/>

82.7.4 PICS Proforma Tables for PCS, type 40GBASE-R and 100GBASE-R

82.7.4.1 Coding rules

Item	Feature	Subclause	Value/Comment	Status	Support
C1	Encoder (and ENCODE function) implements the code as specified	82.2.3 and 82.2.18.2.3		M	Yes [] No []
C2	Decoder (and DECODE function) implements the code as specified	82.2.3 and 82.2.18.2.3		M	Yes [] No []
C3	Only valid block types are transmitted	82.2.3.3		M	Yes [] No []
C4	Invalid block types are treated as an error	82.2.3.3		M	Yes [] No []
C5	Only valid control characters are transmitted	82.2.3.4		M	Yes [] No []
C6	Invalid control characters are treated as an error	82.2.3.4		M	Yes [] No []
C7	Idles do not interrupt data	82.2.3.6		M	Yes [] No []
C8	IDLE control code insertion and deletion	82.2.3.6	Insertion or Deletion in groups of 8 /I/s	M	Yes [] No []
C9	Sequence ordered_set deletion	82.2.3.9	Only one whole ordered_set of two consecutive sequence ordered sets may be deleted	M	Yes [] No []

82.7.4.2 Scrambler and Descrambler

Item	Feature	Subclause	Value/Comment	Status	Support
S1	Scrambler	82.2.5	Performs as shown in Figure 49–8	M	Yes [] No []
S2	Descrambler	82.2.15	Performs as shown in Figure 49–10	M	Yes [] No []

82.7.4.3 Deskew and Reordering

Item	Feature	Subclause	Value/Comment	Status	Support
DR1	Deskew	82.2.12	Able to deskew up to the value in Table 82–5	M	Yes [] No []
DR2	Reordering	82.2.13	Performs reordering.	M	Yes [] No []

82.7.4.4 Alignment Markers

Item	Feature	Subclause	Value/Comment	Status	Support
AM1	Alignment marker insertion	82.2.7	Alignment markers are inserted periodically as described in section 82.2.7	M	Yes [] No []
AM2	Alignment marker form	82.2.7	Alignment markers are formed as described in section 82.2.7	M	Yes [] No []
AM3	Lane mapping	82.2.18.3	PCS lane number is captured	MD:M	Yes [] No []

82.7.5 Test-pattern modes

Item	Feature	Subclause	Value/Comment	Status	Support
JT1	Scrambled idle transmit test-pattern generator is implemented	82.2.10	Performs as in 82.2.10	JTM:M	Yes [] No [] N/A []
JT2	Scrambled idle receive test-pattern checker is implemented	82.2.17	Performs as in 82.2.17	JTM:M	Yes [] No [] N/A []
JT3	Transmit and receive test-pattern modes can operate simultaneously	82.2.1		JTM:M	Yes [] No [] N/A []

82.7.5.1 Bit order

Item	Feature	Subclause	Value/Comment	Status	Support
B1	Transmit bit order	82.2.3.2	Placement of bits into the PCS lanes as shown in Figure 82–3	M	Yes [] No []
B2	Receive bit order	82.2.3.2	Placement of bits into the XLGMII/CGMII as shown in Figure 82–4	M	Yes [] No []

82.7.6 Management

Item	Feature	Subclause	Value/Comment	Status	Support
M1	Alternate access to PCS Management objects is provided	82.3		O	Yes [] No []

82.7.6.1 State diagrams

Item	Feature	Subclause	Value/Comment	Status	Support
SM1	40GBASE-R Block Lock	82.2.18.3	Implements 4 block lock processes as depicted in Figure 82–10	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM2	100GBASE-R Block Lock	82.2.18.3	Implements 20 block lock processes as depicted in Figure 82–10	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM3	The SLIP function evaluates all possible bit positions	82.2.18.2.3		M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM4	40GBASE-R Alignment Marker Lock	82.2.18.3	Implements 4 alignment marker lock processes as depicted in Figure 82–11	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM5	100GBASE-R Alignment Marker Lock	82.2.18.3	Implements 20 alignment marker lock processes as depicted in Figure 82–11	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM6	The AM_SLIP functions evaluates all possible blocks	82.2.18.2.3		M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM7	40GBASE-R PCS deskew state diagram	82.2.18.3	Meets the requirements of Figure 82–12	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM8	100GBASE-R PCS deskew state diagram	82.2.18.3	Meets the requirements of Figure 82–12	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM9	40GBASE-R BER Monitor	82.2.18.3	Meets the requirements of Figure 82–13 with xus_timer_done set to 1.25 ms	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM10	100GBASE-R BER Monitor	82.2.18.3	Meets the requirements of Figure 82–13 with xus_timer_done set to 500 μ s	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM11	40GBASE-R Transmit process	82.2.18.3	Meets the requirements of Figure 82–14	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM12	100GBASE-R Transmit process	82.2.18.3	Meets the requirements of Figure 82–14	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM13	40GBASE-R and 100GBASE-R Receive process	82.2.18.3	Meets the requirements of Figure 82–15	PCS40:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
SM14	100GBASE-R Receive process	82.2.18.3	Meets the requirements of Figure 82–15	PCS100:M	Yes [<input type="checkbox"/> No [<input type="checkbox"/>

82.7.6.2 Loopback

Item	Feature	Subclause	Value/Comment	Status	Support
L1	Supports loopback	82.4	Performs as in 82.4	M	Yes [] No [] N/A []
L2	When in loopback, transmits what it receives from the XLGMII/CGMII	82.4	Performs as in 82.4	M	Yes [] No []

82.7.6.3 Delay constraints

Item	Feature	Subclause	Value/Comment	Status	Support
TIM1	PCS Delay Constraint	82.5	No more than 11264 BT for sum of transmit and receive path delays for 40GBASE-R and 35328 BT for 100GBASE-R	M	Yes [] No []

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	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.indication(link_status), when the PCS is used with 10GBASE-KR PMD	AN1:M	Yes []
AN3	link_status parameter	82.6	Takes the value OK or FAIL, as described in 82.6	AN1:M	Yes []
AN4	Generation of AN_LINK.indication primitive	82.6	Generated when the value of link_status changes	AN1:M	Yes []

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

83.1 Overview

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. The 100GBASE-R PMA(s) can support any of the 100 Gb/s PMDs in Table 80–2. The terms 40GBASE-R and 100GBASE-R are used when referring generally to Physical Layers using the PMA defined in this clause.

40GBASE-R and 100GBASE-R can be extended to support any full duplex medium requiring only that the PMD be compliant with the appropriate PMA interface.

The interfaces for the inputs of the 40GBASE-R and 100GBASE-R PMAs are defined in an abstract manner and do not imply any particular implementation. The optional physical instantiation of the PMD service interfaces for 40GBASE-SR4, 40GBASE-LR4, and 100GBASE-SR10 PMDs, known as XLPI and CPPI, are defined in Annex 86A. The PMD service interfaces for other PMDs are defined in an abstract manner according to 80.3.1. For 40GBASE-R PMAs, electrical interfaces connecting PMA sublayers, known as XLAUI, are defined in Annex 83A and Annex 83B. For 100GBASE-R PMAs, electrical interfaces connecting PMA sublayers, known as CAUI, are defined in Annex 83A and Annex 83B.

83.1.2 Position of the PMA in the 40GBASE-R or 100GBASE-R sublayers

Figure 83–1 shows the relationship of the PMA sublayer (shown shaded) with other sublayers to the ISO Open System Interconnection (OSI) reference model.

83.1.3 Summary of functions

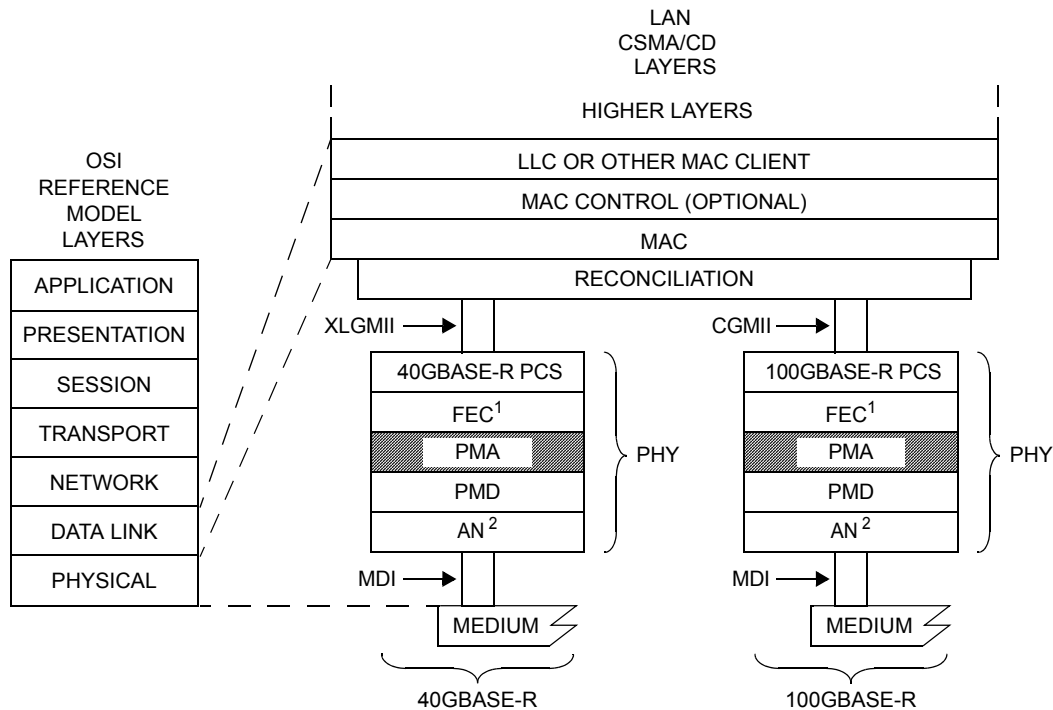
The following is a summary of the principal functions implemented (when required) by the PMA in both the transmit and receive directions:

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

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

83.1.4 PMA sublayer positioning

An implementation may use one or more PMA sublayers to adapt the number and rate of the PCS lanes to the number and rate of the PMD lanes. The number of PMA sublayers required depends on the partitioning of functionality for a particular implementation. An example is illustrated in Figure 83–2. This example illustrates the partitioning that might arise from use of a FEC device that is separate from the PCS.



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

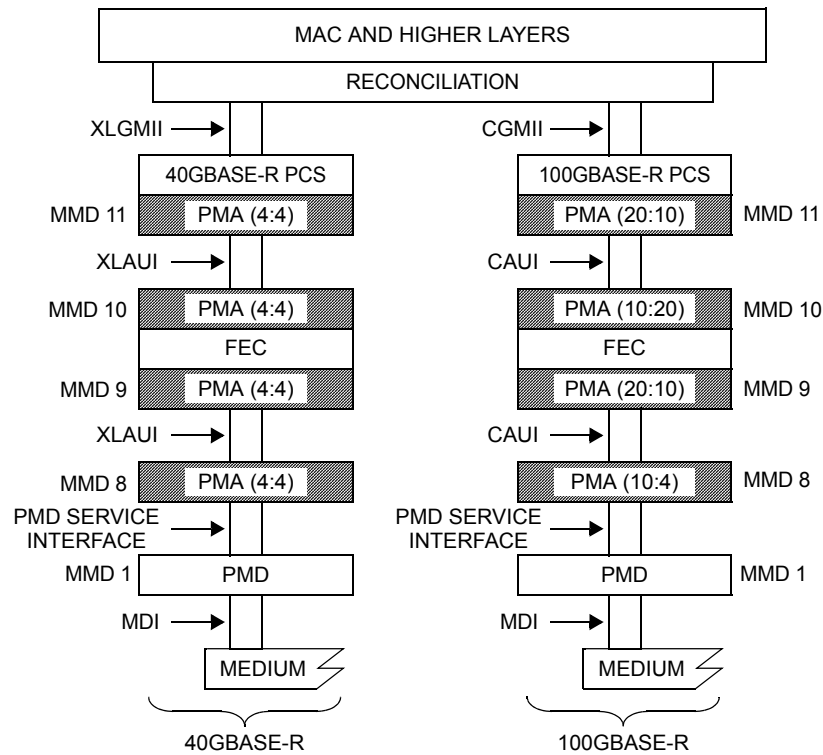
NOTE 1—OPTIONAL OR OMITTED DEPENDING ON PHY TYPE
NOTE 2—CONDITIONAL BASED ON PHY TYPE

Figure 83-1—40GBASE-R and 100GBASE-R PMA relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

Additional examples are illustrated in Annex 83C. Each PMA maps the PCSs from p PMA input lanes to q PMA output lanes in the Tx direction, and from q PMA input lanes to p PMA output lanes in the Rx direction.

Management Data Input/Output (MDIO) Manageable Device (MMD) addresses 1, 8, 9, 10, and 11 are available for addressing multiple instances of PMA sublayers (see Table 45-1 for MMD device addresses). If the PMA sublayer that is closest to the PMD is packaged with the PMD, it shares MMD 1 with the PMD. If the PMD service interface is physically instantiated as nPPI (see Annex 86A), the PMA sublayer that is closest to the PMD will be addressed as MMD 8. More addressable instances of PMA sublayers, each one separated from lower addressable instances by chip-to-chip interfaces, may be implemented and addressed allocating MMD addresses to PMAs in increasing numerical order going from the PMD toward the PCS. The example shown in Figure 83-2 could be implemented with four addressable instances: MMD 8 addressing the lowest PMA sublayer (note that this cannot share MMD 1 with the PMD as they are not packaged together in this example), MMD 9 addressing the PMA sublayer above the XLAUI/CAUI below the FEC, MMD 10 addressing the PMA sublayer below the XLAUI/CAUI above the FEC, and MMD 11 addressing the PMA sublayer closest to the PCS.

The number of input lanes and the number of output lanes for a PMA are always divisors of the number of PCSs. For PMA sublayers supporting 40GBASE-R PMDs, the number of PCSs is 4, and for PMA sublayers supporting 100GBASE-R PMDs, the number of PCSs is 20.



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
 MMD = MDIO MANAGEABLE DEVICE

PCS = PHYSICAL CODING SUBLAYER
 PMA = PHYSICAL MEDIUM ATTACHMENT
 PMD = PHYSICAL MEDIUM DEPENDENT
 XLAUI = 40 Gb/s ATTACHMENT UNIT INTERFACE
 XLMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

Figure 83-2—Example 40GBASE-R and 100GBASE-R PMA layering

The following guidelines apply to the partitioning of PMAs:

- a) The inter-sublayer service interface, defined in 80.3, is used for the PMA, FEC, and PMD service interfaces supporting a flexible architecture with optional FEC and multiple PMA sublayers.
 - 1) An instance of this interface can only connect service interfaces with the same number of lanes, where the lanes operate at the same rate.
- b) XLAUI and CAUI are physical instantiations of the connection between two adjacent PMA sublayers.
 - 1) As a physical instantiation, it defines electrical and timing specification as well as requiring a receive re-timing function
 - 2) XLAUI is a 10.3125 GBd by 4 lane physical instantiation of the respective 40 Gb/s connection
 - 3) CAUI is a 10.3125 GBd by 10 lane physical instantiation of the respective 100 Gb/s connection
- c) The abstract inter-sublayer service interface can be physically instantiated as a XLAUI or CAUI, using associated PMAs to map to the appropriate number of lanes.
- d) Opportunities for optional test-pattern generation, optional test-pattern detection, optional local loopback and optional remote loopback are dependent upon the location of the PMA sublayer in the implementation. See Figure 83-5.
- e) A minimum of one PMA sublayer is required in a PHY.
- f) A maximum of four PMA sublayers are addressable as MDIO MMDs.

83.2 PMA interfaces

All PMA variants for 40GBASE-R and 100GBASE-R signals are based on a generic specification of a bit mux function that applies to all input/output lane counts and each direction of transmission. Each direction of transmission employs one or more such bit muxes to adapt from the appropriate number of input lanes to the appropriate number of output lanes as illustrated in Figure 83–3.

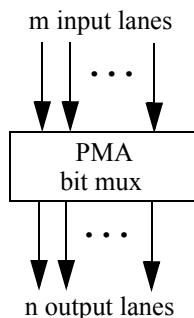


Figure 83–3—PMA bit mux used in both Tx and Rx directions

Conceptually, the PMA bit mux operates in one direction of transmission by demultiplexing PCSs from m PMA input lanes and remultiplexing them into n PMA output lanes. The mapping of PCSs from input to output lanes is not specified. See 83.5.2 and Figure 83–4 for details.

Figure 83–5 provides the functional block diagram of a PMA. The parameters of a PMA include the following:

- The aggregate rate supported (40GBASE-R or 100GBASE-R).
- The numbers of input and output lanes in each direction.
- Whether the PMA is adjacent to a physically instantiated interface (XLAUI/CAUI above or below, or PMD service interface below).
- Whether the PMA is adjacent to the PCS (or adjacent to FEC when FEC is adjacent to the PCS).
- Whether the PMA is adjacent to the PMD.

83.3 PMA service interface

The PMA service interface for 40GBASE-R and 100GBASE-R is an instance of the inter-sublayer service interface defined in 80.3. The PMA service interface primitives are summarized as follows:

```
PMA:IS_UNITDATA_ i.request(tx_bit)
PMA:IS_UNITDATA_ i.indication(rx_bit)
PMA:IS_SIGNAL.indication(SIGNAL_OK)
```

For a PMA with p lanes at the PMA service interface, the primitives are defined for $i = 0$ to $p - 1$.

If the PMA client is the PCS or an FEC sublayer, the PMA (or PMA client) continuously sends four (for 40GBASE-R) or twenty (for 100GBASE-R) parallel bit streams to the PMA client (or PMA), each at the nominal signaling rate of the PCSL.

If the PMA client is another PMA, for a PMA supporting a 40GBASE-R PMD, the number of PCSs $z = 4$ and for a PMA supporting a 100GBASE-R PMD, the number of PCSs $z = 20$. A PMA with p input lanes receives bits on each of its input lanes at z/p times the PCSL rate. Skew may exist between the bits received

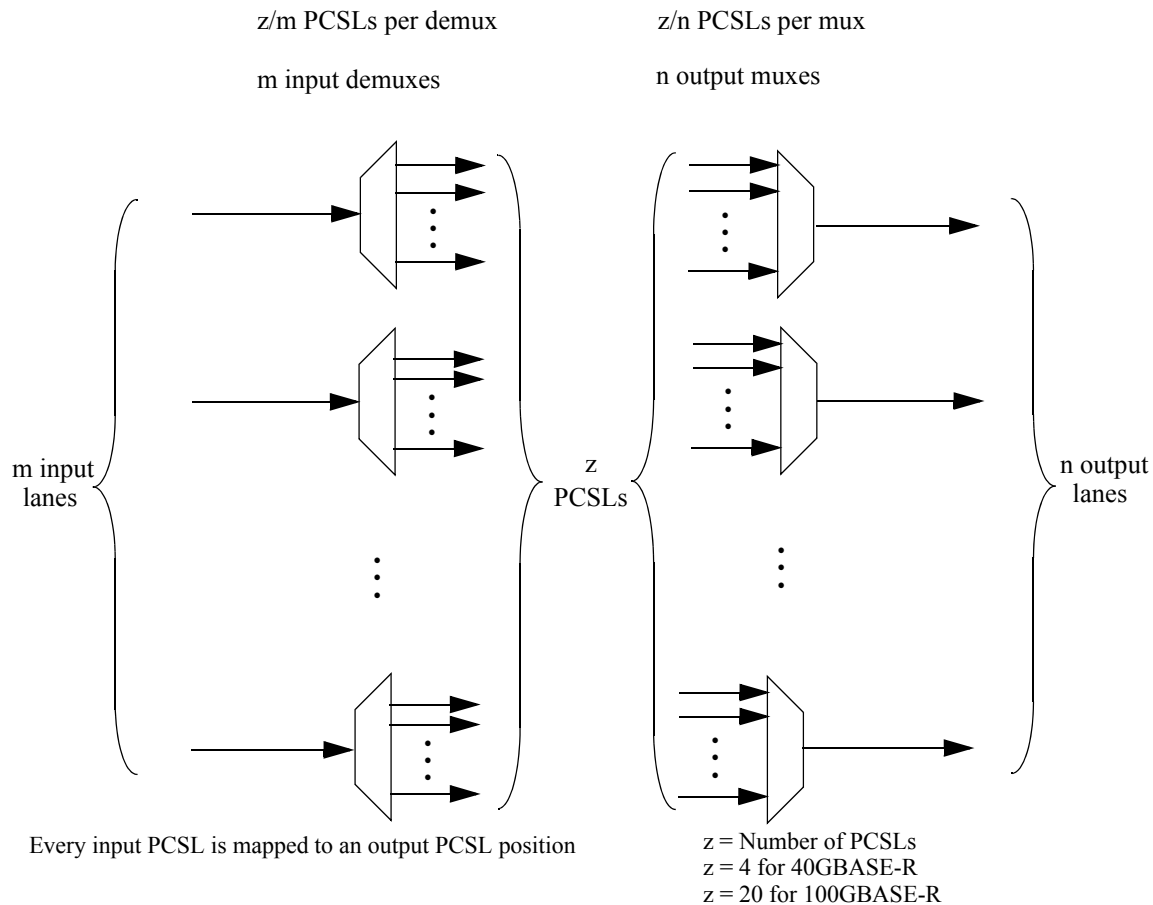


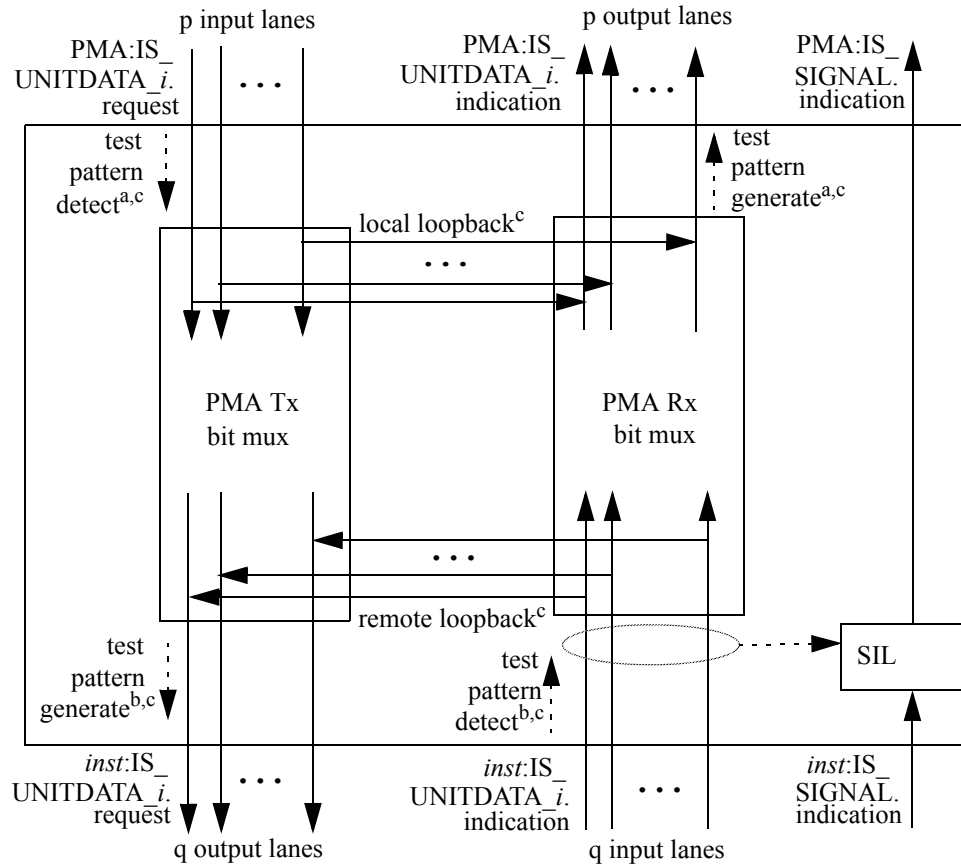
Figure 83–4—PMA bit mux operation used in both Tx and Rx directions

on each lane even though all lanes originate from the same synchronous source, so there is independence of arrival of bits on each lane.

In the Tx direction, if the bit from a PMA:IS_UNITDATA_ i .request primitive is received over a physically instantiated interface (XLAUI/CAUI), clock and data are recovered on the lane receiving the bit. The bit is routed through the PMA to an output lane through a process that may demultiplex PCSLs from the input, perform any necessary buffering to tolerate Skew Variation across input lanes, and multiplex PCSLs to output lanes. The bit is sent on an output lane to the sublayer below using the *inst*:IS_UNITDATA_ k .request (k not necessarily equal to i) primitive (see 83.4).

In the Rx direction, when data is being received from every input lane from the sublayer below the PMA that has a PCSL that is routed to a particular output lane at the PMA service interface, and (if necessary), buffers are filled to allow tolerating the Skew Variation that may appear between the input lanes, PCSLs are demultiplexed from the input lanes, remultiplexed to the output lanes, and bits are transferred over each output lane to the PMA client via the PMA:IS_UNITDATA_ i .indication primitive.

The PMA:IS_SIGNAL.indication primitive is generated through a set of Signal Indication Logic (SIL) that reports signal health based on receipt of the *inst*:IS_SIGNAL.indication from the sublayer below, data being received on all of the input lanes from the sublayer below, buffers filled (if necessary) to accommodate Skew Variation, and bits being sent to the PMA client on all of the output lanes. When these conditions are met, the SIGNAL_OK parameter sent to the PMA client via the PMA:IS_SIGNAL.indication primitive will have the value OK. Otherwise, the SIGNAL_OK primitive will have the value FAIL.



inst PMD, PMA, or FEC, depending on which sublayer is below this PMA
 SIL Signal Indication Logic

^a If physically instantiated interface (XLAUI/CAUI) immediately above this PMA.

^b If physically instantiated interface (XLAUI/CAUI or PMD service interface) immediately below this PMA, or if this is the closest PMA to the PMD.

^c Optional.

Figure 83-5—PMA Functional Block Diagram

83.4 Service interface below PMA

Since the architecture supports multiple PMA sublayers for various PMD lane counts and device partitioning, there are several different sublayers that may appear below a PMA, including FEC, the PMD, or another PMA. The variable *inst* represents whichever sublayer appears below the PMA (e.g., another PMA, FEC, or PMD).

The sublayer below the PMA utilizes the inter-sublayer service interface defined in 80.3. The service interface primitives provided to the PMA are summarized as follows:

```

inst:IS_UNITDATA_i.request(tx_bit)
inst:IS_UNITDATA_i.indication(rx_bit)
inst:IS_SIGNAL.indication(SIGNAL_OK)
    
```

The number of lanes *q* for the service interface matches the number of lanes expected by the PMA. The *inst:IS_UNITDATA_i* primitives are defined for each lane *i* = 0 to *q* - 1 of the service interface below the

PMA. Note that electrical and timing specifications of the service interface are defined if the interface is physically instantiated (e.g., XLAUI/CAUI or nPPI), otherwise the service interface is specified only abstractly. The interface between the PMA and the sublayer below consists of q lanes for data transfer and a status indicating a good signal sent by the sublayer below the PMA (see Figure 83–5).

In the Tx direction, when data is being received via the PMA:IS_UNITDATA_ i .request primitive from every input lane from the PMA client at the PMA service interface (see 83.3) that has a PCSL that is routed to this output lane, and (if necessary), buffers are filled to allow tolerating the Skew Variation that may appear between the input lanes from the PMA client, PCSLs are demultiplexed from the input lanes, remultiplexed to the output lanes, and bits are transferred over each output lane to the sublayer below the PMA.

In the Rx direction, if the bit is received over a physically instantiated interface (XLAUI/CAUI or nPPI), clock and data are recovered on the lane receiving the bit. The bit is routed through the PMA to an output lane toward the PMA client through a process that may demultiplex PCSLs from the input, perform any necessary buffering to tolerate Skew Variation across input lanes, and multiplex PCSLs to output lanes, and finally sending the bit on an output lane to the PMA client using the PMA:IS_UNITDATA_ k .indication (k not necessarily equal to i) primitive at the PMA service interface.

83.5 Functions within the PMA

The purpose of the PMA is to adapt the PCSL formatted signal to an appropriate number of abstract or physical lanes, to recover clock from the received signal (if appropriate), and optionally to provide test signals and loopback. Each input (Tx direction) or output (Rx direction) lane between the PMA and the PMA client carries one or more PCSLs that are bit-multiplexed. All input and output lanes between the PMA and the PMA client carry the same number of PCSLs and operate at the same nominal signaling rate. Likewise, each input (Rx direction) or output (Tx direction) lane between the PMA and the sublayer below the PMA carries one or more PCSLs that are bit-multiplexed. All input and output lanes between the PMA and the sublayer below the PMA carry the same number of PCSLs and operate at the same nominal signaling rate. As described in 83.1.4, the number of input lanes and the number of output lanes for a given PMA are divisors of the number of PCSLs for the interface type supported.

83.5.1 Per input-lane clock and data recovery

If the interface between the PMA client and the PMA is physically instantiated as XLAUI/CAUI, the PMA shall meet the electrical and timing specifications in Annex 83A or Annex 83B as appropriate. If the interface between the sublayer below the PMA and the PMA is physically instantiated as XLAUI/CAUI or nPPI, the PMA shall meet the electrical and timing specifications at the service interface as specified in Annex 83A, Annex 83B, or Annex 86A as appropriate.

83.5.2 Bit-level multiplexing

The PMA provides bit-level multiplexing in both the Tx and Rx directions. In the Tx direction, the function is performed among the bits received from the PMA client via the PMA:IS_UNITDATA_ i .request primitives (for PMA client lanes $i = 0$ to $p - 1$) with the result sent to the service interface below the PMA using the *inst*:IS_UNITDATA_ i .request primitives (for service interface lanes $i = 0$ to $q - 1$), referencing the functional block diagram shown in Figure 83–5. The bit multiplexing behavior is illustrated in Figure 83–4.

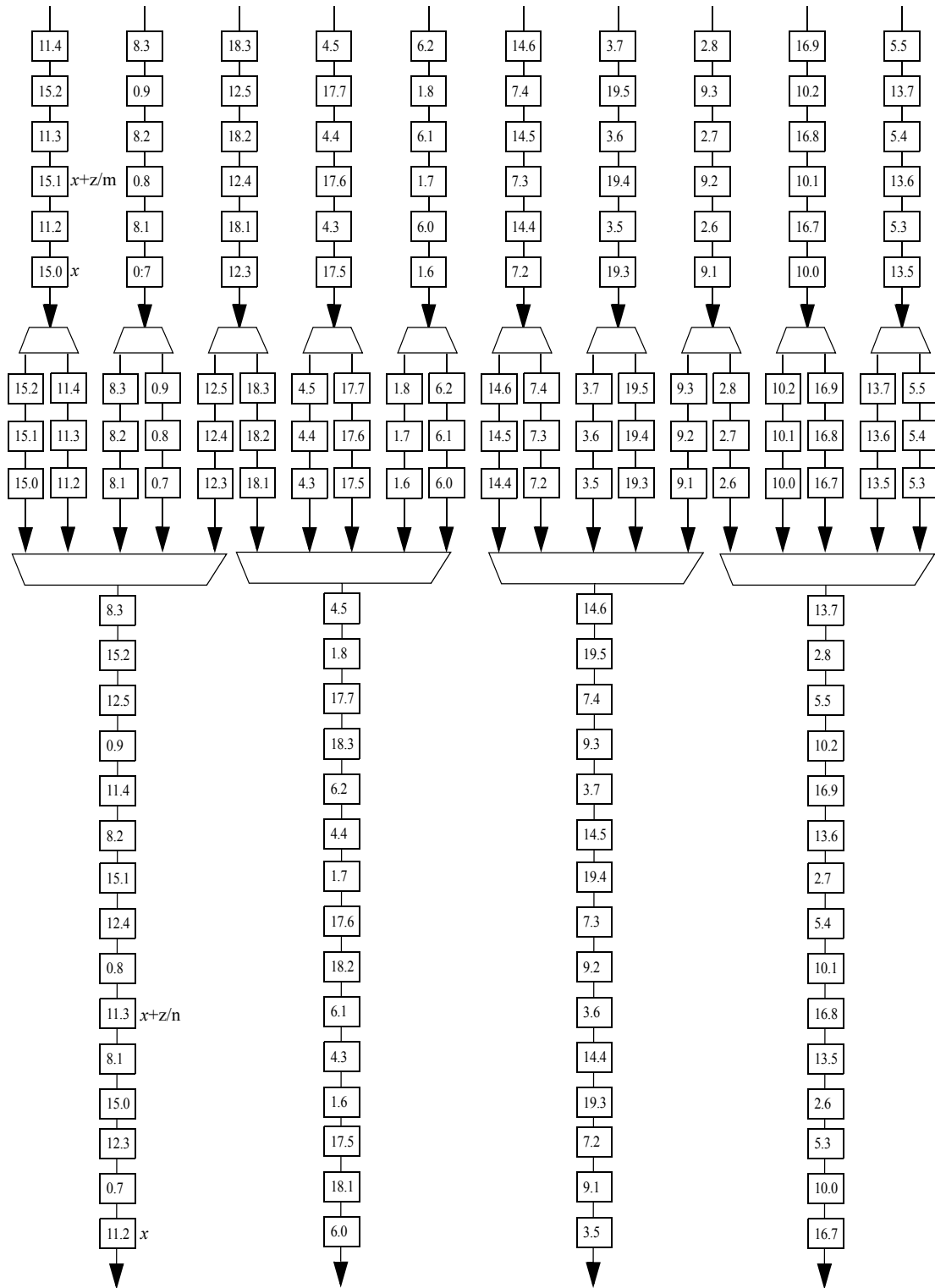
The aggregate signal carried by the group of input lanes or the group of output lanes is arranged as a set of PCSLs. For PMA sublayers supporting 40GBASE-R interfaces, the number of PCSLs z is 4, and the nominal signaling rate R of each PCSL is 10.3125 GBd. For PMA sublayers supporting 100GBASE-R interfaces, the number of PCSLs z is 20, and the nominal signaling rate R of each PCSL is 5.15625 GBd.

For a PMA with m input lanes (Tx or Rx direction), each input lane carries, bit multiplexed, z/m PCSLs. Each input lane has a nominal signaling rate of $R \times z/m$. If bit x received on an input lane belongs to a particular PCSL, the next bit of that same PCSL is received on the same input lane at bit position $x+(z/m)$. The z/m PCSLs may arrive in any sequence on a given input lane.

For a PMA with n output lanes (Tx or Rx direction), each output lane carries, bit multiplexed, z/n PCSLs. Each output lane has a nominal signaling rate of $R \times z/n$. Each PCSL is mapped from a position in the sequence on one of the z/m input lanes to a position in the sequence on one of the z/n output lanes. If bit x sent on an output lane belongs to a particular PCSL, the next bit of that same PCSL is sent on the same output lane at bit position $x + (z/n)$. The PMA shall maintain the chosen sequence of PCSLs on all output lanes while it is receiving a valid stream of bits on all input lanes.

Each PCSL received in any temporal position on an input lane is transferred into a temporal position on an output lane. As the PCS (see Clause 82) has fully flexible receive logic, an implementation is free to perform the mapping of PCSLs from input lanes to output lanes without constraint. Figure 83–6 illustrates one possible bit ordering for a 10:4 PMA bit mux. Other bit orderings are also valid.

10 Lane PMA Input



4 Lane PMA Output

NOTE—i.k indicates bit k on PCSL i. Skew may exist between PCSLs.

Figure 83–6—Example 10: 4 PMA bit mux

83.5.3 Skew and Skew Variation

The Skew (relative delay) between the PCSs must be kept within limits so that the information on the lanes can be reassembled by the PCS.

Any PMA that combines PCSs from different input lanes onto the same output lane must tolerate Skew Variation between the input lanes without changing the PCSL positions on the output. Skew and Skew Variation are defined in 80.5. The limits for Skew and Skew Variation at physically instantiated interfaces are specified at Skew points SP1 and SP2 in the transmit direction and SP5 and SP6 in the receive direction as defined in 80.5 and illustrated in Figure 80–4 and Figure 80–5.

83.5.3.1 Skew generation toward SP1

In an implementation with one or more physically instantiated XLAUI/CAUI interfaces, the PMA that sends data in the transmit direction toward the XLAUI/CAUI that is closest to the PMD (SP1 in Figure 80–4 and Figure 80–5) shall produce no more than 29 ns of Skew between PCSs toward the XLAUI/CAUI, and no more than 200 ps of Skew Variation.

83.5.3.2 Skew tolerance at SP1

In an implementation with one or more physically instantiated XLAUI/CAUI interfaces, the PMA service interface that receives data in the transmit direction from the XLAUI/CAUI (SP1 in Figure 80–4 and Figure 80–5) shall tolerate the maximum amount of Skew Variation allowed at SP1 (200 ps) between input lanes while maintaining the bit ordering and position of each PCSL on each PMA lane in the transmit direction (toward the PMD).

83.5.3.3 Skew generation toward SP2

In an implementation with a physically instantiated PMD service interface, the PMA adjacent to the PMD service interface shall generate no more than 43 ns of Skew, and no more than 400 ps of Skew Variation between output lanes toward the PMD service interface (SP2 in Figure 80–4 and Figure 80–5). If there is a physically instantiated XLAUI/CAUI as well, then the Skew measured at SP1 is limited to no more than 29 ns of Skew and no more than 200 ps of Skew Variation.

83.5.3.4 Skew tolerance at SP5

In an implementation with a physically instantiated PMD service interface, the PMA adjacent to the PMD service interface (SP5) shall tolerate the maximum amount of Skew Variation allowed at SP5 (3.6 ns) between output lanes from the PMD service interface while maintaining the bit ordering and position of each PCSL on each PMA lane in the receive direction (toward the PCS).

83.5.3.5 Skew generation at SP6

In an implementation with one or more physically instantiated XLAUI/CAUI interfaces, at SP6 (the receive direction of the XLAUI/CAUI closest to the PCS), the PMA or group of PMAs between the PMD and the XLAUI/CAUI closest to the PCS shall deliver no more than 160 ns of Skew, and no more than 3.8 ns of Skew Variation between output lanes toward the XLAUI/CAUI in the Rx direction. If there is a physically instantiated PMD service interface as well, the Skew measured at SP5 is limited to no more than 145 ns of Skew and no more than 3.6 ns of Skew Variation. If there is no physically instantiated PMD service interface, the Skew measured at SP4 is limited to no more than 134 ns of Skew, and no more than 3.4 ns of Skew Variation.

83.5.3.6 Skew tolerance at SP6

In an implementation with one or more physically instantiated XLAUI/CAUI interfaces, the PMA between the XLAUI/CAUI closest to the PCS and the PCS shall tolerate the maximum amount of Skew Variation allowed at SP6 (3.8 ns) between input lanes while maintaining the bit order and position of PCSLs on lanes sent in the receive direction towards the PCS.

83.5.4 Delay constraints

The maximum cumulative delay contributed by up to four PMA stages in a PHY (sum of transmit and receive delays at one end of the link) shall meet the values specified in Table 83–1. 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.

Table 83–1—Delay constraints

Sublayer	Maximum (bit time)	Maximum (pause_quanta)	Maximum (ns)
40GBASE-R PMA	4096	8	102.4
100GBASE-R PMA	9216	18	92.16

83.5.5 Clocking architecture

A PMA with m input lanes and n output lanes shall clock the output lanes at m/n times the rate of the input lanes. This applies in both the Tx and Rx directions of transmission. In the case where the interfaces between the PMA client and the PMA and/or the PMA and the sublayer below the PMA are physically instantiated, the PMA may derive its input clock(s) from the electrical interface on one or more of the input lanes, and generate the output clock(s) with an appropriate PLL multiplier/divider circuit.

There is no requirement that the PMA clock all output lanes in unison. Examples of independent clocking of output lanes include the following:

- The case where the number of input and output lanes are equal (the PMA is provided for retiming and regeneration of the signal). This may be implemented without any rearrangement of PCSLs between input lanes and output lanes (although rearrangements are allowed), and such a PMA may be implemented by driving each output lane using the clock recovered from the corresponding input lane.
- If the number of input and output lanes have a common factor, the PMA may be partitioned such that PCSLs from a subset of the input lanes are mapped only to a subset of the output lanes (for example, a 10:4 PMA could be implemented as two 5:2 PMAs). The output clock for one subset may be independent of the output clock for other subset(s).

83.5.6 Signal drivers

For cases where the interface between the PMA client and the PMA, or between the PMA and the sublayer below the PMA represent a physically instantiated interface, the PMA provides electrical signal drivers for that interface. The electrical and jitter/timing specifications for these interfaces appear in

- Annex 83A, which specifies the XLAUI/CAUI interface for chip-to-chip applications.
- Annex 83B, which specifies the XLAUI/CAUI interface for chip-to-module applications.

- 86.2, which specifies the PMD service interface for 40GBASE-SR4 and 100GBASE-SR10 PMDs.
- 87.2, which specifies the PMD service interface for 40GBASE-LR4 PMDs.
- Annex 86A, which specifies the Parallel Physical Interface (XLPPi and CPPI), an optional physical instantiation of the PMD service interface for 40GBASE-SR4, 40GBASE-LR4, and 100GBASE-SR10 PMDs.

83.5.7 Link status

The PMA shall provide link status information to the PMA client using the PMA:IS_SIGNAL.indication primitive. The PMA continuously monitors the link status reported by the service interface below from the *inst*:IS_SIGNAL.indication primitive, and uses this as input to Signal Indication Logic (SIL) to determine the link status to report to the layer above. Other inputs to the SIL may include the status of clock and data recovery on the lanes from the service interface below the PMA and whether buffers/FIFOs have reached the required fill level to accommodate Skew Variation so that data is being sent on the output lanes.

83.5.8 PMA local loopback mode

PMA local loopback shall be provided by the PMA adjacent to the PMD for 40GBASE-KR4, 40GBASE-CR4, and 100BASE-CR10 PMDs. PMA local loopback mode is optional for other PMDs or for PMAs not adjacent to the PMD. If it is implemented, it shall be as described in this subclause (83.5.8).

The PMA sublayer may provide a local loopback function. The function involves looping back each input lane to the corresponding output lane. Each bit received from the PMA:IS_UNITDATA_*i*.request(tx_bit) primitive is looped back in the direction of the PCS using the PMA:IS_UNITDATA_*i*.indication(rx_bit) primitive.

During local loopback, the PMA performs normal bit muxing of PCSs per 83.5.2 onto the lanes in the Tx direction toward the service interface below the PMA.

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).

83.5.9 PMA remote loopback mode (optional)

PMA remote loopback mode is optional. If implemented, it shall be as described in this subclause (83.5.9).

Remote loopback, if provided, should be implemented in a PMA sublayer 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 *inst*:IS_UNITDATA_*i*.indication is looped back to the corresponding output lane toward the PMD via *inst*:IS_UNITDATA_*i*.request. Note that the service interface below the PMA can be provided by the FEC, PMD, or another PMA sublayer.

During remote loopback, the PMA performs normal bit muxing of PCSs per 83.5.2 onto the lanes 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).

83.5.10 PMA test patterns (optional)

Where the output lanes of the PMA appear on a physically instantiated interface XLAUI/CAUI or the PMD service interface (whether or not it is physically instantiated), the PMA may optionally generate and detect test patterns. These test patterns are used to test adjacent layer interfaces for an individual PMA sublayer or to perform testing between a physically instantiated interface of a PMA sublayer and external testing equipment.

The ability to generate each of the respective test patterns in each direction of transmission are indicated by the `PRBS9_Tx_generator_ability`, `PRBS9_Rx_generator_ability`, `PRBS31_Tx_generator_ability`, and `PRBS31_Rx_generator_ability` status variables, which if a Clause 45 MDIO is implemented are accessible through bits 1.1500.5, 1.1500.4, 1.1500.3, and 1.1500.1, respectively (see 45.2.1.99).

The ability to check PRBS31 test patterns in each direction of transmission are indicated by the `PRBS31_Tx_checker_ability` and `PRBS31_Rx_checker_ability` status variables, which, if a Clause 45 MDIO is implemented, are accessible through bits 1.1500.2 and 1.1500.0, respectively (see 45.2.1.99).

If supported, when send Tx PRBS31 test pattern is enabled by the `PRBS31_enable` and `PRBS_Tx_gen_enable` control variables, the PMA shall generate a PRBS31 pattern (as defined in 49.2.8) on each of the lanes toward the service interface below the PMA via the `inst:IS_UNITDATA_i.request` primitive. To avoid correlated crosstalk, it is highly recommended that the PRBS31 patterns generated on each lane be generated from independent, random seeds or at a minimum offset of 20 000 UI between the PRBS31 sequence on any lane and any other lane. When send Tx PRBS31 test pattern is disabled, the PMA returns to normal operation performing bit multiplexing as described in 83.5.2. If a Clause 45 MDIO is implemented, the `PRBS31_enable` and `PRBS_Tx_gen_enable` control variables are accessible through bits 1.1501.7 and 1.1501.3 (see 45.2.1.100).

If supported, when send Rx PRBS31 test pattern is enabled by the `PRBS31_enable` and `PRBS_Rx_gen_enable` control variables, the PMA shall generate a PRBS31 pattern on each of the lanes toward the PMA client via the `PMA:IS_UNITDATA_i.indication` primitive. While this test pattern is enabled, the PMA also generates `PMA:IS_SIGNAL.indication(SIGNAL_OK)` toward the PMA client independent of the link status at the service interface below the PMA. When send Rx PRBS31 test pattern is disabled, the PMA returns to normal operation performing bit multiplexing as described in 83.5.2. If a Clause 45 MDIO is implemented, the `PRBS31_enable` and `PRBS_Rx_gen_enable` control variables are accessible through bits 1.1501.7 and 1.1501.1 (see 45.2.1.100).

If supported, when check Tx PRBS31 test pattern mode is enabled by the `PRBS31_enable` and `PRBS_Tx_check_enable` control variables, the PMA shall check for the PRBS31 pattern on each of the lanes received from the PMA client via the `PMA:IS_UNITDATA_i.request` primitive. The checker shall increment the test-pattern error counter by one for each incoming bit error in the PRBS31 pattern for isolated single bit errors. Implementations should be capable of counting at least one error whenever one or more errors occur in a sliding 1000-bit window. If a Clause 45 MDIO is implemented, the `PRBS31_enable` and `PRBS_Tx_check_enable` control variables are accessible through bits 1.1501.7 and 1.1501.2 (see 45.2.1.100). The Tx test-pattern error counters `Ln0_PRBS_Tx_test_err_counter` through `Ln9_PRBS_Tx_test_err_counter` count, per lane, errors in detecting the PRBS31 pattern on the lanes from the PMA client. If a Clause 45 MDIO is implemented, these counters are accessible through registers 1.1600 through 1.1609 (see 45.2.1.102). When check Tx PRBS31 test pattern is disabled, the PMA expects normal traffic and test-pattern error counting does not continue. While in check Tx PRBS31 test-pattern mode, bit multiplexing continues as described in 83.5.2. Note that bit multiplexing of per-lane PRBS31 may produce a signal which is not meaningful for downstream sublayers.

If supported, when check Rx PRBS31 test-pattern mode is enabled by the `PRBS31_enable` and `PRBS_Rx_check_enable` control variables, the PMA shall check for the PRBS31 pattern on each of the lanes received from the service interface below the PMA via the `inst:IS_UNITDATA_i.indication` primitive.

If a Clause 45 MDIO is implemented, the PRBS31_enable and PRBS_Rx_check_enable control variables are accessible through bits 1.1501.7 and 1.1501.0 (see 45.2.1.100). The Rx test-pattern error counters Ln0_PRBS_Rx_test_err_counter through Ln9_PRBS_Rx_test_err_counter count, per lane, errors in detecting the PRBS31 pattern on the lanes from the service interface below the PMA. If a Clause 45 MDIO is implemented, these counters are accessible through registers 1.1700 through 1.1709 (see 45.2.1.103). While in check Rx PRBS31 test-pattern mode, the PMA:IS_SIGNAL.indication primitive does not indicate a valid signal. When check Rx PRBS31 test pattern is disabled, the PMA returns to normal operation performing bit multiplexing as described in 83.5.2.

If supported, when send Tx PRBS9 test-pattern mode is enabled by the PRBS9_enable and PRBS_Tx_gen_enable control variables, the PMA shall generate a PRBS9 pattern (as defined in Table 68–6) on each lane toward the service interface below the PMA via the *inst*:IS_UNITDATA_i.request primitive. If a Clause 45 MDIO is implemented, the PRBS9_enable and PRBS_Tx_gen_enable control variables are accessible through bits 1.1501.6 and 1.1501.3 (see 45.2.1.100). When send Tx PRBS9 test-pattern mode is disabled, the PMA returns to normal operation performing bit multiplexing as described in 83.5.2.

If supported, when send Rx PRBS9 test-pattern mode is enabled by the PRBS9_enable and PRBS_Rx_gen_enable control variables, the PMA shall generate a PRBS9 pattern on each lane toward the PMA client via the PMA:IS_UNITDATA_i.indication primitive. The PMA will also generate PMA:IS_SIGNAL.indication(SIGNAL_OK) toward the PMA client independent of the link status at the service interface below the PMA. If a Clause 45 MDIO is implemented, the PRBS9_enable and PRBS_Rx_gen_enable control variables are accessible through bits 1.1501.6 and 1.1501.1 (see 45.2.1.100). When send Rx PRBS9 test-pattern mode is disabled, the PMA returns to normal operation performing bit multiplexing as described in 83.5.2.

Note that PRBS9 is intended to be checked by external test gear, and no PRBS9 checking function is provided within the PMA.

Transmit square wave test-pattern mode optionally applies to each lane of the Tx direction PMA towards a physically instantiated XLAUI/CAUI or towards the PMD service interface whether or not it is physically instantiated. The ability to perform this function is indicated by the Square_wave_ability status variable. If a Clause 45 MDIO is implemented, the Square_wave_ability status variable is accessible through the Square wave test ability bit 1.1500.12 (see 45.2.1.99). If implemented, the transmit square wave test-pattern mode is enabled by control variables Square_wave_enable_0 through Square_wave_enable_9. If a Clause 45 MDIO is implemented, these control variables are accessible through the square wave testing control and status register bits 1.1510.0 through 1.1510.9 (limited to the number of lanes of the service interface below the PMA, see 45.2.1.101). When enabled, the PMA shall generate a square wave test pattern (8 ones followed by 8 zeros) on the square wave enabled lanes toward the service interface below the PMA via the *inst*:IS_UNITDATA_i.request primitive. Lanes for which square wave is not enabled will transmit normal data resulting from the bit multiplexing operations described in 83.5.2 or test patterns as determined by other registers. When transmit square wave test pattern is disabled for all lanes, the PMA will perform normal operation performing bit multiplexing as described in 83.5.2 or transmit test patterns as determined by other registers.

83.6 PMA MDIO function mapping

The optional MDIO capability described in Clause 45 describes several variables that provide control and status information for and about the PMA. Since a given implementation may employ more than one PMA sublayer, the PMA control and status information is organized into multiple addressable instances, one for each possible PMA sublayer. See 45.2.1 and 83.1.4 for the allocation of MMD addresses to PMA sublayers. Control and status registers for MMD 8, 9, 10, and 11 will use the Extended PMA control and status registers at identical locations to those for MMD 1.

Mapping of MDIO control variables to PMA control variables is shown in Table 83–2. Mapping of MDIO status variables to PMA status variables is shown in Table 83–3. Mapping of MDIO counter to PMA counters is shown in Table 83–4. These tables provide the register and bit numbers for the PMA addressed as MMD 1. For implementations with multiple PMA sublayers, additional PMA sublayers use the corresponding register and bit numbers in MMDs 8, 9, 10, and 11 as necessary.

Table 83–2—MDIO/PMA control variable mapping

MDIO variable	PMA/PMD register name	Register/ bit number	PMA control variable
PMA remote loopback	PMA/PMD control 1	1.0.1	Remote_loopback_enable
PMA local loopback	PMA/PMD control 1	1.0.0	Local_loopback_enable
PRBS31 pattern enable	PRBS pattern testing control	1.1501.7	PRBS31_enable
PRBS9 pattern enable	PRBS pattern testing control	1.1501.6	PRBS9_enable
Tx generator enable	PRBS pattern testing control	1.1501.3	PRBS_Tx_gen_enable
Tx checker enable	PRBS pattern testing control	1.1501.2	PRBS_Tx_check_enable
Rx generator enable	PRBS pattern testing control	1.1501.1	PRBS_Rx_gen_enable
Rx checker enable	PRBS pattern testing control	1.1501.0	PRBS_Rx_check_enable
Lane 0 SW enable	Square wave testing control	1.1510.0	Square_wave_enable_0
Lane 1 SW enable	Square wave testing control	1.1510.1	Square_wave_enable_1
Lane 2 SW enable	Square wave testing control	1.1510.2	Square_wave_enable_2
Lane 3 SW enable	Square wave testing control	1.1510.3	Square_wave_enable_3
Lane 4 SW enable	Square wave testing control	1.1510.4	Square_wave_enable_4
Lane 5 SW enable	Square wave testing control	1.1510.5	Square_wave_enable_5
Lane 6 SW enable	Square wave testing control	1.1510.6	Square_wave_enable_6
Lane 7 SW enable	Square wave testing control	1.1510.7	Square_wave_enable_7
Lane 8 SW enable	Square wave testing control	1.1510.8	Square_wave_enable_8
Lane 9 SW enable	Square wave testing control	1.1510.9	Square_wave_enable_9

Table 83–3—MDIO/PMA status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMA status variable
PMA remote loopback ability	40G/100G PMA/PMD extended ability register	1.13.15	Remote_loopback_ability
PMA local loopback ability	PMA/PMD status 2 register	1.8.0	Local_loopback_ability
PRBS9 Tx generator ability	Test-pattern ability register	1.1500.5	PRBS9_Tx_generator_ability
PRBS9 Rx generator ability	Test-pattern ability register	1.1500.4	PRBS9_Rx_generator_ability
PRBS31Tx generator ability	Test-pattern ability register	1.1500.3	PRBS31_Tx_generator_ability
PRBS31Tx checker ability	Test-pattern ability register	1.1500.2	PRBS31_Tx_checker_ability
PRBS31 Rx generator ability	Test-pattern ability register	1.1500.1	PRBS31_Rx_generator_ability
PRBS31 Rx checker ability	Test-pattern ability register	1.1500.0	PRBS31_Rx_checker_ability
Square wave test ability	Test-pattern ability register	1.1500.12	Square_wave_ability

Table 83–4—MDIO/PMA counters mapping

MDIO variable	PMA/PMD register name	Register/bit number	PMA status variable
Error counter Tx, lane 0	PRBS Tx pattern testing error counter, lane 0	1.1600	Ln0_PRBS_Tx_test_err_counter
Error counter Tx, lane 1	PRBS Tx pattern testing error counter, lane 1	1.1601	Ln1_PRBS_Tx_test_err_counter
Error counter Tx, lane 2	PRBS Tx pattern testing error counter, lane 2	1.1602	Ln2_PRBS_Tx_test_err_counter
Error counter Tx, lane 3	PRBS Tx pattern testing error counter, lane 3	1.1603	Ln3_PRBS_Tx_test_err_counter
Error counter Tx, lane 4	PRBS Tx pattern testing error counter, lane 4	1.1604	Ln4_PRBS_Tx_test_err_counter
Error counter Tx, lane 5	PRBS Tx pattern testing error counter, lane 5	1.1605	Ln5_PRBS_Tx_test_err_counter
Error counter Tx, lane 6	PRBS Tx pattern testing error counter, lane 6	1.1606	Ln6_PRBS_Tx_test_err_counter
Error counter Tx, lane 7	PRBS Tx pattern testing error counter, lane 7	1.1607	Ln7_PRBS_Tx_test_err_counter
Error counter Tx, lane 8	PRBS Tx pattern testing error counter, lane 8	1.1608	Ln8_PRBS_Tx_test_err_counter
Error counter Tx, lane 9	PRBS Tx pattern testing error counter, lane 9	1.1609	Ln9_PRBS_Tx_test_err_counter

Table 83–4—MDIO/PMA counters mapping (*continued*)

MDIO variable	PMA/PMD register name	Register/bit number	PMA status variable
Error counter Rx, lane 0	PRBS Rx pattern testing error counter, lane 0	1.1700	Ln0_PRBS_Rx_test_err_counter
Error counter Rx, lane 1	PRBS Rx pattern testing error counter, lane 1	1.1701	Ln1_PRBS_Rx_test_err_counter
Error counter Rx, lane 2	PRBS Rx pattern testing error counter, lane 2	1.1702	Ln2_PRBS_Rx_test_err_counter
Error counter Rx, lane 3	PRBS Rx pattern testing error counter, lane 3	1.1703	Ln3_PRBS_Rx_test_err_counter
Error counter Rx, lane 4	PRBS Rx pattern testing error counter, lane 4	1.1704	Ln4_PRBS_Rx_test_err_counter
Error counter Rx, lane 5	PRBS Rx pattern testing error counter, lane 5	1.1705	Ln5_PRBS_Rx_test_err_counter
Error counter Rx, lane 6	PRBS Rx pattern testing error counter, lane 6	1.1706	Ln6_PRBS_Rx_test_err_counter
Error counter Rx, lane 7	PRBS Rx pattern testing error counter, lane 7	1.1707	Ln7_PRBS_Rx_test_err_counter
Error counter Rx, lane 8	PRBS Rx pattern testing error counter, lane 8	1.1708	Ln8_PRBS_Rx_test_err_counter
Error counter Rx, lane 9	PRBS Rx pattern testing error counter, lane 9	1.1709	Ln9_PRBS_Rx_test_err_counter

83.7 Protocol implementation conformance statement (PICS) proforma for Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R⁸

83.7.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R, 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 same, can be found in Clause 21.

83.7.2 Identification

83.7.2.1 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).	

83.7.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 83, Physical Medium Attachment (PMA) sublayer, type 40GBASE-R and 100GBASE-R
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 802.3-2012.)	
Date of Statement	

⁸Copyright 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.

83.7.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*PMA40	PMA for 40GBASE-R	83.1.1		O.1	Yes [] No []
*PMA100	PMA for 100GBASE-R	83.1.1		O.1	Yes [] No []
LANES_UPSTREAM	Number of lanes in direction of PCS	83.1.4	Divisor of number of PCS lanes	PMA40:M	4 []
				PMA100:M	4 [] 10 [] 20 []
LANES_DOWNSTREAM	Number of lanes in direction of PMD	83.1.4	Divisor of number of PCS lanes	PMA40:M	1 [] 4 []
				PMA100:M	4 [] 10 [] 20 []
RX_CLOCK	Signaling rate of output lanes in Rx direction	83.5.5	LANES_DOWNSTREAM / LANES_UPSTREAM times signaling rate of input lanes in Rx direction	M	Yes [] No []
TX_CLOCK	Signaling rate of output lanes in Tx direction	83.5.5	LANES_UPSTREAM / LANES_DOWNSTREAM times signaling rate of input lanes in Tx direction	M	Yes [] No []
LANE_MAPPING	Maintain lane mapping while link is in operation	83.5.2	Maintain sequence of PCSs on all output lanes	M	Yes [] No []
LNKS	PMA link status	83.5.7	Meets the requirements of 83.5.7	M	Yes [] No []
JTP	Supports test-pattern mode	83.5.10		O	Yes [] No [] N/A []
*KRCR	PMA adjacent to the PMD for 40GBASE-KR4, 40GBASE-CR4, or 100GBASE-CR10	83.5.8		O	Yes [] No []
*LBL	PMA local loopback	83.5.8	Supports local loopback	KRCR:M !KRCR:O	Yes [] No [] N/A []
*LBR	PMA remote loopback	83.5.9	Supports remote loopback	O	Yes [] No [] N/A []
MD	MDIO	83.6	Registers and interface supported	O	Yes [] No []

Item	Feature	Subclause	Value/Comment	Status	Support
*USP1SP6	Physically instantiated XLAUI or CAUI above (toward PCS)	83.5.3		O	Yes [] No []
*DSP1SP6	Physically instantiated XLAUI or CAUI below (toward PMD)	83.5.3		O	Yes [] No []
*SP2SP5	Physically instantiated PMD service interface	83.5.3		O	Yes [] No []
*PPI	PMD service interface instantiated as nPPI	83.5.1, 83.5.5		O	Yes [] No []
UNAU1	Electrical and timing requirements of Annex 83A or Annex 83B as appropriate met by upstream XLAUI/CAUI	83.5.1, 83.5.5		USP1SP6: M	Yes [] No []
DNAUI	Electrical and timing requirements of Annex 83A or Annex 83B as appropriate met by downstream XLAUI/CAUI	83.5.1, 83.5.5		DSP1SP6: M	Yes [] No []
PPIET	Electrical and timing requirements of Annex 86A met by PMD service interface	83.5.1, 83.5.5		PPI:M	Yes [] No []
DELAY40	Roundtrip delay limit for 40GBASE-R	83.5.4	No more than 4096 BT or 8 pause quanta	PMA40:M	Yes [] No []
DELAY100	Roundtrip delay limit for 100GBASE-R	83.5.4	No more than 9216 BT or 18 pause quanta	PMA100: M	Yes [] No []

83.7.4 Skew generation and tolerance

Item	Feature	Subclause	Value/Comment	Status	Support
S1	Skew generation toward SP1 in Tx direction	83.5.3.1	≤ 29 ns	DSP1SP6:M	Yes [] No []
S2	Skew variation generation toward SP1 in Tx direction	83.5.3.1	≤ 200 ps	DSP1SP6:M	Yes [] No []
S3	Skew variation tolerance at SP1	83.5.3.2	Minimum 200 ps	USP1SP6:M	Yes [] No []
S4	Skew generation toward SP2 in Tx direction	83.5.3.3	≤ 43 ns	SP2SP5:M	Yes [] No []
S5	Skew variation generation toward SP2 in Tx direction	83.5.3.3	≤ 400 ps	SP2SP5:M	Yes [] No []
S6	Skew variation tolerance at SP5	83.5.3.4	Minimum 3.6 ns	SP2SP5:M	Yes [] No []
S7	Skew generation toward SP6 in Rx direction	83.5.3.5	≤ 160 ns	USP1SP6:M	Yes [] No []
S8	Skew variation generation toward SP6 in Rx direction	83.5.3.5	≤ 3.8 ns	USP1SP6:M	Yes [] No []
S9	Skew variation tolerance at SP6	83.5.3.6	Minimum 3.8 ns	DSP1SP6:M	Yes [] No []

83.7.5 Test patterns

Item	Feature	Subclause	Value/Comment	Status	Support
*JTP1	Physically Instantiated XLAUI/CAUI between PMA and PMA client	83.5.10		O	Yes [] No []
*JTP2	Physically Instantiated XLAUI/CAUI between PMA and sublayer below the PMA, or adjacent to PMD whether or not PMD service interface is physically instantiated as nPPI	83.5.10		O	Yes [] No []
J1	Send PRBS31 Tx	83.5.10		JTP2:O	Yes [] No []
J2	Send PRBS31 Rx	83.5.10		JTP1:O	Yes [] No []
J3	Check PRBS31 Tx	83.5.10		JTP1:O	Yes [] No []
J4	Check PRBS31 Rx	83.5.10		JTP2:O	Yes [] No []
J5	Send PRBS9 Tx	83.5.10		JTP2:O	Yes [] No []

Item	Feature	Subclause	Value/Comment	Status	Support
J6	Send PRBS9 Rx	83.5.10		JTP1:O	Yes [] No []
J7	Send square wave Tx	83.5.10		JTP2:O	Yes [] No []

83.7.6 Loopback modes

Item	Feature	Subclause	Value/Comment	Status	Support
LB1	PMA local loopback implemented	83.5.8	Meets the requirements of 83.5.8	LBL:M	Yes [] No []
LB2	PMA remote loopback implemented	83.5.9	Meets the requirements of 83.5.9	LBR:M	Yes [] No []

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

84.1 Overview

This clause specifies the 40GBASE-KR4 PMD. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 84–1, 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.

Table 84–1—Physical Layer clauses associated with the 40GBASE-KR4 PMD

Associated clause	40GBASE-KR4
81—RS	Required
81—XLGMII ^a	Optional
82—PCS for 40GBASE-R	Required
74—FEC for BASE-R	Optional
83—PMA for 40GBASE-R	Required
83A—XLAUI	Optional
73—Auto-Negotiation	Required

^aThe XLGMII is an optional interface. However, if the XLGMII is not implemented, a conforming implementation must behave functionally as though the RS and XLGMII were present.

Figure 84–1 shows the relationship of the 40GBASE-KR4 PMD and MDI (shown shaded) with other sublayers to the ISO/IEC Open System Interconnection (OSI) reference model.

84.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 40GBASE-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 40GBASE-KR4 PMD has four parallel bit streams, hence $i = 0$ to 3.

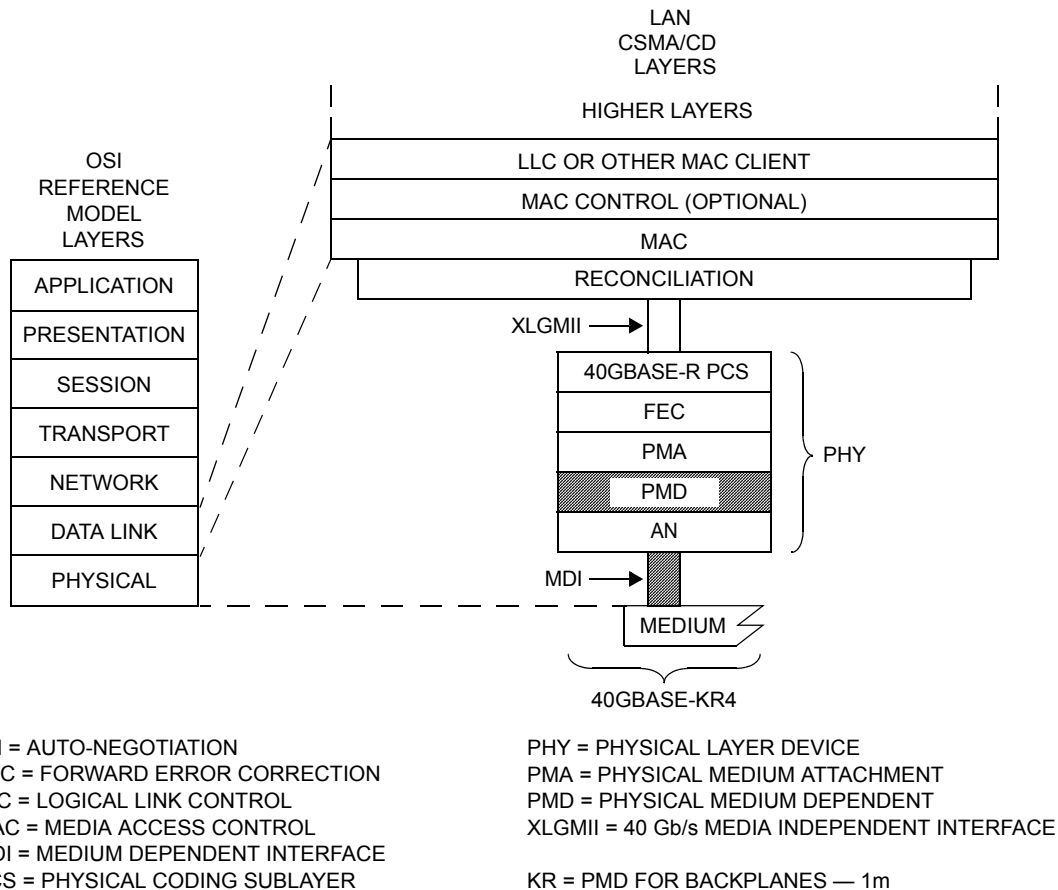


Figure 84–1—40GBASE-KR4 PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

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 10.3125 GBd.

SIGNAL_DETECT in 40GBASE-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.

84.3 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.)

84.4 Delay constraints

The sum of the transmit and the receive delays at one end of the link contributed by the 40GBASE-KR4 PMD, AN, and the medium in one direction shall be no more than 2048 bit times (4 pause_quanta or 51.2 ns). It is assumed that the one way delay through the medium is no more than 320 bit times (8 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.

84.5 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 43 ns and the Skew Variation at SP2 is limited to 400 ps.

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

The Skew at SP4 (the receiver MDI) shall be less than 134 ns and the Skew Variation at SP4 shall be less than 3.4 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 145 ns and the Skew Variation at SP5 shall be less than 3.6 ns.

For more information on Skew and Skew Variation see 80.5.

84.6 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 84–2, and MDIO status variables to PMD status variables as shown in Table 84–3.

84.7 PMD functional specifications

84.7.1 Link block diagram

The 40GBASE-KR4 PMD uses the same Link block diagram as 10GBASE-KX4, as defined in 71.6.1.

84.7.2 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. A positive output voltage of SL<p> minus SL<n> (differential voltage) shall correspond to tx_bit = one.

Table 84–2—MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable register	1.9.0	Global_PMD_transmit_disable
PMD transmit disable 3	PMD transmit disable register	1.9.4	PMD_transmit_disable_3
PMD transmit disable 2	PMD transmit disable register	1.9.3	PMD_transmit_disable_2
PMD transmit disable 1	PMD transmit disable register	1.9.2	PMD_transmit_disable_1
PMD transmit disable 0	PMD transmit disable register	1.9.1	PMD_transmit_disable_0
Restart training	BASE-R PMD control register	1.150.0	mr_restart_training
Training enable	BASE-R PMD control register	1.150.1	mr_training_enable

Table 84–3—MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	Global_PMD_signal_detect
PMD receive signal detect 3	PMD receive signal detect register	1.10.4	PMD_signal_detect_3
PMD receive signal detect 2	PMD receive signal detect register	1.10.3	PMD_signal_detect_2
PMD receive signal detect 1	PMD receive signal detect register	1.10.2	PMD_signal_detect_1
PMD receive signal detect 0	PMD receive signal detect register	1.10.1	PMD_signal_detect_0
Receiver status 3	BASE-R PMD status register	1.151.12	rx_trained_3
Frame lock 3	BASE-R PMD status register	1.151.13	frame_lock_3
Start-up protocol status 3	BASE-R PMD status register	1.151.14	training_3
Training failure 3	BASE-R PMD status register	1.151.15	training_failure_3
Receiver status 2	BASE-R PMD status register	1.151.8	rx_trained_2
Frame lock 2	BASE-R PMD status register	1.151.9	frame_lock_2
Start-up protocol status 2	BASE-R PMD status register	1.151.10	training_2
Training failure 2	BASE-R PMD status register	1.151.11	training_failure_2
Receiver status 1	BASE-R PMD status register	1.151.4	rx_trained_1
Frame lock 1	BASE-R PMD status register	1.151.5	frame_lock_1
Start-up protocol status 1	BASE-R PMD status register	1.151.6	training_1

Table 84–3—MDIO/PMD status variable mapping (continued)

MDIO status variable	PMA/PMD register name	Register/ bit number	PMD status variable
Training failure 1	BASE-R PMD status register	1.151.7	training_failure_1
Receiver status 0	BASE-R PMD status register	1.151.0	rx_trained_0
Frame lock 0	BASE-R PMD status register	1.151.1	frame_lock_0
Start-up protocol status 0	BASE-R PMD status register	1.151.2	training_0
Training failure 0	BASE-R PMD status register	1.151.3	training_failure_0

84.7.3 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<p> minus DL<n> (differential voltage) shall correspond to rx_bit = one.

84.7.4 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 40GBASE-KR4 to indicate the successful completion of the start-up protocol on all lanes.

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.

84.7.5 PMD lane-by-lane signal detect function

If the MDIO interface is implemented, then 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 (see 45.2.1.9).

84.7.6 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 the 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 with TX disabled as specified in Table 72–7.
- b) If a PMD_fault (84.7.9) is detected, then the PMD may turn off the electrical transmitter in all lanes.
- c) Loopback, as defined in 84.7.8, shall not be affected by Global_PMD_transmit_disable.

If the MDIO interface is implemented, then this function shall map to the Global_PMD_transmit_disable bit as specified in 45.2.1.8.7.

84.7.7 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 with TX disabled as specified in Table 72–7.
- b) If a PMD_fault (84.7.9) 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 84.7.8, shall not be affected by PMD_transmit_disable_*i*.

84.7.8 Loopback mode

Local loopback shall be provided by the adjacent PMA (see 83.5.8) for the 40GBASE-KR4 PMD 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 that continues to send data (unless disabled). The method of implementing loopback mode is not defined by this standard.

Control of the loopback function is specified in 45.2.1.1.5.

NOTE 1—The signal paths that are exercised in the loopback mode are implementation specific, but it is recommended that these signal paths 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.

84.7.9 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.

84.7.10 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.

84.7.11 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 PMA/PMD Receive fault bit as specified in 45.2.1.7.5.

84.7.12 PMD control function

Each lane of the 40GBASE-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.

84.8 40GBASE-KR4 electrical characteristics

84.8.1 Transmitter characteristics

Transmitter electrical characteristics at TP1 for 40GBASE-KR4 shall be the same as 10GBASE-KR, as detailed in 72.7.1.1 through 72.7.1.11.

84.8.1.1 Test fixture

The test fixture defined for 10GBASE-KR in 72.7.1.1 shall be used on all lanes.

84.8.2 Receiver characteristics

Receiver electrical characteristics at TP4 for 40GBASE-KR4 shall be the same as 10GBASE-KR, as detailed in 72.7.2.1 through 72.7.2.5.

84.8.2.1 Receiver interference tolerance

The receiver interference tolerance tests shall be the same as those described for 10GBASE-KR in 72.7.2.1 and Annex 69A.

For 40GBASE-KR4, each lane shall be tested individually using the methodology defined in Annex 69A with the transmitters of the unused lanes active and terminated by the reference impedance.

84.9 Interconnect characteristics

Informative interconnect characteristics for 40GBASE-KR4 are provided in Annex 69B.

84.10 Environmental specifications

84.10.1 General safety

All equipment subject to this clause shall conform to applicable sections (including isolation requirements) of IEC 60950-1.

84.10.2 Network safety

The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

84.10.3 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.

84.10.4 Electromagnetic compatibility

A system integrating the 40GBASE-KR4 PHY shall comply with applicable local and national codes for the limitation of electromagnetic interference.

84.10.5 Temperature and humidity

A system integrating the 40GBASE-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.

84.11 Protocol implementation conformance statement (PICS) proforma for Clause 84, Physical Medium Dependent sublayer and baseband medium, type 40GBASE-KR4⁹

84.11.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 84, Physical Medium Dependent sublayer and baseband medium, type 40GBASE-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.

84.11.2 Identification

84.11.2.1 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).	

84.11.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 84, Physical Medium Dependent sublayer and baseband medium, type 40GBASE-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 802.3-2012.)	

Date of Statement	
-------------------	--

⁹Copyright 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.

84.11.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
XGE	XLGMII	84.1	Interface is supported	O	Yes [] No []
PCS	Support of 40GBASE-R PCS	84.1		M	Yes []
PMA	Support of 40GBASE-R PMA	84.1		M	Yes []
XLAUI	XLAUI	84.1	Interface is supported	O	Yes [] No []
AN	Auto-Negotiation	84.1	Device implements Auto-Negotiation	M	Yes []
FEC	Forward Error Correction	84.1	Device implements BASE-R Forward Error Correction	O	Yes []
DC	Delay constraints	84.4	Device conforms to delay constraints	M	Yes []
DSC	Skew constraints	84.5	Device conforms to Skew and Skew Variation constraints	M	Yes []
*MD	MDIO interface	84.6	Device implements MDIO	O	Yes [] No []
*TD	Global_PMD_transmit_disable	84.7.6		O	Yes [] No []

84.11.4 PICS proforma tables for Clause 84, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-KR4**84.11.4.1 PMD functional specifications**

Item	Feature	Subclause	Value/Comment	Status	Support
FS1	Transmit function	84.7.2	Conveys bits from PMD service interface to MDI	M	Yes []
FS2	Transmitter signal	84.7.2	A positive differential voltage corresponds to tx_bit = one	M	Yes []
FS3	Receive function	84.7.3	Conveys bits from MDI to PMD service interface	M	Yes []
FS4	Receiver signal	84.7.3	A positive differential voltage corresponds to rx_bit = one	M	Yes []
FS5	Signal detect	84.7.4	Report to PMD service interface	M	Yes []
FS6	SIGNAL_DETECT value	84.7.4	Set to FAIL on reset	M	Yes []
FS7	SIGNAL_DETECT value	84.7.4	Set to OK when training completes successfully	M	Yes []
FS8	SIGNAL_DETECT value	84.7.4	Set to OK when training disabled	M	Yes []
FS9	Transmit disable requirements	84.7.6	Requirements of 84.7.6 and 84.7.7	TD:M	Yes [] N/A []
FS10	Loopback support	84.7.8	Provided for 40GBASE-KR4 PMD by transmitter and receiver	M	Yes []
FS11	Control function	84.7.12	The same control function as 10GBASE-KR is used	M	Yes []
FS12	Random seed	84.7.12	The random seed shall be different for each lane	M	Yes []

84.11.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
MF1	MDIO Variable Mapping	84.6	Per Table 84–2, and Table 84–3	MD:M	Yes [] N/A []
MF2	Global signal detect	84.7.4	Value described in 45.2.1.9.7	MD:M	Yes [] N/A []
MF3	Lane-by-lane signal detect	84.7.5	Value as described in 84.7.5	MD:M	Yes [] N/A []
MF4	PMD_transmit_fault function	84.7.10	Sets PMD_transmit_fault as specified in 45.2.1.7.4	MD:M	Yes [] N/A []
MF5	PMD_receive_fault function	84.7.11	Sets PMD_receive_fault as specified in 45.2.1.7.5	MD:M	Yes [] N/A []
MF6	PMD training status	84.7.12	Sets training status as specified in 45.2.1.80	MD:M	Yes [] N/A []

84.11.4.3 Transmitter electrical characteristics

Item	Feature	Subclause	Value/Comment	Status	Support
TC1	Transmitter characteristics	84.8.1	The same transmitter characteristics as 10GBASE-KR are used	M	Yes []
TC2	Test fixture	84.8.1.1	The same test fixture as 10GBASE-KR is used	M	Yes []

84.11.4.4 Receiver electrical characteristics

Item	Feature	Subclause	Value/Comment	Status	Support
RC1	Receiver characteristics	84.8.2	The same receiver characteristics as 10GBASE-KR are used	M	Yes []
RC2	Receiver interference tolerance	84.8.2.1	The same receiver interference tolerance test as 10GBASE-KR is used	M	Yes []
RC3	Receiver interference tolerance testing	84.8.2.1	Each lane shall be tested individually using the methodology defined in Annex 69A with the transmitters of the unused lanes active and terminated by the reference impedance	M	Yes []

84.11.4.5 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
ES1	General safety	84.10.1	Complies with applicable section of IEC 60950-1	M	Yes []
ES2	Electromagnetic interference	84.10.4	Complies with applicable local and national codes	M	Yes []

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

85.1 Overview

This clause specifies the 40GBASE-CR4 PMD and the 100GBASE-CR10 PMD (including MDI) and the baseband medium. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 85–1, 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.

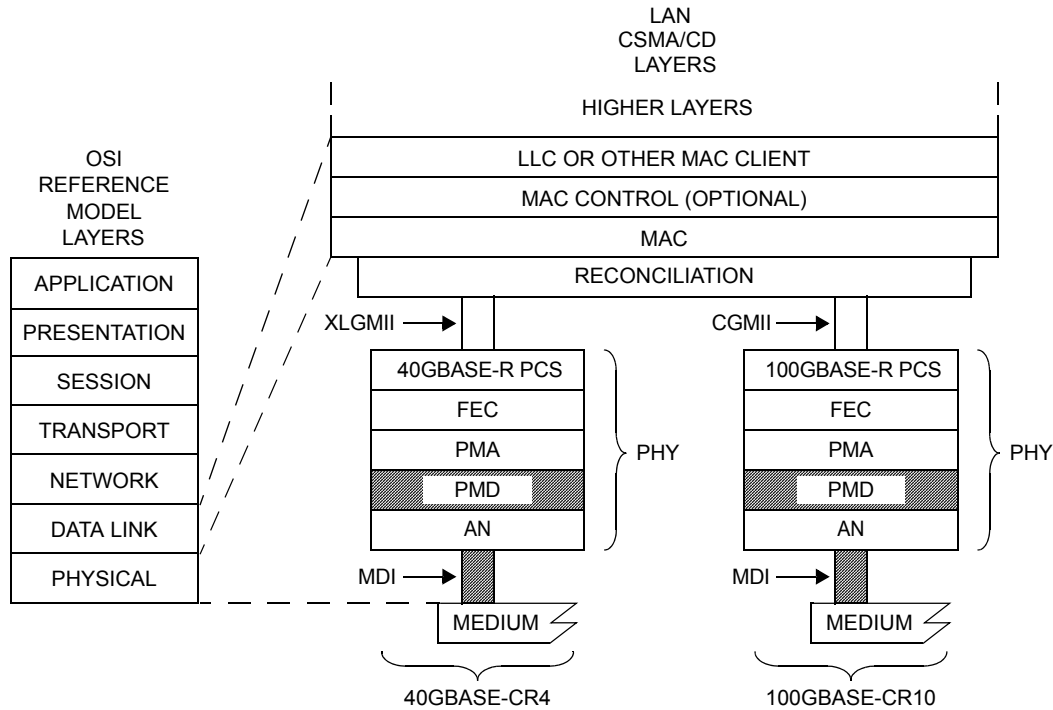
Table 85–1—Physical Layer clauses associated with the 40GBASE-CR4 and 100GBASE-CR10 PMDs

Associated clause	40GBASE-CR4	100GBASE-CR10
81—RS	Required	Required
81—XLGMII ^a	Optional	Not applicable
81—CGMII ^b	Not applicable	Optional
82—PCS for 40GBASE-R	Required	Not applicable
82—PCS for 100GBASE-R	Not applicable	Required
74—FEC	Optional	Optional
83—PMA for 40GBASE-R	Required	Not applicable
83—PMA for 100GBASE-R	Not applicable	Required
83A—XLAUI	Optional	Not applicable
83A—CAUI	Not applicable	Optional
73—Auto-Negotiation	Required	Required

^aThe XLGMII is an optional interface. However, if the XLGMII is not implemented, a conforming implementation must behave functionally as though the RS and XLGMII were present.

^bThe 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 85–1 shows the relationship of the 40GBASE-CR4 and 100GBASE-CR10 PMD sublayers and MDI to the ISO/IEC Open System Interconnection (OSI) reference model.



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
 CR = PMD FOR COPPER — 7 m

Figure 85–1—40GBASE-CR4 and 100GBASE-CR10 PMDs relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

85.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 40GBASE-CR4 and 100GBASE-CR10 PMDs. The service interfaces for these PMDs are described in an abstract manner and do not imply any particular implementation. The PMD service interface supports the exchange of encoded data between the PMA and PMD entities. The PMD translates the encoded data to and from signals suitable for the specified 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 40GBASE-CR4 PMD has four parallel bit streams, hence $i = 0$ to 3 for 40GBASE-CR4, and the 100GBASE-CR10 PMD has ten parallel bit streams, hence $i = 0$ to 9 for 100GBASE-CR10. The PMA (or the PMD) continuously sends four or ten parallel bit streams to the PMD (or the PMA), one per lane, each at a nominal signaling rate of 10.3125 GBd.

For 40GBASE-CR4, SIGNAL_DETECT indicates the successful completion of the start-up protocol on all four lanes. For 100GBASE-CR10, SIGNAL_DETECT indicates the successful completion of the start-up protocol on all ten 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 defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication primitive.

85.3 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.)

85.4 Delay constraints

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.

The sum of the transmit and the receive delays at one end of the link contributed by the 40GBASE-CR4 PMD and AN shall be no more than 4096 bit times (8 pause_quanta or 102.4 ns). The delay through the medium is not included.

The sum of the transmit and the receive delays at one end of the link contributed by the 100GBASE-CR10 PMD and AN shall be no more than 9728 bit times (19 pause_quanta or 97.28 ns). The delay through the medium is not included.

85.5 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 are 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 43 ns and the Skew Variation at SP2 is limited to 400 ps.

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

The Skew at SP4 (the receiver MDI) shall be less than 134 ns and the Skew Variation at SP4 shall be less than 3.4 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 145 ns and the Skew Variation at SP5 shall be less than 3.6 ns. For more information on Skew and Skew Variation, see 80.5.

85.6 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. Mapping of MDIO control variables to PMD control variables is illustrated in Table 85–2 for 40GBASE-CR4 and 100GBASE-CR10. Mapping of MDIO status variables to PMD status variables is illustrated in Table 85–3 for 40GBASE-CR4 and 100GBASE-CR10.

Table 85–2—40GBASE-CR4 and 100GBASE-CR10 MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/ bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable register	1.9.0	Global_PMD_transmit_disable
PMD transmit disable 9 to PMD transmit disable 0	PMD transmit disable register	1.9.10 to 1.9.1	PMD_transmit_disable_9 to PMD_transmit_disable_0
Restart training	BASE-R PMD control register	1.150.0	mr_restart_training
Training enable	BASE-R PMD control register	1.150.1	mr_training_enable

Table 85–3—40GBASE-CR4 and 100GBASE-CR10 MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/ bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	Global_PMD_signal_detect
PMD receive signal detect 9 to PMD receive signal detect 0	PMD receive signal detect register	1.10.10 to 1.10.1	PMD_signal_detect_9 to PMD_signal_detect_0
Receiver status 9	BASE-R PMD status 3 register	1.157.4	rx_trained_9
Frame lock 9	BASE-R PMD status 3 register	1.157.5	frame_lock_9
Start-up protocol status 9	BASE-R PMD status 3 register	1.157.6	training_9
Training failure 9	BASE-R PMD status 3 register	1.157.7	training_failure_9
Receiver status 8	BASE-R PMD status 3 register	1.157.0	rx_trained_8
Frame lock 8	BASE-R PMD status 3 register	1.157.1	frame_lock_8
Start-up protocol status 8	BASE-R PMD status 3 register	1.157.2	training_8
Training failure 8	BASE-R PMD status 3 register	1.157.3	training_failure_8
Receiver status 7	BASE-R PMD status 2 register	1.156.12	rx_trained_7
Frame lock 7	BASE-R PMD status 2 register	1.156.13	frame_lock_7

Table 85–3—40GBASE-CR4 and 100GBASE-CR10 MDIO/PMD status variable mapping (continued)

MDIO status variable	PMA/PMD register name	Register/ bit number	PMD status variable
Start-up protocol status 7	BASE-R PMD status 2 register	1.156.14	training_7
Training failure 7	BASE-R PMD status 2 register	1.156.15	training_failure_7
Receiver status 6	BASE-R PMD status 2 register	1.156.8	rx_trained_6
Frame lock 6	BASE-R PMD status 2 register	1.156.9	frame_lock_6
Start-up protocol status 6	BASE-R PMD status 2 register	1.156.10	training_6
Training failure 6	BASE-R PMD status 2 register	1.156.11	training_failure_6
Receiver status 5	BASE-R PMD status 2 register	1.156.4	rx_trained_5
Frame lock 5	BASE-R PMD status 2 register	1.156.5	frame_lock_5
Start-up protocol status 5	BASE-R PMD status 2 register	1.156.6	training_5
Training failure 5	BASE-R PMD status 2 register	1.156.7	training_failure_5
Receiver status 4	BASE-R PMD status 2 register	1.156.0	rx_trained_4
Frame lock 4	BASE-R PMD status 2 register	1.156.1	frame_lock_4
Start-up protocol status 4	BASE-R PMD status 2 register	1.156.2	training_4
Training failure 4	BASE-R PMD status 2 register	1.156.3	training_failure_4
Receiver status 3	BASE-R PMD status register	1.151.12	rx_trained_3
Frame lock 3	BASE-R PMD status register	1.151.13	frame_lock_3
Start-up protocol status 3	BASE-R PMD status register	1.151.14	training_3
Training failure 3	BASE-R PMD status register	1.151.15	training_failure_3
Receiver status 2	BASE-R PMD status register	1.151.8	rx_trained_2
Frame lock 2	BASE-R PMD status register	1.151.9	frame_lock_2
Start-up protocol status 2	BASE-R PMD status register	1.151.10	training_2
Training failure 2	BASE-R PMD status register	1.151.11	training_failure_2
Receiver status 1	BASE-R PMD status register	1.151.4	rx_trained_1
Frame lock 1	BASE-R PMD status register	1.151.5	frame_lock_1
Start-up protocol status 1	BASE-R PMD status register	1.151.6	training_1
Training failure 1	BASE-R PMD status register	1.151.7	training_failure_1
Receiver status 0	BASE-R PMD status register	1.151.0	rx_trained_0
Frame lock 0	BASE-R PMD status register	1.151.1	frame_lock_0
Start-up protocol status 0	BASE-R PMD status register	1.151.2	training_0
Training failure 0	BASE-R PMD status register	1.151.3	training_failure_0

85.7 PMD functional specifications

85.7.1 Link block diagram

A 40GBASE-CR4 or 100GBASE-CR10 link in one direction is illustrated in Figure 85–2. 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 85–5 are made at TP2 utilizing the test fixture specified in 85.8.3.5. Unless specified otherwise, all receiver measurements and tests defined in 85.8.4 are made at TP3 utilizing the test fixture specified in 85.8.3.5. A mated connector pair has been included in both the transmitter and receiver specifications defined in 85.8.3 and 85.8.4. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is specified in 85.8.3.4.

The 40GBASE-CR4 and 100GBASE-CR10 channels are 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 85–2. Annex 85A 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 85–2. The cable assembly test fixture of Figure 85–13 or its functional equivalent, is required for measuring the cable assembly specifications in 85.10 at TP1 and TP4. Two mated connector pairs and the cable assembly test fixture have been included in the cable assembly specifications defined in 85.10. Transmitter and receiver differential controlled impedance printed circuit board insertion losses defined between TP0–TP1 and TP4–TP5 respectively are provided informatively in Annex 85A.

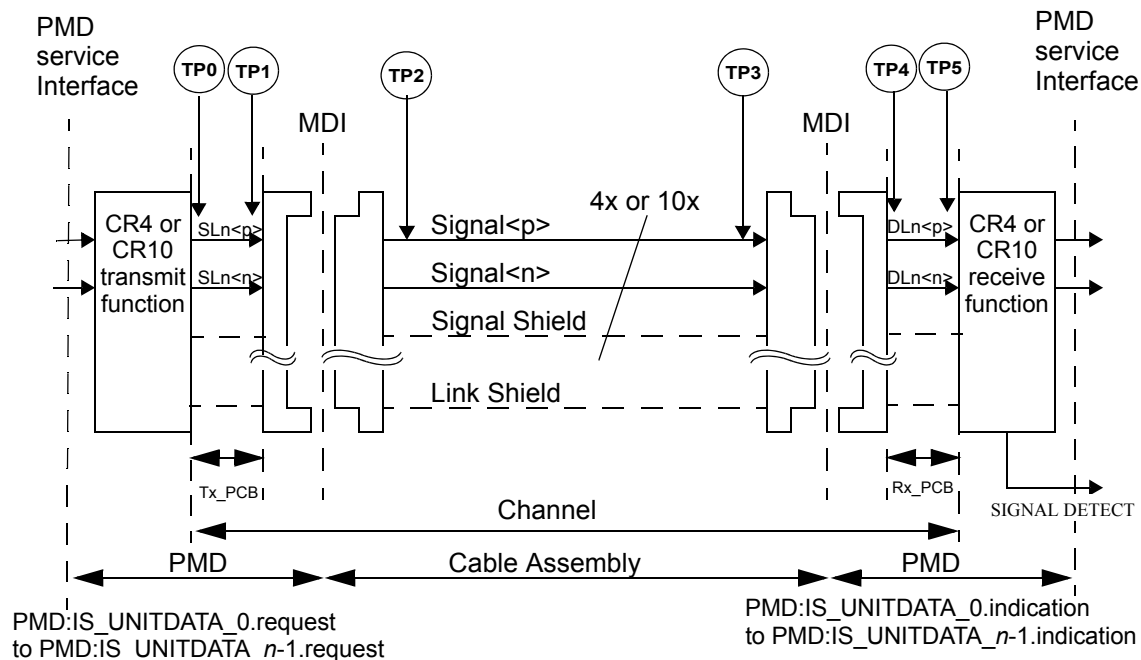


Figure 85–2—40GBASE-CR4 or 100GBASE-CR10 link (half link is illustrated)

Note that the source lanes (SL), signals $SL_n\langle p \rangle$, and $SL_n\langle n \rangle$ are the positive and negative sides of the transmitters differential signal pairs and the destination lanes (DL) signals, $DL_n\langle p \rangle$, and $DL_n\langle n \rangle$ are the positive and negative sides of the receivers differential signal pairs for lane n ($n = 0, 1, 2, 3$ or $n = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$).

Table 85–4 describes the defined test points illustrated in Figure 85–2.

Table 85–4—40GBASE-CR4 and 100GBASE-CR10 test points

Test points	Description
TP0 to TP5	The 40GBASE-CR4 and 100GBASE-CR10 channels 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 85–2. The cable assembly test fixture of Figure 85–13 or its functional equivalent, is required for measuring the cable assembly specifications in 85.10 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 85.8.3 and 85.8.4. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is specified in 85.8.3.4.
TP2	Unless specified otherwise, all transmitter measurements and tests defined in Table 85–5 are made at TP2 utilizing the test fixture specified in 85.8.3.5.
TP3	Unless specified otherwise, all receiver measurements and tests defined in 85.8.4 are made at TP3 utilizing the test fixture specified in 85.8.3.5.

85.7.2 PMD Transmit function

The 40GBASE-CR4 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 85.8.3. A positive output voltage of $SL_{<p>} - SL_{<n>}$ (differential voltage) shall correspond to $tx_bit = one$. The 100GBASE-CR10 PMD Transmit function shall convert the ten bit streams requested by the PMD service interface messages PMD:IS_UNITDATA_0.request to PMD:IS_UNITDATA_9.request. The ten electrical signal streams shall then be delivered to the MDI, all according to the transmit electrical specifications in 85.8.3. A positive output voltage of $SL_{<p>} - SL_{<n>}$ (differential voltage) shall correspond to $tx_bit = one$.

85.7.3 PMD Receive function

The 40GBASE-CR4 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 $DL_{<p>} - DL_{<n>}$ (differential voltage) shall correspond to $rx_bit = one$. The 100GBASE-CR10 PMD Receive function shall convert the ten electrical streams from the MDI into ten bit streams for delivery to the PMD service interface using the messages PMD:IS_UNITDATA_0.indication to PMD:IS_UNITDATA_9.indication. A positive input voltage of $DL_{<p>} - DL_{<n>}$ (differential voltage) shall correspond to $rx_bit = one$.

85.7.4 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 40GBASE-CR4 and 100GBASE-CR10 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.

85.7.5 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 0:3 for 40GBASE-CR4 and 0:9 for 100GBASE-CR10, shall be continuously updated as described in the following two paragraphs.

For 40GBASE-CR4 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.

For 100GBASE-CR10 PMD_signal_detect_0 (1.10.1), PMD_signal_detect_1 (1.10.2), PMD_signal_detect_2 (1.10.3), PMD_signal_detect_3 (1.10.4), PMD_signal_detect_4 (1.10.5), PMD_signal_detect_5 (1.10.6), PMD_signal_detect_6 (1.10.7), PMD_signal_detect_7 (1.10.8), PMD_signal_detect_8 (1.10.9), and PMD_signal_detect_9 (1.10.10) 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.

85.7.6 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 85–5.
- 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.

85.7.7 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_{*i*} function (where *i* represents the lane number in the range 0:3 or 0:9) 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 85–5.
- b) If a PMD_fault (85.7.9) 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 85.7.8, shall not be affected by PMD_transmit_disable_{*i*}.

85.7.8 Loopback mode

Local loopback mode shall be provided by the adjacent PMA (see 83.5.8) for the 40GBASE-CR4 and 100GBASE-CR10 PMDs 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.

85.7.9 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. The PMD fault function shall be mapped to bit 1.1.7 as listed in Table 85–3.

85.7.10 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.

85.7.11 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.

85.7.12 PMD control function

Each lane of the 40GBASE-CR4 or 100GBASE-CR10 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.

85.8 MDI Electrical specifications for 40GBASE-CR4 and 100GBASE-CR10

85.8.1 Signal levels

The 40GBASE-CR4 and 100GBASE-CR10 MDI is a low-swing AC coupled differential interface. AC coupling at the receiver, as defined in 85.8.4.5, allows for interoperability between components operating from different supply voltages. Low-swing differential signaling provides noise immunity and improved electromagnetic interference (EMI).

85.8.2 Signal paths

The 40GBASE-CR4 and 100GBASE-CR10 MDI signal paths are point-to-point connections. Each path corresponds to a 40GBASE-CR4 or 100GBASE-CR10 MDI lane and comprises two complementary signals, which form a balanced differential pair. For 40GBASE-CR4, there are four differential paths in each direction for a total of eight pairs, or sixteen connections. For 100GBASE-CR10, there are ten differential paths in each direction for a total of 20 pairs, or forty connections. The signal paths are intended to operate on twinaxial cable assemblies ranging from 0.5 m to 7 m in length, as described in 85.10.

85.8.3 Transmitter characteristics

Transmitter characteristics shall meet specifications summarized in Table 85–5 at TP2 unless otherwise noted. The transmitter specifications at TP0 are provided informatively in Annex 85A, Table 85A–1.

Table 85–5—Transmitter characteristics at TP2 summary

Parameter	Subclause reference	Value	Units
Signaling rate, per lane	85.8.3.9	10.3125 ± 100 ppm	GBd
Unit interval nominal	85.8.3.9	96.969697	ps
Differential peak-to-peak output voltage (max) with Tx disabled		30	mV
Common-mode voltage limits	72.7.1.4	0 to 1.9	V
Differential output return loss (min)	85.8.3.1	See Equation (85–1)	dB
Common-mode AC output voltage (max., RMS)		30	mV
Amplitude peak-to-peak (max)	72.7.1.4	1200 ^a	mV
Transmitter DC amplitude ^b	85.8.3.3	0.34 min, 0.6 max	V
Linear fit pulse (min) ^c	85.8.3.3	0.63 × Transmitter DC amplitude	V
Transmitted waveform max RMS normalized error (linear fit), “e”	85.8.3.3	0.037	
abs coefficient step size	85.8.3.3.2	0.0083 min, 0.05 max	
minimum precursor fullscale range	85.8.3.3.3	1.54	
minimum post cursor fullscale range	85.8.3.3.3	4	
Far-end transmit output noise (max) Low insertion loss channel High insertion loss channel	85.8.3.2	2 See Equation (85–2) 1 See Equation (85–3)	mV
Max output jitter (peak-to-peak) Random jitter ^d Duty Cycle Distortion ^e Total jitter excluding data dependent jitter ^f		0.15 0.035 0.25	UI UI UI

^aThe 40GBASE-CR4 Style-1 connector may support 40GBASE-CR4 or XLPPI interfaces. For implementations that support both interfaces, the transmitter should not exceed the XLPPI voltage maximum until a 40GBASE-CR4 cable assembly has been identified. The 100GBASE-CR10 connector may support 100GBASE-CR10 or CPPI interfaces. For implementations that support both interfaces, the transmitter should not exceed the CPPI voltage maximum until a 100GBASE-CR10 cable assembly has been identified.

^bThe transmitter DC amplitude is the sum of linear fit pulse response $p(k)$ from step 3) divided by M from step 3).

^cThe peak of the linear fit pulse response $p(k)$ from step 3).

^dRandom jitter is specified at a BER of 10^{-12}

^e See 72.7.1.9 for duty cycle distortion definition.

^fTotal jitter at a BER of 10^{-12} measured per 83A.5.1 excluding data dependent jitter (DDJ). DDJ is a jitter component where jitter that is not correlated to the data pattern has been removed. DDJ is measured with PRBS9 as specified in 83.5.10.

85.8.3.1 Transmitter differential output return loss

The differential output return loss, in dB, of the transmitter shall meet Equation (85–1). This output impedance requirement applies to all valid output levels. The reference impedance for differential return loss measurements shall be 100Ω

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.05 \leq f < 4.11 \\ 6.3 - 13 \log_{10}(f/5.5) & 4.11 \leq f \leq 10 \end{array} \right\} \quad (\text{dB}) \quad (85-1)$$

where

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

85.8.3.2 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 85.10.7. 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 $6 \text{ dB} \pm 1.5 \text{ dB}$ at 5.15625 GHz and cable assembly integrated crosstalk noise (ICN) meeting the requirements of 85.10.7 and a "high-loss" cable assembly with insertion loss on the reference pair of $15.7 \text{ dB} \pm 1.5 \text{ dB}$ at 5.15625 GHz and cable assembly ICN meeting the requirements of 85.10.7. 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 85.10.7 and the parameters in Table 85–11.
- 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 85.10.8.
- 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 (85–2).

$$RMSI_{dev} \leq \sqrt{\sigma_l^2 + 2^2} \quad (\text{mV}) \quad (85-2)$$

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 (85–3).

$$RMS h_{dev} \leq \sqrt{\sigma_h^2 + 1^2} \text{ (mV)} \quad (85-3)$$

85.8.3.3 Transmitter output waveform

The 40GBASE-CR4 and 100GBASE-CR10 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 85-3.

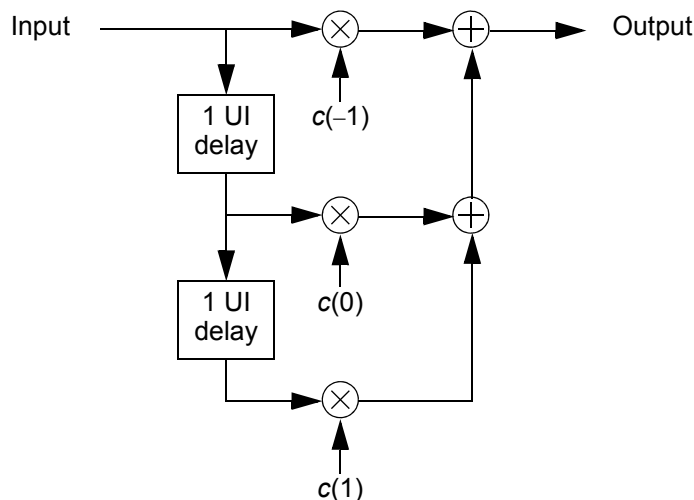


Figure 85-3— Transmit equalizer functional model

The state of the transmit equalizer and hence the transmitted output waveform may be manipulated via the PMD control function defined in 85.7.12 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 40GBASE-CR4 and 100GBASE-CR10 transmit equalizer are intended to be similar to the requirements for 10GBASE-KR specified in 72.7.1.10. 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 85.8.3.3.4.
- 3) Compute the linear fit to the captured waveform and the linear fit pulse response $p(k)$ per 85.8.3.3.5.
- 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 85.8.3.3.6.

The parameters of the pulse fit and the equalizing filter are given in Table 85–6. The DC amplitude, the sum of linear fit pulse response, $p(k)$, from step 3) divided by M from step 3), shall be greater than 0.34 V and less than or equal to 0.6 V. The peak of the linear fit pulse response from step 3) shall be greater than $0.63 \times$ DC amplitude. 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 85–6—Normalized transmit pulse template

Description	Symbol	Value	Units
Linear fit pulse length	N_p	8	UI
Linear fit pulse delay	D_p	2	UI
Equalizer length	N_w	7	UI
Equalizer delay	D_w	1	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 85.8.3.3.4.
- 9) Compute the linear fit to the captured waveform and the linear fit pulse response $p(k)$ per 85.8.3.3.5.
- 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 85.8.3.3.6 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.

85.8.3.3.1 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 $1.29 \pm 10\%$ and the ratio $(c(0)-c(1)+c(-1))/(c(0)+c(1)+c(-1))$ is $2.57 \pm 10\%$. These requirements apply upon the assertion a coefficient status report of “updated” for all coefficients.

85.8.3.3.2 Coefficient step size

The change in the normalized amplitude of coefficient $c(i)$ corresponding to a request to “increment” that coefficient shall be between 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.

85.8.3.3.3 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 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 a coefficient preset request and then manipulating the other coefficients as required by the test.

85.8.3.3.4 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 7. 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.

85.8.3.3.5 Linear fit to the waveform measurement at TP2

Given the captured waveform $y(k)$ and corresponding aligned symbols $x(n)$ derived from the procedure defined in 85.8.3.3.4, define the M -by- N waveform matrix Y as shown in Equation (85–4).

$$Y = \begin{bmatrix} y(1) & y(M+1) & \dots & y(M(N-1)+1) \\ y(2) & y(M+2) & \dots & y(M(N-1)+2) \\ \dots & \dots & \dots & \dots \\ y(M) & y(2M) & \dots & y(MN) \end{bmatrix} \quad (85-4)$$

Rotate the symbols vector x by the specified pulse delay D_p to yield x_r as shown in Equation (85-5).

$$x_r = \begin{bmatrix} x(D_p + 1) & x(D_p + 2) & \dots & x(N) & x(1) & \dots & x(D_p) \end{bmatrix} \quad (85-5)$$

Define the matrix X to be an N -by- N matrix derived from x_r as shown in Equation (85-6).

$$X = \begin{bmatrix} x_r(1) & x_r(2) & \dots & x_r(N) \\ x_r(N) & x_r(1) & \dots & x_r(N-1) \\ \dots & \dots & \dots & \dots \\ x_r(2) & x_r(3) & \dots & x_r(1) \end{bmatrix} \quad (85-6)$$

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 (85-7). The superscript “ T ” denotes the matrix transpose operator.

$$P = YX_1^T(X_1X_1^T)^{-1} \quad (85-7)$$

The error waveform, $e(k)$, is then read column-wise from the elements of E as shown in Equation (85-8).

$$E = PX_1 - Y = \begin{bmatrix} e(1) & e(M+1) & \dots & e(M(N-1)+1) \\ e(2) & e(M+2) & \dots & e(M(N-1)+2) \\ \dots & \dots & \dots & \dots \\ e(M) & e(2M) & \dots & e(MN) \end{bmatrix} \quad (85-8)$$

Define P_1 to be a matrix consisting of the first N_p columns of the matrix P as shown in Equation (85-9). The linear fit pulse response, $p(k)$, is then read column-wise from the elements of P_1 .

$$P_1 = \begin{bmatrix} p(1) & p(M+1) & \dots & p(M(N_p-1)+1) \\ p(2) & p(M+2) & \dots & p(M(N_p-1)+2) \\ \dots & \dots & \dots & \dots \\ p(M) & p(2M) & \dots & p(MN_p) \end{bmatrix} \quad (85-9)$$

85.8.3.3.6 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 (85-10).

$$p_r = \begin{bmatrix} p_i(D_w + 1) & p_i(D_w + 2) & \dots & p_i(N_p) & p_i(1) & \dots & p_i(D_w) \end{bmatrix} \quad (85-10)$$

Define the matrix P_2 to be an N_p -by- N_p matrix derived from p_r as shown in Equation (85-11).

$$P_2 = \begin{bmatrix} p_r(1) & p_r(N_p) & \dots & p_r(2) \\ p_r(2) & p_r(1) & \dots & p_r(3) \\ \dots & \dots & \dots & \dots \\ p_r(N_p) & p_r(N_p - 1) & \dots & p_r(1) \end{bmatrix} \quad (85-11)$$

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 (85-12).

$$w = (P_3^T P_3)^{-1} P_3^T x_p \quad (85-12)$$

Given the column vector of equalizer coefficients, w , the equalized pulse response q_i is determined by Equation (85-13).

$$q_i = P_3 w \quad (85-13)$$

85.8.3.4 Insertion loss TP0 to TP2 or TP3 to TP5

Transmitter measurements and tests defined in Table 85-5 are made at TP2 or TP3 using the test fixture of Figure 85-5, or its functional equivalent. The maximum insertion loss from TP0 to TP2 or TP3 to TP5 including the test fixture is determined using Equation (85-14).

$$Insertion_loss(f) \leq \left\{ \begin{array}{ll} 0.114 + 0.8914\sqrt{f} + 0.846f & 0.01 \leq f < 7 \\ -35.91 + 6.3291f & 7 \leq f < 8 \\ 14.72 & 8 \leq f \leq 10 \end{array} \right\} \text{ (dB)} \quad (85-14)$$

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 85-4.

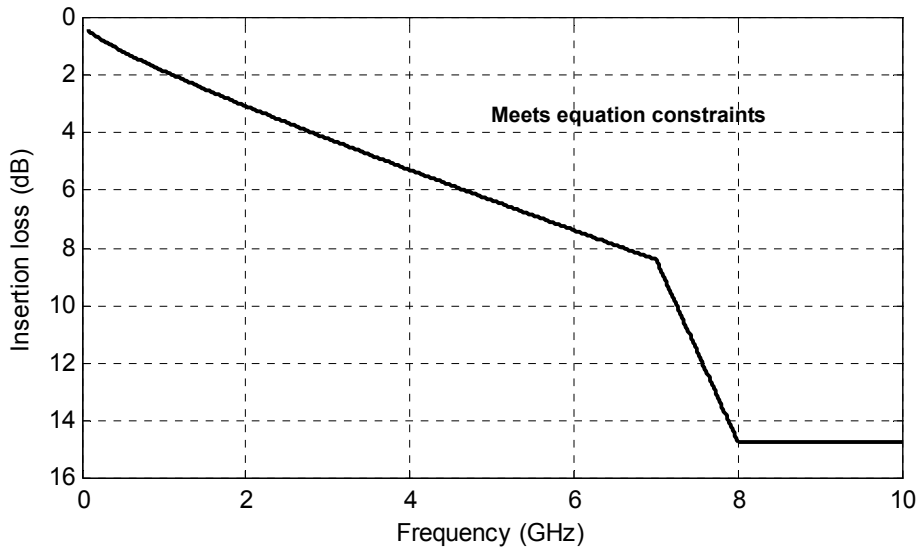


Figure 85-4—Maximum insertion loss TP0 to TP2 or TP3 to TP5

85.8.3.5 Test fixture

The test fixture of Figure 85-5, or its functional equivalent, is required for measuring the transmitter specifications in 85.8.3 at TP2 and the receiver return loss at TP3. TP2 and TP3 are illustrated in Figure 85-2. Figure 85-5 illustrates the test fixture attached to TP2 or TP3.

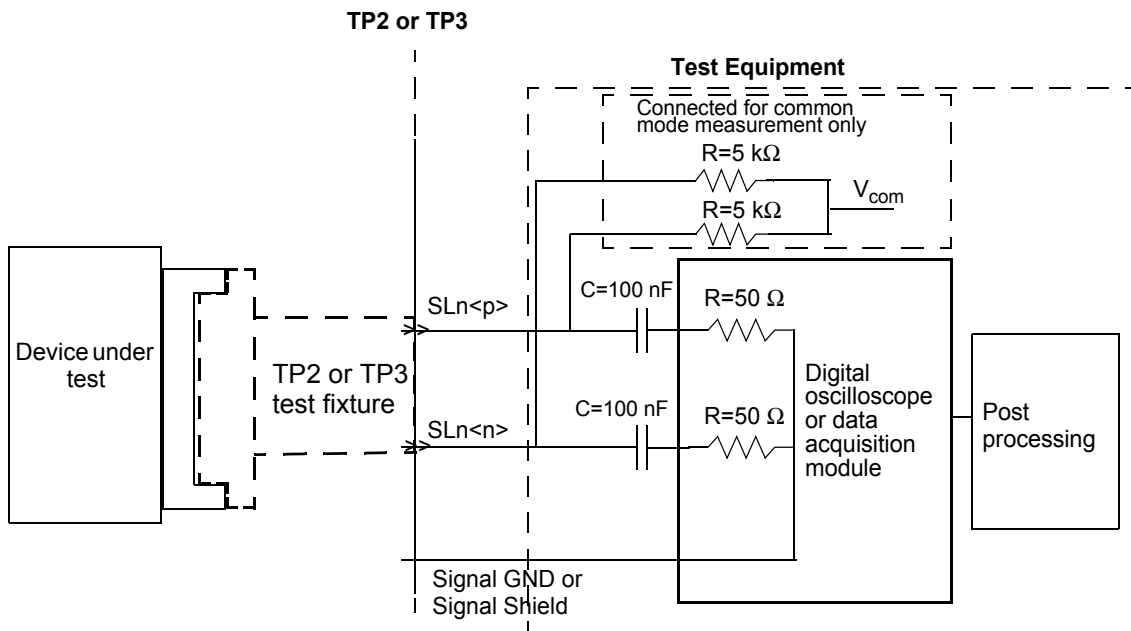


Figure 85-5—Transmitter test fixture

85.8.3.6 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 85.10.9.2.

85.8.3.7 Test fixture insertion loss

The reference test fixture printed circuit board insertion loss shall meet the values determined using Equation (85–15). 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.

$$IL_{tref}(f) = 0.01 + 0.3\sqrt{f} + 0.11f \text{ (dB)} \quad (85-15)$$

for $0.01 \text{ GHz} \leq f \leq 10 \text{ GHz}$

where

f is the frequency in GHz

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

85.8.3.8 Data dependent jitter (DDJ)

An oscilloscope, time interval analyzer, or other instrument with equivalent capability may be used to measure DDJ. A repeating PRBS9 pseudo-random test pattern, 511 bits long, is used. For DDJ measurements, the measurement bandwidth should be at least 12 GHz. If the measurement bandwidth affects the result, it can be corrected for by post-processing.

Establish a crossing level equal to the average value of the entire waveform being measured. Synchronize the instrument to the pattern repetition frequency and average the waveforms or the crossing times sufficiently to remove the effects of random jitter and noise in the system. The PRBS9 pattern has 128 positive-going transitions and 128 negative-going transitions. The mean time of each crossing is then compared to the expected time of the crossing, and a set of 256 timing variations is determined. Crossings earlier than expected give a negative variation. Crossings later than expected give a positive variation. DDJ is the range (maximum minus minimum) of the timing variations. Note that it may be convenient to align the expected time of one of the crossings with the measured mean crossing.

Figure 85–6 illustrates the method. The vertical axis is in arbitrary units, and the horizontal axis is plotted in UI. The waveform is AC coupled to an average value of 0, therefore 0 is the appropriate crossing level. The rectangular waveform shows the ideal crossing times, and the other is the waveform with jitter that is being measured. Only 16 UI are shown (out of 511). The waveforms have been arbitrarily aligned with $(\Delta t_2 = 0)$ at 5 UI.

Data dependent jitter is defined as in Equation (85–16).

$$DDJ = \max(\Delta t_1, \Delta t_2, \dots, \Delta t_{256}) - \min(\Delta t_1, \Delta t_2, \dots, \Delta t_{256}) \quad (85-16)$$

85.8.3.9 Signaling rate range

The 40GBASE-CR4 and 100GBASE-CR10 MDI signaling rate shall be 10.3125 GBd \pm 100 ppm per lane. The corresponding unit interval is nominally 96.969697 ps.

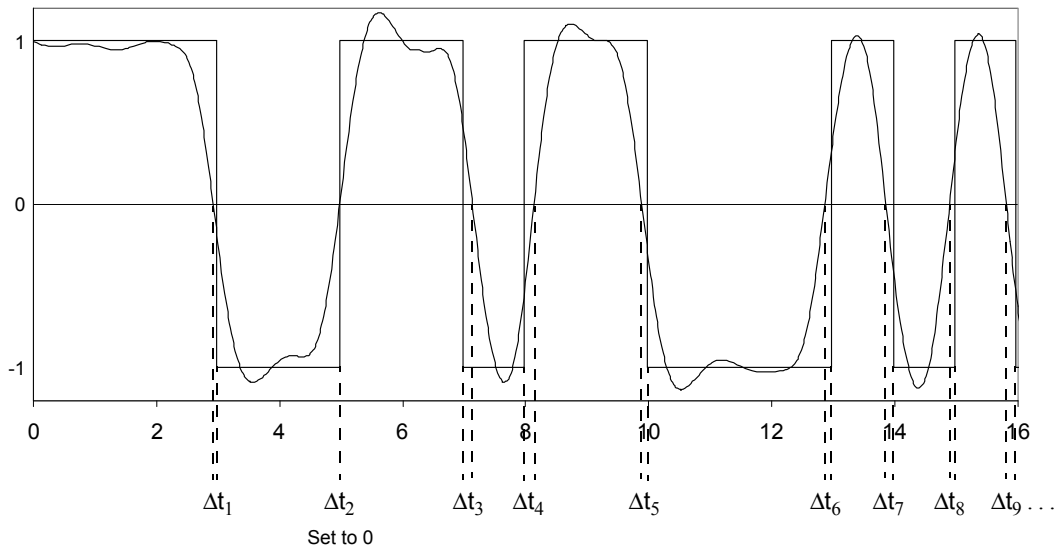


Figure 85-6—Data dependent jitter test method

85.8.4 Receiver characteristics at TP3 summary

The receiver characteristics shall meet the specifications summarized in Table 85-7 at TP3 unless otherwise noted.

Table 85-7—Receiver characteristics at TP3 summary

Parameter	Subclause reference	Value	Units
Bit error ratio	85.8.4.3	10^{-12} or better	
Signaling rate, per lane	85.8.4.4	10.3125 ± 100 ppm	GBd
Unit interval (UI) nominal	85.8.4.4	96.969697	ps
Differential peak-to-peak input amplitude tolerance (max)	72.7.2.4	1200	mV
Differential input return loss (min) ^a	85.8.4.1	Equation (85-17)	dB
Differential to common-mode input return loss		10 min from 10 MHz to 10 GHz	dB

^aRelative to 100 Ω differential.

85.8.4.1 Receiver differential input return loss

The differential input return loss, in dB, of the receiver shall meet Equation (85-17). This return loss requirement applies at all valid input levels. The reference impedance for differential return loss measurements is 100 Ω.

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.01 \leq f < 4.11 \\ 6.3 - 13 \log_{10}(f/5.5) & 4.11 \leq f \leq 10 \end{array} \right\} \quad (\text{dB}) \quad (85-17)$$

where

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

85.8.4.2 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 85–8 when measured according to the requirements of 85.8.4.2.1 to 85.8.4.2.5.

Table 85–8—40GBASE-CR4 and 100GBASE-CR10 interference tolerance parameters

Parameter	Test 1 values	Test 2 values	Units
Maximum BER	10^{-12}	10^{-12}	
Fitted insertion loss coefficients	$a_1 = 2.15$ $a_2 = 0.78$ $a_4 = 0.03$	$a_1 = 6.04$ $a_2 = 0.94$ $a_4 = 0.08$	dB/ $\sqrt{\text{GHz}}$ dB/GHz dB/GHz ²
Applied SJ ^a (peak-to-peak)	0.115	0.115	UI
Applied RJ ^b (peak-to-peak)	0.13	0.13	UI
Applied DCD (peak-to-peak)	0.035	0.035	UI
Calibrated far-end crosstalk (RMS)	6.3	2.2	mV
Calibrated ICN (RMS) – σ_{nx}	3.7	3.7	mV

^aApplied SJ frequency >15 MHz, specified at TP0.

^bApplied random jitter at TP0 is specified at 10^{-12} .

85.8.4.2.1 Test setup

The interference tolerance test is performed with the setup shown in Figure 85–7. The requirements of this subclause are verified at the pattern generator connection (PGC) or test references in Figure 85–7 and Figure 85–8. The lanes under test (LUT) are illustrated in Figure 85–7 and Figure 85–8. The cable assembly receive lanes are terminated in 100 Ω differentially.

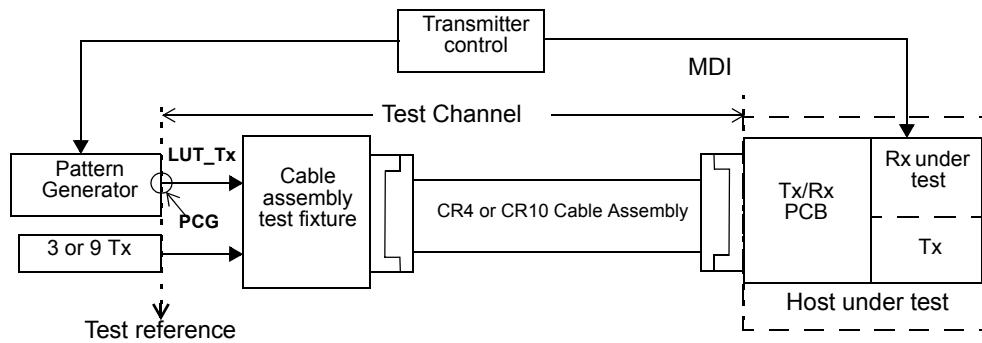


Figure 85-7—Interference tolerance test setup

85.8.4.2.2 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

85.8.4.2.3 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 85-7 using the cable assembly test fixtures specified in 85.10.8.

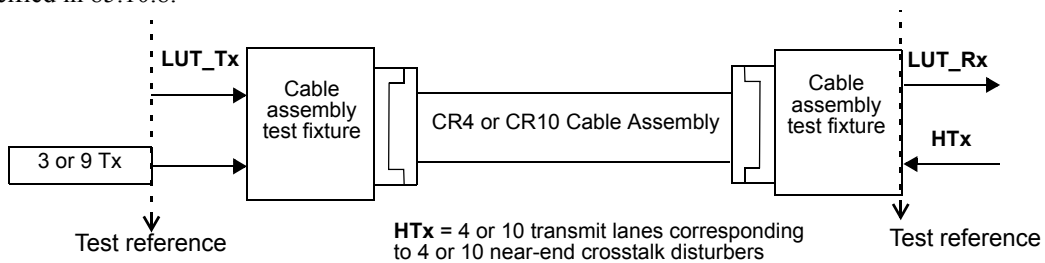


Figure 85-8—Test channel calibration

The fitted insertion loss coefficients of the lane under test (LUT), derived using the fitting procedure in 85.10.2, shall meet the test values in Table 85-8. 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 85-8.

The MDNEXT is measured from points HTx to adjacent point LUT_Rx in Figure 85-8. HTx is the set of 4 or 10 transmit lanes of the device under test corresponding to the 4 or 10 near-end crosstalk disturbers with the parameters given in Table 85-11. The RMS value of the integrated MDNEXT crosstalk noise, determined using Equation (85-30) through Equation (85-34), shall meet the test values in Table 85-8.

The far-end crosstalk disturbers consist of 40GBASE-CR4 or 100GBASE-CR10 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 40GBASE-CR4 test channels, the crosstalk that is coupled into a receive lane is from three transmitters. For 100GBASE-CR10 test channels, the crosstalk that is coupled into a receiver lane is from nine 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 85–8 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.

85.8.4.2.4 Pattern generator

The pattern generator transmits data to the device under test. Its output amplitude shall be no more than 800 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 47 ps. If the rise and fall times of the pattern generator, T_r , are less than 47 ps the value of da_4 in Table 85–8 is increased by da_4 from Equation (85–18).

$$da_4 = 60.51 \times 10^{-6} (47^2 - T_r^2) \quad (85-18)$$

where T_r is the rise time in ps.

The pattern generator shall meet the jitter specification in Table 85–8. The output waveform of the pattern generator shall comply to 72.7.1.11.

85.8.4.2.5 Test procedure

For 40GBASE-CR4 or 100GBASE-CR10 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 85–8. 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).

85.8.4.3 Bit error ratio

The receiver shall operate with a BER 10^{-12} or better when receiving a compliant transmit signal, as defined in 85.8.3, through a compliant cable assembly as defined in 85.10.

85.8.4.4 Signaling rate range

A 40GBASE-CR4 and 100GBASE-CR10 receiver shall comply with the requirements of 85.8.4.3 for any signaling rate in the range 10.3125 GBd \pm 100 ppm. The corresponding unit interval is nominally 96.969697 ps.

85.8.4.5 AC coupling

The 40GBASE-CR4 and 100GBASE-CR10 receivers are AC coupled. AC coupling shall be part of the receive function for Style-2 40GBASE-CR4 connectors. For Style-1 40GBASE-CR4 and 100GBASE-CR10 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 50 kHz.

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

85.9 Channel characteristics

The 40GBASE-CR4 and 100GBASE-CR10 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 85–2. 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 85A.4 through 85A.7.

85.10 Cable assembly characteristics

The 40GBASE-CR4 and 100GBASE-CR10 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 7 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 85.10.8 and illustrated in Figure 85–13. These cable assembly specifications are based upon twinaxial cable characteristics, but other cable types are acceptable if the specifications of 85.10 are met.

Table 85–9 provides a summary of the cable assembly differential characteristics at 5.15625 GHz and references to the subclauses addressing each parameter.

Table 85–9—Cable assembly differential characteristics summary

Description	Reference	Value	Unit
Maximum insertion loss at 5.15625 GHz	85.10.2	17.04	dB
Minimum insertion loss at 5.15625 GHz		3	dB
Insertion loss deviation at 5.15625 GHz	85.10.3	max = 1.73 min = –1.73	dB
Minimum return loss at 5.15625 GHz	85.10.4	6.66	dB
MDNEXT loss	85.10.5	Equation (85–26)	dB
MDFEXT loss	85.10.6	Equation (85–27)	dB
Maximum integrated crosstalk noise	85.10.7	Equation (85–33)	mV

85.10.1 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 .

85.10.2 Cable assembly insertion loss

The fitted cable assembly insertion loss $IL_{fitted}(f)$ as a function of frequency f is defined in Equation (85–19).

$$IL_{\text{fitted}}(f) = a_1\sqrt{f} + a_2f + a_4f^2 \quad (\text{dB}) \quad (85-19)$$

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 7500 MHz with a maximum frequency spacing of 10 MHz, the coefficients of the fitted insertion loss are determined using Equation (85-20) and Equation (85-21).

Define the frequency matrix F as shown in Equation (85-20).

$$F = \begin{bmatrix} \sqrt{f_1} & f_1 & f_1^2 \\ \sqrt{f_2} & f_2 & f_2^2 \\ \dots & \dots & \dots \\ \sqrt{f_N} & f_N & f_N^2 \end{bmatrix} \quad (85-20)$$

The polynomial coefficients a_1 , a_2 , and a_4 are determined using Equation (85-21). In Equation (85-21), 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 .

$$\begin{bmatrix} a_1 \\ a_2 \\ a_4 \end{bmatrix} = (F^T F)^{-1} F^T IL \quad (85-21)$$

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 40GBASE-CR4 and 100GBASE-CR10 in Equation (85-19) and the maximum insertion loss at 5.15625 GHz shall meet the specifications summarized in Table 85-10 unless otherwise noted. The fitted insertion loss corresponding to one example of the maximum insertion loss at 5.15625 GHz and the maximum allowed values of a_1 , a_2 , and a_4 is illustrated in Figure 85-9.

Table 85-10—Maximum cable assembly insertion loss characteristics

Description	Value	Unit
Maximum insertion loss at 5.15625 GHz	17.04 ^a	dB
Maximum fitted insertion loss coefficient a_1	6	dB/ $\sqrt{\text{GHz}}$
Maximum fitted insertion loss coefficient a_2	1	dB/GHz
Maximum fitted insertion loss coefficient a_4	0.08	dB/GHz ²

^aThe limit on the maximum insertion loss at 5.15625 GHz precludes the coefficients a_1 , a_2 , and a_4 from simultaneous maximum values.

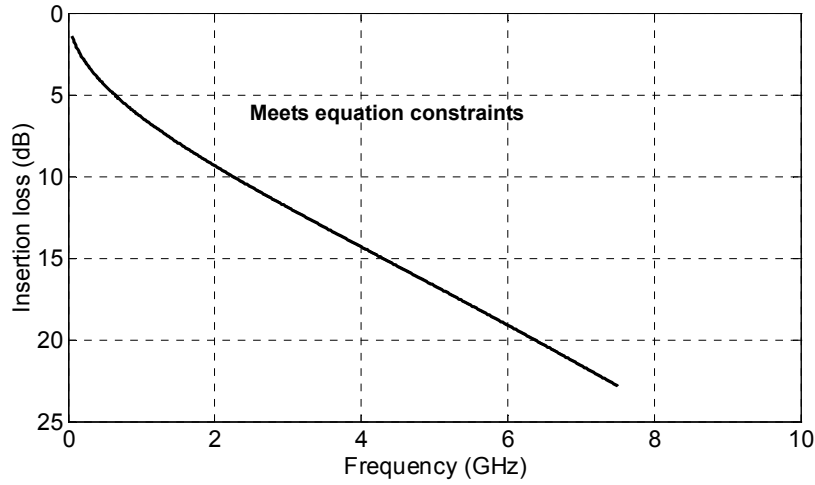


Figure 85-9—Example maximum cable assembly insertion loss

85.10.3 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 (85-22).

$$ILD(f) = IL(f) - IL_{fitted}(f) \tag{85-22}$$

where

f is the frequency in MHz

$ILD(f)$ is the cable assembly insertion loss deviation at frequency f

The ILD shall be within the region defined by Equation (85-23) and Equation (85-24). This includes the insertion loss of the differential cabling pairs and the cable assembly connectors.

$$ILD(f) \geq ILD_{min}(f) = -0.7 - 0.2 \times 10^{-3} f \text{ (dB)} \tag{85-23}$$

$$ILD(f) \leq ILD_{max}(f) = 0.7 + 0.2 \times 10^{-3} f \text{ (dB)} \tag{85-24}$$

for $50 \text{ MHz} \leq f \leq 7500 \text{ MHz}$

where

f is the frequency in MHz

The insertion loss deviation limits are illustrated in Figure 85–10.

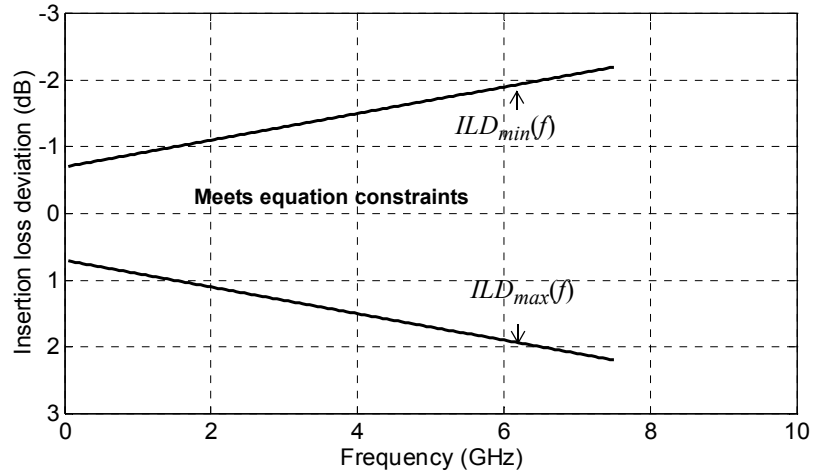


Figure 85–10—Maximum cable assembly insertion loss deviation

85.10.4 Cable assembly return loss

The return loss of each pair of the 40GBASE-CR4 and 100GBASE-CR10 cable assembly shall meet the values determined using Equation (85–25).

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.05 \leq f < 4.1 \\ 6.3 - 13 \log_{10}(f/5.5) & 4.1 \leq f \leq 10 \end{array} \right\} \text{ (dB)} \quad (85-25)$$

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 85–11.

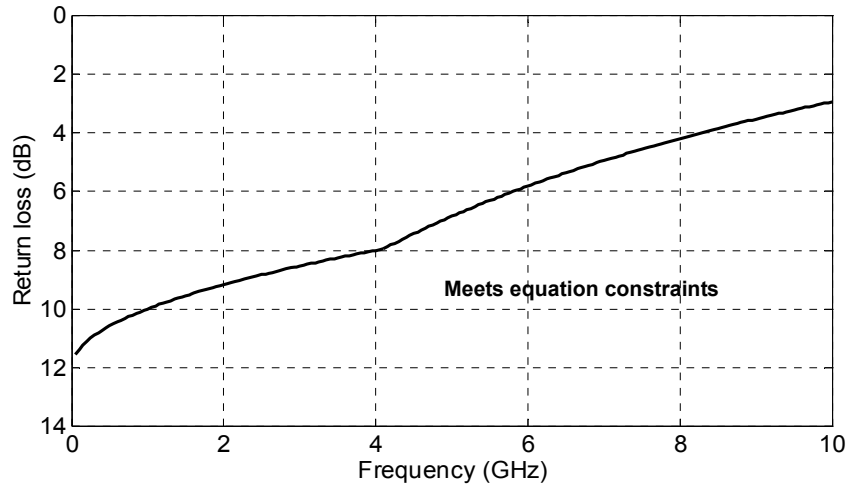


Figure 85-11—Minimum cable assembly return loss

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

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

MDNEXT loss is determined from the four or ten individual pair-to-pair differential NEXT loss values using Equation (85-26).

$$MDNEXT_loss(f) = -10\log_{10}\left(\sum_{i=0}^{i=3\text{ or }9} 10^{-NL_i(f)/10}\right) \text{ (dB)} \quad (85-26)$$

for $50 \text{ MHz} \leq f \leq 10000 \text{ MHz}$

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 MHz

i is the 0 to 3 (pair-to-pair combination) or 0 to 9 (pair-to-pair combination)

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

Since four lanes or ten 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 or nine individual pair-to-pair differential FEXT loss values using Equation (85-27).

$$MDFEXT_loss(f) = -10\log_{10}\left(\sum_{i=0}^{i=2\text{ or }8} 10^{-NL_i(f)/10}\right) \text{ (dB)} \quad (85-27)$$

for $50 \text{ MHz} \leq f \leq 10000 \text{ MHz}$

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 MHz
- i is the 0 to 2 (pair-to-pair combination) or 0 to 8 (pair-to-pair combination)

85.10.7 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 10000 MHz with a maximum frequency spacing of 10 MHz, the RMS value of the integrated crosstalk noise shall be determined using Equation (85–28) through Equation (85–32). 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 (85–28) and Equation (85–29) as weighting functions to the multiple disturber crosstalk in Equation (85–30) and Equation (85–31). The sinc function is defined by $\text{sinc}(x) = \sin(\pi x)/(\pi x)$.

Define the weight at each frequency f_n using Equation (85–28) and Equation (85–29).

$$W_{nt}(f_n) = (A_{nt}^2/f_b) \text{sinc}(f_n/f_b)^2 \left[\frac{1}{1 + (f_n/f_{nt})^4} \right] \left[\frac{1}{1 + (f_n/f_r)^8} \right] \quad (85-28)$$

$$W_{ft}(f_n) = (A_{ft}^2/f_b) \text{sinc}(f_n/f_b)^2 \left[\frac{1}{1 + (f_n/f_{ft})^4} \right] \left[\frac{1}{1 + (f_n/f_r)^8} \right] \quad (85-29)$$

where the equation parameters are given in Table 85–11.

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 7.5 GHz.

The near-end integrated crosstalk noise σ_{nx} is calculated using Equation (85–30).

$$\sigma_{nx} = \left[2\Delta f \sum_n W_{nt}(f_n) 10^{-MDNEXT_{loss}(f_n)/10} \right]^{1/2} \quad (85-30)$$

The far-end integrated crosstalk noise σ_{fx} is calculated using Equation (85–31).

$$\sigma_{fx} = \left[2\Delta f \sum_n W_{ft}(f_n) 10^{-MDFEXT_{loss}(f_n)/10} \right]^{1/2} \quad (85-31)$$

where Δf is the uniform frequency step of f_n .

The total integrated crosstalk noise σ_x is calculated using Equation (85–32).

$$\sigma_x = \sqrt{\sigma_{nx}^2 + \sigma_{fx}^2} \tag{85-32}$$

The total integrated crosstalk noise for the cable assembly shall be computed using the parameters shown in Table 85-11.

Table 85-11—Cable assembly integrated crosstalk parameters

Description	Symbol	Value	Units
Symbol rate	f_b	10.3125	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}	24	ps
Far-end disturber 20% to 80% rise and fall times	T_{ft}	24	ps

The total integrated crosstalk RMS noise voltage shall meet the values determined by Equation (85-33) illustrated in Figure 85-12.

$$\sigma_{x, ca} \leq \left\{ \begin{array}{ll} 10 & 3 \leq IL \leq 5.3 \\ 12.4 - 0.45IL & 5.3 < IL \leq 17.04 \end{array} \right\} \text{ (mV)} \tag{85-33}$$

where IL is the value of the cable assembly insertion loss in dB at 5.15625 GHz.

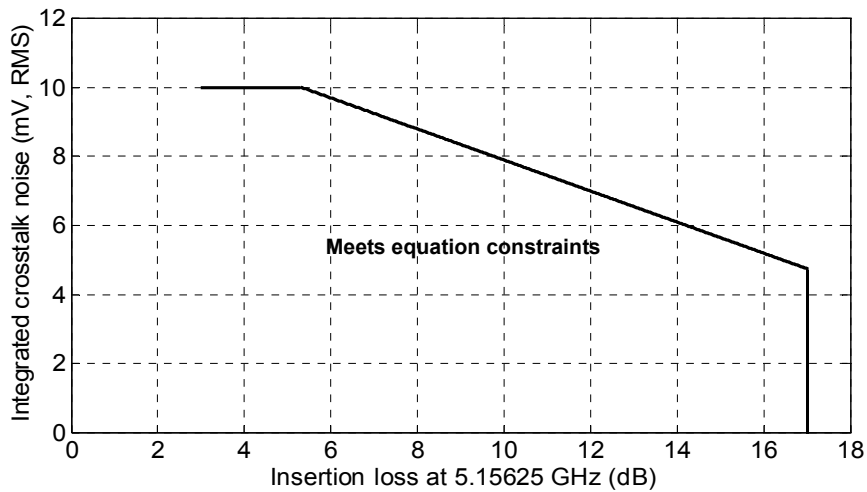


Figure 85-12—Integrated crosstalk noise limits

85.10.8 Cable assembly test fixture

The test fixture of Figure 85–13 or its functional equivalent, is required for measuring the cable assembly specifications in 85.10 at TP1 and TP4. TP1 and TP4 are illustrated in Figure 85–2 and Figure 85–13. The test fixture return loss is equivalent to the test fixture return loss specified in 85.8.3.6. The reference test fixture printed circuit board insertion loss is given in Equation (85–34). 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.

$$IL_{\text{catf}}(f) = 0.0006 + 0.16\sqrt{f} + 0.0587f \text{ (dB)} \quad (85-34)$$

for $0.01 \text{ GHz} \leq f \leq 10 \text{ GHz}$

where

f is the frequency in GHz

$IL_{\text{catf}}(f)$ is the reference test fixture printed circuit board insertion loss at frequency f

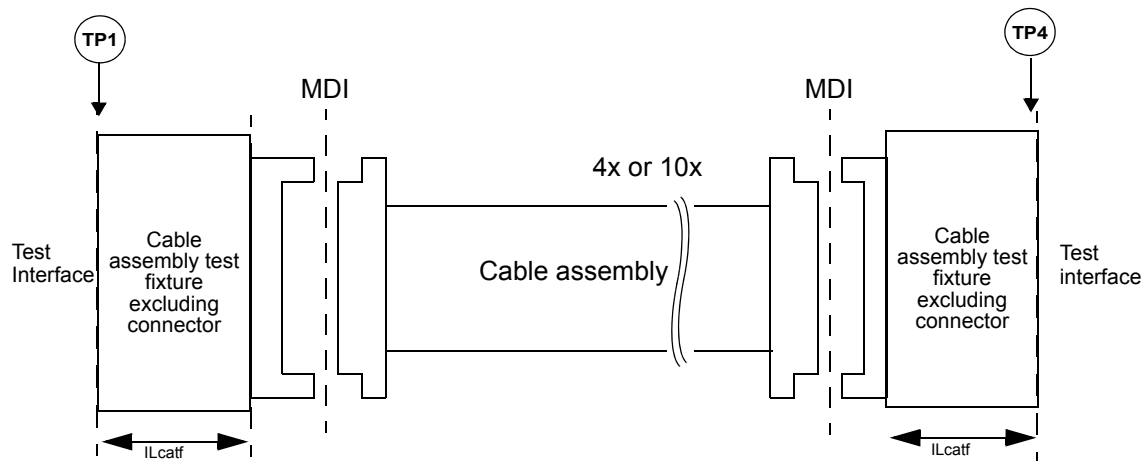


Figure 85–13—Cable assembly test fixtures

85.10.9 Mated test fixtures

The test fixtures of Figure 85–5 and Figure 85–13 are specified in a mated state illustrated in Figure 85–14.

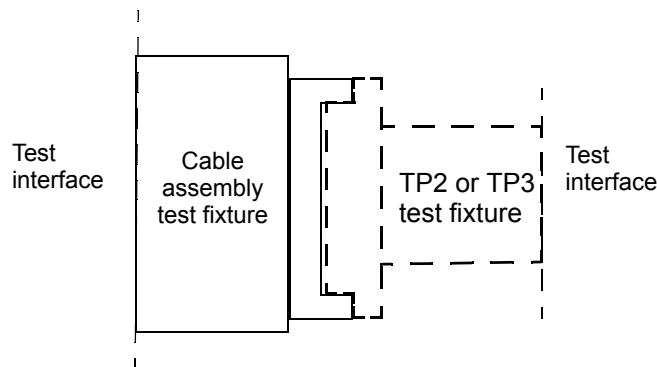


Figure 85–14—Mated test fixtures

85.10.9.1 Mated test fixtures insertion loss

The insertion loss of the mated test fixtures shall meet the values determined using Equation (85–35) and Equation (85–36).

$$IL(f) \geq IL_{MTFmin}(f) = -0.11 + 0.46\sqrt{f} + 0.16f \quad 0.01 \leq f \leq 10 \quad (\text{dB}) \tag{85-35}$$

$$IL(f) \leq IL_{MTEmax}(f) = \left\{ \begin{array}{ll} 0.029 + 0.861\sqrt{f} + 0.158f & 0.01 \leq f < 5.5 \\ 0.2 + 0.65f & 5.5 \leq f \leq 10 \end{array} \right\} (\text{dB}) \tag{85-36}$$

where

f is the frequency in GHz
 $IL(f)$ is the mated test fixture insertion loss at frequency f

The mated test fixtures insertion loss limits are illustrated in Figure 85–15.

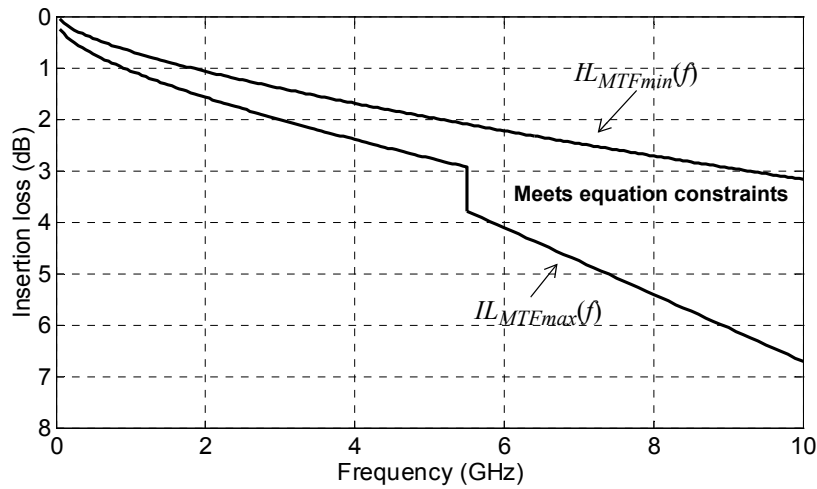


Figure 85–15—Mated test fixtures Insertion loss

85.10.9.2 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 (85–37).

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 20 - 2f & 0.01 \leq f < 2.5 \\ 15 & 2.5 \leq f < 5 \\ 13.8 - 28.85\log_{10}(f/5.5) & 5 \leq f \leq 10 \end{array} \right\} (\text{dB}) \tag{85-37}$$

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 85–16.

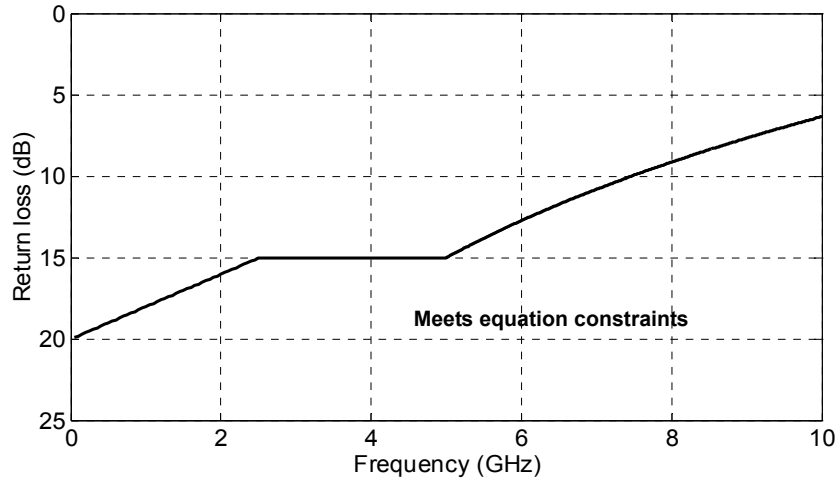


Figure 85–16—Mated test fixtures return loss

85.10.9.3 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 (85–38).

$$Conversion_loss(f) \geq \begin{cases} 30 - 2.91f & 0.01 \leq f < 5.5 \\ 14 & 5.5 \leq f \leq 10 \end{cases} \quad (\text{dB}) \quad (85-38)$$

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 85–17.

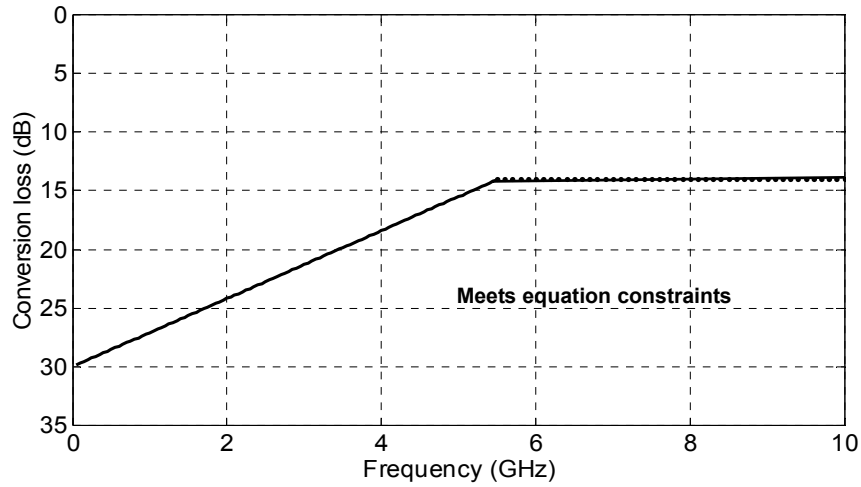


Figure 85-17—Common-mode conversion loss

85.10.9.4 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 (85-28) through Equation (85-32) 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 (85-28) through Equation (85-32) 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 85-12.

Table 85-12—Mated test fixtures integrated crosstalk noise

Parameter	40GBASE-CR4	100GBASE-CR10	Units
Near-end integrated crosstalk noise voltage (max, RMS)	0.7	3	mV
Far-end integrated crosstalk noise voltage (max, RMS)	2.5	4	mV
MDNEXT integrated crosstalk noise voltage (max, RMS)	1	3	mV
MDFEXT integrated crosstalk noise voltage (max, RMS)	3.5	5	mV

85.10.10 Shielding

The cable assembly shall provide Class 2 or better shielding in accordance with IEC 61196-1.

85.10.11 Crossover function

The cable assembly shall be wired in a crossover fashion as illustrated in Figure 85-18, with each of the four or ten pairs being attached to the transmitter contacts at one end and the receiver contacts at the other end.

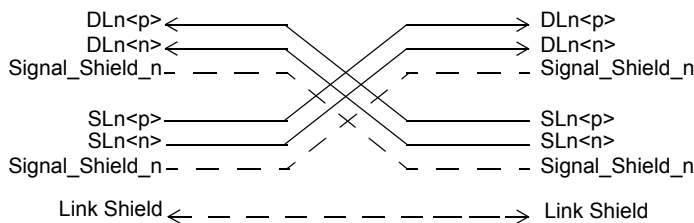


Figure 85-18—Cable assembly wiring

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 or n = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Signal_Shield_n is the signal shield of the differential signal pair for Lane n.

85.11 MDI specification

This subclause defines the Media Dependent Interface (MDI). The 40GBASE-CR4 and 100GBASE-CR10 PMD, as per 85.7, is coupled to the cable assembly, as per 85.10, by the MDI.

85.11.1 40GBASE-CR4 MDI connectors

Connectors meeting the requirements of 85.11.1.1 (Style-1) or 85.11.1.2 (Style-2) shall be used as the mechanical interface between the PMD of 85.7 and the cable assembly of 85.10. The plug connector shall be used on the cable assembly and the receptacle on the PHY. Style-1 or Style-2 connectors may be used as the MDI interface.

85.11.1.1 Style-1 40GBASE-CR4 MDI connectors

The connector for each end of the cable assembly shall be the quad small form factor pluggable (QSFP+) with the mechanical mating interface defined by SFF-8436 and illustrated in Figure 85-19. The MDI connector shall be the quad small form factor pluggable (QSFP+) receptacle with the mechanical mating interface defined by SFF-8436 and illustrated in Figure 85-20. These connectors have contact assignments matching that in Table 85-13, and electrical performance consistent with the signal quality and electrical requirements of 85.8 and 85.9.

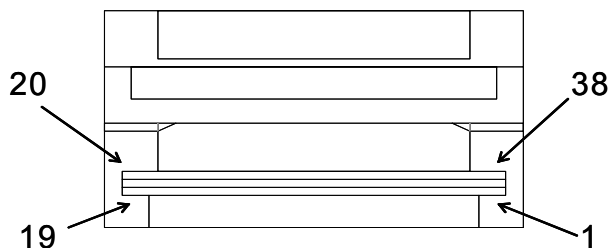


Figure 85-19—Example Style-1 cable assembly plug

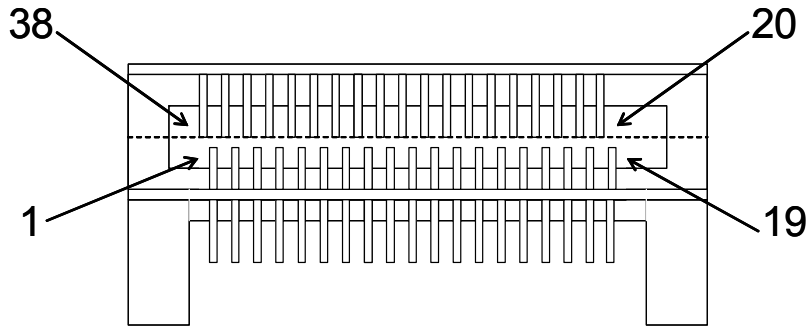


Figure 85–20—Example Style-1 MDI board receptacle

The Style-1 MDI connector of the 40GBASE-CR4 PMD comprises 38 signal connections. The Style-1 40GBASE-CR4 MDI connector contact assignments shall be as defined in Table 85–13.

Table 85–13—Style-1 40GBASE-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

NOTE—Although the 40GBASE-CR4 Style-1 MDI supports 38 connections only the transmitter and receiver contact assignments are specified.

85.11.1.1.1 Style-1 AC coupling

For Style-1 40GBASE-CR4 plug connectors the receive lanes are AC coupled; the coupling capacitors are contained within the plug connectors as specified in 85.8.4.5.

85.11.1.2 Style-2 40GBASE-CR4 MDI connectors

The connector for each end of the cable assembly shall be the latch-type plug with the mechanical mating interface defined by IEC 61076-3-113 and illustrated in Figure 85–21. The MDI connector shall be the latch-type receptacle with the mechanical mating interface defined by IEC 61076-3-113 and illustrated in Figure 85–22. These connectors have a pinout matching that in Table 85–14, and electrical performance consistent with the signal quality and electrical requirements of 85.8 and 85.9. Note that support of compatibility with 10GBASE-CX4 is at the Style-2 40GBASE-CR4 MDI.

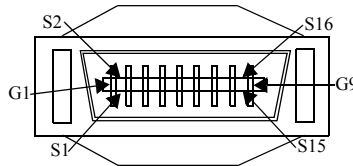


Figure 85–21—Example Style-2 cable assembly plug

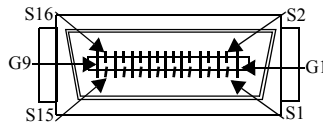


Figure 85–22—Example Style-2 MDI board receptacle

85.11.1.2.1 Style-2 40GBASE-CR4 Connector pin assignments

The MDI connector of the PMD comprises sixteen signal connections, eight signal shield connections, and one link shield connection. The 40GBASE-CR4 MDI connector pin assignments shall be as defined in Table 85–14.

Table 85–14—Style-2 40GBASE-CR4 lane to MDI connector pin mapping

Rx lane	MDI connector pin	Tx lane	MDI connector pin
DL0<p>	S1	SL0<p>	S16
DL0<n>	S2	SL0<n>	S15
DL1<p>	S3	SL1<p>	S14
DL1<n>	S4	SL1<n>	S13
DL2<p>	S5	SL2<p>	S12
DL2<n>	S6	SL2<n>	S11
DL3<p>	S7	SL3<p>	S10
DL3<n>	S8	SL3<n>	S9
Signal Shield	G1	Signal Shield	G5
Signal Shield	G2	Signal Shield	G6
Signal Shield	G3	Signal Shield	G7
Signal Shield	G4	Signal Shield	G8
—	—	Link Shield	G9

85.11.2 100GBASE-CR10 MDI connectors

The connector for each end of the cable assembly shall be the plug with the mechanical mating interface defined in SFF-8642 and illustrated in Figure 85–23. The MDI connector shall be the receptacle with the mechanical mating interface defined by SFF-8642 and illustrated in Figure 85–24. These connectors have contact assignments matching that in Table 85–15, and electrical performance consistent with the signal quality and electrical requirements of 85.8 and 85.9.

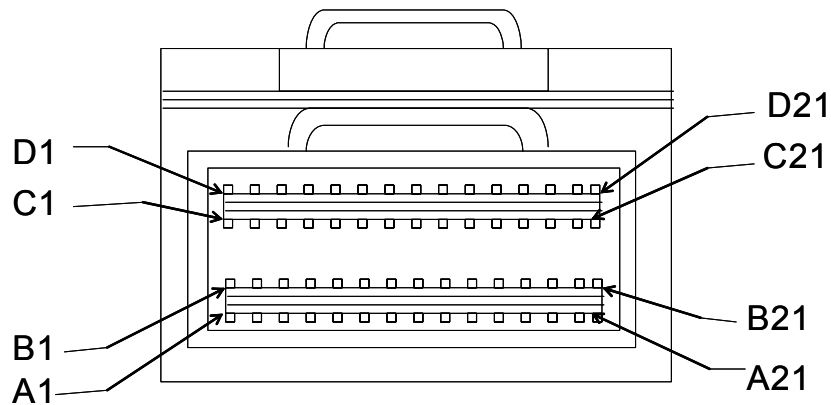


Figure 85–23—Example cable assembly plug

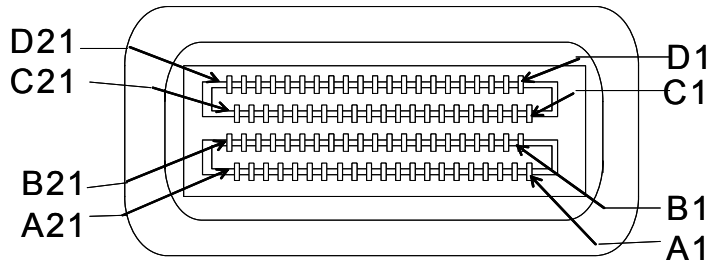


Figure 85–24—Example MDI board receptacle

The MDI connector of the PMD comprises 84 connections. The 100GBASE-CR10 PMD MDI connector contact assignments shall be as defined in Table 85–15.

Table 85–15—100GBASE-CR10 lane to MDI connector contact mapping

Tx lane	MDI connector contact	Tx lane	MDI connector contact	Rx lane	MDI connector contact	Rx lane	MDI connector contact
signal gnd	A1	signal gnd	B1	signal gnd	C1	signal gnd	D1
SL0<p>	A2	—	B2	DL0<p>	C2	—	D2
SL0<n>	A3	—	B3	DL0<n>	C3	—	D3
signal gnd	A4	signal gnd	B4	signal gnd	C4	signal gnd	D4
SL2<p>	A5	SL1<p>	B5	DL2<p>	C5	DL1<p>	D5
SL2<n>	A6	SL1<n>	B6	DL2<n>	C6	DL1<n>	D6
signal gnd	A7	signal gnd	B7	signal gnd	C7	signal gnd	D7
SL4<p>	A8	SL3<p>	B8	DL4<p>	C8	DL3<p>	D8
SL4<n>	A9	SL3<n>	B9	DL4<n>	C9	DL3<n>	D9
signal gnd	A10	signal gnd	B10	signal gnd	C10	signal gnd	D10
SL6<p>	A11	SL5<p>	B11	DL6<p>	C11	DL5<p>	D11
SL6<n>	A12	SL5<n>	B12	DL6<n>	C12	DL5<n>	D12
signal gnd	A13	signal gnd	B13	signal gnd	C13	signal gnd	D13
SL8<p>	A14	SL7<p>	B14	DL8<p>	C14	DL7<p>	D14
SL8<n>	A15	SL7<n>	B15	DL8<n>	C15	DL7<n>	D15
signal gnd	A16	signal gnd	B16	signal gnd	C16	signal gnd	D16
—	A17	SL9<p>	B17	—	C17	DL9<p>	D17
—	A18	SL9<n>	B18	—	C18	DL9<n>	D18
signal gnd	A19	signal gnd	B19	signal gnd	C19	signal gnd	D19

NOTE—Although the 100GBASE-CR10 MDI supports 84 connections only the transmitter and receiver contact assignments are specified.

85.11.2.1 100GBASE-CR10 MDI AC coupling

For 100GBASE-CR10 plug connectors, the receive lanes are AC coupled; the coupling capacitors are contained within the plug connectors as specified in 85.8.4.5.

85.11.3 Electronic keying

Electronic keying can be used to enable the detection of Style-1 40GBASE-CR4 MDI connectors or 100GBASE-CR10 MDI cable assembly plugs versus fiber modules or no modules present. Specifications of electronic keying are beyond the scope of this standard.

85.12 Environmental specifications

All equipment subject to this clause shall conform to the applicable requirements of 14.7.

85.13 Protocol implementation conformance statement (PICS) proforma for Clause 85, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10¹⁰

85.13.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 85 Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10, 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.

85.13.2 Identification

85.13.2.1 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	
<p>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).</p>	

85.13.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 85, Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10
Identification of amendments and corrigenda to this PICS proforma that have been completed as part of this PICS	
<p>Have any Exception items been required? No [] Yes [] (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3-2012.)</p>	

Date of Statement	
-------------------	--

¹⁰Copyright 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.

85.13.3 Major capabilities/options

Item ^a	Feature	Subclause	Value/Comment	Status	Support
XLGMII	XLGMII	85.1	Interface is supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CGMII	CGMII	85.1	Interface is supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
XLAUI	XLAUI	85.1		O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CAUI	CAUI	85.1		O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CR4	40GBASE-CR4 PMD	85.1	Can operate as 40GBASE-CR4 PMD	O.1	Yes [<input type="checkbox"/>
CR10	100GBASE-CR10 PMD	85.1	Can operate as 100GBASE-CR10 PMD	O.1	Yes [<input type="checkbox"/>
PCS	Support of 40GBASE-R PCS	85.1		CR4:M	Yes [<input type="checkbox"/>
PCS	Support of 100GBASE-R PCS	85.1		CR10:M	Yes [<input type="checkbox"/>
PMA	Support of 40GBASE-R PMA	85.1		CR4:M	Yes [<input type="checkbox"/>
PMA	Support of 100GBASE-R PMA	85.1		CR10:M	Yes [<input type="checkbox"/>
FEC	Forward error correction	85.1	Device implements BASE-R Forward Error Correction	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
AN	Auto-negotiation	85.1	Device implements Auto-Negotiation	M	Yes [<input type="checkbox"/>
DC	Delay constraints	85.4	Device conforms to delay constraints specified in 85.4	M	Yes [<input type="checkbox"/>
DSC	Skew constraints	85.5	Device conforms to Skew and Skew Variation constraints specified in 85.5	M	Yes [<input type="checkbox"/>
*MD	MDIO capability	85.6	Registers and interface supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*CBL	Cable assembly	85.10	Items marked with CBL include cable assembly specifications not applicable to a PHY manufacturer	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CA401	40GBASE-CR4 Style-1 cable assembly	85.10	Cable assembly supports 40GBASE-CR4 Style-1	CBL:O.3	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CA402	40GBASE-CR4 Style-2 cable assembly	85.10	Cable assembly supports 40GBASE-CR4 Style-2	CBL:O.3	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
CA100	100GBASE-CR10 cable assembly	85.10	Cable assembly supports 100GBASE-CR10	CBL:O.3	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
MDIST1	Style-1 MDI connector	85.11.1.1	40GBASE-CR4 device uses Style-1 MDI	O:2	Yes [<input type="checkbox"/> N/A [<input type="checkbox"/>
MDIST2	Style-2 MDI connector	85.11.1.2	40GBASE-CR4 device uses Style-2 MDI	O:2	Yes [<input type="checkbox"/> N/A [<input type="checkbox"/>

^aA "*" preceding an "Item" identifier indicates there are other PICS that depend on whether or not this item is supported.

85.13.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and baseband medium, type 40GBASE-CR4 and 100GBASE-CR10

85.13.4.1 PMD functional specifications

Item	Feature	Sub clause	Value/Comment	Status	Support
PF1	Transmit function	85.7.2	Converts four or ten logical bit streams into four or ten separate electrical streams	M	Yes []
PF2	Transmit function	85.7.2	Conveys four or ten bits from PMD service interface to MDI lanes	M	Yes []
PF3	Transmitter signal	85.7.2	A positive differential voltage corresponds to tx_bit = one	M	Yes []
PF4	Receive function	85.7.3	Converts four or ten electrical signal streams from the MDI into four or ten logical bit streams	M	Yes []
PF5	Receive function	85.7.3	Conveys four or ten logical bit streams from the MDI lanes to the PMD service interface	M	Yes []
PF6	Receiver signal	85.7.3	A positive differential voltage corresponds to rx_bit = one	M	Yes []
PF7	Global PMD Signal Detect function	85.7.4	Report state via PMD:IS_SIGNAL.indication (SIGNAL_DETECT)	M	Yes []
PF8	Signal_Detect value	85.7.4	Set to FAIL following reset	M	Yes []
PF9	Signal_Detect value	85.7.4	Set to OK on completion of training	M	Yes []
PF10	Signal_Detect value	85.7.4	Set to OK if training disabled by management	M	Yes []
PF11	Global_PMD_transmit_disable	85.7.6	Disables all transmitters by forcing a constant output state	O	Yes [] No []
PF12	Global_PMD_transmit_disable	85.7.6	Loopback not affected	O	Yes [] No []
PF13	Lane-by-lane_PMD_transmit_disable	85.7.7	Allows each lane transmitter to be selectively disabled	O	Yes [] No []
PF14	Lane-by-lane_PMD_transmit_disable	85.7.7	Disables transmitters by forcing a constant output state	O	Yes [] No []
PF15	Lane-by-lane_PMD_transmit_disable	85.7.7	Loopback not affected	O	Yes [] No []
PF16	Loopback	85.7.8	Loopback function provided	M	Yes []
PF17	PMD fault function	85.7.9	Mapped to bit 1.1.7 as listed in Table 85-3	M	Yes []
PF18	PMD control function	85.7.12	Pattern described in 72.6.10 different for each of the lanes	M	Yes []

85.13.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
MF1	Global_PMD_signal_detect	85.7.4	Set to the value described in 45.2.1.9.7	O	Yes []
MF2	Global_PMD_signal_detect	85.7.4	Set defined by the training state diagram in Figure 72–5	O	Yes [] No []
MF3	Lane-by-Lane Signal Detect function	85.7.5	Sets PMD_signal_detect_n values on a lane-by-lane basis per requirements of 85.7.5	MD:M	Yes [] N/A []
MF4	PMD_transmit_fault function	85.7.10	Mapped to the PMD_transmit_fault bit as specified in 45.2.1.7.4	MD:M	Yes [] N/A []
MF5	PMD_receive_fault function	85.7.11	Contributes to the PMA/PMD receive fault bit as specified in 45.2.1.7.5	MD:M	Yes [] N/A []

85.13.4.3 Transmitter specifications

Item	Feature	Subclause	Value/Comment	Status	Support
DS1	Meets specifications at TP2	85.8.3	Unless otherwise noted per Table 85–5	M	Yes []
DS2	Test load	85.8.3.6	100 Ω differential load with return loss using Equation (85–1)	M	Yes []
DS3	Test fixture return loss	85.8.3.6	Equation (85–37)	M	Yes []
DS4	Test fixture insertion loss	85.8.3.7	Per Equation (85–15)	M	Yes []
DS5	Signaling rate, per lane	85.8.3.9	10.3125 GBd \pm 100 ppm	M	Yes []

85.13.4.4 Receiver specifications

Item	Feature	Subclause	Value/Comment	Status	Support
RS1	Receiver tolerance	85.8.4.2	BER of better than 10^{-12}	M	Yes []
RS2	Receiver tolerance	85.8.4.2	Maximum fitted insertion loss coefficients	M	Yes []
RS3	Receiver tolerance	85.8.4.2	MDNEXT crosstalk noise	M	Yes []
RS4	Receiver tolerance	85.8.4.2	Pattern generator output amplitude	M	Yes []
RS5	Receiver tolerance	85.8.4.2	Pattern generator jitter specification	M	Yes []
RS6	Bit Error Ratio	85.8.4.2.5	BER of better than 10^{-12}	M	Yes []
RS7	Meets specifications at TP3	85.8.4	Unless otherwise noted per Table 85-7	M	Yes []
RS8	Signaling rate, per lane	85.8.4.4	10.3125 GBd \pm 100 ppm	M	Yes []
RS9	AC coupling	85.8.4.5	3 dB cutoff	M	Yes []

85.13.4.5 Cable assembly specifications

Item	Feature	Subclause	Value/Comment	Status	Support
CA1	Differential reference impedance	85.10.1	100 Ω	CBL:M	Yes [] N/A []
CA2	Insertion loss	85.10.2	Per Equation (85–19) and Table 85–10	CBL:M	Yes [] N/A []
CA3	Insertion loss deviation	85.10.3	Per Equation (85–23) and Equation (85–24)	CBL:M	Yes [] N/A []
CA4	Return loss	85.10.4	Per Equation (85–25)	CBL:M	Yes [] N/A []
CA5	Integrated crosstalk noise	85.10.7	Integrated crosstalk noise voltage per Equation (85–28) through Equation (85–33) and Table 85–11	CBL:M	Yes [] N/A []
CA6	Cable assembly test fixture insertion loss	85.10.8	Per Equation (85–34)	CBL:M	Yes [] N/A []
CA7	Mated test fixture insertion loss	85.10.9.1	Per Equation (85–35) and Equation (85–36)	CBL:M	Yes [] N/A []
CA8	Mated test fixture return loss	85.10.9.2	Per Equation (85–38)	CBL:M	Yes [] N/A []
CA9	Mated test fixtures integrated crosstalk noise	85.10.9.4	Per Equation (85–28), through Equation (85–32) and Table 85–12	CBL:M	Yes [] N/A []
CA10	Shielding	85.10.10	Class 2 or better in accordance with IEC 61196-1	CBL:M	Yes [] N/A []
CA11	Crossover function	85.10.11	Per Figure 85–18	CBL:M	Yes [] N/A []
CA12	Cable assembly connector type	85.11.1	40GBASE-CR4 Style-1 plug (SFF-8436 plug)	CA401:M	Yes [] N/A []
CA13	Pin assignments	85.11.1.1	Per Table 85–13		Yes [] N/A []
CA14	Cable assembly connector type	85.11.1	IEC 61076-3-113 latch-type plug	CA402:M	Yes [] N/A []
CA15	Pin assignments	85.11.1.2.1	Per Table 85–14		Yes [] N/A []
CA16	Cable assembly connector type	85.11.2	100GBASE-CR10 plug (SFF-8642 plug)	CA100:M	Yes [] N/A []
CA17	Pin assignments	85.11.2	Per Table 85–15		Yes [] N/A []
CA18	AC coupling	85.8.4.5	3 dB cutoff	M	Yes []

85.13.4.6 MDI connector specifications

Item	Feature	Subclause	Value/Comment	Status	Support
MDC1	MDI connector type	85.11.1.1	40GBASE-CR4 Style-1 receptacle (SFF-8436 receptacle)	CR4*MDIST1:M	Yes [] N/A []
MDC2	MDI connector type	85.11.1.2	IEC 61076-3-113 latch-type receptacle	CR4*MDIST2:M	Yes []
MDC3	MDI connector type	85.11.2.1	100GBASE-CR10 receptacle (SFF-8642 receptacle)	CR10:M	Yes [] N/A []

85.13.4.7 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
CC1	Environmental specifications	85.12		M	Yes []

86. Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE–SR4 and 100GBASE–SR10

86.1 Overview

This clause specifies the 40GBASE–SR4 PMD and the 100GBASE–SR10 PMD together with the multimode fiber medium. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 86–1, to the medium through the MDI, and optionally to the management functions that are accessible through the management interface defined in Clause 45, or equivalent.

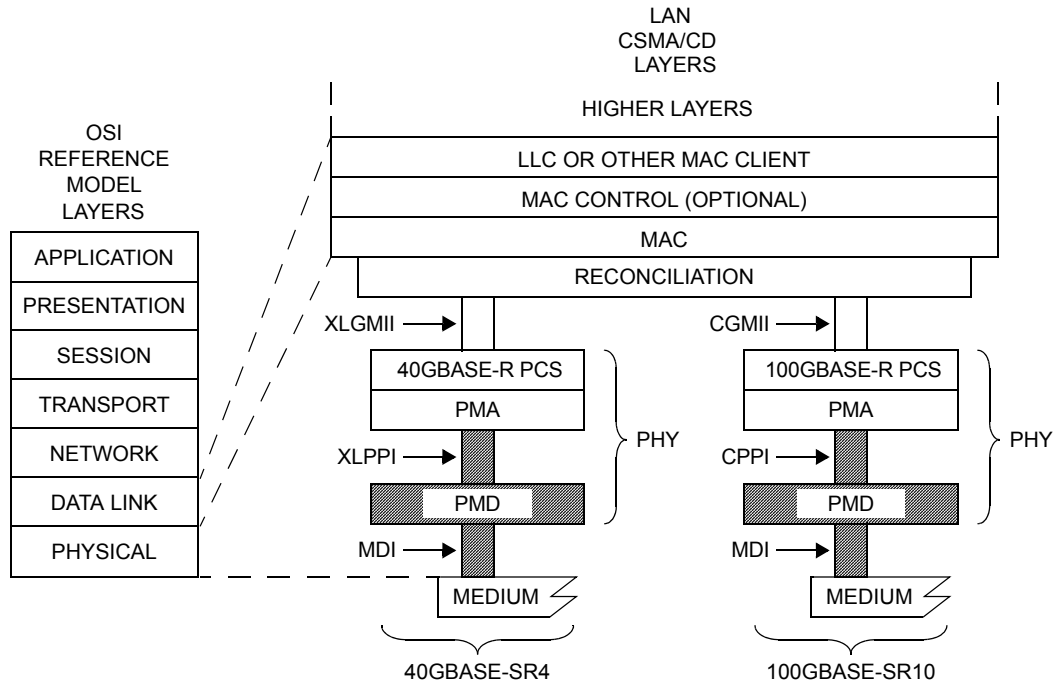
Table 86–1—Physical Layer clauses associated with the 40GBASE-SR4 and 100GBASE-SR10 PMDs

Associated clause	40GBASE–SR4	100GBASE–SR10
81—RS	Required	Required
81—XLGMII ^a	Optional	Not applicable
81—CGMII ^a	Not applicable	Optional
82—PCS for 40GBASE-R	Required	Not applicable
82—PCS for 100GBASE-R	Not applicable	Required
83—PMA for 40GBASE-R4	Required	Not applicable
83—PMA for 100GBASE-R10	Not applicable	Required
83A—XLAUI ^b	Optional	Not applicable
83A—CAUI ^b	Not applicable	Optional
83B—Chip to module XLAUI ^b	Optional	Not applicable
83B—Chip to module CAUI ^b	Not applicable	Optional
86A—XLPPi	Optional	Not applicable
86A—CPPI	Not applicable	Optional

^a XLGMII and CGMII are optional interfaces. However, if the appropriate interface is not implemented, a conforming implementation must behave functionally as though the RS, and XLGMII or CGMII, were present.

^b If XLAUI or CAUI is present, there is at least a PMA between the XLAUI or CAUI and the PMD.

Figure 86–1 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC Open System Interconnection (OSI) reference model. 40 Gb/s and 100 Gb/s Ethernet is introduced in Clause 80 and the purpose of each PHY sublayer is summarized in 80.2. Further relevant information may be found in Clause 1 (terminology and conventions, references, definitions and abbreviations) and Annex A (bibliography, referenced as [B1], [B2], etc.).



CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
 CPPI = 100 Gb/s PARALLEL PHYSICAL 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
 XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE
 XLPPi = 40 Gb/s PARALLEL PHYSICAL INTERFACE
 SR = PMD FOR MULTIMODE FIBER

Figure 86–1—40GBASE-SR4 and 100GBASE-SR10 PMDs relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

The 40GBASE-SR4 and 100GBASE-SR10 PMD sublayers provide point-to-point 40 Gb/s and 100 Gb/s Ethernet links over four or ten pairs of multimode fiber, up to at least 100 m. Table 86–2 shows the primary attributes of each PMD type.

Table 86–2—Summary of 40GBASE-SR4 and 100GBASE-SR10

PMD Type	40GBASE-SR4	100GBASE-SR10	Unit
Fiber type	50/125 μm multimode, type A1a.2 ^a (OM3) or A1a.3 ^b (OM4)		
Number of fiber pairs	4	10	
Nominal wavelength	850		nm
Required operating range	0.5 to 100 for OM3		m
	0.5 to 150 for OM4 ^c		
Signaling rate, each lane	10.3125 ±100 ppm		GBd

^a Type A1a.2 (OM3) specified in IEC 60793-2-10. See 86.10.2.1.

^b Type A1a.3 (OM4) specified in IEC 60793-2-10. See 86.10.2.1.

^c This is an engineered link with maximum 1 dB connection and splice loss.

40GBASE-SR4 uses four identical lanes, while 100GBASE-SR10 uses ten of the same lanes. In this clause, where there are four or ten items (depending on PMD type) such as lanes, the items are numbered from 0 to $n - 1$, and an example item is numbered i . Thus n is 4 or 10.

The connection to the PMA may use the optional physical instantiation of the PMD service interface called XLPPi (four lanes, for 40GBASE-SR4) or CPPI (ten lanes, for 100GBASE-SR10). The term “nPPI” is used to denote either XLPPi or CPPI, or both.

This clause is arranged as follows: following the overview and an abstract description of the PMD service interface, delay and Skew specifications, control and status variables and registers, a block diagram and high-level specification of the PMD functions, and lane assignments, 86.7 (four parts) contains the optical specifications for 40GBASE-SR4 and 100GBASE-SR10. 86.8 defines test points and optical and dual-use parameters. 86.9 addresses safety, installation, environment and labeling, 86.10 defines the optical channel, and 86.11 contains the PICS. Annex 86A contains the electrical specifications for XLPPi and CPPI, electrical compliance boards, definitions of electrical parameters, and a recommended PCB response for the host (PMA).

86.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 40GBASE-SR4 and 100GBASE-SR10 PMDs. The service interfaces for these PMDs are described in an abstract manner and do not imply any particular implementation, although an optional implementation of the PMD service interface, the Parallel Physical Interface (nPPI), is specified in 86A.4.1 and 86A.4.2. The PMD service interface supports the exchange of encoded data between the PMA entity that resides just above the PMD, and the PMD entity. The PMD translates the encoded data to and from signals suitable for the specified medium.

The PMD service interface 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 40GBASE-SR4 PMD has four parallel bit streams, hence $i = 0$ to 3 for 40GBASE-SR4 and the 100GBASE-SR10 PMD has ten parallel bit streams, hence $i = 0$ to 9 for 100GBASE-SR10.

The PMA (or the PMD) continuously sends four or ten parallel bit streams to the PMD (or the PMA), one per lane, each at a nominal signaling rate of 10.3125 GBd (see 83.4).

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

The PMD:IS_UNITDATA_ i .indication primitive corresponds to one of the signals received from the MDI. This primitive is received by the client (the PMA), as described in 83.4.

The PMD:IS_SIGNAL.indication(SIGNAL_DETECT) primitive is generated by the PMD to report the parameter SIGNAL_DETECT, which indicates the status of the signals being received from the MDI (see 86.5.4). There is one parameter and primitive, reporting on all the lanes as a group. This primitive is received by the client (the PMA) as PMD:IS_SIGNAL.indication(SIGNAL_OK), as described in 83.4.

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

86.3 Delay and Skew

86.3.1 Delay constraints

The sum of the transmit and receive delays at one end of the link contributed by the 40GBASE-SR4 PMD including 2 m of fiber in one direction shall be no more than 1024 bit times (2 `pause_quanta` or 25.6 ns). The sum of the transmit and receive delays at one end of the link contributed by the 100GBASE-SR10 PMD including 2 m of fiber in one direction shall be no more than 2048 bit times (4 `pause_quanta` or 20.48 ns). A description of overall system delay constraints and the definitions for bit times and `pause_quanta` can be found in 80.4 and its references.

86.3.2 Skew and Skew Variation 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. Skew points as they relate to the nPPI are shown in Figure 86–3.

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 43 ns and the Skew Variation at SP2 is limited to 400 ps.

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

The Skew at SP4 (the receiver MDI) shall be less than 134 ns and the Skew Variation at SP4 shall be less than 3.4 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 145 ns and the Skew Variation at SP5 shall be less than 3.6 ns.

For more information on Skew and Skew Variation see 80.5. The measurements of Skew and Skew Variation are defined in 86.8.3.1.

86.4 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that may provide control and status information for and about the PMD. If MDIO is implemented, the mapping of MDIO control variables to PMD control variables shall be as shown in Table 86–3, and the mapping of MDIO status variables to PMD status variables shall be as shown in Table 86–4.

Table 86–3—MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable register	1.9.0	PMD_global_transmit_disable
PMD transmit disable 9 ^a	PMD transmit disable register	1.9.10	PMD_transmit_disable_9
PMD transmit disable 8 ^a to PMD transmit disable 0	PMD transmit disable register	1.9.9 to 1.9.1	PMD_transmit_disable_8 to PMD_transmit_disable_0

^a For 40GBASE-SR4, the highest-numbered six of the ten lane-by-lane transmit disables do not apply.

Table 86–4—MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	PMD_global_signal_detect
PMD receive signal detect 9 ^a	PMD receive signal detect register	1.10.10	PMD_signal_detect_9
PMD receive signal detect 8 ^a to PMD receive signal detect 0	PMD receive signal detect register	1.10.9 to 1.10.1	PMD_signal_detect_8 to PMD_signal_detect_0

^a For 40GBASE-SR4, the highest-numbered six of the ten lane-by-lane signal detects do not apply.

86.5 PMD functional specifications

The 40GBASE–SR4 and 100GBASE–SR10 PMDs perform the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

86.5.1 PMD block diagram

The PMD block diagram is shown in Figure 86–2. Figure 86–3 shows the test points. It is not required that the PMD service interface be exposed or measurable (nPPI as defined in Annex 86A with compliance points TP1, TP1a, TP4, TP4a). However, if it is not, a conforming implementation must behave as though the interface were compliant.

For purposes of system conformance, the PMD sublayer is standardized at the test points described in 86.8.1. The transmit side electrical signal (PMA output and PMD electrical input) is defined at TP1 and TP1a; see Annex 86A. The optical transmit signal is defined at the output end of a 50 μm multimode fiber patch cord (TP2), between 2 m and 5 m in length. Unless specified otherwise, all optical transmitter measurements and tests defined in 86.8 are made at TP2. The optical receive signal is defined at the output

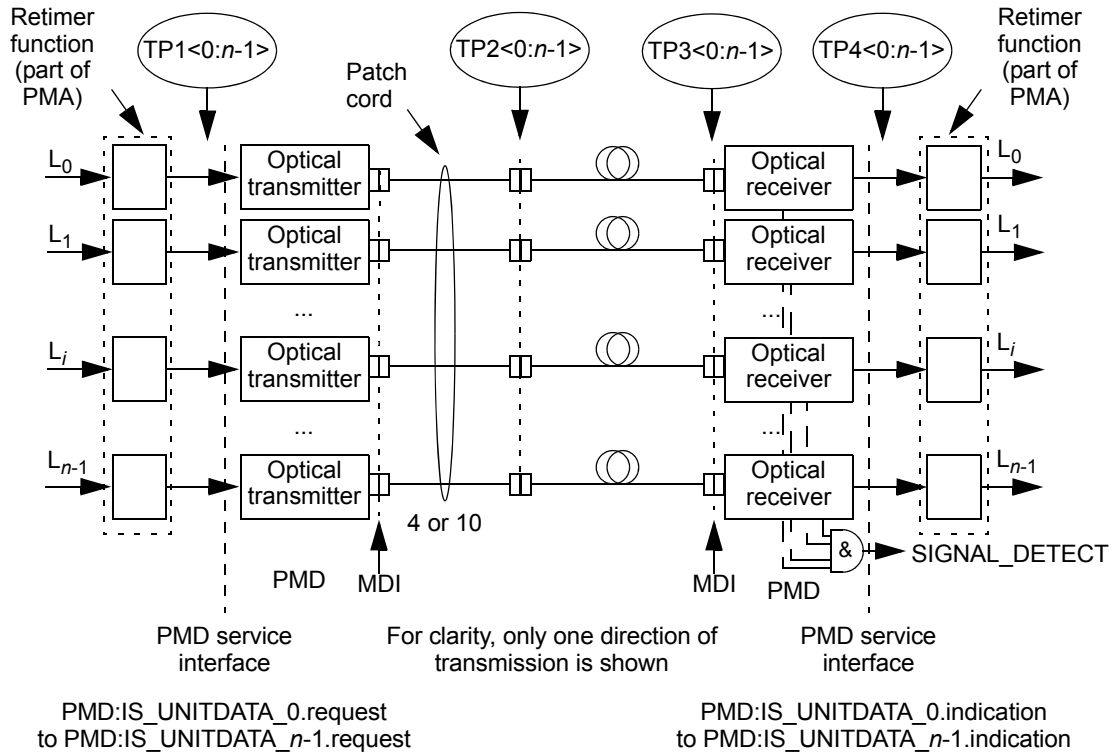


Figure 86–2—Block diagram for 40GBASE–SR4 and 100GBASE–SR10 transmit/receive paths

of the fiber optic cabling (TP3) at the MDI (see 86.10.3). Unless specified otherwise, all optical receiver measurements and tests defined in 86.8 are made at TP3. The receive side electrical signal (PMD electrical output and PMA input) is defined at TP4 and TP4a; see Annex 86A.

86.5.2 PMD transmit function

The PMD Transmit function shall convert the four or ten electronic bit streams requested by the PMD service interface messages `PMD:IS_UNITDATA_0.request` to `PMD:IS_UNITDATA_n-1.request` into the same number of optical signal streams. The optical signal streams are delivered to the MDI, which contains four or ten parallel light paths for transmit, according to the transmit optical specifications in this clause. The higher optical power level in each signal stream shall correspond to `tx_bit = one`.

86.5.3 PMD receive function

The PMD Receive function shall convert the four or ten parallel optical signal streams received from the MDI into separate electronic bit streams for delivery to the PMD service interface using the messages `PMD:IS_UNITDATA_0.indication` to `PMD:IS_UNITDATA_n-1.indication`, all according to the receive optical specifications in this clause. The higher optical power level in each signal stream shall correspond to `rx_bit = one`.

86.5.4 PMD global signal detect function

The PMD global signal detect function shall report the state of `SIGNAL_DETECT` via the PMD service interface. The `SIGNAL_DETECT` parameter is signaled continuously, while the `PMD:IS_SIGNAL.indication` message is generated when a change in the value of `SIGNAL_DETECT` occurs.

SIGNAL_DETECT shall be a global indicator of the presence of optical signals on all lanes. The value of the SIGNAL_DETECT parameter shall be generated according to the conditions defined in Table 86–5. The PMD receiver is not required to verify whether a compliant 40GBASE-SR4 or 100GBASE-SR10 signal is being received. This standard imposes no response time requirements on the generation of the SIGNAL_DETECT parameter.

Table 86–5—SIGNAL_DETECT value definition

Receive conditions	SIGNAL_DETECT value
For any lane; Average optical power at TP3 ≤ -30 dBm	FAIL
For all lanes; [(Optical power at TP3 \geq Minimum OMA, each lane, in Table 86–7) and (compliant 40GBASE–SR4 or 100GBASE–SR10 signal input as appropriate)]	OK
All other conditions	Unspecified

As an unavoidable consequence of the requirements for the setting of the SIGNAL_DETECT parameter, implementations must provide adequate margin between the input optical power level at which the SIGNAL_DETECT parameter is set to OK, and the inherent noise level of the PMD including the effects of crosstalk, power supply noise, etc.

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

86.5.5 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each PMD_signal_detect_{*i*}, where *i* represents the lane number in the range 0:*n*-1, shall be continuously set in response to the optical signal on its associated lane, according to the requirements of Table 86–5.

86.5.6 PMD reset function

If the MDIO interface is implemented, and if PMD_reset is asserted, the PMD shall be reset as defined in 45.2.1.1.1.

86.5.7 PMD global transmit disable function (optional)

The PMD_global_transmit_disable function is optional and allows all of the optical transmitters to be disabled.

- a) When the PMD_global_transmit_disable variable is set to one, this function shall turn off all of the optical transmitters so that each transmitter meets the requirements of the average launch power of the OFF transmitter in Table 86–6.
- b) If a PMD_fault is detected, then the PMD may set the PMD_global_transmit_disable to one, turning off the optical transmitter in each lane.

86.5.8 PMD lane-by-lane transmit disable function (optional)

The PMD_transmit_disable_*i* function (where *i* represents the lane number in the range 0:*n*-1) is optional and allows the optical transmitter in each lane to be selectively disabled.

- a) When a PMD_transmit_disable_*i* variable is set to one, this function shall turn off the optical transmitter associated with that variable so that the transmitter meets the requirements of the average launch power of the OFF transmitter in Table 86–6.
- b) If a PMD_fault is detected, then the PMD may set each PMD_transmit_disable_*i* to one, turning off the optical transmitter in each lane.

If the optional PMD_transmit_disable_*i* function is not implemented in MDIO, an alternative method may be provided to independently disable each transmit lane.

86.5.9 PMD fault function (optional)

If the PMD has detected a local fault on any of the transmit or receive paths, the PMD shall set PMD_fault to one. If the MDIO interface is implemented, PMD_fault shall be mapped to the PMA/PMD fault bit as specified in 45.2.1.2.1.

86.5.10 PMD transmit fault function (optional)

If the PMD has detected a local fault on any transmit lane, the PMD shall set the PMD_transmit_fault variable to one. If the MDIO interface is implemented, PMD_transmit_fault shall be mapped to the PMA/PMD transmit fault bit as specified in 45.2.1.7.4.

86.5.11 PMD receive fault function (optional)

If the PMD has detected a local fault on any receive lane, the PMD shall set the PMD_receive_fault variable to one. If the MDIO interface is implemented, PMD_receive_fault shall be mapped to the PMA/PMD receive fault bit as specified in 45.2.1.7.5.

86.6 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for 40GBASE-SR4 or 100GBASE-SR10. While it is expected that a PMD will map electrical lane *i* to optical lane *i* and vice versa, there is no need to define the physical ordering of the lanes, as the PCS is capable of receiving the lanes in any arrangement. The positioning of transmit and receive lanes at the MDI is specified in 86.10.3.

86.7 PMD to MDI specifications for 40GBASE-SR4 or 100GBASE-SR10

The required operating range for the 40GBASE-SR4 and 100GBASE-SR10 PMD is defined in Table 86–2. A compliant PMD operates on 50/125 μm multimode fibers according to the specifications of Table 86–14. A PMD which exceeds the operating range requirement while meeting all other optical specifications is considered compliant (e.g., operating at 125 m meets the operating range requirement of 0.5 m to 100 m). The signaling rate for a lane of a 40GBASE-SR4 or 100GBASE-SR10 PMD shall be as defined in Table 86–2. The optical signal at the transmit and receive side of the MDI is specified in 86.7.1 and 86.7.3. The range of optical signals within the optical medium is defined in 86.7.2, and an illustrative link power budget is provided in 86.7.4. Test points are defined in 86.8.1.

86.7.1 Transmitter optical specifications

Each lane of a 40GBASE-SR4 or 100GBASE-SR10 optical transmitter shall meet the specifications of Table 86-6 per the definitions in 86.8.

Table 86-6—40GBASE-SR4 or 100GBASE-SR10 optical transmit characteristics

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

^a RMS spectral width is the standard deviation of the spectrum.

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

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

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

Table 86–7 gives the characteristics of a signal within, and at the receiving end of, a lane of a compliant 40GBASE–SR4 or 100GBASE–SR10 optical channel, and the aggregate signal. A signal with power in OMA or average power not within the ranges given cannot be compliant. However, a signal with power values within the ranges is not necessarily compliant.

Table 86–7—Characteristics of signal within, and at the receiving end of, a compliant optical channel

Description	Minimum		Maximum	Unit
	OM3	OM4		
Fiber type	OM3	OM4		
Total average power for 40GBASE–SR4	–3.5	–3.1	+8.4	dBm
Total average power for 100GBASE–SR10	+0.5	+0.9	+12.4	dBm
Average power, each lane	–9.5	–9.1	+2.4	dBm
Optical Modulation Amplitude (OMA), each lane	–7.5	–7.1	+3	dBm

NOTE—Table 86–7 provides information for diagnostic purposes that is needed by network operators in maintenance. There is no need to assure compliance to it in normal circumstances.

86.7.3 40GBASE–SR4 or 100GBASE–SR10 receiver optical specifications

Each lane of a 40GBASE–SR4 or 100GBASE–SR10 optical receiver shall meet the specifications defined in Table 86–8 per the definitions in 86.8.

Table 86–8—40GBASE–SR4 or 100GBASE–SR10 optical receiver characteristics

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

^a The receiver shall be able to tolerate, without damage, continuous exposure to a modulated optical input signal having this power level on one lane. The receiver does not have to operate correctly at this input power.

^b Measured with conformance test signal at TP3 (see 86.8.4.7).

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

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

86.7.4 40GBASE–SR4 or 100GBASE–SR10 illustrative link power budget

Illustrative power budgets and penalties for 40GBASE–SR4 or 100GBASE–SR10 optical channels are shown in Table 86–9.

Table 86–9—40GBASE–SR4 or 100GBASE–SR10 illustrative link power budgets

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

^a Per IEC 60793-2-10.

^b The channel insertion loss is calculated using the maximum distances specified in Table 86–2 and cabled optical fiber attenuation of 3.5 dB/km at 850 nm plus an allocation for connection and splice loss given in 86.10.2.2.1.

^c Link penalties are used for link budget calculations. They are not requirements and are not meant to be tested.

^d This unallocated margin is not available for use.

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

In this section, the test points and the parameters applicable to optical signals and for dual use (both optical and electrical) are defined. Test points are defined in 86.8.1, test patterns in 86.8.2, and parameters in 86.8.3 and 86.8.4. Further parameters for electrical use, and the use of electrical compliance boards, are defined in 86A.5.

86.8.1 Test points and compliance boards

Figure 86–3 shows the six test points for 40GBASE-SR4 and 100GBASE-SR10. These are TP1, TP1a, TP2, TP3, TP4, and TP4a; four of these are Skew points SP2, SP3, SP4, and SP5 as shown. Figure 86–3 also shows the substitution of compliance boards for PMD or PMA. This is explained in Annex 86A. Table 86–10 shows the parameters or signals measured at each point, including the electrical compliance points.

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

86.8.2 Test patterns and related subclauses

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some parameters. Table 86–11 lists the defined test patterns, and Table 86–12 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subclauses in which each parameter is defined. As Pattern 3 is more demanding than

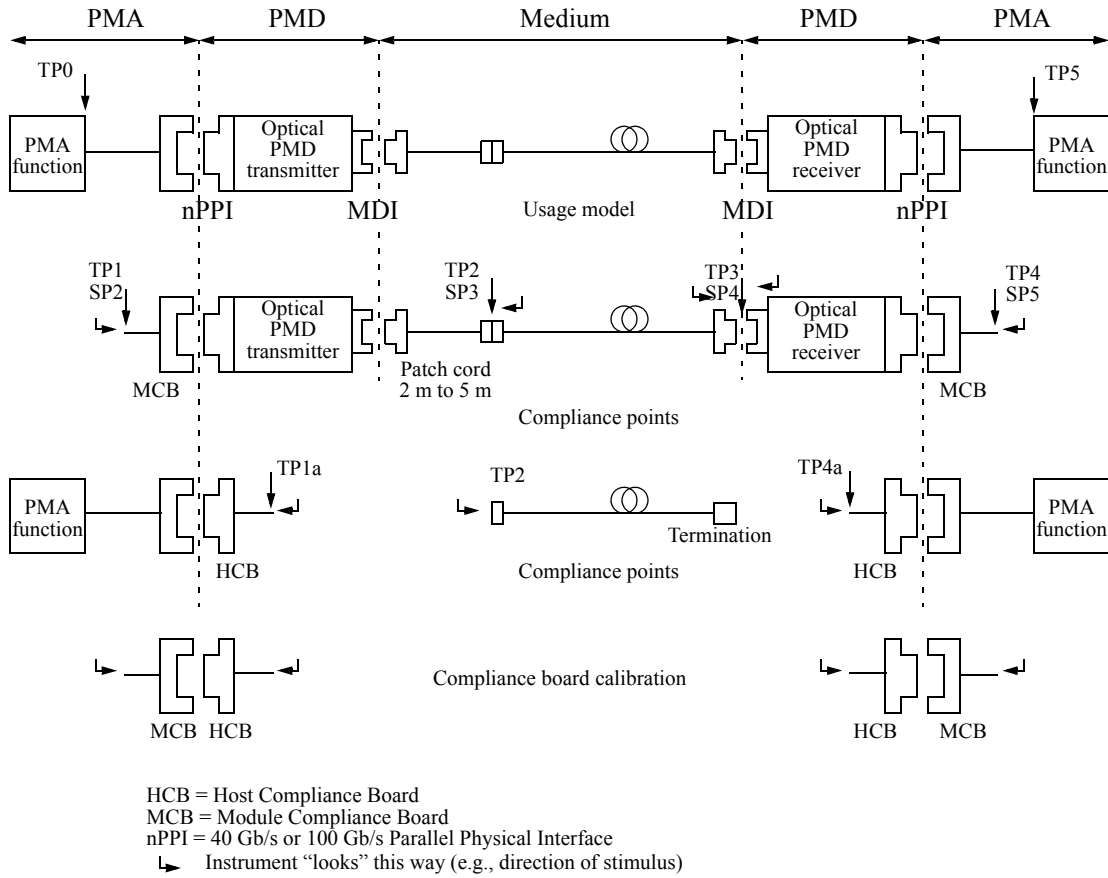


Figure 86-3—Test points for 40GBASE-SR4 and 100GBASE-SR10

Table 86-10—Parameters defined at each test point

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

Table 86–11—Test patterns

Pattern no.	Pattern	Pattern defined in
Square wave	Square wave (8 ones, 8 zeros)	83.5.10
3	PRBS31	83.5.10
4	PRBS9	83.5.10
5	Scrambled idle	82.2.10

Pattern 5 (which itself is the same or more demanding than other 40GBASE-R or 100GBASE-R bit streams) an item that is compliant using Pattern 5 is considered compliant even if it does not meet the required limit using Pattern 3. Test patterns for further electrical parameters are given in Table 86A–6.

Table 86–12—Test patterns and related subclauses

Parameter	Pattern	Related subclause
Wavelength, spectral width	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.4.1
Average optical power	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.4.2
Transmitter OMA (modulated optical power)	Square wave or 4	86.8.4.3
Extinction ratio	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.4.5
Transmitted optical waveform (eye mask)	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.3.2, 86.8.4.6
TDP (transmitter and dispersion penalty)	3 or 5	86.8.4.4
Stressed receiver sensitivity	3 or 5	86.8.4.7
Receiver jitter tolerance	3 or 5	86.8.4.8
Calibration of OMA for receiver tests	Square wave or 4	52.9.9
Vertical eye closure penalty calibration	3 or 5	52.9.9
J2 Jitter	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.3.3.1
J9 Jitter	3 or 5	86.8.3.3.2
Data Dependent Pulse Width Shrinkage (DDPWS)	4	86A.5.3.4

Table 86–12—Test patterns and related subclauses (continued)

Parameter	Pattern	Related subclause
AC common-mode voltage	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86A.5.3.1
Transition time	Square wave or 4	86A.5.3.3
Electrical waveform (eye mask)	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.3.2, 86A.5.3.6

86.8.2.1 Multi-lane testing considerations

TDP is defined for each lane, at a BER of 10^{-12} on that lane. Stressed receiver sensitivity, receiver jitter tolerance and host input signal tolerance (in Annex 86A) are defined for an interface BER of 10^{-12} . The interface BER is the average of the four or ten BERs of the receive lanes when they are stressed.

Measurements with Pattern 3 (PRBS31) allow lane-by-lane BER measurements. Measurements with Pattern 5 (scrambled idle) give the interface BER if all lanes are stressed at the same time. If each lane is stressed in turn, the BER is diluted by the three or nine unstressed lanes, and the BER for that stressed lane alone must be found, e.g., by multiplying by 4 or 10 if the unstressed lanes have low BER. To allow TDP measurement with Pattern 5, unstressed lanes for the error detector may be created by setting the power at the reference receivers well above their sensitivities, or by copying the contents of the transmit lanes not under BER test to the error detector by other means. In stressed receiver sensitivity and receiver jitter tolerance measurements, unstressed lanes may be created by setting the power at the receiver under test well above its sensitivity and/or not stressing those lanes with ISI and jitter, or by other means. Each receive lane is stressed in turn while all are operated. All aggressor lanes are operated as specified. To find the interface BER, the BERs of all the lanes when stressed are averaged.

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

86.8.3 Parameters applicable to both electrical and optical signals

86.8.3.1 Skew and Skew Variation

Skew and Skew Variation are defined in 80.5 and are required to remain within the limits given in 86.3.2 over the time that the link is in operation. Skew points as they relate to the nPPI are shown in Figure 86–3. The measurement of Skew and Skew Variation is made by acquiring the data on each lane using a clock and data recovery unit with a high-frequency corner bandwidth and slope as specified in 86.8.3.2. The arrival times of the one to zero transition of the alignment marker sync bits on each lane are then compared. This arrangement ensures that any high-frequency jitter that is present on the signals is not included in the Skew measurement.

86.8.3.2 Eye diagrams

Eye diagrams can be used to assess both electrical and optical signals. The measurement of the electrical eye is defined in 86A.5.3.6 and the measurement of the optical transmitter waveform is defined in 86.8.4.6. Whether electrical or optical eye diagrams, all co-propagating and counter-propagating lanes are active as crosstalk sources, using one of patterns 3, 5, or a valid 40GBASE-R or 100GBASE-R signal. The input lanes of the item under test are receiving signals that are asynchronous to those being output.

Normalized times of 0 and 1 on the unit interval scale are determined by the eye crossing means measured at the average value of the eye pattern. A clock recovery unit (CRU) is used to trigger the oscilloscope for mask measurements, as shown in Figure 52–9. It has a high-frequency corner bandwidth of 4 MHz and a slope of –20 dB/decade. The CRU tracks acceptable levels of low-frequency jitter and wander.

Consideration should be given as to whether a correction is needed for actual instrument properties.

86.8.3.2.1 Eye mask acceptable hit count examples

An example calculation relating hit count to hit ratio for an eye mask measurement using an oscilloscope is detailed below.

If an oscilloscope records 1350 samples/screen, and the time-base is set to 0.2 UI per division with 10 divisions across the screen, and the measurement is continued for 200 waveforms, then a signal that averages less than 6.75 hits is compliant to a hit ratio specification of 5×10^{-5} , i.e.,

$$\frac{5 \times 10^{-5} \times 200 \times 1350}{0.2 \times 10} = 6.75 \quad (86-1)$$

Likewise, if a measurement is continued for 1000 waveforms, then an average of less than 33.75 hits is compliant. An extended measurement is expected to give a more accurate result, and a single reading of 6 hits in 200 waveforms would not give a statistically significant pass or fail.

The hit ratio limit has been chosen to avoid misleading results due to signal and oscilloscope noise.

86.8.3.3 Jitter

J2 Jitter and J9 Jitter are specified with all co-propagating and counter-propagating lanes active as crosstalk sources, using one of patterns 3, 5, or a valid 40GBASE-R or 100GBASE-R signal. The input lanes of the item under test are receiving signals that are asynchronous to those being output.

86.8.3.3.1 J2 Jitter

J2 Jitter is defined as the time interval that includes all but 10^{-2} of the jitter distribution, which is the time interval from the 0.5th to the 99.5th percentile of the jitter histogram. This may be measured using an oscilloscope, or if measured by plotting BER vs. decision time, J2 is the time interval between the two points with a BER of 2.5×10^{-3} . Oscilloscope histograms should include at least 10 000 hits, and should be taken over about 1% of the signal amplitude. Test patterns are given in Table 86–12.

86.8.3.3.2 J9 Jitter

J9 Jitter is defined as the time interval that includes all but 10^{-9} of the jitter distribution. If measured by plotting BER vs. decision time, it is the time interval between the two points with a BER of 2.5×10^{-10} . Test patterns are given in Table 86–12.

86.8.4 Optical parameter definitions

86.8.4.1 Wavelength and spectral width

The wavelength of each optical lane shall be within the range given in Table 86–6 if measured using the method given in TIA–455–127-A. The lane under test is modulated using the test pattern defined in Table 86–12.

86.8.4.2 Average optical power

Average optical power is defined by the methods given in IEC 61280-1-1.

86.8.4.3 Optical Modulation Amplitude (OMA)

OMA shall be as defined in 52.9.5 for measurement with a square wave (8 ones, 8 zeros) test pattern or 68.6.2 (from the variable MeasuredOMA in 68.6.6.2) for measurement with a PRBS9 test pattern, with the exception that each optical lane is tested individually. See 86.8.2 for test pattern information.

86.8.4.4 Transmitter and dispersion penalty (TDP)

Transmitter and dispersion penalty (TDP) shall be as defined for 10GBASE-S in 52.9.10 with the following exceptions:

- a) Each optical lane is tested individually with all other lanes in operation.
- b) The test pattern is as defined in Table 86–12.
- c) The transmitter is tested using an optical channel with an optical return loss of 12 dB.
- d) The reference receiver (including the effect of the decision circuit) has a fourth order Bessel-Thomson filter response with a bandwidth of 6.1 GHz. The transversal filter of 52.9.10.3 is not used.
- e) The reference sensitivity S and the measurement P_{DUT} are both measured with the sampling instant displaced from the eye center by ± 0.15 UI. For each of the two cases (early and late), if $P_{DUT}(i)$ is larger than $S(i)$, the $TDP(i)$ for the transmitter under test is the difference between $P_{DUT}(i)$ and $S(i)$, $TDP(i) = P_{DUT}(i) - S(i)$. Otherwise, $TDP(i) = 0$. The TDP is the larger of the two $TDP(i)$.
- f) The test setup illustrated in Figure 52–12 shows the reference method. Other measurement implementations may be used with suitable calibration.
- g) The BER of 10^{-12} is for the lane under test on its own. See 86.8.2.1 for multi-lane pattern considerations.

NOTE—Because practical receivers and decision circuits have noise and timing impairments, the sampling instant offsets have to be calibrated. One method of doing this is via a jitter bathtub method using a known low-jitter signal.

86.8.4.5 Extinction ratio

Extinction ratio is defined by the methods of IEC 61280–2–2 using the test pattern defined in Table 86–12.

NOTE—Extinction ratio and OMA are defined with different test patterns (see Table 86–12).

86.8.4.6 Transmitter optical waveform (transmit eye)

The required optical transmitter pulse shape characteristics are specified in the form of a mask of the transmitter eye diagram specified in Table 86–6 and shown in Figure 86–4. The transmitter optical waveform of a port transmitting the test pattern specified in Table 86–12 shall meet specifications according to the methods specified in 86.8.4.6.1 with the filter nominal reference frequency f_r of 7.5 GHz and filter tolerances as specified for STM-64 in ITU-T G.691.

86.8.4.6.1 Optical transmitter eye mask

The eye is measured with respect to the mask specified in Table 86–6 and shown in Figure 86–4 using a receiver with the fourth-order Bessel-Thomson response having a transfer function given by Equation (86–2) and Equation (86–3):

$$H(y) = \frac{105}{105 + 105y + 45y^2 + 10y^3 + y^4} \tag{86-2}$$

where:

$$y = 2.114p; \quad p = \frac{j\omega}{\omega_r}; \quad \omega_r = 2\pi f_r; \quad f_r = \text{Reference frequency in GHz} \tag{86-3}$$

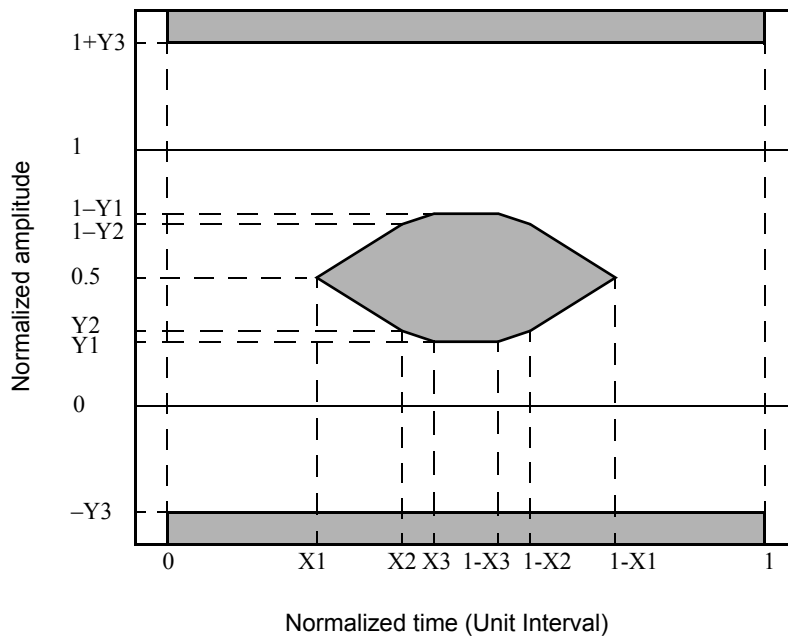


Figure 86–4—Transmitter eye mask definition

The Bessel-Thomson receiver is not intended to represent the noise filter used within a compliant optical receiver, but is intended to provide uniform measurement conditions at the transmitter. Compensation may be made for variation of the reference receiver filter response from an ideal fourth-order Bessel-Thomson response.

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

Further requirements are given in 86.8.3.2.

The transmitter shall achieve a hit ratio lower than the limit of hits per sample specified in the appropriate table or 5×10^{-5} hits per sample if not otherwise specified. “Hits” are the number of samples within the grey areas of Figure 86–4, and the sample count is the total number of samples from 0 UI to 1 UI. Some illustrative examples are provided in 86.8.3.2.1.

Further information on optical eye pattern measurement procedures may be found in IEC 61280-2-2.

86.8.4.7 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in Table 86–8 if measured using the method defined by 52.9.9 with the conformance test signal at TP3 and with the following exceptions:

- a) The reference test procedure for a single lane is defined in 52.9.9. See 86.8.2.1 and below for multi-lane considerations.
- b) The sinusoidal amplitude interferer is replaced by a Gaussian noise generator.
- c) The fourth-order Bessel-Thomson filter is replaced by a low-pass filter followed by a limiter and a fourth-order Bessel-Thomson filter.
- d) The sinusoidal jitter is at a fixed 80 MHz frequency and between 0 and 0.05 UI peak-to-peak amplitude.
- e) The Gaussian noise generator, the amplitude of the sinusoidal jitter, and the Bessel-Thomson filter are adjusted so that the VECP, J2 Jitter and J9 Jitter specifications given in Table 86–8 are simultaneously met (the random noise effects such as RIN, random clock jitter do not need to be minimized).
- f) The pattern for the received compliance signal is specified in Table 86–12.
- g) The interface BER of the PMD receiver is the average of the BER of all receive lanes while stressed and at the specified receive OMA.
- h) Where nPPI or XLAUI/CAUI is exposed, a PMD receiver is considered compliant if it meets the module electrical output specifications at TP4 given in Table 86A–3 for nPPI, or the requirements in Table 83B–3 for XLAUI/CAUI.

Stressed receiver sensitivity is defined with all transmit and receive lanes in operation. Each receive lane is tested in turn while all aggressor receive lanes are operated as specified in Table 86–8. Pattern 3 or Pattern 5, or a valid 40GBASE–R4 or 100GBASE–R10 signal is sent from the transmit section of the receiver under test. The signal being transmitted is asynchronous to the received signal. If Pattern 3 is used for the transmit and receive lanes not under test, there is at least 31 UI delay between the PRBS31 patterns generated on one lane and any other lane.

For 40GBASE-SR4 and 100GBASE-SR10, the relevant BER is the interface BER. The interface BER is the average of the four or ten BERs of the receive lanes when stressed: see 86.8.2.1.

86.8.4.8 Receiver jitter tolerance

Receiver jitter tolerance shall be as defined as in 68.6.11, with the following differences:

- a) The reference test procedure for a single lane is defined in 68.6.11. See 86.8.2.1 for multi-lane considerations.
- b) The pattern to be received is specified in Table 86–12.
- c) The parameters of the signal are specified in Table 86–8 and the power in OMA at the receiver is set to the maximum for receiver jitter tolerance in OMA given in Table 86–8.
- d) Each receive lane is tested in turn while all are operated. All aggressor lanes are operated as specified in Table 86–8.
- e) The receive lanes not being tested are receiving Pattern 3, Pattern 5, or a valid 40GBASE-R4 or 100GBASE-R10 signal.
- f) The transmitter is transmitting one of these signals using all lanes.
- g) The transmitter and the receiver are not synchronous.
- h) The interface BER of the PMD receiver is the average of the BER of all receive lanes when stressed.
- i) The mode-conditioning patch cord suitable for 62.5/125 μm fiber is not used.

86.9 Safety, installation, environment, and labeling

86.9.1 General safety

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

86.9.2 Laser safety

40GBASE-SR4 and 100GBASE-SR10 optical transceivers shall conform to Hazard Level 1M laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service. This documentation explicitly defines requirements and usage restrictions on the host system necessary to meet these safety certifications.¹¹

86.9.3 Installation

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

86.9.4 Environment

The 40GBASE-SR4 and 100GBASE-SR10 operating environment specifications are as defined in 52.11, as defined in 52.11.1 for electromagnetic emission, and as defined in 52.11.2 for temperature, humidity, and handling.

86.9.5 PMD labeling

The 40GBASE-SR4 and 100GBASE-SR10 labeling recommendations and requirements are as defined in 52.12.

86.10 Optical channel

The fiber optic cabling (channel) contains 4 or 10 optical fibers for each direction to support 40GBASE-SR4 or 100GBASE-SR10, respectively. The fiber optic cabling interconnects the transmitters at the MDI on one end of the channel to the receivers at the MDI on the other end of the channel. As defined in 86.10.3, the optical lanes appear in defined locations at the MDI but the locations are not assigned specific lane numbers within this standard because any transmitter lane may be connected to any receiver lane.

86.10.1 Fiber optic cabling model

The fiber optic cabling model is shown in Figure 86-5.

The channel shall comply with the specifications in Table 86-13.

¹¹A host system that fails to meet the manufacturers requirements and/or usage restrictions may emit laser radiation in excess of the safety limits of one or more safety standards. In such a case, the host manufacturer is required to obtain its own laser safety certification.

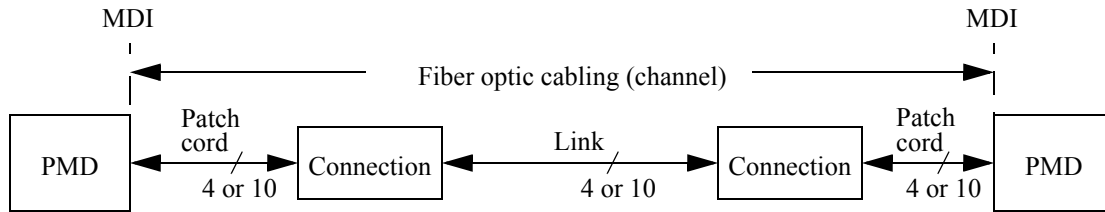


Figure 86-5—Fiber optic cabling model

Table 86-13—Fiber optic cabling (channel) characteristics at 850 nm

Description	Type	OM3	OM4	Unit
Operating distance	Max	100	150	m
Cabling Skew	Max	79		ns
Cabling Skew Variation ^a	Max	2.5		ns
Channel insertion loss	Min	0		dB
Channel insertion loss ^b	Max	1.9 ^c	1.5 ^d	dB

^a An additional 300 ps of Skew Variation could be caused by wavelength changes, which are attributable to the transmitter not the channel.

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

^c 1.5 dB allocated for connection and splice loss.

^d 1 dB allocated for connection and splice loss.

A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, modal dispersion, reflections and losses of all connectors and splices meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with IEC 61280-4-1:2009. As OM4 optical fiber meets the requirements for OM3, a channel compliant to the “OM3” column may use OM4 optical fiber, or a combination of OM3 and OM4.

86.10.2 Characteristics of the fiber optic cabling (channel)

86.10.2.1 Optical fiber cable

The fiber contained within the 40GBASE-SR4 or 100GBASE-SR10 fiber optic cabling shall comply with the specifications and parameters of Table 86-14. A variety of multimode cable types may satisfy these requirements, provided the resulting channel also meets the specifications of Table 86-13. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

86.10.2.2 Optical fiber connection

An optical fiber connection, as shown in Figure 86-5, consists of a mated pair of optical connectors.

Table 86–14—Optical fiber and cable characteristics

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

^a IEC 60793-2-10 type A1a.2^b IEC 60793-2-10 type A1a.3^c When measured with the launch conditions specified in Table 86–6.**86.10.2.2.1 Connection insertion loss**

The maximum operating distances are based on an allocation of 1.5 dB or 1 dB total connection and splice loss. For example, these allocations support two connections, each with an insertion loss of 0.75 dB or 0.5 dB respectively. Connections with lower loss characteristics may be used provided the requirements of Table 86–13 are met. However, the loss of a single connection shall not exceed 0.75 dB.

86.10.2.2.2 Maximum discrete reflectance

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

86.10.3 Medium Dependent Interface (MDI)

The 40GBASE–SR4 or 100GBASE–SR10 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in Figure 86–5). The 40GBASE–SR4 PMD is coupled to the fiber optic cabling through one connector plug into the MDI optical receptacle as shown in Figure 86–6. The 100GBASE–SR10 PMD is coupled to the fiber optic cabling through one or two connector plugs into the MDI optical receptacle(s), depending on choice of implementation, as shown in Figure 86–7. Example constructions of the MDI include the following:

- a) PMD with a connectorized fiber pigtail plugged into an adapter;
- b) PMD receptacle.

86.10.3.1 Optical lane assignments for 40GBASE-SR4

The four transmit and four receive optical lanes of 40GBASE-SR4 shall occupy the positions depicted in Figure 86–6 when looking into the MDI receptacle with the connector keyway feature on top. The interface contains eight active lanes within twelve total positions. The transmit optical lanes occupy the leftmost four positions. The receive optical lanes occupy the rightmost four positions. The four center positions are unused.

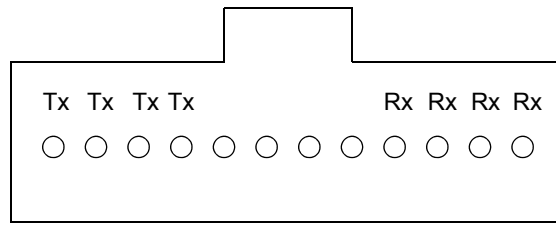


Figure 86–6—40GBASE-SR4 optical lane assignments

86.10.3.2 Optical lane assignments for 100GBASE-SR10

The ten transmit and ten receive optical lanes of 100GBASE-SR10 shall occupy the positions depicted in Figure 86–7 when looking into the MDI optical receptacle(s) with the connector keyway feature(s) on top. The single-receptacle Option A is recommended, the two-receptacle Option B and Option C are alternatives. The interface contains 20 active lanes within up to 24 total positions arranged in two rows of 10 or 12 positions. One row is dedicated to transmit optical lanes and the other row to receive optical lanes. For the depicted 12-position rows, the optical lanes occupy the center ten positions of each row with the outermost positions unused.

86.10.3.3 Medium Dependent Interface (MDI) requirements

The MDI adapter or receptacle shall meet the dimensional specifications of IEC 61754-7 interface 7-3, the MPO adapter interface. The plug terminating the optical fiber cabling shall meet the dimensional specifications of IEC 61754-7 interface 7-4, MPO female plug connector flat interface. The MDI shall optically mate with the plug on the optical fiber cabling. Figure 86–8 shows an MPO female plug connector with flat interface, and an MDI as a PMD receptacle using an MPO adapter interface.

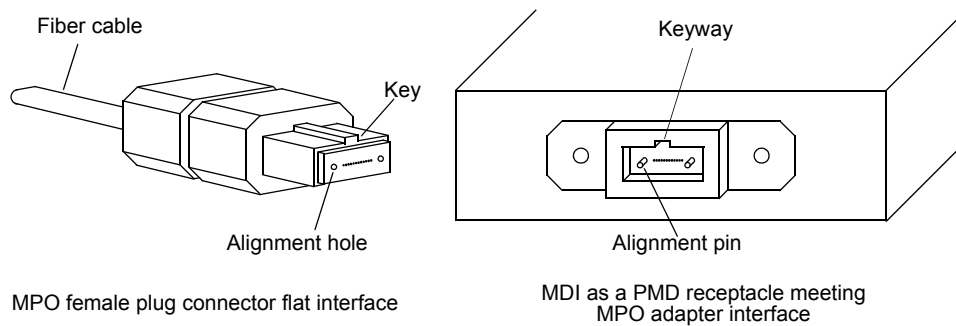


Figure 86–8—MPO female plug and MDI as a PMD receptacle using MPO adapter

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

NOTE—Transmitter compliance testing is performed at TP2 as defined in 86.5.1, not at the MDI.

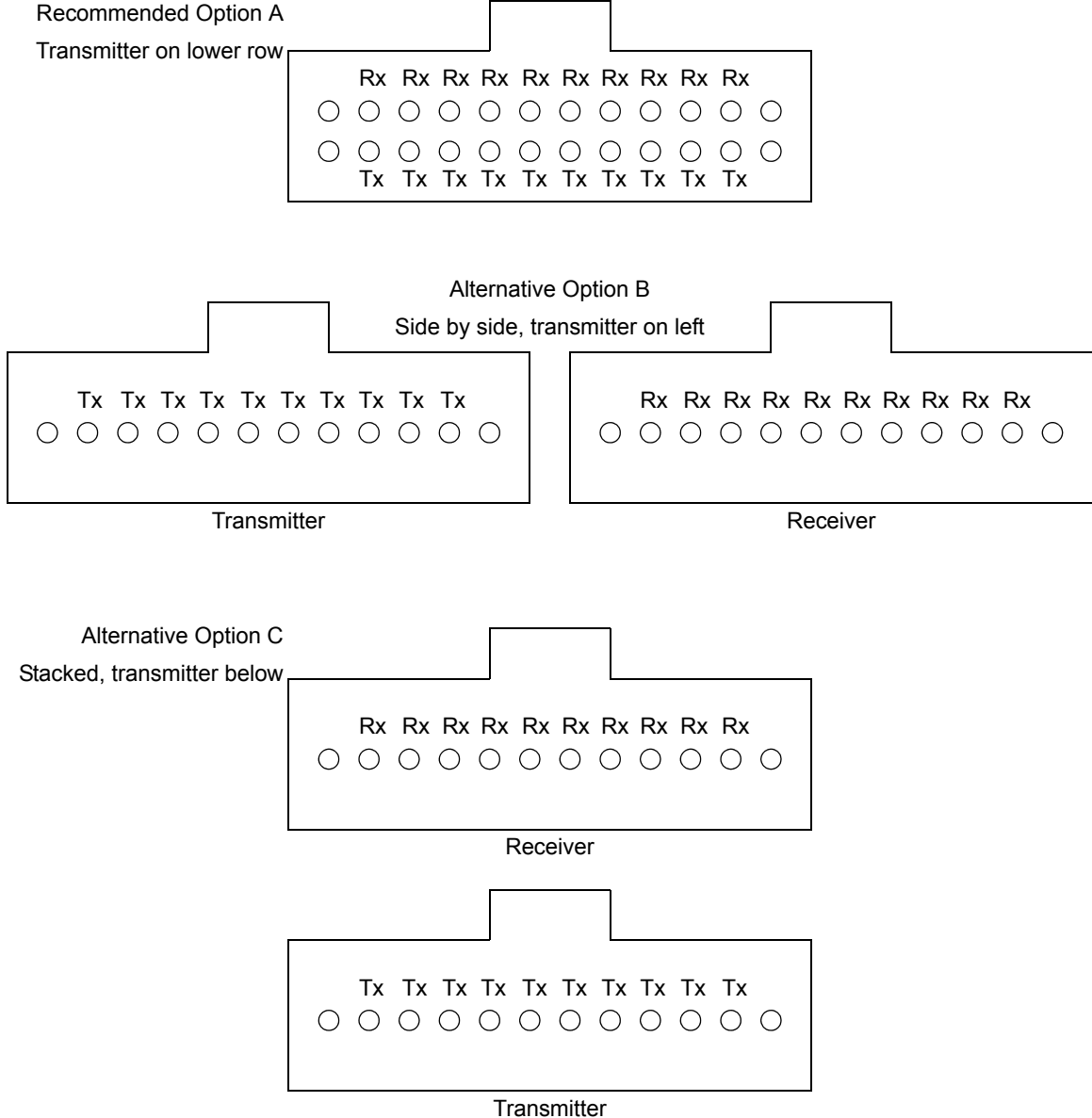


Figure 86-7—100GBASE-SR10 optical lane assignments

86.11 Protocol implementation conformance statement (PICS) proforma for Clause 86, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE–SR4 and 100GBASE–SR10¹²

86.11.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 86, Physical Medium Dependent sublayer and medium, type 40GBASE–SR4 and 100GBASE–SR10, 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.

86.11.2 Identification

86.11.2.1 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).	

86.11.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 86, Physical Medium Dependent sublayer and medium, type 40GBASE–SR4 and 100GBASE–SR10
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 802.3-2012.)	

Date of Statement	
-------------------	--

¹²Copyright 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.

86.11.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*SR4	40GBASE-SR4 PMD	86.7	Can operate as 40GBASE-SR4 PMD	O.1	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*SR10	100GBASE-SR10 PMD	86.7	Can operate as 100GBASE-SR10 PMD	O.1	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*INS	Installation / cable	86.10.1	Items marked with INS include installation practices and cable specifications not applicable to a PHY manufacturer	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*TP1	Compliance point TP1 exposed and available for testing	86.5.1	See text	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*TP4	Compliance point TP4 exposed and available for testing	86.5.1	See text	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*PIT	nPPI Tx interface	86.1	Uses XLPPi or CPPI host to module (see Annex 86A)	TP1:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*PIR	nPPI Rx interface	86.1	Uses XLPPi or CPPI module to host (see Annex 86A)	TP4:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>
*MD	MDIO capability	86.4	Registers and interface supported	O	Yes [<input type="checkbox"/> No [<input type="checkbox"/>

86.11.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, types 40GBASE-SR4 and 100GBASE-SR10**86.11.4.1 PMD functional specifications**

Item	Feature	Subclause	Value/Comment	Status	Support
SF1	Compatible with 40GBASE-R or 100GBASE-R PCS and PMA	86.1	See text	M	Yes []
SF2	Integration with management functions	86.1	See text	O	Yes [] No []
D	Delay constraints	86.3.1	For SR4, max 25.6 ns. For SR10, max 20.48 ns.	M	Yes []
SF3	Skew and Skew Variation at SP3 and SP4	86.3.2	At SP3, less than 54 ns, 600 ps. At SP4, less than 134 ns, 3.4 ns.	M	Yes []
SF4	Skew and Skew Variation at SP5 (TP4)	86.3.2	If measurable, less than 145 ns, 3.6 ns.	TP4: M	Yes [] N/A []
SF5	Transmit function	86.5.2	Conveys bits from PMD service interface to MDI	M	Yes []
SF6	Delivery to the MDI	86.5.2	4 or 10 optical signal streams for delivery to the MDI	M	Yes []
SF7	Mapping between optical signal and logical signal for transmitter	86.5.2	Higher optical power is a one	M	Yes []
SF8	Receive function	86.5.3	Conveys bits from MDI to PMD service interface	M	Yes []
SF9	Conversion of optical signals to electrical signals	86.5.3	For delivery to the PMD service interface	M	Yes []
SF10	Mapping between optical signal and logical signal for receiver	86.5.3	Higher optical power is a one	M	Yes []
SF11	Global Signal Detect function	86.5.4	Report to the PMD service interface the message PMD:IS_SIGNAL.indication(SIGNAL_DETECT)	M	Yes []
SF12	Global Signal Detect behavior	86.5.4	Global indicator of the presence of optical signals on all lanes per Table 86-5	M	Yes []
SF13	Lane-by-lane Signal Detect function	86.5.5	Sets PMD_signal_detect _i values on a lane-by-lane basis per Table 86-5	MD:O	Yes [] No [] N/A []
SF14	PMD_reset function	86.5.6	Resets the PMD sublayer	MD:O	Yes [] No [] N/A []

86.11.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
SM1	Management register set	86.4	See 86.4	MD:M	Yes [] N/A []
SM2	Global transmit disable function	86.5.7	Disables all of the optical transmitters with Global_PMD_transmit_disable	MD:O	Yes [] No [] N/A []
SM3	PMD_transmit_disable_ <i>i</i> function	86.5.8	Disables the optical transmitter on the lane associated with PMD_transmit_disable_ <i>i</i>	MD:O	Yes [] No [] N/A []
SM4	PMD lane-by-lane transmit disable	86.5.8	Disables each optical transmitter independently if MD = No	!MD:O	Yes [] No [] N/A []
SM5	PMD_fault function	86.5.9	Sets PMD_fault to one if any local fault is detected	O	Yes [] No []
SM6	PMD_fault mapping to MDIO	86.5.9	Mapped to PMA/PMD fault bit	SM5*MD :M	Yes [] N/A []
SM7	PMD_transmit_fault function	86.5.10	Sets PMD_transmit_fault to one if a local fault is detected on any transmit lane	O	Yes [] No []
SM8	PMD_transmit_fault mapping to MDIO	86.5.10	Mapped to PMA/PMD transmit fault bit	SM7*MD :M	Yes [] N/A []
SM9	PMD_receive_fault function	86.5.11	Sets PMD_receive_fault to one if a local fault is detected on any receive lane	O	Yes [] No []
SM10	PMD_receive_fault mapping to MDIO	86.5.11	Mapped to PMA/PMD receive fault bit	SM9*MD :M	Yes [] N/A []

86.11.4.3 Optical specifications for 40GBASE-SR4 or 100GBASE-SR10

Item	Feature	Subclause	Value/Comment	Status	Support
S1	Signaling rate per Table 86-2	86.7	10.3125 ±100 ppm	M	Yes [] N/A []
S2	Optical transmitter meets specifications in Table 86-6	86.7.1	Per definitions in 86.8	M	Yes [] N/A []
S3	Optical receiver meets specifications in Table 86-8	86.7.3	Per definitions in 86.8	M	Yes [] N/A []

86.11.4.4 Definitions of parameters and measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
SOM1	Measurement cable	86.8.1	2 m to 5 m in length	M	Yes []
SOM2	Center wavelength	86.8.4.1	Per TIA-455-127-A	M	Yes []
SOM3	Average optical power	86.8.4.2	Per IEC 61280-1-1	M	Yes []
SOM4	OMA measurements	86.8.4.3	Each lane, per methods of 52.9.5 or 68.6.2	M	Yes []
SOM5	TDP	86.8.4.4	Each lane	M	Yes []
SOM6	Extinction ratio	86.8.4.5	Per IEC 61280-2-2	M	Yes []
SOM7	Transmit eye	86.8.4.6	Per 86.8.4.6.1 and 86.8.3.2, hit ratio lower than specified limit	M	Yes []
SOM8	Stressed receiver conformance	86.8.4.7	See 86.8.4.7	M	Yes []
SOM9	Receiver jitter tolerance	86.8.4.8	Per 68.6.11 as modified	M	Yes []

86.11.4.5 Environmental and safety specifications

Item	Feature	Subclause	Value/Comment	Status	Support
SES1	General safety	86.9.1	Conforms to IEC 60950-1	M	Yes []
SES2	Laser safety	86.9.2	Conforms to Hazard Level 1M laser requirements defined in IEC 60825-1 and IEC 60825-2	M	Yes []
SES3	Electromagnetic interference	86.9.4	As 52.11. Complies with codes for the limitation of electromagnetic interference	M	Yes []

86.11.4.6 Optical channel and MDI

Item	Feature	Subclause	Value/Comment	Status	Support
SOC1	Fiber optic cabling	86.10.1	Specified in Table 86–13	INS:M	Yes [] N/A []
SOC2	Optical fiber characteristics	86.10.2.1	Per Table 86–14	INS:M	Yes [] N/A []
SOC3	Maximum loss per connection	86.10.2.2.1	0.75 dB	INS:M	Yes [] N/A []
SOC4	Maximum discrete reflectance	86.10.2.2.2	Less than –20 dB	INS:M	Yes [] N/A []
SOC5	40G MDI layout	86.10.3.1	Optical lane assignments per Figure 86–6	SR4:M	Yes [] N/A []
SOC6	100G MDI layout	86.10.3.2	Optical lane assignments per Figure 86–7	SR10:M	Yes [] N/A []
SOC7	MDI dimensions	86.10.3.3	Per IEC 61754-7 interface 7-3	M	Yes []
SOC8	Cabling connector dimensions	86.10.3.3	Per IEC 61754-7 interface 7-4	INS:M	Yes [] N/A []
SOC9	MDI mating	86.10.3.3	MDI optically mates with plug on the cabling	M	Yes []
SOC10	MDI requirements	86.10.3.3	Meets IEC 61753-1-1 and IEC 61753-022-2	M INS:M	Yes [] N/A []

87. Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4

87.1 Overview

This clause specifies the 40GBASE-LR4 PMD together with the single-mode fiber medium. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 87-1, 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.

Table 87-1—Physical Layer clauses associated with the 40GBASE-LR4 PMD

Associated clause	40GBASE-LR4
81—RS	Required
81—XLGMII ^a	Optional
82—PCS for 40GBASE-R	Required
83—PMA for 40GBASE-R	Required
83A—XLAUI	Optional
83B—Chip to module XLAUI	Optional
86A—XLPPi	Optional

^aThe XLGMII is an optional interface. However, if the XLGMII is not implemented, a conforming implementation must behave functionally as though the RS and XLGMII were present.

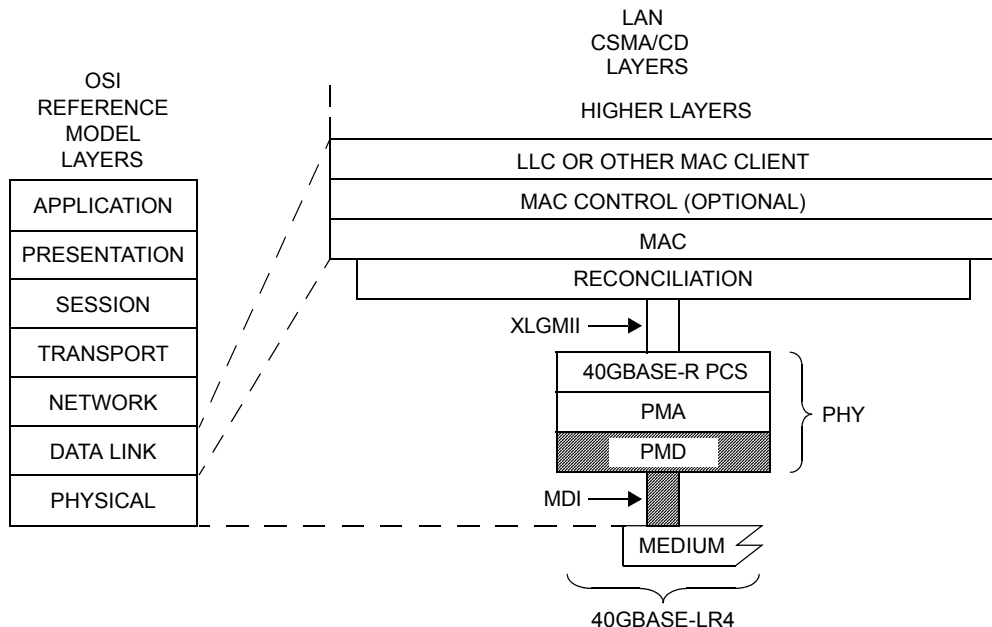
Figure 87-1 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC Open System Interconnection (OSI) reference model. 40 Gb/s and 100 Gb/s Ethernet is introduced in Clause 80 and the purpose of each PHY sublayer is summarized in 80.2.

87.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 40GBASE-LR4 PMD. The service interface for this PMD is described in an abstract manner and does not imply any particular implementation, although an optional implementation of the PMD service interface, the 40 Gb/s Parallel Physical Interface (XLPPi), is specified in Annex 86A. The PMD service interface supports the exchange of encoded data between the PMA entity that resides just above the PMD, and the PMD entity. The PMD translates the encoded data to and from signals suitable for the specified medium.

The PMD service interface 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
```



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
 LR = PMD FOR SINGLE-MODE FIBER — 10 km

Figure 87–1—40GBASE-LR4 PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

The 40GBASE-LR4 PMD has four parallel bit streams, hence $i = 0$ to 3.

In the transmit direction, the PMA continuously sends four parallel bit streams to the PMD, one per lane, each at a nominal signaling rate of 10.3125 GBd. The PMD converts these streams of bits into appropriate signals on the MDI.

In the receive direction, the PMD continuously sends four parallel bit streams to the PMA corresponding to the signals received from the MDI, one per lane, each at a nominal signaling rate of 10.3125 GBd.

The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication(SIGNAL_OK) inter-sublayer service primitive defined in 80.3.

The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the rx_bit parameters are undefined.

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

87.3 Delay and Skew

87.3.1 Delay constraints

The sum of the transmit and the receive delays at one end of the link contributed by the 40GBASE-LR4 PMD including 2 m of fiber in one direction shall be no more than 1024 bit times (2 pause_quanta or 25.6 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.

87.3.2 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 43 ns and the Skew Variation at SP2 is limited to 400 ps.

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

The Skew at SP4 (the receiver MDI) shall be less than 134 ns and the Skew Variation at SP4 shall be less than 3.4 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 145 ns and the Skew Variation at SP5 shall be less than 3.6 ns.

For more information on Skew and Skew Variation see 80.5. The measurements of Skew and Skew Variation are defined in 87.8.2.

87.4 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that may provide control and status information for and about the PMD. If the MDIO interface is implemented, the mapping of MDIO control variables to PMD control variables shall be as shown in Table 87–2 and the mapping of MDIO status variables to PMD status variables shall be as shown in Table 87–3.

87.5 PMD functional specifications

The 40GBASE–LR4 PMD performs the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

87.5.1 PMD block diagram

The PMD block diagram is shown in Figure 87–2. For purposes of system conformance, the PMD sublayer is standardized at the points described in this subclause. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2 m and 5 m in length. Unless specified otherwise, all transmitter measurements and tests defined in 87.8 are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see 87.11.3). Unless specified otherwise, all receiver measurements and tests defined in 87.8 are made at TP3.

Table 87–2—MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/ bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable register	1.9.0	PMD_global_transmit_disable
PMD transmit disable 3	PMD transmit disable register	1.9.4	PMD_transmit_disable_3
PMD transmit disable 2	PMD transmit disable register	1.9.3	PMD_transmit_disable_2
PMD transmit disable 1	PMD transmit disable register	1.9.2	PMD_transmit_disable_1
PMD transmit disable 0	PMD transmit disable register	1.9.1	PMD_transmit_disable_0

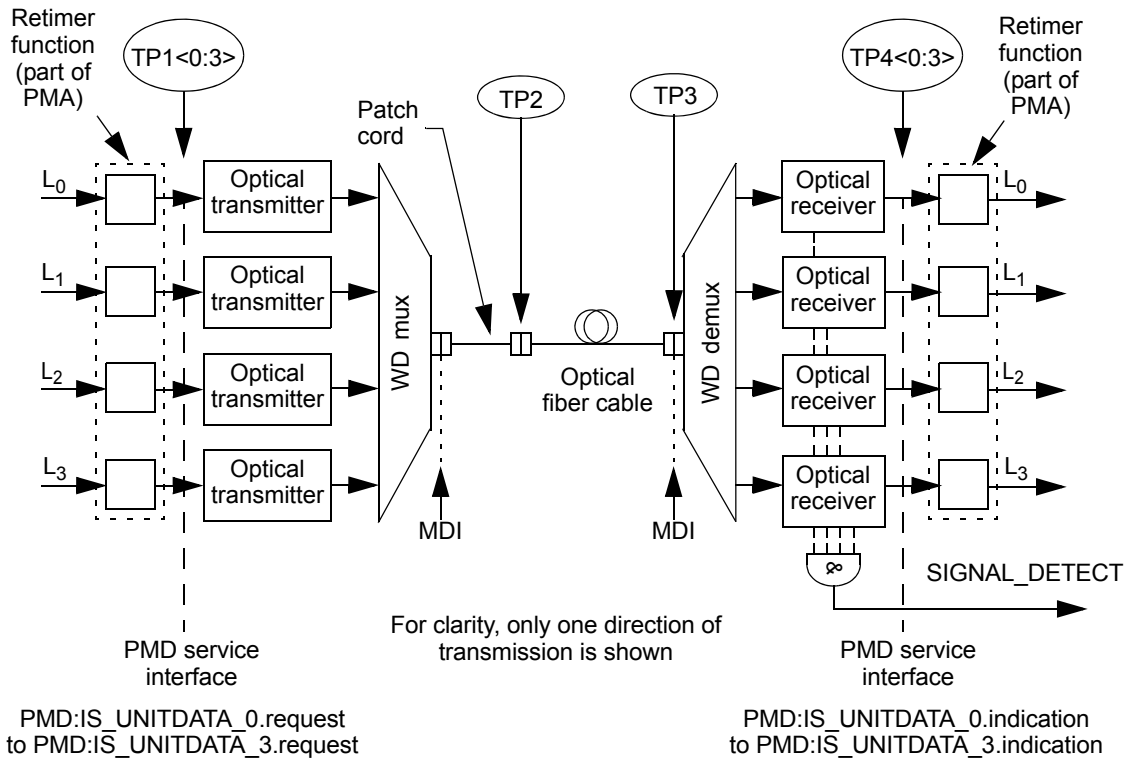
Table 87–3—MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/ bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	PMD_global_signal_detect
PMD receive signal detect 3	PMD receive signal detect register	1.10.4	PMD_signal_detect_3
PMD receive signal detect 2	PMD receive signal detect register	1.10.3	PMD_signal_detect_2
PMD receive signal detect 1	PMD receive signal detect register	1.10.2	PMD_signal_detect_1
PMD receive signal detect 0	PMD receive signal detect register	1.10.1	PMD_signal_detect_0

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

87.5.2 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 optical signal streams. The four optical signal streams shall then be wavelength division multiplexed and delivered to the MDI, all according to the transmit optical specifications in this clause. The higher optical power level in each signal stream shall correspond to tx_bit = one.



WD = Wavelength division

NOTE—Specification of the retimer function is beyond the scope of this standard.

Figure 87–2—Block diagram for 40GBASE-LR4 transmit/receive paths

87.5.3 PMD receive function

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

87.5.4 PMD global signal detect function

The PMD global signal detect function shall report the state of SIGNAL_DETECT via the PMD service interface. The SIGNAL_DETECT parameter is signaled continuously, while the PMD:IS_SIGNAL.indication message is generated when a change in the value of SIGNAL_DETECT occurs. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the inter-sublayer service interface primitives defined in 80.3.

SIGNAL_DETECT shall be a global indicator of the presence of optical signals on all four lanes. The value of the SIGNAL_DETECT parameter shall be generated according to the conditions defined in Table 87–4. The PMD receiver is not required to verify whether a compliant 40GBASE-R signal is being received. This standard imposes no response time requirements on the generation of the SIGNAL_DETECT parameter.

Table 87–4—SIGNAL_DETECT value definition

Receive conditions	SIGNAL_DETECT value
For any lane; Average optical power at TP3 \leq -30 dBm	FAIL
For all lanes; [(Optical power at TP3 \geq receiver sensitivity (max) in OMA in Table 87–8) AND (compliant 40GBASE-R signal input)]	OK
All other conditions	Unspecified

As an unavoidable consequence of the requirements for the setting of the SIGNAL_DETECT parameter, implementations must provide adequate margin between the input optical power level at which the SIGNAL_DETECT parameter is set to OK, and the inherent noise level of the PMD including the effects of crosstalk, power supply noise, etc.

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

87.5.5 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each PMD_signal_detect_ *i*, where *i* represents the lane number in the range 0:3, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of Table 87–4.

87.5.6 PMD reset function

If the MDIO interface is implemented, and if PMD_reset is asserted, the PMD shall be reset as defined in 45.2.1.1.1.

87.5.7 PMD global transmit disable function (optional)

The PMD_global_transmit_disable function is optional and allows all of the optical transmitters to be disabled.

- a) When the PMD_global_transmit_disable variable is set to one, this function shall turn off all of the optical transmitters so that each transmitter meets the requirements of the average launch power of the OFF transmitter in Table 87–7.
- b) If a PMD_fault is detected, then the PMD may set the PMD_global_transmit_disable to one, turning off the optical transmitter in each lane.

87.5.8 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_ *i* (where *i* represents the lane number in the range 0:3) function is optional and allows the optical transmitters in each lane to be selectively disabled.

- a) When a PMD_transmit_disable_ *i* variable is set to one, this function shall turn off the optical transmitter associated with that variable so that the transmitter meets the requirements of the average launch power of the OFF transmitter in Table 87–7.
- b) If a PMD_fault is detected, then the PMD may set each PMD_transmit_disable_ *i* to one, turning off the optical transmitter in each lane.

If the optional `PMD_transmit_disable_i` function is not implemented in MDIO, an alternative method shall be provided to independently disable each transmit lane for testing purposes.

87.5.9 PMD fault function (optional)

If the PMD has detected a local fault on any of the transmit or receive paths, the PMD shall set `PMD_fault` to one. If the MDIO interface is implemented, `PMD_fault` shall be mapped to the fault bit as specified in 45.2.1.2.1.

87.5.10 PMD transmit fault function (optional)

If the PMD has detected a local fault on any transmit lane, the PMD shall set the `PMD_transmit_fault` variable to one. If the MDIO interface is implemented, `PMD_transmit_fault` shall be mapped to the transmit fault bit as specified in 45.2.1.7.4.

87.5.11 PMD receive fault function (optional)

If the PMD has detected a local fault on any receive lane, the PMD shall set the `PMD_receive_fault` variable to one. If the MDIO interface is implemented, `PMD_receive_fault` shall be mapped to the receive fault bit as specified in 45.2.1.7.5.

87.6 Wavelength-division-multiplexed lane assignments

The wavelength range for each lane of the 40GBASE-LR4 PMD is defined in Table 87-5. The center wavelengths are members of the CWDM wavelength grid defined in ITU-T G.694.2 and are spaced at 20 nm.

Table 87-5—Wavelength-division-multiplexed lane assignments

Lane	Center wavelength	Wavelength range
L ₀	1271 nm	1264.5 to 1277.5 nm
L ₁	1291 nm	1284.5 to 1297.5 nm
L ₂	1311 nm	1304.5 to 1317.5 nm
L ₃	1331 nm	1324.5 to 1337.5 nm

NOTE—There is no requirement to associate a particular electrical lane with a particular optical lane, as the PCS is capable of receiving lanes in any arrangement.

87.7 PMD to MDI optical specifications for 40GBASE-LR4

The operating range for the 40GBASE-LR4 PMD is defined in Table 87-6. A 40GBASE-LR4 compliant PMD operates on type B1.1, B1.3 or B6_a single-mode fibers according to the specifications defined in Table 87-14. A PMD that exceeds the required operating range while meeting all other optical specifications is considered compliant (e.g., operating at 12.5 km meets the operating range requirement of 2 m to 10 km).

Table 87–6—40GBASE–LR4 operating range

PMD type	Required operating range
40GBASE–LR4	2 m to 10 km

87.7.1 40GBASE–LR4 transmitter optical specifications

The 40GBASE–LR4 transmitter shall meet the specifications defined in Table 87–7 per the definitions in 87.8.

Table 87–7—40GBASE-LR4 transmit characteristics

Description	Value	Unit
Signaling rate, each lane (range)	10.3125 ± 100 ppm	GBd
Lane wavelengths (range)	1264.5 to 1277.5 1284.5 to 1297.5 1304.5 to 1317.5 1324.5 to 1337.5	nm
Side-mode suppression ratio (SMSR), (min)	30	dB
Total average launch power (max)	8.3	dBm
Average launch power, each lane (max)	2.3	dBm
Average launch power, each lane ^a (min)	–7	dBm
Optical Modulation Amplitude (OMA), each lane (max)	3.5	dBm
Optical Modulation Amplitude (OMA), each lane (min) ^b	–4	dBm
Difference in launch power between any two lanes (OMA) (max)	6.5	dB
Launch power in OMA minus TDP, each lane (min)	–4.8	dBm
Transmitter and dispersion penalty (TDP), each lane (max)	2.6	dB
Average launch power of OFF transmitter, each lane (max)	–30	dBm
Extinction ratio (min)	3.5	dB
RIN ₂₀ OMA (max)	–128	dB/Hz
Optical return loss tolerance (max)	20	dB
Transmitter reflectance ^c (max)	–12	dB
Transmitter eye mask definition {X1, X2, X3, Y1, Y2, Y3}	{0.25, 0.4, 0.45, 0.25, 0.28, 0.4}	

^aAverage launch power, each lane (min) is informative and not the principal indicator of signal strength. A transmitter with launch power below this value cannot be compliant; however, a value above this does not ensure compliance.

^bEven if the TDP < 0.8dB, the OMA (min) must exceed this value.

^cTransmitter reflectance is defined looking into the transmitter.

87.7.2 40GBASE-LR4 receive optical specifications

The 40GBASE-LR4 receiver shall meet the specifications defined in Table 87-8 per the definitions in 87.8.

Table 87-8—40GBASE-LR4 receive characteristics

Description	Value	Unit
Signaling rate, each lane (range)	10.3125 ± 100 ppm	GBd
Lane wavelengths (range)	1264.5 to 1277.5 1284.5 to 1297.5 1304.5 to 1317.5 1324.5 to 1337.5	nm
Damage threshold ^a (min)	3.3	dBm
Average receive power, each lane (max)	2.3	dBm
Average receive power, each lane ^b (min)	-13.7	dBm
Receive power, each lane (OMA) (max)	3.5	dBm
Difference in receive power between any two lanes (OMA) (max)	7.5	dB
Receiver reflectance (max)	-26	dB
Receiver sensitivity (OMA), each lane ^c (max)	-11.5	dBm
Receiver 3 dB electrical upper cutoff frequency, each lane (max)	12.3	GHz
Stressed receiver sensitivity (OMA), each lane ^d (max)	-9.6	dBm
Conditions of stressed receiver sensitivity test:		
Vertical eye closure penalty, ^e each lane	1.9	dB
Stressed eye J2 Jitter, ^e each lane	0.3	UI
Stressed eye J9 Jitter, ^e each lane	0.47	UI

^aThe receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level

^bAverage receive power, each lane (min) is informative and not the principal indicator of signal strength. A received power below this value cannot be compliant; however, a value above this does not ensure compliance.

^cReceiver sensitivity (OMA), each lane (max) is informative.

^dMeasured with conformance test signal at TP3 (see 87.8.11) for BER = 10⁻¹².

^eVertical eye closure penalty, stressed eye J2 Jitter, and stressed eye J9 Jitter are test conditions for measuring stressed receiver sensitivity. They are not characteristics of the receiver.

87.7.3 40GBASE-LR4 illustrative link power budget

An illustrative power budget and penalties for 40GBASE-LR4 channels are shown in Table 87-9.

Table 87–9—40GBASE–LR4 illustrative link power budget

Parameter	Value	Unit
Power budget (for max TDP)	9.3	dB
Operating distance	10	km
Channel insertion loss ^a	6.7	dB
Maximum discrete reflectance	–26	dB
Allocation for penalties ^b (for max TDP)	2.6	dB
Additional insertion loss allowed	0	dB

^aThe channel insertion loss is calculated using the maximum distance specified in Table 87–6 and cabled optical fiber attenuation of 0.47 dB/km at 1264.5 nm plus an allocation for connection and splice loss given in 87.11.2.1.

^bLink penalties are used for link budget calculations. They are not requirements and are not meant to be tested.

87.8 Definition of optical parameters and measurement methods

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

87.8.1 Test patterns for optical parameters

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some parameters. Table 87–11 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subclauses in which each parameter is defined. Any of the test patterns given for a particular test in Table 87–11 may be used to perform that test. The test patterns used in this clause are shown in Table 87–10.

Table 87–10—Test patterns

Pattern	Pattern description	Defined in
Square wave	Square wave (8 ones, 8 zeros)	83.5.10
3	PRBS31	83.5.10
4	PRBS9	83.5.10
5	Scrambled idle	82.2.10

87.8.2 Skew and Skew Variation

Skew and Skew Variation are defined in 80.5 and are required to remain within the limits given in 87.3.2 over the time that the link is in operation. The measurement of Skew and Skew Variation is made by separating optical lanes with an optical demultiplexer and then acquiring the data on each lane using clock and data recovery units with high-frequency corner bandwidths as specified in 86.8.3.2 and a slope of –20 dB/decade. The arrival times of the one to zero transition of the alignment marker sync bits on each lane are

Table 87–11—Test-pattern definitions and related subclauses

Parameter	Pattern	Related subclause
Wavelength	3, 5 or valid 40GBASE-LR signal	87.8.3
Side mode suppression ratio	3, 5 or valid 40GBASE-LR signal	—
Average optical power	3, 5 or valid 40GBASE-LR signal	87.8.4
Optical modulation amplitude (OMA)	Square wave or 4	87.8.5
Transmitter and dispersion penalty (TDP)	3 or 5	87.8.6
Extinction ratio	3, 5 or valid 40GBASE-LR signal	87.8.7
RIN ₂₀ OMA	Square wave or 4	87.8.8
Transmitter optical waveform	3, 5 or valid 40GBASE-LR signal	87.8.9
Stressed receiver sensitivity	3 or 5	87.8.11
Calibration of OMA for receiver tests	Square wave or 4	87.8.11
Vertical eye closure penalty calibration	3 or 5	87.8.11
Receiver 3 dB electrical upper cutoff frequency	3, 5 or valid 40GBASE-LR signal	87.8.12

then compared. This arrangement ensures that any high frequency jitter that is present on the signals is not included in the Skew measurement.

87.8.3 Wavelength

The wavelength of each optical lane shall be within the ranges given in Table 87–5 if measured per TIA/EIA–455–127–A or IEC 61280–1–3. The lane under test is modulated using the test pattern defined in Table 87–11.

87.8.4 Average optical power

The average optical power of each lane shall be within the limits given in Table 87–7 for 40GBASE–LR4 if measured using the methods given in IEC 61280–1–1, with the sum of the optical power from all of the lanes not under test below –30 dBm, per the test setup in Figure 53–6.

87.8.5 Optical Modulation Amplitude (OMA)

OMA shall be as defined in 52.9.5 for measurement with a square wave (8 ones, 8 zeros) test pattern or 68.6.2 (from the variable MeasuredOMA in 68.6.6.2) for measurement with a PRBS9 test pattern, with the exception that each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below –30 dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test.

87.8.6 Transmitter and dispersion penalty

Transmitter and dispersion penalty (TDP) shall be as defined in 52.9.10 with the exception that each optical lane is tested individually using an optical filter to separate the lane under test from the others. The measurement procedure for 40GBASE-LR4 is detailed in 87.8.6.1 to 87.8.6.4.

The optical filter passband ripple shall be limited to 0.5 dB peak-to-peak and the isolation is chosen such that the ratio of the power in the lane being measured to the sum of the powers of all of the other lanes is greater than 20 dB (see ITU-T G.959.1 Annex B). The lanes not under test shall be operating with PRBS31 or valid 40GBASE-LR4 bit streams.

87.8.6.1 Reference transmitter requirements

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

- a) Rise/fall times of less than 25 ps at 20% to 80%.
- b) The output optical eye is symmetric and passes the transmitter optical waveform test of 87.8.9.
- c) In the center 20% region of the eye, the worst-case vertical eye closure penalty as defined in 52.9.9.2 is less than 0.5 dB.
- d) Total Jitter less than 0.2 UI peak-to-peak.
- e) RIN of less than -136 dB/Hz.

87.8.6.2 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in Table 87–12.

Table 87–12—Transmitter compliance channel specifications

PMD type	Dispersion ^a (ps/nm)		Insertion loss ^b	Optical return loss ^c
	Minimum	Maximum		
40GBASE-LR4	$0.2325 \times \lambda \times [1 - (1324 / \lambda)^4]$	$0.2325 \times \lambda \times [1 - (1300 / \lambda)^4]$	Minimum	20 dB

^aThe dispersion is measured for the wavelength of the device under test (λ in nm). The coefficient assumes 10 km for 40GBASE-LR4.

^bThere is no intent to stress the sensitivity of the BERT’s optical receiver.

^cThe optical return loss is applied at TP2.

A 40GBASE-LR4 transmitter is to be compliant with a total dispersion at least as negative as the “minimum dispersion” and at least as positive as the “maximum dispersion” columns specified in Table 87–12 for the wavelength of the device under test. This may be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify that the fiber has the correct amount of dispersion, the measurement method defined in IEC 60793-1-42 may be used. The measurement is made in the linear power regime of the fiber.

The channel provides an optical return loss specified in Table 87–12. The state of polarization of the back reflection is adjusted to create the greatest RIN.

87.8.6.3 Reference receiver requirements

The reference receiver is required to have the bandwidth given in 87.8.9. The sensitivity of the reference receiver is limited by Gaussian noise. The receiver has minimal threshold offset, deadband, hysteresis, baseline wander, deterministic jitter, or other distortions. Decision sampling has minimal uncertainty and setup/hold times.

The nominal sensitivity of the reference receiver, S , is measured in OMA using the setup of Figure 52–12 without the test fiber and with the transversal filter removed. The sensitivity S must be corrected for any significant reference transmitter impairments including any vertical eye closure. It is measured while sampling at the eye center or corrected for off-center sampling. It is calibrated at the wavelength of the transmitter under test.

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

The clock recovery unit (CRU) used in the TDP measurement has a corner frequency of 4 MHz and a slope of 20 dB/decade. When using a clock recovery unit as a clock for BER measurements, passing of low-frequency jitter from the data to the clock removes this low-frequency jitter from the measurement.

87.8.6.4 Test procedure

The test procedure is as defined in 52.9.10.4 with the exception that all lanes are operational in both directions (transmit and receive), each lane is tested individually using an optical filter to separate the lane under test from the others, and the BER of 1×10^{-12} is for the lane under test on its own.

87.8.7 Extinction ratio

The extinction ratio of each lane shall be within the limits given in Table 87–7 for 40GBASE–LR4 if measured using the methods specified in IEC 61280–2–2, with the sum of the optical power from all of the lanes not under test below –30 dBm. The extinction ratio is measured using the test pattern defined in Table 87–11.

NOTE—Extinction ratio and OMA are defined with different test patterns (see Table 87–11).

87.8.8 Relative Intensity Noise (RIN_{20OMA})

RIN shall be as defined by the measurement methodology of 52.9.6 with the following exceptions:

- a) The optical return loss is 20 dB.
- b) Each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below –30 dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test.
- c) The upper –3 dB limit of the measurement apparatus is to be approximately equal to the signaling rate (i.e., 10.3 GHz).

87.8.9 Transmitter optical waveform (transmit eye)

The required optical transmitter pulse shape characteristics are specified in the form of a mask of the transmitter eye diagram as shown in Figure 86–4. The transmitter optical waveform of a port transmitting the test pattern specified in Table 87–11 shall meet specifications according to the methods specified in 86.8.4.6.1 with the filter nominal reference frequency f_r of 7.5 GHz and filter tolerances as specified for

STM-64 in ITU-T G.691. Compensation may be made for variation of the reference receiver filter response from an ideal fourth-order Bessel-Thomson response.

87.8.10 Receiver sensitivity

Receiver sensitivity, which is defined for an ideal input signal, is informative and compliance is not required. If measured, the test signal should have negligible impairments such as intersymbol interference (ISI), rise/fall times, jitter, and RIN. Instead, the normative requirement for receivers is stressed receiver sensitivity.

87.8.11 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in Table 87–8 for 40GBASE–LR4 if measured using the method described in 87.8.11.1 and 87.8.11.5 with the conformance test signal at TP3 as described in 87.8.11.2. The BER is required to be met for the lane under test on its own.

For each lane, the stressed receiver sensitivity is defined with the transmit section in operation on all four lanes and with the receive lanes not under test also in operation. Pattern 3 or Pattern 5, or a valid 40GBASE–R signal, is sent from the transmit section of the PMD under test. The signal being transmitted is asynchronous to the received signal.

87.8.11.1 Stressed receiver conformance test block diagram

A block diagram for the receiver conformance test is shown in Figure 87–3. The patterns used for testing the receiver are specified in Table 87–11. The optical test signal is conditioned (stressed) using the stressed receiver methodology defined in 87.8.11.2, and has sinusoidal jitter applied as specified in 87.8.11.4. A suitable test set is needed to characterize and verify that the signal used to test the receiver has the appropriate characteristics.

The low-pass filter is used to create ISI-induced vertical eye closure penalty (VECP). The low-pass filter, when combined with the E/O converter, should have a frequency response that results in the appropriate level of initial vertical eye closure before the sinusoidal terms are added.

The sinusoidal amplitude interferer 1 causes jitter that is intended to emulate instantaneous bit shrinkage that can occur with DDJ. This type of jitter cannot be created by simple phase modulation. The sinusoidal amplitude interferer 2 causes additional eye closure, but in conjunction with the finite edge rates from the limiter, also causes some jitter. The sinusoidally jittered clock represents other forms of jitter and also verifies that the receiver under test can track low-frequency jitter. The sinusoidal amplitude interferers may be set at any frequency between 100 MHz and 2 GHz, although care should be taken to avoid harmonic relationships between the sinusoidal interferers, the sinusoidal jitter, the signaling rate, and the pattern repetition rate. The Gaussian noise generator, the amplitude of the sinusoidal interferers, and the low-pass filter are adjusted so that the VECP, stressed eye J2 Jitter, and stressed eye J9 Jitter specifications are met simultaneously.

For improved visibility for calibration, all elements in the signal path (cables, DC blocks, E/O converter, etc.) should have wide and smooth frequency response, and linear phase response, throughout the spectrum of interest. Baseline wander and overshoot and undershoot should be minimized.

The stressed receiver conformance signal verification is described in 87.8.11.3.

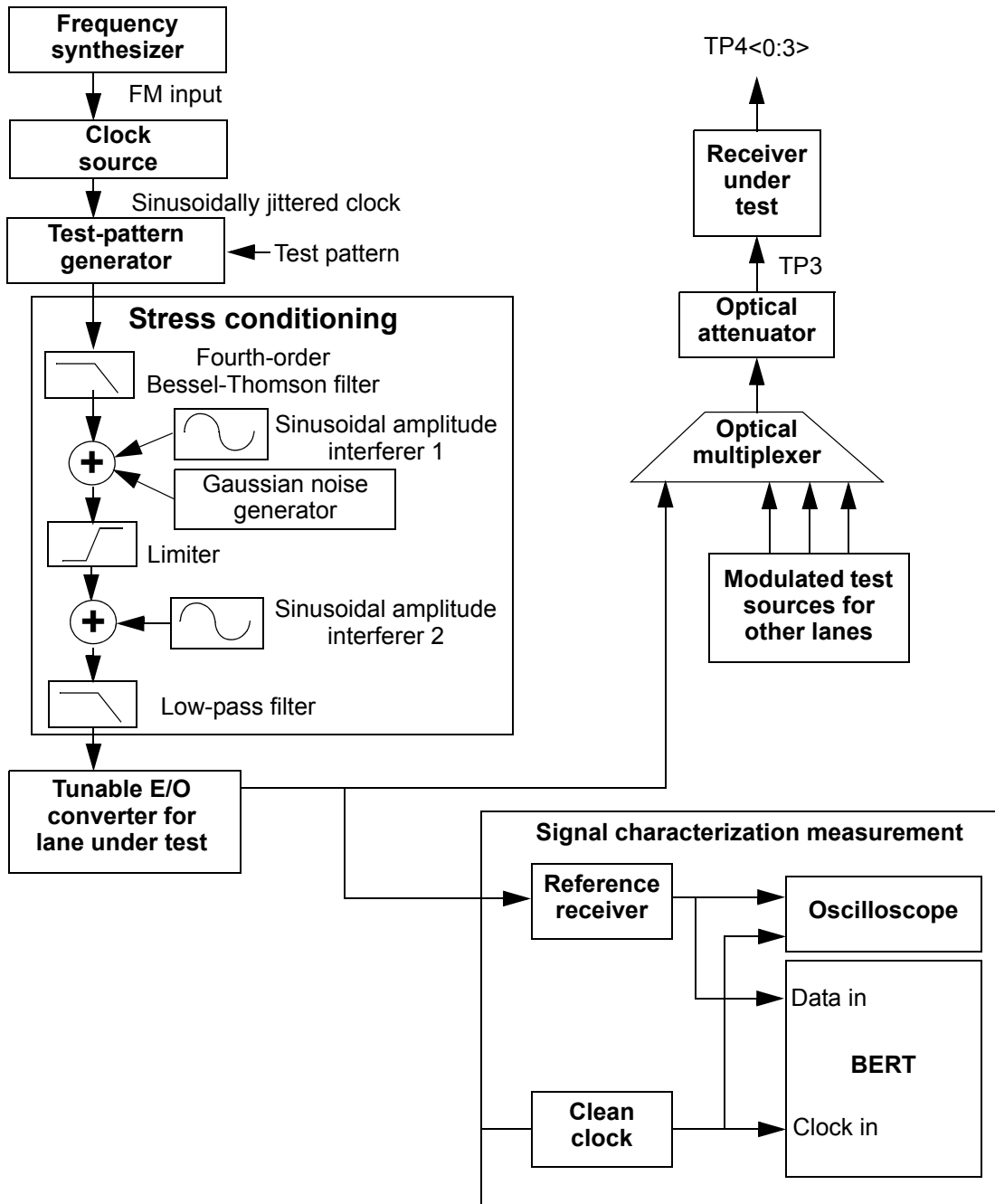


Figure 87-3—Stressed receiver conformance test block diagram

87.8.11.2 Stressed receiver conformance test signal characteristics and calibration

The conformance test signal is used to validate that the PMD receiver of the lane under test meets BER requirements with near worst-case waveforms at TP3.

The primary parameters of the conformance test signal are vertical eye closure penalty (VECP), stressed eye J2 Jitter, and stressed eye J9 Jitter. VECP is measured at the time center of the eye (halfway between 0 and 1

on the unit interval scale as defined in 52.9.7). Stressed eye J2 Jitter is defined in 86.8.3.3.1 and stressed eye J9 Jitter is defined in 86.8.3.3.2. The values of these components are defined by their histogram results, as measured at the average optical power, which can be obtained with AC coupling.

The vertical eye closure penalty is given by Equation (87-1).

$$\text{Vertical eye closure penalty} = 10\log_{10}\frac{OMA}{A_O} \quad (\text{dB}) \quad (87-1)$$

where

A_O is the amplitude of the eye opening from the 99.95th percentile of the lower histogram to the 0.05th percentile of the upper histogram

OMA is the optical modulation amplitude as defined in 87.8.5

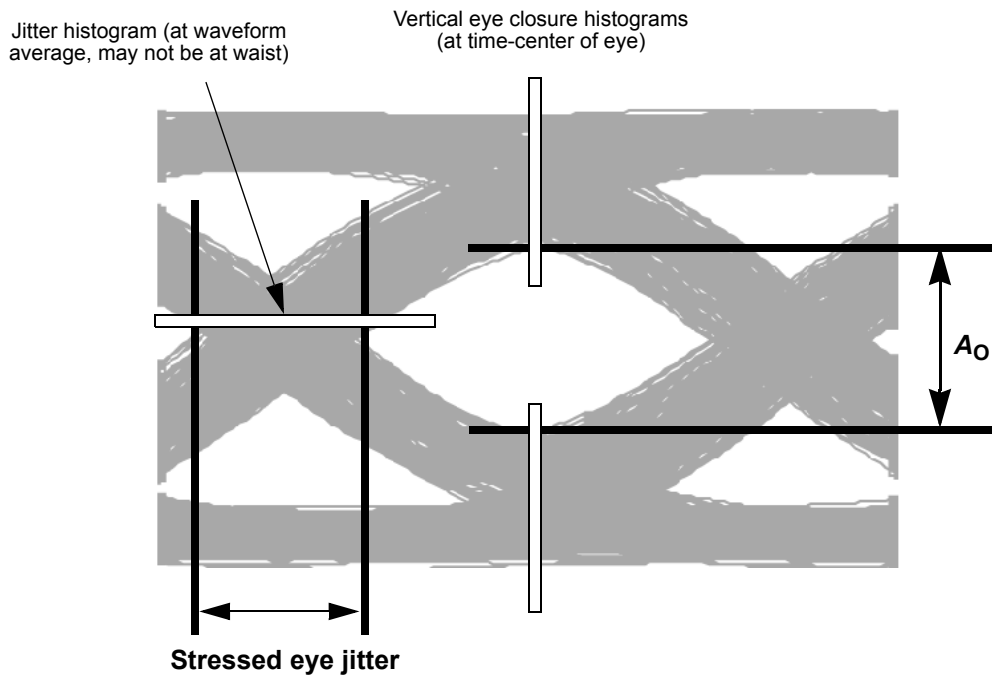


Figure 87-4—Required characteristics of the conformance test signal at TP3

An example stressed receiver conformance test setup is shown in Figure 87-3; however, any approach that modulates or creates the appropriate levels and frequencies of the VECP and jitter components is acceptable.

Residual low-probability noise and jitter should be minimized so that the outer slopes of the final amplitude histograms are as steep as possible.

The following steps describe a possible method for setting up and calibrating a stressed eye conformance signal when using a stressed receiver conformance test setup as shown in Figure 87-3:

- 1) Set the signaling rate of the test-pattern generator to meet the requirements in Table 87-8.
- 2) With sinusoidal interferers and sinusoidal jitter turned off, set the extinction ratio of the E/O to approximately the minimum specified in Table 87-7.

- 3) The required values of vertical eye closure penalty, stressed eye J2 Jitter, and stressed eye J9 Jitter of the stressed receiver conformance signal are given in Table 87–8.

With the sinusoidal interferers and sinusoidal jitter turned off, greater than two thirds of the dB value of the VECP should be created by the selection of the appropriate bandwidth for the low-pass filter. Any remaining VECP must be created with sinusoidal interferer 2 or sinusoidal jitter.

The sinusoidal amplitude interferers may be set at any frequency between 100 MHz and 2 GHz, although care should be taken to avoid harmonic relationships between the sinusoidal interferers, the sinusoidal jitter, the signaling rate and the pattern repetition rate.

Sinusoidal jitter is added as specified in Table 87–13. When calibrating the conformance signal, the sinusoidal jitter frequency should be well within the 4 MHz to 10 times LB as defined in Table 87–13 and illustrated in Figure 87–5.

Iterate the adjustments of sinusoidal interferers and Gaussian noise generator until the values of VECP, stressed eye J2 Jitter and stressed eye J9 Jitter meet the requirements in Table 87–8, and sinusoidal jitter above 4 MHz is as specified in Table 87–13. The resulting stressed eye conformance signal is required to have at least 0.05 UI of pulse width shrinkage.

Figure 87–3 shows the stress conditioned signal being applied to a tunable E/O converter. However, any optical source may be used that can meet the OMA and wavelength requirements for the lane under test as described in 87.8.11.5. Similarly, the other test sources that supply modulated signals to the other lanes may use any tunable or fixed sources that meet the OMA and wavelength requirements described in 87.8.11.5.

Each receiver lane is conformance tested in turn. The source for the lane under test is adjusted to supply a signal at the input to the receiver under test at the stressed receiver sensitivity OMA specified in Table 87–8, and the test sources for the other lanes are set to the required OMA and wavelength as described in 87.8.11.5.

87.8.11.3 Stressed receiver conformance test signal verification

The stressed receiver conformance test signal can be verified using an optical reference receiver with an ideal fourth-order Bessel-Thomson response with a reference frequency f_r of 7.5 GHz. Use of G.691 tolerance filters may significantly degrade this calibration. The clock output from the clock source in Figure 87–3 will be modulated with the sinusoidal jitter. To use an oscilloscope to calibrate the final stressed eye J2 Jitter and stressed eye J9 Jitter that includes the sinusoidal jitter component, a separate clock source (clean clock of Figure 87–3) is required that is synchronized to the source clock, but not modulated with the jitter source.

Care should be taken when characterizing the test signal because excessive noise/jitter in the measurement system will result in an input signal that does not fully stress the receiver under test. Running the receiver tolerance test with a signal that is under-stressed may result in the deployment of non-compliant receivers. Care should be taken to minimize the noise/jitter introduced by the reference O/E, filters, and BERT and/or to correct for this noise. While the details of a BER scan measurement and test equipment are beyond the scope of this standard, it is recommended that the implementor fully characterize the test equipment and apply appropriate guard bands to ensure that the stressed receiver conformance input signal meets the stress and sinusoidal jitter specified in 87.8.11.2 and 87.8.11.4.

87.8.11.4 Sinusoidal jitter for receiver conformance test

The sinusoidal jitter is used to test receiver jitter tolerance. The amplitude of the applied sinusoidal jitter is dependent on frequency as specified in Table 87–13 and is illustrated in Figure 87–5.

Table 87–13—Applied sinusoidal jitter

Frequency range	Sinusoidal jitter, peak-to-peak (UI)
$f < 40 \text{ kHz}$	Not specified
$40 \text{ kHz} < f \leq 4 \text{ MHz}$	$2 \times 10^5 / f$
$4 \text{ MHz} < f < 10 \text{ LB}^a$	0.05

^a LB = loop bandwidth; upper frequency bound for added sine jitter should be at least 10 times the loop bandwidth of the receiver being tested.

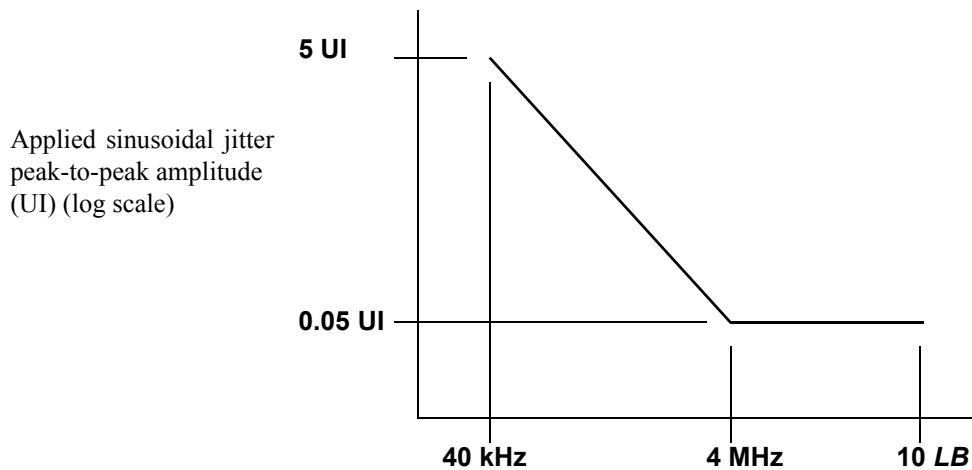


Figure 87–5—Illustration of the mask of the sinusoidal component of jitter tolerance

87.8.11.5 Stressed receiver conformance test procedure for WDM conformance testing

The receiver tests requiring the TP3 conformance test signal are performed on a per lane basis. For each lane, the stressed receiver sensitivity is defined with the transmit section in operation on all four lanes and with the receive lanes not under test also in operation. Pattern 3 or Pattern 5, or a valid 40GBASE–R signal is sent from the transmit section of the PMD under test. The signal being transmitted is asynchronous to the received signal. All test sources are modulated simultaneously, using valid 40GBASE–LR4 signals. The transmitter of the transceiver under test is operating with test patterns as defined in Table 87–11.

A rigorous method for testing WDM conformance using tunable test sources is described as follows:

- a) The OMAs of the test sources in the lanes other than the lane under test are set to the highest OMA relative to the test source in the lane under test allowed by the “difference in receive power between any two lanes (OMA) (max)” parameter in Table 87–8.
- b) The test sources in the lanes other than the lane under test are tuned to wavelengths within their wavelength range corresponding to the worst-case crosstalk to the lane under test.
- c) The test source for the lane under test is tuned to the wavelength within its wavelength range corresponding to the worst-case sensitivity for the receiver under test.

It is recognized that a test setup that uses multiple tunable sources, while allowing rigorous worst-case measurements to be made, is likely to be onerous in practice.

A more practical, alternative method for testing WDM conformance is described below. The alternative WDM conformance test avoids the need for tunable sources but may result in some over-stressing of the receiver under test:

- 1) The test sources in each lane can be at any wavelength within each lane's wavelength range.
- 2) The OMAs of the test sources in the lanes other than the lane under test are set to the highest OMA relative to the test source in the lane under test allowed by the "difference in receive power between any two lanes (OMA) (max)" parameter in Table 87–8, plus an increment corresponding to insertion loss variation within the lane under test (NOTE 1), plus an increment corresponding to the isolation variation of the lane in question (NOTE 2).

NOTE 1—The increment corresponding to the variation of insertion loss with wavelength within the lane under test is equal to the dB variation of measured receiver sensitivity as the test source wavelength is swept across the wavelength range of the lane under test.

NOTE 2—For each of the test sources in the lanes not under test, an increment corresponding to the isolation variation is applied that is equal to the dB variation in optical crosstalk measured in the lane under test as the wavelength of each test source in the lanes not under test is swept across its respective wavelength range.

There are many ways to determine the size of the increments required, two example methods are given in the paragraphs that follow. Note that each lane will have one insertion loss variation value, and three values of optical crosstalk variation (one for each of the other lanes).

Example method 1: If a measure of the received signal strength is available for each lane, scan a tunable laser across the wavelength range of each lane while simultaneously recording the received signal in all lanes. As the laser sweeps across the wavelength range of a particular lane the dB variation in observed signal in that lane is equal to the insertion loss variation. The dB variation in signal observed in each of the other lanes is equal to the optical crosstalk variation.

Example method 2: If a discrete optical demultiplexer and discrete receivers are used, scan a tunable laser across the wavelength range of each channel while simultaneously recording the optical power of all four demultiplexer outputs. As the laser sweeps across the wavelength range of a particular channel, the dB variation in observed optical power in that lane's output is equal to the insertion loss variation. The dB variation in optical power observed in each of the other lanes outputs is equal to the optical crosstalk variation.

87.8.12 Receiver 3 dB electrical upper cutoff frequency

The receiver 3dB electrical upper cutoff frequency shall be within the limits given in Table 87–8 if measured as described in 52.9.11. Each optical lane is measured using an optical signal or signals with its/their wavelength within the specified wavelength range of the lane to be tested. The test may use an electrical combiner and one optical source as in 53.9.13.

87.9 Safety, installation, environment, and labeling

87.9.1 General safety

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

87.9.2 Laser safety

40GBASE-LR4 optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825–1 and IEC 60825–2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service. This documentation explicitly defines requirements and usage restrictions on the host system necessary to meet these safety certifications.¹³

87.9.3 Installation

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

87.9.4 Environment

Normative specifications in this clause shall be met by a system integrating a 40GBASE-LR4 PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate in the literature associated with the PHY the operating environmental conditions to facilitate selection, installation, and maintenance.

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

87.9.4.1 Electromagnetic emission

A system integrating a 40GBASE-LR4 PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

87.9.4.2 Temperature, humidity, and handling

The optical link is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

87.9.5 PMD labeling requirements

It is recommended that each PHY (and supporting documentation) be labeled in a manner visible to the user, with at least the applicable safety warnings and the applicable port type designation (e.g., 40GBASE-LR4).

Labeling requirements for Hazard Level 1 lasers are given in the laser safety standards referenced in 87.9.2.

¹³A host system that fails to meet the manufacturers requirements and/or usage restrictions may emit laser radiation in excess of the safety limits of one or more safety standards. In such a case, the host manufacturer is required to obtain its own laser safety certification.

87.10 Fiber optic cabling model

The fiber optic cabling model is shown in Figure 87–6.

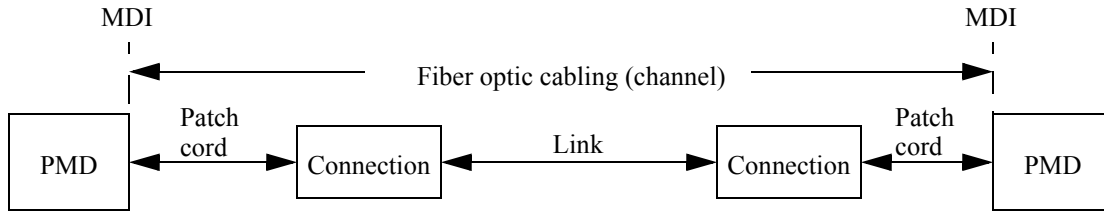


Figure 87–6—Fiber optic cabling model

The channel insertion loss is given in Table 87–14. A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with ANSI/TIA/EIA–526–7/method A–1. The fiber optic cabling model (channel) defined here is the same as a simplex fiber optic link segment. The term channel is used here for consistency with generic cabling standards.

Table 87–14—Fiber optic cabling (channel) characteristics for 40GBASE–LR4

Description	Value	Unit
Operating distance (max)	10	km
Channel insertion loss ^{a, b} (max)	6.7	dB
Channel insertion loss (min)	0	dB
Positive dispersion ^b (max)	33.5	ps/nm
Negative dispersion ^b (min)	–59.5	ps/nm
DGD_max ^c	10	ps
Optical return loss (min)	21	dB

^aThese channel insertion loss values include cable, connectors, and splices.

^bOver the wavelength range 1264.5 nm to 1337.5 nm.

^cDifferential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal. DGD_max is the maximum differential group delay that the system must tolerate.

87.11 Characteristics of the fiber optic cabling (channel)

The 40GBASE–LR4 fiber optic cabling shall meet the specifications defined in Table 87–14. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

87.11.1 Optical fiber cable

The fiber optic cable requirements are satisfied by cables containing IEC 60793–2–50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) fibers or the requirements in Table 87–15 where they differ.

Table 87–15—Optical fiber and cable characteristics for 40GBASE–LR4

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.47 ^a or 0.5 ^b	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm ² km

^aThe 0.47 dB/km at 1264.5 nm attenuation for optical fiber cables is derived from Appendix I of ITU-T G.695.

^bThe 0.5 dB/km attenuation is provided for Outside Plant cable as defined in ANSI/TIA 568-C.3. Using 0.5 dB/km may not support operation at 10 km for 40GBASE–LR4.

87.11.2 Optical fiber connection

An optical fiber connection, as shown in Figure 87–6, consists of a mated pair of optical connectors.

87.11.2.1 Connection insertion loss

The maximum link distances for single-mode fiber are calculated based on an allocation of 2 dB total connection and splice loss. For example, this allocation supports four connections with an average insertion loss per connection of 0.5 dB. Connections with different loss characteristics may be used provided the requirements of Table 87–14 are met.

87.11.2.2 Maximum discrete reflectance

The maximum discrete reflectance shall be less than –26 dB.

87.11.3 Medium Dependent Interface (MDI) requirements

The 40GBASE–LR4 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in Figure 87–6). Examples of an MDI include the following:

- a) Connectorized fiber pigtail
- b) PMD receptacle

When the MDI is a connector plug and receptacle connection, it shall meet the interface performance specifications of IEC 61753–1–1 and IEC 61753–021–2.

NOTE—Transmitter compliance testing is performed at TP2 as defined in 87.5.1, not at the MDI.

87.12 Protocol implementation conformance statement (PICS) proforma for Clause 87, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4¹⁴

87.12.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 87, Physical Medium Dependent sublayer and medium, type 40GBASE-LR4, 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.

87.12.2 Identification

87.12.2.1 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).	

87.12.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 87, Physical Medium Dependent sublayer and medium, type 40GBASE-LR4
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 802.3-2012.)	

Date of Statement	
-------------------	--

¹⁴Copyright 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.

87.12.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
LR4	40GBASE-LR4 PMD	87.7	Device supports requirements for 40GBASE-LR4 PHY	O	Yes [] No []
*INS	Installation/cable	87.10	Items marked with INS include installation practices and cable specifications not applicable to a PHY manufacturer	O	Yes [] No []
XLTP1	Reference point TP1 exposed and available for testing	87.5.1	This point may be made available for use by implementors to certify component conformance	O	Yes [] No []
XLTP4	Reference point TP4 exposed and available for testing	87.5.1	This point may be made available for use by implementors to certify component conformance	O	Yes [] No []
XLDC	Delay constraints	87.3.1	Device conforms to delay constraints	M	Yes []
XLSC	Skew constraints	87.3.2	Device conforms to Skew and Skew Variation constraints	M	Yes []
*MD	MDIO capability	87.4	Registers and interface supported	O	Yes [] No []

87.12.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-LR4**87.12.4.1 PMD functional specifications**

Item	Feature	Subclause	Value/Comment	Status	Support
XLF1	Compatible with 40GBASE-R PCS and PMA	87.1		M	Yes []
XLF2	Integration with management functions	87.1		O	Yes [] No []
XLF3	Transmit function	87.5.2	Conveys bits from PMD service interface to MDI	M	Yes []
XLF4	Optical multiplexing and delivery to the MDI	87.5.2	Optically multiplexes the four optical signal streams for delivery to the MDI	M	Yes []
XLF5	Mapping between optical signal and logical signal for transmitter	87.5.2	Higher optical power is a one	M	Yes []
XLF6	Receive function	87.5.3	Conveys bits from MDI to PMD service interface	M	Yes []
XLF7	Conversion of four optical signals to four electrical signals	87.5.3	For delivery to the PMD service interface	M	Yes []
XLF8	Mapping between optical signal and logical signal for receiver	87.5.3	Higher optical power is a one	M	Yes []
XLF9	Global Signal Detect function	87.5.4	Report to the PMD service interface the message PMD:IS_SIGNAL.indication(SIGNAL_DETECT)	M	Yes []
XLF10	Global Signal Detect behavior	87.5.4	SIGNAL_DETECT is a global indicator of the presence of optical signals on all four lanes	M	Yes []
XLF11	Lane-by-Lane Signal Detect function	87.5.5	Sets PMD_signal_detect _i values on a lane-by-lane basis per requirements of Table 87-4	MD:O	Yes [] No [] N/A []
XLF12	PMD_reset function	87.5.6	Resets the PMD sublayer	MD:O	Yes [] No [] N/A []

87.12.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
XLM1	Management register set	87.4		MD:M	Yes [] N/A []
XLM2	Global transmit disable function	87.5.7	Disables all of the optical transmitters with the PMD_global_transmit_disable variable	MD:O	Yes [] No [] N/A []
XLM3	PMD_lane_by_lane_transmit_disable function	87.5.8	Disables the optical transmitter on the lane associated with the PMD_transmit_disable_i variable	MD:O.2	Yes [] No [] N/A []
XLM4	PMD_lane_by_lane_transmit_disable	87.5.8	Disables each optical transmitter independently if XLM3 = No	O.2	Yes [] No []
XLM5	PMD_fault function	87.5.9	Sets PMD_fault to one if any local fault is detected	MD:O	Yes [] No [] N/A []
XLM6	PMD_transmit_fault function	87.5.10	Sets PMD_transmit_fault to one if a local fault is detected on any transmit lane	MD:O	Yes [] No [] N/A []
XLM7	PMD_receive_fault function	87.5.11	Sets PMD_receive_fault to one if a local fault is detected on any receive lane	MD:O	Yes [] No [] N/A []

87.12.4.3 PMD to MDI optical specifications for 40GBASE-LR4

Item	Feature	Subclause	Value/Comment	Status	Support
XLLR1	Transmitter meets specifications in Table 87-7	87.7.1	Per definitions in 87.8	M	Yes [] N/A []
XLLR2	Receiver meets specifications in Table 87-8	87.7.2	Per definitions in 87.8	M	Yes [] N/A []

87.12.4.4 Optical measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
XLOM1	Measurement cable	87.8	2 m to 5 m in length	M	Yes []
XLOM2	Center wavelength	87.8.3	Per TIA-455-127-A or IEC 61280-1-3 under modulated conditions	M	Yes []
XLOM3	Average optical power	87.8.4	Per IEC 61280-1-1	M	Yes []
XLOM4	Transmitter and dispersion penalty	87.8.6	Each lane	M	Yes []
XLOM5	OMA	87.8.5	Each lane	M	Yes []
XLOM6	Extinction ratio	87.8.7	Per IEC 61280-2-2	M	Yes []
XLOM7	RIN ₂₀ OMA	87.8.8	Each lane	M	Yes []
XLOM8	Transmit eye	87.8.9	Each lane	M	Yes []
XLOM9	Stressed receiver sensitivity	87.8.11	Each lane	M	Yes []
XLOM10	Receiver 3 dB electrical upper cutoff frequency	87.8.12	Each lane	M	Yes []

87.12.4.5 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
XLES1	General safety	87.9.1	Conforms to IEC 60950-1	M	Yes []
XLES2	Laser safety—IEC Hazard Level 1	87.9.2	Conforms to Hazard Level 1 laser requirements defined in IEC 60825-1 and IEC 60825-2	M	Yes []
XLES3	Electromagnetic interference	87.9.4.1	Complies with applicable local and national codes for the limitation of electromagnetic interference	M	Yes []

87.12.4.6 Characteristics of the fiber optic cabling and MDI

Item	Feature	Subclause	Value/Comment	Status	Support
XLOC1	Fiber optic cabling	87.11	Meets the requirements specified in Table 87-14	INS:M	Yes [] N/A []
XLOC2	Maximum discrete reflectance	87.11.2.2	Less than -26 dB	INS:M	Yes [] N/A []
XLOC3	MDI requirements	87.11.3	Meets the interface performance specifications listed	INS:M	Yes [] N/A []

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

88.1 Overview

This clause specifies the 100GBASE-LR4 PMD and the 100GBASE-ER4 PMD together with the single-mode fiber medium. When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 88-1, 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.

Table 88-1—Physical Layer clauses associated with the 100GBASE-LR4 and 100GBASE-ER4 PMDs

Associated clause	100GBASE-LR4	100GBASE-ER4
81—RS	Required	Required
81—CGMII ^a	Optional	Optional
82—PCS for 100GBASE-R	Required	Required
83—PMA for 100GBASE-R	Required	Required
83A—CAUI	Optional	Optional
83B—Chip to module CAUI	Optional	Optional

^aThe 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 88-1 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC Open System Interconnection (OSI) reference model. 40 Gb/s and 100 Gb/s Ethernet is introduced in Clause 80 and the purpose of each PHY sublayer is summarized in 80.2.

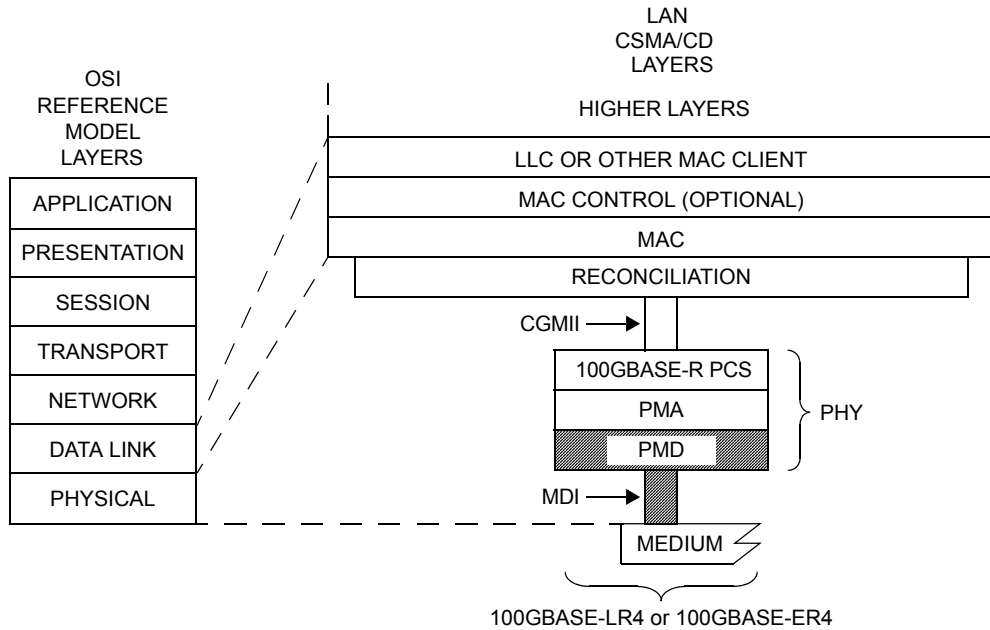
88.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 100GBASE-LR4 and 100GBASE-ER4 PMDs. The service interfaces for these PMDs are described in an abstract manner and do not imply any particular implementation. The PMD service interface supports the exchange of encoded data between the PMA entity that resides just above the PMD, and the PMD entity. The PMD translates the encoded data to and from signals suitable for the specified medium.

The PMD service interface 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-LR4 and 100GBASE-ER4 PMDs have four parallel bit streams, hence $i = 0$ to 3.



CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE PMA = PHYSICAL MEDIUM ATTACHMENT
 LLC = LOGICAL LINK CONTROL PMD = PHYSICAL MEDIUM DEPENDENT
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIUM DEPENDENT INTERFACE LR = PMD FOR SINGLE-MODE FIBER — 10 km
 PCS = PHYSICAL CODING SUBLAYER ER = PMD FOR SINGLE-MODE FIBER — 40 km
 PHY = PHYSICAL LAYER DEVICE

Figure 88–1—100GBASE-LR4 and 100GBASE-ER4 PMDs relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

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

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

The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication(SIGNAL_OK) inter-sublayer service interface primitive defined in 80.3.

The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the rx_bit parameters are undefined.

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

88.3 Delay and Skew

88.3.1 Delay constraints

The sum of the transmit and receive delays at one end of the link contributed by the 100GBASE-LR4 or 100GBASE-ER4 PMD including 2 m of fiber in one direction shall be no more than 2048 bit times (4 pause_quanta or 20.48 ns). A description of overall system delay constraints and the definitions for bit times and pause_quanta can be found in 80.4 and its references.

88.3.2 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 43 ns and the Skew Variation at SP2 is limited to 400 ps.

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

The Skew at SP4 (the receiver MDI) shall be less than 134 ns and the Skew Variation at SP4 shall be less than 3.4 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 145 ns and the Skew Variation at SP5 shall be less than 3.6 ns.

For more information on Skew and Skew Variation see 80.5. The measurements of Skew and Skew Variation are defined in 87.8.2 with the exception that the clock and data recovery units' high-frequency corner bandwidths are 10 MHz.

88.4 PMD MDIO function mapping

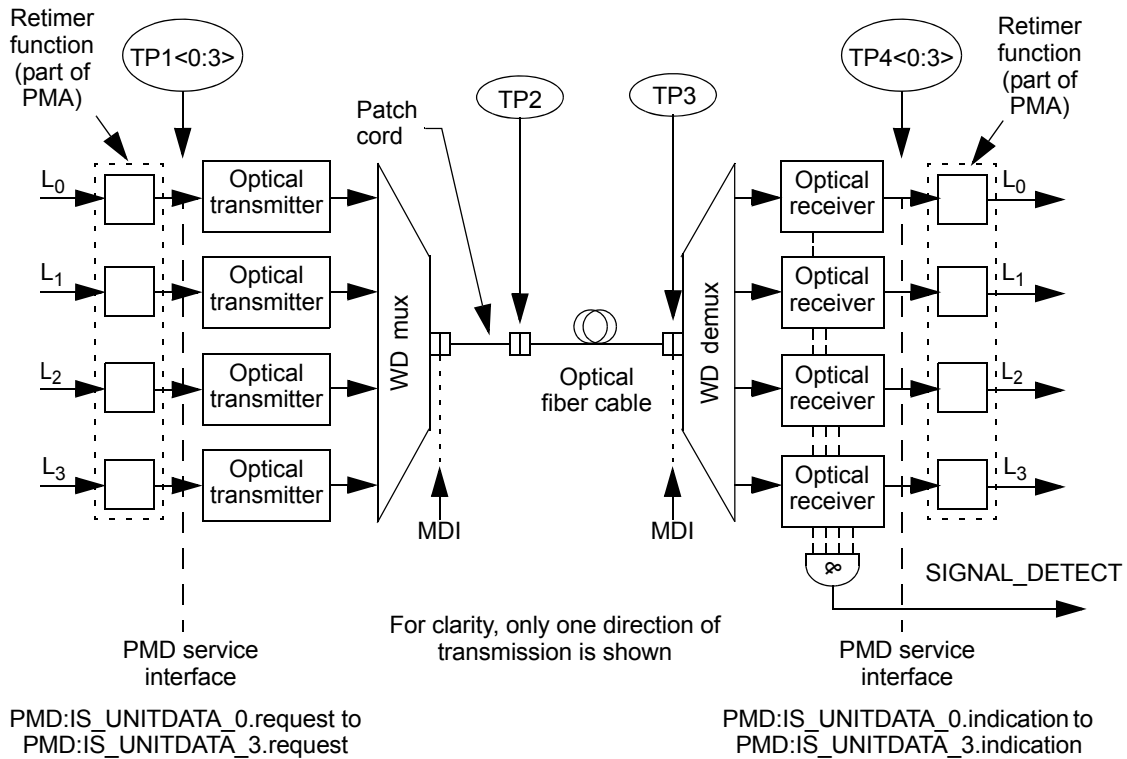
The optional MDIO capability described in Clause 45 defines several variables that may provide control and status information for and about the PMD. If the MDIO interface is implemented, the mapping of MDIO control variables to PMD control variables shall be as shown in Table 88-2 and the mapping of MDIO status variables to PMD status variables shall be as shown in Table 88-3.

88.5 PMD functional specifications

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

88.5.1 PMD block diagram

The PMD block diagram is shown in Figure 88-2. For purposes of system conformance, the PMD sublayer is standardized at the points described in this subclause. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2 m and 5 m in length. Unless specified otherwise, all transmitter measurements and tests defined in 88.8 are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see 88.11.3). Unless specified otherwise, all receiver measurements and tests defined in 88.8 are made at TP3.



WD = Wavelength division

NOTE—Specification of the retimer function and the electrical implementation of the PMD service interface is beyond the scope of this standard.

Figure 88–2—Block diagram for 100GBASE-LR4 and 100GBASE-ER4 transmit/receive paths

Table 88–2—MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD transmit disable register	1.9.0	PMD_global_transmit_disable
PMD transmit disable 3	PMD transmit disable register	1.9.4	PMD_transmit_disable_3
PMD transmit disable 2	PMD transmit disable register	1.9.3	PMD_transmit_disable_2
PMD transmit disable 1	PMD transmit disable register	1.9.2	PMD_transmit_disable_1
PMD transmit disable 0	PMD transmit disable register	1.9.1	PMD_transmit_disable_0

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

Table 88–3—MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/ bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	PMD_global_signal_detect
PMD receive signal detect 3	PMD receive signal detect register	1.10.4	PMD_signal_detect_3
PMD receive signal detect 2	PMD receive signal detect register	1.10.3	PMD_signal_detect_2
PMD receive signal detect 1	PMD receive signal detect register	1.10.2	PMD_signal_detect_1
PMD receive signal detect 0	PMD receive signal detect register	1.10.1	PMD_signal_detect_0

88.5.2 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 optical signal streams. The four optical signal streams shall then be wavelength division multiplexed and delivered to the MDI, all according to the transmit optical specifications in this clause. The higher optical power level in each signal stream shall correspond to tx_bit = one.

88.5.3 PMD receive function

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

88.5.4 PMD global signal detect function

The PMD global signal detect function shall report the state of SIGNAL_DETECT via the PMD service interface. The SIGNAL_DETECT parameter is signaled continuously, while the PMD:IS_SIGNAL.indication message is generated when a change in the value of SIGNAL_DETECT occurs. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the inter-sublayer service interface primitives defined in 80.3.

SIGNAL_DETECT shall be a global indicator of the presence of optical signals on all four lanes. The value of the SIGNAL_DETECT parameter shall be generated according to the conditions defined in Table 88–4. The PMD receiver is not required to verify whether a compliant 100GBASE-R signal is being received. This standard imposes no response time requirements on the generation of the SIGNAL_DETECT parameter.

As an unavoidable consequence of the requirements for the setting of the SIGNAL_DETECT parameter, implementations must provide adequate margin between the input optical power level at which the SIGNAL_DETECT parameter is set to OK, and the inherent noise level of the PMD including the effects of crosstalk, power supply noise, etc.

Table 88–4—SIGNAL_DETECT value definition

Receive conditions	SIGNAL_DETECT value
For any lane; Average optical power at TP3 \leq -30 dBm	FAIL
For all lanes; [(Optical power at TP3 \geq receiver sensitivity (max) in OMA in Table 88–8) AND (compliant 100GBASE-R signal input)]	OK
All other conditions	Unspecified

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

88.5.5 PMD lane-by-lane signal detect function

Various implementations of the Signal Detect function are permitted by this standard. When the MDIO is implemented, each PMD_signal_detect_{*i*}, where *i* represents the lane number in the range 0:3, shall be continuously set in response to the magnitude of the optical signal on its associated lane, according to the requirements of Table 88–4.

88.5.6 PMD reset function

If the MDIO interface is implemented, and if PMD_reset is asserted, the PMD shall be reset as defined in 45.2.1.1.1.

88.5.7 PMD global transmit disable function (optional)

The PMD_global_transmit_disable function is optional and allows all of the optical transmitters to be disabled.

- a) When the PMD_global_transmit_disable variable is set to one, this function shall turn off all of the optical transmitters so that each transmitter meets the requirements of the average launch power of the OFF transmitter in Table 88–7.
- b) If a PMD_fault is detected, then the PMD may set the PMD_global_transmit_disable to one, turning off the optical transmitter in each lane.

88.5.8 PMD lane-by-lane transmit disable function

The PMD_transmit_disable_{*i*} (where *i* represents the lane number in the range 0:3) function is optional and allows the optical transmitters in each lane to be selectively disabled.

- a) When a PMD_transmit_disable_{*i*} variable is set to one, this function shall turn off the optical transmitter associated with that variable so that the transmitter meets the requirements of the average launch power of the OFF transmitter in Table 88–7.
- b) If a PMD_fault is detected, then the PMD may set each PMD_transmit_disable_{*i*} to one, turning off the optical transmitter in each lane.

If the optional `PMD_transmit_disable_i` function is not implemented in MDIO, an alternative method shall be provided to independently disable each transmit lane for testing purposes.

88.5.9 PMD fault function (optional)

If the PMD has detected a local fault on any of the transmit or receive paths, the PMD shall set `PMD_fault` to one.

If the MDIO interface is implemented, `PMD_fault` shall be mapped to the fault bit as specified in 45.2.1.2.1.

88.5.10 PMD transmit fault function (optional)

If the PMD has detected a local fault on any transmit lane, the PMD shall set the `PMD_transmit_fault` variable to one.

If the MDIO interface is implemented, `PMD_transmit_fault` shall be mapped to the transmit fault bit as specified in 45.2.1.7.4.

88.5.11 PMD receive fault function (optional)

If the PMD has detected a local fault on any receive lane, the PMD shall set the `PMD_receive_fault` variable to one.

If the MDIO interface is implemented, `PMD_receive_fault` shall be mapped to the receive fault bit as specified in 45.2.1.7.5.

88.6 Wavelength-division-multiplexed lane assignments

The wavelength range for each lane of the 100GBASE-LR4 and 100GBASE-ER4 PMDs is defined in Table 88-5. The center frequencies are members of the frequency grid for 100 GHz spacing and above defined in ITU-T G.694.1 and are spaced at 800 GHz.

Table 88-5—Wavelength-division-multiplexed lane assignments

Lane	Center frequency	Center wavelength	Wavelength range
L ₀	231.4 THz	1295.56 nm	1294.53 to 1296.59 nm
L ₁	230.6 THz	1300.05 nm	1299.02 to 1301.09 nm
L ₂	229.8 THz	1304.58 nm	1303.54 to 1305.63 nm
L ₃	229 THz	1309.14 nm	1308.09 to 1310.19 nm

NOTE—There is no requirement to associate a particular electrical lane with a particular optical lane, as the PCS is capable of receiving lanes in any arrangement.

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

The operating ranges for the 100GBASE-LR4 and 100GBASE-ER4 PMDs are defined in Table 88-6. A 100GBASE-LR4 or 100GBASE-ER4 compliant PMD operates on type B1.1, B1.3, or B6_a single-mode fibers according to the specifications defined in Table 88-14. A PMD that exceeds the operating range

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

Table 88-6—100GBASE-LR4 and 100GBASE-ER4 operating ranges

PMD type	Required operating range
100GBASE-LR4	2 m to 10 km
100GBASE-ER4	2 m to 30 km
	2 m to 40 km ^a

^aLinks longer than 30 km for the same link power budget are considered engineered links. Attenuation for such links needs to be less than the worst case specified for B1.1, B1.3, or B6_a single-mode fiber.

88.7.1 100GBASE-LR4 and 100GBASE-ER4 transmitter optical specifications

The 100GBASE-LR4 transmitter shall meet the specifications defined in Table 88-7 per the definitions in 88.8. The 100GBASE-ER4 transmitter shall meet the specifications defined in Table 88-7 per the definitions in 88.8.

Table 88–7—100GBASE-LR4 and 100GBASE-ER4 transmit characteristics

Description	100GBASE-LR4	100GBASE-ER4	Unit
Signaling rate, each lane (range)	25.78125 ± 100 ppm		GBd
Lane wavelengths (range)	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		nm
Side-mode suppression ratio (SMSR), (min)	30		dB
Total average launch power (max)	10.5	8.9	dBm
Average launch power, each lane (max)	4.5	2.9	dBm
Average launch power, each lane ^a (min)	−4.3	−2.9	dBm
Optical Modulation Amplitude (OMA), each lane (max)	4.5		dBm
Optical Modulation Amplitude (OMA), each lane (min)	−1.3 ^b	0.1	dBm
Difference in launch power between any two lanes (OMA) (max)	5	—	dB
Difference in launch power between any two lanes (Average and OMA) (max)	—	3.6	
Launch power in OMA minus TDP, each lane (min)	−2.3	—	dBm
Transmitter and dispersion penalty (TDP), each lane (max)	2.2	2.5	dB
Average launch power of OFF transmitter, each lane (max)	−30		dBm
Extinction ratio (min)	4	8	dB
RIN ₂₀ OMA (max)	−130		dB/Hz
Optical return loss tolerance (max)	20		dB
Transmitter reflectance ^c (max)	−12		dB
Transmitter eye mask definition {X1, X2, X3, Y1, Y2, Y3}	{0.25, 0.4, 0.45, 0.25, 0.28, 0.4}		

^aAverage launch power, each lane (min) is informative and not the principal indicator of signal strength. A transmitter with launch power below this value cannot be compliant; however, a value above this does not ensure compliance.

^bEven if the TDP < 1 dB, the OMA (min) must exceed this value.

^cTransmitter reflectance is defined looking into the transmitter.

88.7.2 100GBASE-LR4 and 100GBASE-ER4 receive optical specifications

The 100GBASE-LR4 receiver shall meet the specifications defined in Table 88-8 per the definitions in 88.8. The 100GBASE-ER4 receiver shall meet the specifications defined in Table 88-8 per the definitions in 88.8.

Table 88-8—100GBASE-LR4 and 100GBASE-ER4 receive characteristics

Description	100GBASE-LR4	100GBASE-ER4	Unit
Signaling rate, each lane (range)	25.78125 ± 100 ppm		GBd
Lane wavelengths (range)	1294.53 to 1296.59 1299.02 to 1301.09 1303.54 to 1305.63 1308.09 to 1310.19		nm
Damage threshold ^a	5.5		dBm
Average receive power, each lane (max)	4.5 ^b		dBm
Average receive power, each lane ^c (min)	-10.6	-20.9	dBm
Receive power, each lane (OMA) (max)	4.5		dBm
Difference in receive power between any two lanes (OMA) (max)	5.5	—	dB
Difference in receive power between any two lanes (Average and OMA) (max)	—	4.5	
Receiver reflectance (max)	-26		dB
Receiver sensitivity (OMA), each lane ^d (max)	-8.6	-21.4	dBm
Receiver 3 dB electrical upper cutoff frequency, each lane (max)	31		GHz
Stressed receiver sensitivity (OMA), each lane ^e (max)	-6.8	-17.9	dBm
Conditions of stressed receiver sensitivity test			
Vertical eye closure penalty, ^f each lane	1.8	3.5	dB
Stressed eye J2 Jitter, ^f each lane	0.3		UI
Stressed eye J9 Jitter, ^f each lane	0.47		UI

^aThe receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level.

^bThe average receive power, each lane (max) for 100GBASE-ER4 is larger than the 100GBASE-ER4 transmitter value to allow compatibility with 100GBASE-LR4 units at short distances.

^cAverage receive power, each lane (min) is informative and not the principal indicator of signal strength. A received power below this value cannot be compliant; however, a value above this does not ensure compliance.

^dReceiver sensitivity (OMA), each lane (max) is informative.

^eMeasured with conformance test signal at TP3 (see 88.8.10) for BER = 10⁻¹².

^fVertical eye closure penalty, stressed eye J2 Jitter, and stressed eye J9 Jitter are test conditions for measuring stressed receiver sensitivity. They are not characteristics of the receiver.

88.7.3 100GBASE-LR4 and 100GBASE-ER4 illustrative link power budgets

Illustrative power budgets and penalties for 100GBASE-LR4 and 100GBASE-ER4 channels are shown in Table 88-9.

Table 88-9—100GBASE-LR4 and 100GBASE-ER4 illustrative link power budgets

Parameter	100GBASE-LR4	100GBASE-ER4		Unit
Power budget (for maximum TDP)	8.5	—		dB
Power budget	—	21.5		dB
Operating distance	10	30	40 ^a	km
Channel insertion loss	6.3 ^b	15	18	dB
Maximum discrete reflectance	−26	−26		dB
Allocation for penalties ^c (for maximum TDP)	2.2	—		dB
Allocation for penalties ^c	—	3.5		
Additional insertion loss allowed	0	3	0	dB

^aLinks longer than 30 km are considered engineered links. Attenuation for such links needs to be less than the worst case for B1.1, B1.3, or B6—a single-mode cabled optical fiber

^bThe channel insertion loss is calculated using the maximum distance specified in Table 88-6 for 100GBASE-LR4 and fiber attenuation of 0.43 dB/km at 1295 nm plus an allocation for connection and splice loss given in 88.11.2.1.

^cLink penalties are used for link budget calculations. They are not requirements and are not meant to be tested.

88.8 Definition of optical parameters and measurement methods

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

88.8.1 Test patterns for optical parameters

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some parameters. Table 88-11 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subclauses in which each parameter is defined. Any of the test patterns given for a particular test in Table 88-11 may be used to perform that test. The test patterns used in this clause are shown in Table 88-10.

88.8.2 Wavelength

The wavelength of each optical lane shall be within the ranges given in Table 88-5 if measured per TIA/EIA-455-127-A or IEC 61280-1-3. The lane under test is modulated using the test pattern defined in Table 88-11.

Table 88–10—Test patterns

Pattern	Pattern description	Defined in
Square wave	Square wave (8 ones, 8 zeros)	83.5.10
3	PRBS31	83.5.10
4	PRBS9	83.5.10
5	Scrambled idle	82.2.10

Table 88–11—Test-pattern definitions and related subclauses

Parameter	Pattern	Related subclause
Wavelength	3, 5 or valid 100GBASE-R signal	88.8.2
Side mode suppression ratio	3, 5 or valid 100GBASE-R signal	—
Average optical power	3, 5 or valid 100GBASE-R signal	88.8.3
Optical modulation amplitude (OMA)	Square wave or 4	88.8.4
Transmitter and dispersion penalty (TDP)	3 or 5	88.8.5
Extinction ratio	3, 5 or valid 100GBASE-R signal	88.8.6
$RIN_{20}OMA$	Square wave or 4	88.8.7
Transmitter optical waveform	3, 5 or valid 100GBASE-R signal	88.8.8
Stressed receiver sensitivity	3 or 5	88.8.10
Calibration of OMA for receiver tests	Square wave or 4	87.8.11
Vertical eye closure penalty calibration	3 or 5	87.8.11
Receiver 3 dB electrical upper cutoff frequency	3, 5 or valid 100GBASE-R signal	88.8.11

88.8.3 Average optical power

The average optical power of each lane shall be within the limits given in Table 88–7 if measured using the methods given in IEC 61280–1–1, with the sum of the optical power from all of the lanes not under test below –30 dBm, per the test setup in Figure 53–6.

88.8.4 Optical Modulation Amplitude (OMA)

OMA shall be as defined in 52.9.5 for measurement with a square wave (8 ones, 8 zeros) test pattern or 68.6.2 (from the variable MeasuredOMA in 68.6.6.2) for measurement with a PRBS9 test pattern, with the exception that each lane may be tested individually with the sum of the optical power from all of the lanes not under test being below –30 dBm, or if other lanes are operating, a suitable optical filter may be used to separate the lane under test.

88.8.5 Transmitter and dispersion penalty (TDP)

Transmitter and dispersion penalty (TDP) shall be as defined in 52.9.10 with the exception that each optical lane is tested individually using an optical filter to separate the lane under test from the others. The measurement procedure for 100GBASE-LR4 and 100GBASE-ER4 is detailed in 88.8.5.1 to 88.8.5.4.

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

88.8.5.1 Reference transmitter requirements

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

- a) Rise/fall times of less than 12 ps at 20% to 80%.
- b) The output optical eye is symmetric and passes the transmitter optical waveform test of 88.8.8.
- c) In the center 20% region of the eye, the worst-case vertical eye closure penalty as defined in 87.8.11.2 is less than 0.5 dB.
- d) Total Jitter less than 0.2 UI peak-to-peak.
- e) RIN of less than -138 dB/Hz.

88.8.5.2 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in Table 88-12.

Table 88-12—Transmitter compliance channel specifications

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

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

^bThere is no intent to stress the sensitivity of the BERT’s optical receiver.

^cThe optical return loss is applied at TP2.

A 100GBASE-LR4 or 100GBASE-ER4 transmitter is to be compliant with a total dispersion at least as negative as the “minimum dispersion” and at least as positive as the “maximum dispersion” columns specified in Table 88-12 for the wavelength of the device under test. This may be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify that the fiber has the correct amount of dispersion, the measurement method defined in IEC 60793-1-42 may be used. The measurement is made in the linear power regime of the fiber.

The channel provides an optical return loss specified in Table 88-12. The state of polarization of the back reflection is adjusted to create the greatest RIN.

The mean DGD of the channel is to be less than the value specified in Table 88-12.

88.8.5.3 Reference receiver requirements

The reference receiver is required to have the bandwidth given in 88.8.8. The sensitivity of the reference receiver is limited by Gaussian noise. The receiver has minimal threshold offset, deadband, hysteresis, baseline wander, deterministic jitter, or other distortions. Decision sampling has minimal uncertainty and setup/hold times.

The nominal sensitivity of the reference receiver, S , is measured in OMA using the setup of Figure 52–12 without the test fiber and with the transversal filter removed. The sensitivity S must be corrected for any significant reference transmitter impairments including any vertical eye closure. It is measured while sampling at the eye center or corrected for off-center sampling. It is calibrated at the wavelength of the transmitter under test.

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

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

88.8.5.4 Test procedure

The test procedure is as defined in 52.9.10.4 with the exception that all lanes are operational in both directions (transmit and receive), each lane is tested individually using an optical filter to separate the lane under test from the others and the BER of 1×10^{-12} is for the lane under test on its own.

88.8.6 Extinction ratio

The extinction ratio of each lane shall be within the limits given in Table 88–7 if measured using the methods specified in IEC 61280–2–2, with the sum of the optical power from all of the lanes not under test below -30 dBm. The extinction ratio is measured using the test pattern defined in Table 88–11.

NOTE—Extinction ratio and OMA are defined with different test patterns (see Table 88–11).

88.8.7 Relative Intensity Noise (RIN_{20OMA})

RIN shall be as defined by the measurement methodology of 52.9.6 with the following exceptions:

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

88.8.8 Transmitter optical waveform (transmit eye)

The required optical transmitter pulse shape characteristics are specified in the form of a mask of the transmitter eye diagram as shown in Figure 86–4. The transmitter optical waveform of a port transmitting the test pattern specified in Table 88–11 shall meet specifications according to the methods specified in 86.8.4.6.1 with the exception that the clock recovery unit's high-frequency corner bandwidth is 10 MHz. The filter nominal reference frequency f_c is 19.34 GHz and the filter tolerances are as specified for STM-64

in ITU-T G.691. Compensation may be made for variation of the reference receiver filter response from an ideal fourth-order Bessel-Thomson response.

88.8.9 Receiver sensitivity

Receiver sensitivity, which is defined for an ideal input signal, is informative and compliance is not required. If measured, the test signal should have negligible impairments such as intersymbol interference (ISI), rise/fall times, jitter and RIN. Instead, the normative requirement for receivers is stressed receiver sensitivity.

88.8.10 Stressed receiver sensitivity

Stressed receiver sensitivity shall be within the limits given in Table 88–8 if measured using the method defined in 87.8.11 with the following exceptions:

- a) Added sinusoidal jitter is as specified in Table 88–13.
- b) The stressed eye J2 Jitter, stressed eye J9 Jitter, and vertical eye closure penalty are as given in Table 88–8.
- c) The test pattern is as given in Table 88–11.
- d) The reference receiver used to verify the conformance test signal is required to have the bandwidth given in 88.8.8.

Table 88–13—Applied sinusoidal jitter

Frequency range	Sinusoidal jitter, peak-to-peak (UI)
$f < 100$ kHz	Not specified
100 kHz $< f \leq 10$ MHz	$5 \times 10^5 / f$
10 MHz $< f < 10$ LB ^a	0.05

^aLB = loop bandwidth; upper frequency bound for added sine jitter should be at least 10 times the loop bandwidth of the receiver being tested.

88.8.11 Receiver 3 dB electrical upper cutoff frequency

The receiver 3 dB electrical upper cutoff frequency shall be within the limits given in Table 88–8 if measured as described in 52.9.11. Each optical lane is measured using an optical signal or signals with its/their wavelength within the specified wavelength range of the lane to be tested. The test may use an electrical combiner and one optical source as in 53.9.13.

88.9 Safety, installation, environment, and labeling

88.9.1 General safety

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

88.9.2 Laser safety

100GBASE-LR4 and 100GBASE-ER4 optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825–1 and IEC 60825–2, under any condition of operation. This includes single fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service. This documentation explicitly defines requirements and usage restrictions on the host system necessary to meet these safety certifications.¹⁵

88.9.3 Installation

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

88.9.4 Environment

Normative specifications in this clause shall be met by a system integrating a 100GBASE-LR4 or 100GBASE-ER4 PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate in the literature associated with the PHY the operating environmental conditions to facilitate selection, installation, and maintenance.

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

88.9.5 Electromagnetic emission

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

88.9.6 Temperature, humidity, and handling

The optical link is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

88.9.7 PMD labeling requirements

It is recommended that each PHY (and supporting documentation) be labeled in a manner visible to the user, with at least the applicable safety warnings and the applicable port type designation (e.g., 100GBASE-LR4).

Labeling requirements for Hazard Level 1 lasers are given in the laser safety standards referenced in 88.9.2.

¹⁵A host system that fails to meet the manufacturer's requirements and/or usage restrictions may emit laser radiation in excess of the safety limits of one or more safety standards. In such a case, the host manufacturer is required to obtain its own laser safety certification.

88.10 Fiber optic cabling model

The fiber optic cabling model is shown in Figure 88–3.

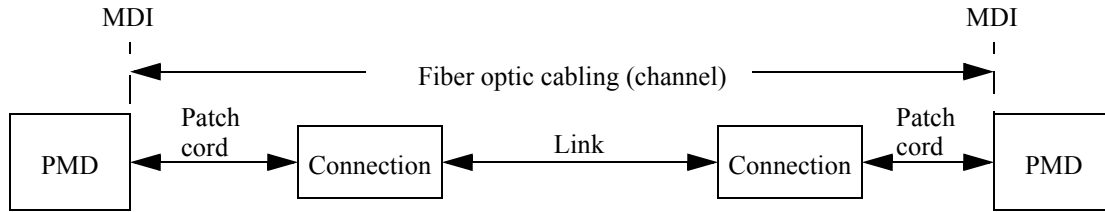


Figure 88–3—Fiber optic cabling model

The channel insertion loss is given in Table 88–14. A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with ANSI/TIA/EIA–526–7/method A–1. The fiber optic cabling model (channel) defined here is the same as a simplex fiber optic link segment. The term channel is used here for consistency with generic cabling standards.

Table 88–14—Fiber optic cabling (channel) characteristics

Description	100GBASE–LR4	100GBASE–ER4		Unit
Operating distance (max)	10	30	40	km
Channel insertion loss ^{a, b} (max)	6.3	18	18	dB
Channel insertion loss (min)	0	0		dB
Positive dispersion ^b (max)	9.5	28	36	ps/nm
Negative dispersion ^b (min)	–28.5	–85	–114	ps/nm
DGD_max ^c	8	10.3	10.3	ps
Optical return loss (min)	21	21	21	dB

^aThese channel insertion loss values include cable, connectors, and splices.

^bOver the wavelength range 1294.53 nm to 1310.19 nm

^cDifferential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal. DGD_max is the maximum differential group delay that the system must tolerate.

88.11 Characteristics of the fiber optic cabling (channel)

The 100GBASE–LR4 and 100GBASE–ER4 fiber optic cabling shall meet the specifications defined in Table 88–14. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

88.11.1 Optical fiber cable

The fiber optic cable requirements are satisfied by cables containing IEC 60793–2–50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) fibers or the requirements in Table 88–15 where they differ.

Table 88–15—Optical fiber and cable characteristics

Description	Value	Unit
Nominal fiber specification wavelength	1310	nm
Cabled optical fiber attenuation (max)	0.43 ^a or 0.5 ^b	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm ² km

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

^bThe 0.5 dB/km attenuation is provided for Outside Plant cable as defined in ANSI/TIA 568-C.3. Using 0.5 dB/km may not support operation at 10 km for 100GBASE-LR4 or 40 km for 100GBASE-ER4.

88.11.2 Optical fiber connection

An optical fiber connection, as shown in Figure 88–3, consists of a mated pair of optical connectors.

88.11.2.1 Connection insertion loss

The maximum link distance is based on an allocation of 2 dB total connection and splice loss. For example, this allocation supports four connections with an average insertion loss per connection of 0.5 dB. Connections with different loss characteristics may be used provided the requirements of Table 88–14 are met.

88.11.2.2 Maximum discrete reflectance

The maximum discrete reflectance shall be less than –26 dB.

88.11.3 Medium Dependent Interface (MDI) requirements

The 100GBASE-LR4 or 100GBASE-ER4 PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in Figure 88–3). Examples of an MDI include the following:

- a) Connectorized fiber pigtail
- b) PMD receptacle

When the MDI is a connector plug and receptacle connection, it shall meet the interface performance specifications of IEC 61753–1–1 and IEC 61753–021–2.

NOTE—Transmitter compliance testing is performed at TP2 as defined in 88.5.1, not at the MDI.

88.12 Protocol implementation conformance statement (PICS) proforma for Clause 88, Physical Medium Dependent (PMD) sublayer and medium, type 100GBASE-LR4 and 100GBASE-ER4¹⁶

88.12.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 88, Physical Medium Dependent sublayer and medium, type 100GBASE-LR4 and 100GBASE-ER4, 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.

88.12.2 Identification

88.12.2.1 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).	

88.12.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 88, Physical Medium Dependent sublayer and medium, type 100GBASE-LR4 and 100GBASE-ER4
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 802.3-2012.)	

Date of Statement	
-------------------	--

¹⁶Copyright 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.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54

88.12.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*LR4	100GBASE-LR4 PMD	88.7	Device supports requirements for 100GBASE-LR4 PHY	O.1	Yes [] No []
*ER4	100GBASE-ER4 PMD	88.7	Device supports requirements for 100GBASE-ER4 PHY	O.1	Yes [] No []
*INS	Installation / cable	88.10	Items marked with INS include installation practices and cable specifications not applicable to a PHY manufacturer	O	Yes [] No []
CTP1	Reference point TP1 exposed and available for testing	88.5.1	This point may be made available for use by implementors to certify component conformance	O	Yes [] No []
CTP4	Reference point TP4 exposed and available for testing	88.5.1	This point may be made available for use by implementors to certify component conformance	O	Yes [] No []
CDC	Delay constraints	88.3.1	Device conforms to delay constraints	M	Yes []
CSC	Skew constraints	88.3.2	Device conforms to Skew and Skew Variation constraints	M	Yes []
*MD	MDIO capability	88.4	Registers and interface supported	O	Yes [] No []

88.12.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, types 100GBASE-LR4 and 100GBASE-ER4**88.12.4.1 PMD functional specifications**

Item	Feature	Subclause	Value/Comment	Status	Support
CF1	Compatible with 100GBASE-R PCS and PMA	88.1		M	Yes []
CF2	Integration with management functions	88.1		O	Yes [] No []
CF3	Transmit function	88.5.2	Conveys bits from PMD service interface to MDI	M	Yes []
CF4	Optical multiplexing and delivery to the MDI	88.5.2	Optically multiplexes the four optical signal streams for delivery to the MDI	M	Yes []
CF5	Mapping between optical signal and logical signal for transmitter	88.5.2	Higher optical power is a one	M	Yes []
CF6	Receive function	88.5.3	Conveys bits from MDI to PMD service interface	M	Yes []
CF7	Conversion of four optical signals to four electrical signals	88.5.3	For delivery to the PMD service interface	M	Yes []
CF8	Mapping between optical signal and logical signal for receiver	88.5.3	Higher optical power is a one	M	Yes []
CF9	Global Signal Detect function	88.5.4	Report to the PMD service interface the message PMD:IS_SIGNAL.indication(SIGNAL_DETECT)	M	Yes []
CF10	Global Signal Detect behavior	88.5.4	SIGNAL_DETECT is a global indicator of the presence of optical signals on all four lanes	M	Yes []
CF11	Lane-by-lane Signal Detect function	88.5.5	Sets PMD_signal_detect _{<i>i</i>} values on a lane-by-lane basis per requirements of Table 88-4	MD:O	Yes [] No [] N/A []
CF12	PMD reset function	88.5.6	Resets the PMD sublayer	MD:O	Yes [] No [] N/A []

88.12.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
CM1	Management register set	88.4		MD:M	Yes [] N/A []
CM2	Global transmit disable function	88.5.7	Disables all of the optical transmitters with the PMD_global_transmit_disable variable	MD:O	Yes [] No [] N/A []
CM3	PMD_lane_by_lane_transmit_disable function	88.5.8	Disables the optical transmitter on the lane associated with the PMD_transmit_disable_i variable	MD:O.2	Yes [] No [] N/A []
CM4	PMD lane-by-lane transmit disable	88.5.8	Disables each optical transmitter independently if CM3 = No	O.2	Yes [] No []
CM5	PMD_fault function	88.5.9	Sets PMD_fault to one if any local fault is detected	MD:O	Yes [] No [] N/A []
CM6	PMD_transmit_fault function	88.5.10	Sets PMD_transmit_fault to one if a local fault is detected on any transmit lane	MD:O	Yes [] No [] N/A []
CM7	PMD_receive_fault function	88.5.11	Sets PMD_receive_fault to one if a local fault is detected on any receive lane	MD:O	Yes [] No [] N/A []

88.12.4.3 PMD to MDI optical specifications for 100GBASE-LR4

Item	Feature	Subclause	Value/Comment	Status	Support
CLR1	Transmitter meets specifications in Table 88-7	88.7.1	Per definitions in 88.8	LR4:M	Yes [] N/A []
CLR2	Receiver meets specifications in Table 88-8	88.7.2	Per definitions in 88.8	LR4:M	Yes [] N/A []

88.12.4.4 PMD to MDI optical specifications for 100GBASE-ER4

Item	Feature	Subclause	Value/Comment	Status	Support
CER1	Transmitter meets specifications in Table 88-7	88.7.1	Per definitions in 88.8	ER4:M	Yes [] N/A []
CER2	Receiver meets specifications in Table 88-8	88.7.2	Per definitions in 88.8	ER4:M	Yes [] N/A []

88.12.4.5 Optical measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
COM1	Measurement cable	88.8	2 m to 5 m in length	M	Yes []
COM2	Center wavelength	88.8.2	Per TIA/EIA-455-127-A or IEC 61280-1-3 under modulated conditions	M	Yes []
COM3	Average optical power	88.8.3	Per IEC 61280-1-1	M	Yes []
COM4	OMA measurements	88.8.4	Each lane	M	Yes []
COM5	Transmitter and dispersion penalty	88.8.5	Each lane	M	Yes []
COM6	Extinction ratio	88.8.6	Per IEC 61280-2-2	M	Yes []
COM7	RIN ₂₀ OMA measurement procedure	88.8.7	Each lane	M	Yes []
COM8	Transmit eye	88.8.8	Each lane	M	Yes []
COM9	Stressed receiver sensitivity	88.8.10	Each lane	M	Yes []
COM10	Receiver 3 dB electrical upper cutoff frequency	88.8.11	Each lane	M	Yes []

88.12.4.6 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
CES1	General safety	88.9.1	Conforms to IEC 60950-1	M	Yes []
CES2	Laser safety—IEC Hazard Level 1	88.9.2	Conforms to Hazard Level 1 laser requirements defined in IEC 60825-1 and IEC 60825-2	M	Yes []
CES3	Electromagnetic interference	88.9.5	Complies with applicable local and national codes for the limitation of electromagnetic interference	M	Yes []

88.12.4.7 Characteristics of the fiber optic cabling and MDI

Item	Feature	Subclause	Value/Comment	Status	Support
COC1	Fiber optic cabling	88.11	Meets requirements specified in Table 88-14	INS:M	Yes [] N/A []
COC2	Maximum discrete reflectance	88.11.2.2	Less than -26 dB	INS:M	Yes [] N/A []
COC3	MDI requirements	88.11.3	Meets IEC 61753-1-1 and IEC 61753-021-2	INS:M	Yes [] N/A []

89. Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR

89.1 Overview

This clause specifies the 40GBASE-FR PMD together with the single-mode fiber medium. It is intended that devices compliant with this specification could be implemented to also be used for application VSR2000-3R2 as defined in ITU-T G.693 [B50] (but that is beyond the scope of this standard). The 40GBASE-FR PMD is defined using a specification and test methodology that is similar to that used in ITU-T G.693 [B50], which is different from the specification and test methodologies used for other optical PMDs at 1000 Mb/s and above in IEEE Std 802.3. For example, the transmitter specifications for 40GBASE-FR do not include OMA or TDP parameters but do include a dispersion penalty limit, and the receiver specifications for 40GBASE-FR do not include stressed receiver sensitivity parameters but do include a part-stressed sensitivity limit and a separate jitter tolerance limit. See 89.6.4.

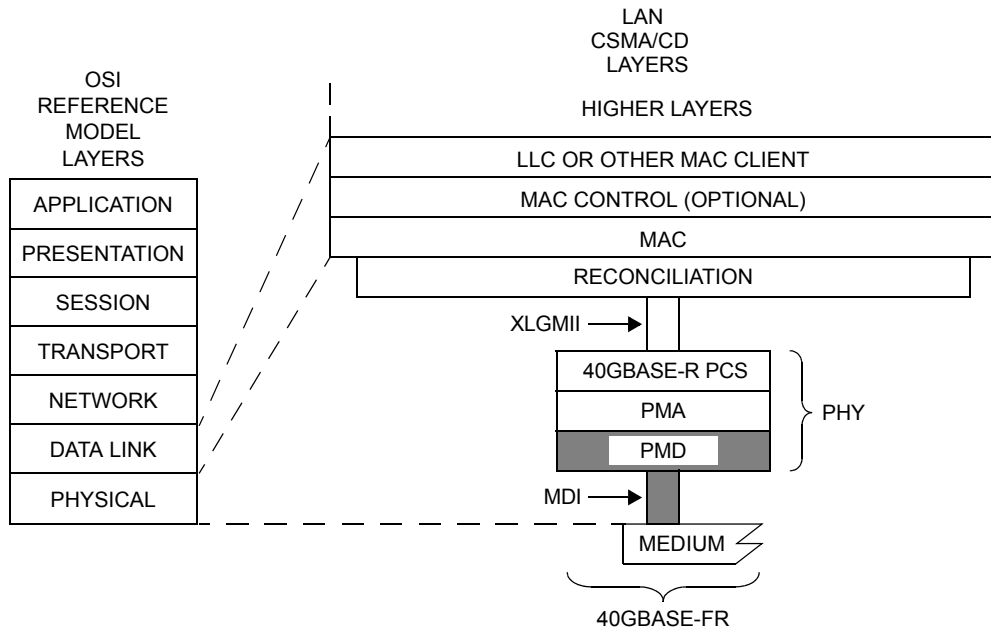
When forming a complete Physical Layer, a PMD shall be connected to the appropriate PMA as shown in Table 89–1, to the medium through the MDI and to the management functions that are optionally accessible through the management interface defined in Clause 45, or equivalent.

Table 89–1—Physical Layer clauses associated with the 40GBASE-FR PMD

Associated clause	40GBASE-FR
81—RS	Required
81—XLGMII ^a	Optional
82—PCS for 40GBASE-R	Required
83—PMA for 40GBASE-R	Required
83A—XLAUI	Optional
83B—Chip to module XLAUI	Optional

^aThe XLGMII is an optional interface. However, if the XLGMII is not implemented, a conforming implementation must behave functionally as though the RS and XLGMII were present.

Figure 89–1 shows the relationship of the PMD and MDI (shown shaded) with other sublayers to the ISO/IEC Open System Interconnection (OSI) reference model. 40 Gb/s and 100 Gb/s Ethernet is introduced in Clause 80 and the purpose of each PHY sublayer is summarized in 80.2.



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
 FR = PMD FOR SINGLE-MODE FIBER — 2 km

Figure 89–1—40GBASE-FR PMD relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 CSMA/CD LAN model

89.2 Physical Medium Dependent (PMD) service interface

This subclause specifies the services provided by the 40GBASE-FR 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 between the PMA entity that resides just above the PMD, and the PMD entity. The PMD translates the encoded data to and from signals suitable for the specified medium.

The PMD service interface 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_0.request
PMD:IS_UNITDATA_0.indication
PMD:IS_SIGNAL.indication
```

In the transmit direction, the PMA continuously sends a single serial stream of bits to the PMD at a nominal signaling rate of 41.25 GBd. The PMD converts this stream of bits into an appropriate signal on the MDI.

In the receive direction, the PMD continuously sends a single serial stream of bits to the PMA corresponding to the signal received from the MDI at a nominal signaling rate of 41.25 GBd.

The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the PMD:IS_SIGNAL.indication(SIGNAL_OK) inter-sublayer service primitive defined in 80.3.

The SIGNAL_DETECT parameter can take on one of two values: OK or FAIL. When SIGNAL_DETECT = FAIL, the rx_bit parameter is undefined.

NOTE—SIGNAL_DETECT = OK does not guarantee that the rx_bit parameter is known to be good. It is possible for a poor quality link to provide sufficient light for a SIGNAL_DETECT = OK indication and still not meet the 10^{-12} BER objective.

89.3 Delay and skew

89.3.1 Delay constraints

The sum of the transmit and the receive delays at one end of the link contributed by the 40GBASE-FR PMD including 2 m of fiber in one direction shall be no more than 1024 bit times (2 pause_quanta or 25.6 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.

89.3.2 Skew constraints

The Skew (relative delay) between the PCS 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 43 ns as defined by 83.5.3.3. Since the signal at the PMD service interface represents a serial bit stream, there is no Skew Variation at this point.

The Skew at SP3 (the transmitter MDI) shall be less than 54 ns. Since the signal at the MDI represents a serial bit stream, there is no Skew Variation at this point.

The Skew at SP4 (the receiver MDI) shall be less than 134 ns. Since the signal at the MDI represents a serial bit stream, there is no Skew Variation at this point.

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 145 ns. Since the signal at the PMD service interface represents a serial bit stream, there is no Skew Variation at this point.

For more information on Skew and Skew Variation, see 80.5. The measurements of Skew and Skew Variation are defined in 89.7.2.

89.4 PMD MDIO function mapping

The optional MDIO capability described in Clause 45 defines several variables that may provide control and status information for and about the PMD. If the MDIO interface is implemented, the mapping of MDIO control variables to PMD control variables shall be as shown in Table 89-2, and the mapping of MDIO status variables to PMD status variables shall be as shown in Table 89-3.

Table 89–2—MDIO/PMD control variable mapping

MDIO control variable	PMA/PMD register name	Register/bit number	PMD control variable
Reset	PMA/PMD control 1 register	1.0.15	PMD_reset
Global PMD transmit disable	PMD Transmit disable register	1.9.0	PMD_global_transmit_disable

Table 89–3—MDIO/PMD status variable mapping

MDIO status variable	PMA/PMD register name	Register/bit number	PMD status variable
Fault	PMA/PMD status 1 register	1.1.7	PMD_fault
Transmit fault	PMA/PMD status 2 register	1.8.11	PMD_transmit_fault
Receive fault	PMA/PMD status 2 register	1.8.10	PMD_receive_fault
Global PMD receive signal detect	PMD receive signal detect register	1.10.0	PMD_global_signal_detect

89.5 PMD functional specifications

The 40GBASE-FR PMD performs the Transmit and Receive functions, which convey data between the PMD service interface and the MDI.

89.5.1 PMD block diagram

The PMD block diagram is shown in Figure 89–2. For purposes of system conformance, the PMD sublayer is standardized at the points described in this subclause. The optical transmit signal is defined at the output end of a single-mode fiber patch cord (TP2), between 2 m and 5 m in length. Unless specified otherwise, all transmitter measurements and tests defined in 89.7 are made at TP2. The optical receive signal is defined at the output of the fiber optic cabling (TP3) at the MDI (see 89.10.3). Unless specified otherwise, all receiver measurements and tests defined in 89.7 are made at TP3.

89.5.2 PMD transmit function

The PMD Transmit function shall convert the bit stream requested by the PMD service interface message `PMD:IS_UNITDATA_0.request` into an optical signal stream and deliver it to the MDI, according to the transmit optical specifications in this clause. The higher optical power level in the signal stream shall correspond to `tx_bit = one`.

89.5.3 PMD receive function

The optical signal stream received from the MDI shall be converted into a bit stream for delivery to the PMD service interface using the message `PMD:IS_UNITDATA_0.indication`, according to the receive optical specifications in this clause. The higher optical power level in the signal stream shall correspond to `rx_bit = one`.

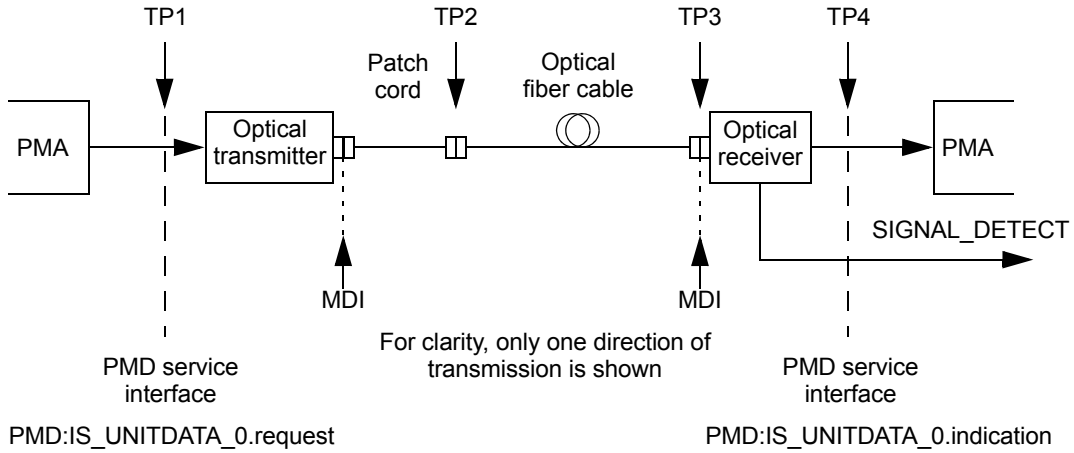


Figure 89–2—Block diagram for 40GBASE-FR transmit/receive paths

89.5.4 PMD global signal detect function

The PMD global signal detect function shall report the state of SIGNAL_DETECT via the PMD service interface. The SIGNAL_DETECT parameter is signaled continuously, while the PMD:IS_SIGNAL.indication message is generated when a change in the value of SIGNAL_DETECT occurs. The SIGNAL_DETECT parameter defined in this clause maps to the SIGNAL_OK parameter in the inter-sublayer service interface primitives defined in 80.3.

SIGNAL_DETECT is an indicator of the presence of an optical signal. The value of the SIGNAL_DETECT parameter shall be generated according to the conditions defined in Table 89–4. The PMD receiver is not required to verify whether a compliant 40GBASE-R signal is being received. This standard imposes no response time requirements on the generation of the SIGNAL_DETECT parameter.

Table 89–4—SIGNAL_DETECT value definition

Receive conditions	SIGNAL_DETECT value
Average optical power at TP3 ≤ -30 dBm	FAIL
(Optical power at TP3 \geq receiver sensitivity (average power) (max) in Table 89–7) AND (compliant 40GBASE–R signal input)	OK
All other conditions	Unspecified

As an unavoidable consequence of the requirements for the setting of the SIGNAL_DETECT parameter, implementations must provide adequate margin between the input optical power level at which the SIGNAL_DETECT parameter is set to OK and the inherent noise level of the PMD including the effects of crosstalk, power supply noise, etc.

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

89.5.5 PMD reset function

If the MDIO interface is implemented and if PMD_reset is asserted, the PMD shall be reset as defined in 45.2.1.1.1.

89.5.6 PMD global transmit disable function (optional)

The PMD_global_transmit_disable function is optional and allows the optical transmitter to be disabled.

- a) When the PMD_global_transmit_disable variable is set to one, this function shall turn off the optical transmitter so that it meets the requirements of the average launch power of the OFF transmitter in Table 89–6.
- b) If a PMD_fault is detected, then the PMD may set the PMD_global_transmit_disable to one, turning off the optical transmitter.

If the MDIO interface is implemented, then this function shall map to the PMD_global_transmit_disable bit as specified in 45.2.1.8.7.

NOTE—The PMD lane-by-lane transmit disable function is not used for serial PMDs.

89.5.7 PMD fault function (optional)

If the PMD has detected a local fault on either the transmit or receive path, the PMD shall set PMD_fault to one. If the MDIO interface is implemented, PMD_fault shall be mapped to the fault bit as specified in 45.2.1.2.1.

89.5.8 PMD transmit fault function (optional)

If the PMD has detected a local fault on the transmitter, the PMD shall set the PMD_transmit_fault variable to one. If the MDIO interface is implemented, PMD_transmit_fault shall be mapped to the transmit fault bit as specified in 45.2.1.7.4.

89.5.9 PMD receive fault function (optional)

If the PMD has detected a local fault on the receiver, the PMD shall set the PMD_receive_fault variable to one. If the MDIO interface is implemented, PMD_receive_fault shall be mapped to the receive fault bit as specified in 45.2.1.7.5.

89.6 PMD to MDI optical specifications for 40GBASE-FR

The required operating range for the 40GBASE-FR PMD is defined in Table 89–5. A 40GBASE-FR compliant PMD operates on IEC 60793-2-50 type B1.1, B1.3 or B6_a single-mode fibers according to the specifications defined in Table 89–13. A PMD that exceeds the required operating range while meeting all other optical specifications is considered compliant (e.g., operating at 2.5 km meets the operating range requirement of 2 m to 2 km).

Table 89–5—40GBASE-FR required operating range

PMD type	Required operating range
40GBASE-FR	2 m to 2 km

89.6.1 40GBASE-FR transmitter optical specifications

The 40GBASE-FR transmitter shall meet the specifications defined in Table 89–6 per the definitions in 89.7.

Table 89–6—40GBASE-FR transmit characteristics

Description	Value	Unit
Signaling rate (range)	41.25 ± 100 ppm	GBd
Center wavelength (range)	1530 to 1565	nm
Side-mode suppression ratio (SMSR), (min)	35	dB
Average launch power (max)	3	dBm
Average launch power (min)	0	dBm
Dispersion penalty (max)	2	dB
Average launch power of OFF transmitter (max)	–30	dBm
Extinction ratio (min)	8.2	dB
RIN ₂₀ OMA (max)	–132	dB/Hz
Optical return loss tolerance (max)	20	dB
Transmitter reflectance ^a (max)	–12	dB
Transmitter eye mask definition {X1, X2, X3, Y1, Y2, Y3}	{0.25, 0.4, 0.45, 0.22, 0.25, 0.4}	

^aTransmitter reflectance is defined looking into the transmitter.

89.6.2 40GBASE-FR receive optical specifications

The 40GBASE-FR receiver shall meet the specifications defined in Table 89–7 per the definitions in 89.7.

Table 89–7—40GBASE-FR receive characteristics

Description	Value	Unit
Signaling rate (range)	41.25 ± 100 ppm	GBd
Center wavelength (range)	1290 to 1330 and 1530 to 1565	nm
Damage threshold ^a (min)	4	dBm
Average receive power (max)	3	dBm
Receiver reflectance (max)	–26	dB
Receiver sensitivity (average power) ^b (max)	–6	dBm
Receiver jitter tolerance (max)	1	dB
Receiver 3 dB electrical upper cutoff frequency (max)	49	GHz

^aThe receiver shall be able to tolerate, without damage, continuous exposure to an optical input signal having this average power level.

^bReceiver sensitivity (average power) is defined in 89.7.9 and is to be met with a transmitter with worst-case transmit eye, extinction ratio, transmitter reflectance, and RIN₂₀OMA. This is a different definition of receiver sensitivity from that used in other clauses (e.g., that in Clause 38). See 89.6.4 for a comparison.

89.6.3 40GBASE-FR illustrative link power budget

An illustrative power budget and penalties for 40GBASE-FR channels are shown in Table 89–8.

Table 89–8—40GBASE-FR illustrative power budget

Parameter	Value	Unit
Power budget ^a	6	dB
Operating distance	2	km
Channel insertion loss ^b	4	dB
Maximum discrete reflectance	–26	dB
Path penalty ^{a, c}	2	dB
Additional insertion loss allowed	0	dB

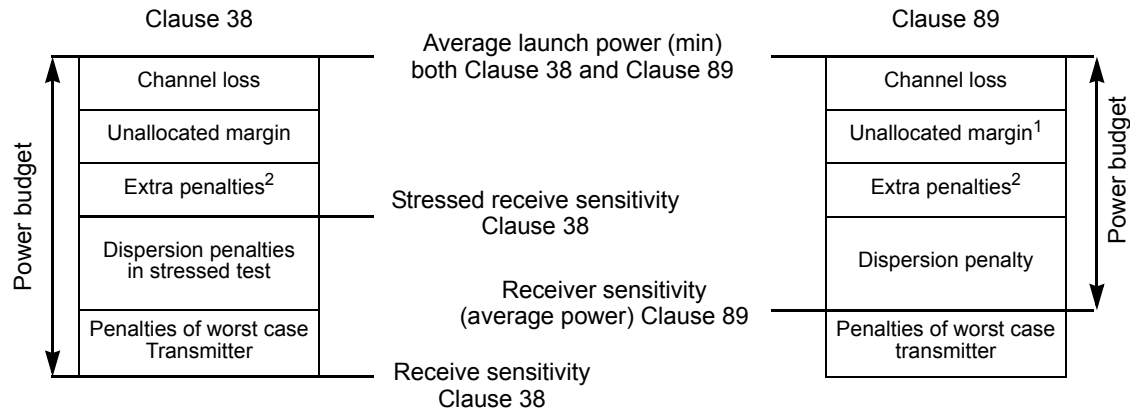
^aThis budget does not include the effect of a non-ideal transmitter waveform. See 89.6.4.

^bThe channel insertion loss is calculated using the maximum distance specified in Table 89–5 and cabled optical fiber attenuation of 0.5 dB/km at 1530 nm plus an allocation for connection and splice loss given in 89.10.2.1.

^cPath penalty is the combined penalty caused by chromatic dispersion and polarization mode dispersion.

89.6.4 Comparison of power budget methodology

This clause uses the budgeting methodology that is used for application VSR2000-3R2 in ITU-T G.693 [B50], which is different from the methodology used in other clauses of this standard (e.g., Clause 38, Clause 52, Clause 86, Clause 87, Clause 88). Figure 89–3 compares the terminology used in this clause with Clause 38. Receiver sensitivity in this clause is specified with a worst-case transmitter input whereas in the other clauses it is specified with a perfect signal without penalties. Stressed receiver sensitivity is not specified in this clause but is specified as the key requirement in the other clauses with a signal that includes both transmitter and link penalties.



NOTE 1—For Clause 89 unallocated margin is zero and not mentioned in the rest of this clause.

NOTE 2—For Clause 38 extra penalties are any penalties that are not included in the stressed test while for Clause 89 extra penalties are any penalties that are not included in the dispersion penalty test.

Figure 89–3—Optical power budget methodology comparison with Clause 38

89.7 Definition of optical parameters and measurement methods

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

89.7.1 Test patterns for optical parameters

Compliance is to be achieved in normal operation. Table 89–10 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subclauses in which each parameter is defined. Any of the test patterns given for a particular test in Table 89–10 may be used to perform that test. The test patterns used in this clause are shown in Table 89–9.

Table 89–9—Test patterns

Pattern	Pattern description	Defined in
Square wave	Square wave (8 ones, 8 zeros)	83.5.10
3	PRBS31	83.5.10
4	PRBS9	83.5.10
5	Scrambled idle	82.2.10

Table 89–10—Test-pattern definitions and related subclauses

Parameter	Pattern	Related subclause
Wavelength	3, 5 or valid 40GBASE-R signal	89.7.3
Side mode suppression ratio	3, 5 or valid 40GBASE-R signal	—
Average optical power	3, 5 or valid 40GBASE-R signal	89.7.4
Dispersion penalty	3 or 5	89.7.5
Extinction ratio	3, 5 or valid 40GBASE-R signal	89.7.6
RIN ₂₀ OMA	Square wave or 4	89.7.7
Transmitter optical waveform	3, 5 or valid 40GBASE-R signal	89.7.8
Receiver sensitivity	3 or 5	89.7.9
Receiver jitter tolerance	3 or 5	89.7.10
Receiver 3 dB electrical upper cutoff frequency	3, 5 or valid 40GBASE-R signal	89.7.11

89.7.2 Skew and Skew Variation

Skew and Skew Variation are defined in 80.5 and are required to remain within the limits given in 89.3.2 over the time that the link is in operation. Since the signal for this PMD is a serial bit stream at the PMD service interface and the MDI, there is no Skew Variation at skew points SP2, SP3, SP4, and SP5. The measurement of Skew is made by acquiring the data on the single lane using a clock and data recovery unit with high-frequency corner bandwidth of 16 MHz and a slope of -20 dB/decade. The arrival times of the one to zero transition of the alignment marker sync bits on each PCS lane are then compared. This arrangement ensures that any high-frequency jitter that is present on the signals is not included in the Skew measurement.

89.7.3 Wavelength

The center wavelength shall be within the range given in Table 89–6 if measured per TIA/EIA-455-127-A or IEC 61280-1-3 using the test pattern referenced in Table 89–10.

89.7.4 Average optical power

The average optical power shall be within the limits given in Table 89–6 and Table 89–7 if measured using the methods given in IEC 61280-1-1 using the test pattern referenced in Table 89–10.

89.7.5 Dispersion penalty

The dispersion penalty of a transmitter is defined as the difference in sensitivity (at a BER of 10^{-12}) of a reference receiver as defined in 89.7.5.2 when receiving the signal from that transmitter via the channel defined in 89.7.5.1 compared to the sensitivity obtained when receiving the signal via an attenuator only.

89.7.5.1 Channel requirements

The transmitter is tested using an optical channel that meets the requirements listed in Table 89–11.

Table 89–11—Transmitter compliance channel specifications

PMD type	Dispersion ^a (ps/nm)		Insertion loss ^b	Optical return loss ^c	Max mean DGD
	Minimum	Maximum			
40GBASE-FR	$0.0465 \times \lambda \times [1 - (1324 / \lambda)^4]$	$0.0465 \times \lambda \times [1 - (1300 / \lambda)^4]$	Minimum	20 dB	0.5 ps

^aThe dispersion is measured for the wavelength of the device under test (λ in nm). The coefficient assumes a maximum operating distance of 2 km.

^bThere is no intent to stress the sensitivity of the BERT’s optical receiver.

^cThe optical return loss is applied at TP2.

A 40GBASE-FR transmitter is to be compliant with a total dispersion at least as negative as the “minimum dispersion” and at least as positive as the “maximum dispersion” columns specified in Table 89–11 for the wavelength of the device under test. This may be achieved with channels consisting of fibers with lengths chosen to meet the dispersion requirements.

To verify that the fiber has the correct amount of dispersion, the measurement method defined in IEC 60793-1-42 may be used. The measurement is made in the linear power regime of the fiber.

The channel provides an optical return loss specified in Table 89–11. The state of polarization of the back reflection is adjusted to create the greatest RIN.

The mean DGD of the channel is to be less than the value specified in Table 89–11.

89.7.5.2 Reference receiver requirements

The reference receiver is required to have the bandwidth given in 89.7.8. The sensitivity of the reference receiver is limited by Gaussian noise. The receiver has minimal threshold offset, deadband, hysteresis, baseline wander, deterministic jitter, or other distortions. Decision sampling has minimal uncertainty and setup/hold times.

The clock recovery unit (CRU) used in the dispersion penalty measurement has a corner frequency of 16 MHz and a slope of 20 dB/decade. When using a clock recovery unit as a clock for BER measurements, passing of low-frequency jitter from the signal to the clock removes this low-frequency jitter from the measurement.

89.7.5.3 Test procedure

To measure the dispersion penalty, the following procedure shall be used:

- Configure the test equipment as described above and illustrated in Figure 89–4.
- With a short patch cable rather than the test fiber and with no reflection, adjust the attenuation of the optical attenuator to obtain a BER of 10^{-12} .
- Record the optical power at the input to the reference receiver, P_{ND} , in dBm.
- With the test fiber and reflection in place, adjust the attenuation of the optical attenuator to obtain a BER of 10^{-12} .
- Record the optical power at the input to the reference receiver, P_D , in dBm.
- If P_D is larger than P_{ND} , the dispersion penalty for the transmitter under test is $P_D - P_{ND}$. Otherwise, the dispersion penalty is zero.

It is to be ensured that the measurements are made in the linear power regime of the fiber.

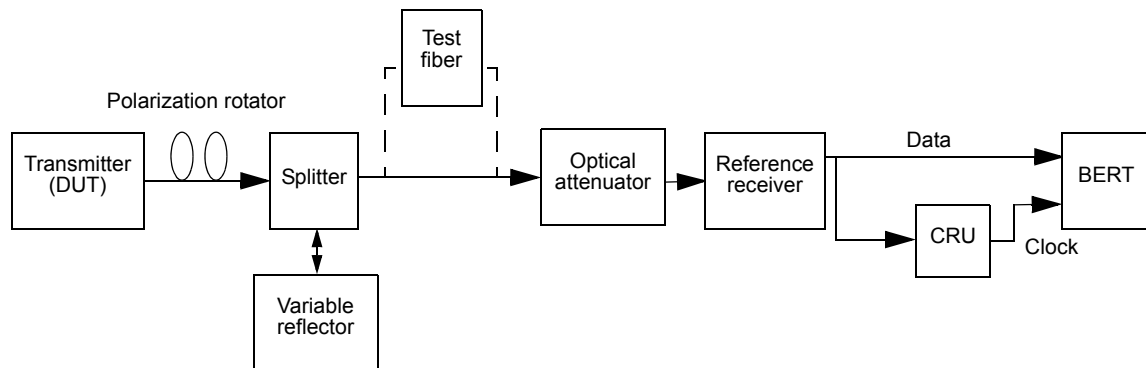


Figure 89–4—Test setup for measurement of dispersion penalty

89.7.6 Extinction ratio

The extinction ratio shall be within the limits given in Table 89–6 if measured using the methods specified in IEC 61280-2-2 using the test pattern referenced in Table 89–10.

NOTE—Extinction ratio and OMA are defined with different test patterns (see Table 89–10).

89.7.7 Relative Intensity Noise (RIN_{20OMA})

The RIN measurement methodology shall be as defined in 52.9.6 with the following exceptions:

- a) The optical return loss is 20 dB.
- b) The upper -3 dB limit of the measurement apparatus is to be approximately equal to the signaling rate (i.e., 41.25 GHz).

89.7.8 Transmitter optical waveform (transmit eye)

The required optical transmitter pulse shape characteristics are specified in the form of a mask of the transmitter eye diagram as shown in Figure 86-4. The transmitter optical waveform of a port transmitting the test pattern specified in Table 89-10 shall meet specifications according to the methods specified in 86.8.4.6.1 with the filter nominal reference frequency f_r of 30.94 GHz and filter tolerances as specified for STM-64 in ITU-T G.691. Compensation may be made for variation of the reference receiver filter response from an ideal fourth-order Bessel-Thomson response.

89.7.9 Receiver sensitivity

Receiver sensitivity is defined as the minimum value of mean received power at TP3 to achieve a BER of 10^{-12} . This shall be met with a transmitter with worst-case transmit eye, extinction ratio, transmitter reflectance, and RIN_{20OMA}. The receiver sensitivity does not have to be met in the presence of dispersion, as this effect is specified separately as the dispersion penalty.

Receiver sensitivity is defined with the transmit section in operation. Pattern 3, or Pattern 5, or a valid 40GBASE-R signal is sent from the transmit section of the PMD under test. The signal being transmitted is asynchronous to the received signal.

89.7.10 Receiver jitter tolerance

Receiver jitter tolerance is defined as the largest increase in received optical power required to maintain a BER of 10^{-10} when the sinusoidal jitter defined in Table 89-12 is applied to the signal from a reference transmitter across the defined range of jitter frequencies.

The reference transmitter is a high-quality instrument-grade device, which can be implemented by a CW laser modulated by a high-performance modulator. The following basic requirements apply:

- a) Rise/fall times are less than 10 ps at 20% to 80%.
- b) The output optical eye is symmetric and passes the transmitter optical waveform test of 89.7.8.
- c) In the center 20% region of the eye, the worst-case vertical eye closure penalty as defined in 52.9.9.2 is less than 0.5 dB.
- d) Total Jitter is less than 0.2 UI peak-to-peak.

The receiver jitter tolerance shall be within the limits given in Table 89-7 if measured according to the above definition.

NOTE—Receiver jitter tolerance as defined here is made without additional amplitude stress.

89.7.11 Receiver 3 dB electrical upper cutoff frequency

The receiver 3dB electrical upper cutoff frequency shall be within the limits given in Table 89-7 if measured as described in 52.9.11 with the exception that the optical power level used in 52.9.11 b) is approximately equal to the receiver sensitivity level in Table 89-7.

Table 89–12—Applied sinusoidal jitter

Frequency range	Sinusoidal jitter, peak-to-peak (UI)
$f < 480 \text{ kHz}$	Not specified
$480 \text{ kHz} < f \leq 16 \text{ MHz}$	$2.88 \times 10^6 / f$
$16 \text{ MHz} < f < 10 \text{ LB}^a$	0.18

^a LB = loop bandwidth; upper frequency bound for added sinusoidal jitter should be at least 10 times the loop bandwidth of the receiver being tested.

89.8 Safety, installation, environment, and labeling

89.8.1 General safety

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

89.8.2 Laser safety

40GBASE-FR optical transceivers shall conform to Hazard Level 1 laser requirements as defined in IEC 60825-1 and IEC 60825-2, under any condition of operation. This includes single-fault conditions whether coupled into a fiber or out of an open bore.

Conformance to additional laser safety standards may be required for operation within specific geographic regions.

Laser safety standards and regulations require that the manufacturer of a laser product provide information about the product's laser, safety features, labeling, use, maintenance, and service. This documentation explicitly defines requirements and usage restrictions on the host system necessary to meet these safety certifications.¹⁷

89.8.3 Installation

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

89.8.4 Environment

Normative specifications in Clause 89 shall be met by a system integrating a 40GBASE-FR PMD over the life of the product while the product operates within the manufacturer's range of environmental, power, and other specifications.

It is recommended that manufacturers indicate, in the literature associated with the PHY, the operating environmental conditions to facilitate selection, installation, and maintenance.

It is recommended that manufacturers indicate, in the literature associated with the components of the optical link, the distance and operating environmental conditions over which the specifications of Clause 89 will be met.

¹⁷A host system that fails to meet the manufacturers requirements and/or usage restrictions may emit laser radiation in excess of the safety limits of one or more safety standards. In such a case, the host manufacturer is required to obtain its own laser safety certification.

89.8.4.1 Electromagnetic emission

A system integrating a 40GBASE-FR PMD shall comply with applicable local and national codes for the limitation of electromagnetic interference.

89.8.4.2 Temperature, humidity, and handling

The optical link is expected to operate over a reasonable range of environmental conditions related to temperature, humidity, and physical handling (such as shock and vibration). Specific requirements and values for these parameters are considered to be beyond the scope of this standard.

89.8.5 PMD labeling requirements

It is recommended that each PHY (and supporting documentation) be labeled in a manner visible to the user, with at least the applicable safety warnings and the applicable port type designation (e.g., 40GBASE-FR).

Labeling requirements for Hazard Level 1 lasers are given in the laser safety standards referenced in 89.8.2.

89.9 Fiber optic cabling model

The fiber optic cabling model is shown in Figure 89–5.

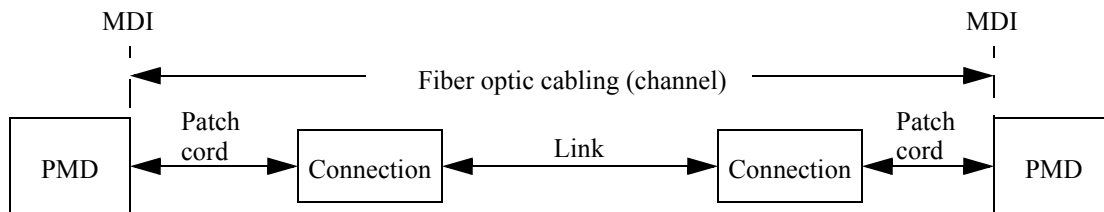


Figure 89–5—Fiber optic cabling model

The channel insertion loss is given in Table 89–13. A channel may contain additional connectors as long as the optical characteristics of the channel, such as attenuation, dispersion, reflections, and polarization mode dispersion, meet the specifications. Insertion loss measurements of installed fiber cables are made in accordance with ANSI/TIA/EIA-526-7 method A–1, chromatic dispersion is measured in accordance with IEC 60793-1-42, and polarization mode dispersion is measured in accordance with IEC 60793-1-48. The fiber optic cabling model (channel) defined here is the same as a simplex fiber optic link segment. The term *channel* is used here for consistency with generic cabling standards.

89.10 Characteristics of the fiber optic cabling (channel)

The 40GBASE-FR fiber optic cabling shall meet the specifications defined in Table 89–13. The fiber optic cabling consists of one or more sections of fiber optic cable and any intermediate connections required to connect sections together.

89.10.1 Optical fiber cable

The fiber optic cable requirements are satisfied by cables containing IEC 60793-2-50 type B1.1 (dispersion un-shifted single-mode), type B1.3 (low water peak single-mode), or type B6_a (bend insensitive) fibers or the requirements in Table 89–14 where they differ.

Table 89–13—Fiber optic cabling (channel) characteristics for 40GBASE-FR

Description	Value	Unit
Operating distance (max)	2	km
Channel insertion loss ^{a, b} (max)	4	dB
Channel insertion loss (min)	0	dB
Dispersion ^b (max)	38	ps/nm
DGD_max ^c	3	ps
Optical return loss (min)	21	dB

^aThese channel insertion loss values include cable, connectors, and splices.

^bOver the wavelength range 1530 nm to 1565 nm.

^cDifferential Group Delay (DGD) is the time difference at reception between the fractions of a pulse that were transmitted in the two principal states of polarization of an optical signal. DGD_max is the maximum differential group delay that the system must tolerate.

Table 89–14—Optical fiber and cable characteristics for 40GBASE-FR

Description	Value	Unit
Nominal fiber specification wavelength	1550	nm
Cabled optical fiber attenuation (max)	0.5 ^a	dB/km
Zero dispersion wavelength (λ_0)	$1300 \leq \lambda_0 \leq 1324$	nm
Dispersion slope (max) (S_0)	0.093	ps/nm ² km

^aThe 0.5 dB/km attenuation is provided for Outside Plant cable as defined in ANSI/TIA-568-C.3-2008.

89.10.2 Optical fiber connection

An optical fiber connection, as shown in Figure 89–5, consists of a mated pair of optical connectors.

89.10.2.1 Connection insertion loss

The maximum link distances for single-mode fiber are calculated based on an allocation of 3 dB total connection and splice loss. For example, this allocation supports six connections with an average insertion loss per connection of 0.5 dB. Connections with different loss characteristics may be used provided the requirements of Table 89–13 are met.

89.10.2.2 Maximum discrete reflectance

The maximum discrete reflectance shall be less than –26 dB.

89.10.3 Medium Dependent Interface (MDI) requirements

The 40GBASE-FR PMD is coupled to the fiber optic cabling at the MDI. The MDI is the interface between the PMD and the “fiber optic cabling” (as shown in Figure 89–5). Examples of an MDI include the following:

- a) Connectorized fiber pigtail
- b) PMD receptacle

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

NOTE—Transmitter compliance testing is performed at TP2 as defined in 89.5.1, not at the MDI.

89.11 Protocol implementation conformance statement (PICS) proforma for Clause 89, Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR¹⁸

89.11.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 89, Physical Medium Dependent sublayer and medium, type 40GBASE-FR, 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.

89.11.2 Identification

89.11.2.1 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).	

89.11.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 89, Physical Medium Dependent sublayer and medium, type 40GBASE-FR
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 802.3-2012.)	

Date of Statement	
-------------------	--

¹⁸Copyright 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.

89.11.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*INS	Installation / cable	89.9	Items marked with INS include installation practices and cable specifications not applicable to a PHY manufacturer	O	Yes [] No []
XLDC	Delay constraints	89.3.1	Device conforms to delay constraints	M	Yes []
XLSC	Skew constraints	89.3.2	Device conforms to Skew and Skew Variation constraints	M	Yes []
*MD	MDIO capability	89.4	Registers and interface supported	O	Yes [] No []

89.11.4 PICS proforma tables for Physical Medium Dependent (PMD) sublayer and medium, type 40GBASE-FR

89.11.4.1 PMD functional specifications

Item	Feature	Subclause	Value/Comment	Status	Support
XLF1	Compatible with 40GBASE-R PCS and PMA	89.1		M	Yes []
XLF2	Integration with management functions	89.1		O	Yes [] No []
XLF3	Transmit function	89.5.2	Conveys bits from PMD service interface to MDI	M	Yes []
XLF4	Mapping between optical signal and logical signal for transmitter	89.5.2	Higher optical power is a one	M	Yes []
XLF5	Receive function	89.5.3	Conveys bits from MDI to PMD service interface	M	Yes []
XLF6	Mapping between optical signal and logical signal for receiver	89.5.3	Higher optical power is a one	M	Yes []
XLF7	Global Signal Detect function	89.5.4	Report to the PMD service interface the message PMD:IS_SIGNAL.indication(SIGNAL_DETECT)	M	Yes []
XLF8	Global Signal Detect behavior	89.5.4	SIGNAL_DETECT is an indicator of the presence of an optical signal	M	Yes []
XLF9	PMD_reset function	89.5.5	Resets the PMD sublayer	MD:O	Yes [] No [] N/A []

89.11.4.2 Management functions

Item	Feature	Subclause	Value/Comment	Status	Support
XLM1	Management register set	89.4		MD:M	Yes [<input type="checkbox"/> N/A [<input type="checkbox"/>
XLM2	Global transmit disable function	89.5.6	Disables the optical transmitter with the PMD_global_transmit_disable variable	MD:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/> N/A [<input type="checkbox"/>
XLM3	PMD_fault function	89.5.7	Sets PMD_fault to one if any local fault is detected	MD:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/> N/A [<input type="checkbox"/>
XLM4	PMD_transmit_fault function	89.5.8	Sets PMD_transmit_fault to one if a local transmitter fault is detected	MD:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/> N/A [<input type="checkbox"/>
XLM5	PMD_receive_fault function	89.5.9	Sets PMD_receive_fault to one if a local receiver fault is detected	MD:O	Yes [<input type="checkbox"/> No [<input type="checkbox"/> N/A [<input type="checkbox"/>

89.11.4.3 PMD to MDI optical specifications for 40GBASE-FR

Item	Feature	Subclause	Value/Comment	Status	Support
XLLR1	Transmitter meets specifications in Table 89-6	89.6.1	Per definitions in 89.7	M	Yes [<input type="checkbox"/> N/A [<input type="checkbox"/>
XLLR2	Receiver meets specifications in Table 89-7	89.6.2	Per definitions in 89.7	M	Yes [<input type="checkbox"/> N/A [<input type="checkbox"/>

89.11.4.4 Optical measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
XLOM1	Measurement cable	89.7	2 m to 5 m in length	M	Yes []
XLOM2	Center wavelength	89.7.3	Per TIA-455-127-A or IEC 61280-1-3	M	Yes []
XLOM3	Average optical power	89.7.4	Per IEC 61280-1-1	M	Yes []
XLOM4	Dispersion penalty	89.7.5.3		M	Yes []
XLOM5	Extinction ratio	89.7.6	Per IEC 61280-2-2	M	Yes []
XLOM6	RIN ₂₀ OMA	89.7.7		M	Yes []
XLOM7	Transmit eye	89.7.8		M	Yes []
XLOM8	Receiver sensitivity	89.7.9		M	Yes []
XLOM9	Receiver jitter tolerance	89.7.10		M	Yes []
XLOM10	Receiver 3 dB electrical upper cutoff frequency	89.7.11		M	Yes []

89.11.4.5 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
XLES1	General safety	89.8.1	Conforms to IEC 60950-1	M	Yes []
XLES2	Laser safety—IEC Hazard Level 1	89.8.2	Conforms to Hazard Level 1 laser requirements defined in IEC 60825-1 and IEC 60825-2	M	Yes []
XLES3	Electromagnetic interference	89.8.4.1	Complies with applicable codes for the limitation of electromagnetic interference	M	Yes []

89.11.4.6 Characteristics of the fiber optic cabling and MDI

Item	Feature	Subclause	Value/Comment	Status	Support
XLOC1	Fiber optic cabling	89.10	Meets the requirements specified in Table 89-13	INS:M	Yes [] N/A []
XLOC2	Maximum discrete reflectance	89.10.2.2	Less than -26 dB	INS:M	Yes [] N/A []
XLOC3	MDI requirements	89.10.3	Meets the interface performance specifications listed	INS:M	Yes [] N/A []

90. Ethernet support for time synchronization protocols

90.1 Introduction

This clause specifies the optional Time Synchronization Service Interface (TSSI). The TSSI can be used to support protocols that require knowledge of packet egress and ingress time.

The TSSI is defined for the full-duplex mode of operation only. It supports MAC operation at various data rates. The MII (Clause 22), GMII (Clause 35), XGMII (Clause 46), XLGMII (Clause 81) and CGMII (Clause 81) specifications are all compatible with the gRS sublayer defined in 90.5.

90.2 Overview

The goal of this clause is to provide an accurate indication of the transmission and reception initiation times of all packets, as required to support various time synchronization protocols, e.g., IEEE Std 1588-2008 [B47], and IEEE Std 802.1AS [B44].

The specific goals are to:

- a) define the Time Synchronization Service Interface (TSSI), with associated service primitives;
- b) add management registers to indicate the maximum and minimum data delays for estimation of link latency at the Time Synchronization Protocol (TimeSync) Client;

90.3 Relationship with other IEEE standards

Per 90.2, the TimeSync capability provides support for various time synchronization protocols, including e.g., IEEE Std 1588-2008 or IEEE Std 802.1AS. The definition of TimeSync Client, its capabilities and functions is outside the scope of IEEE Std 802.3.

90.4 Time Synchronization Service Interface (TSSI)

90.4.1 Introduction

This subclause specifies services provided by an extension to the Reconciliation Sublayers specified elsewhere in this standard.

90.4.1.1 Interlayer service interfaces

Figure 90–1 depicts the TimeSync Client and the RS interlayer service interfaces.

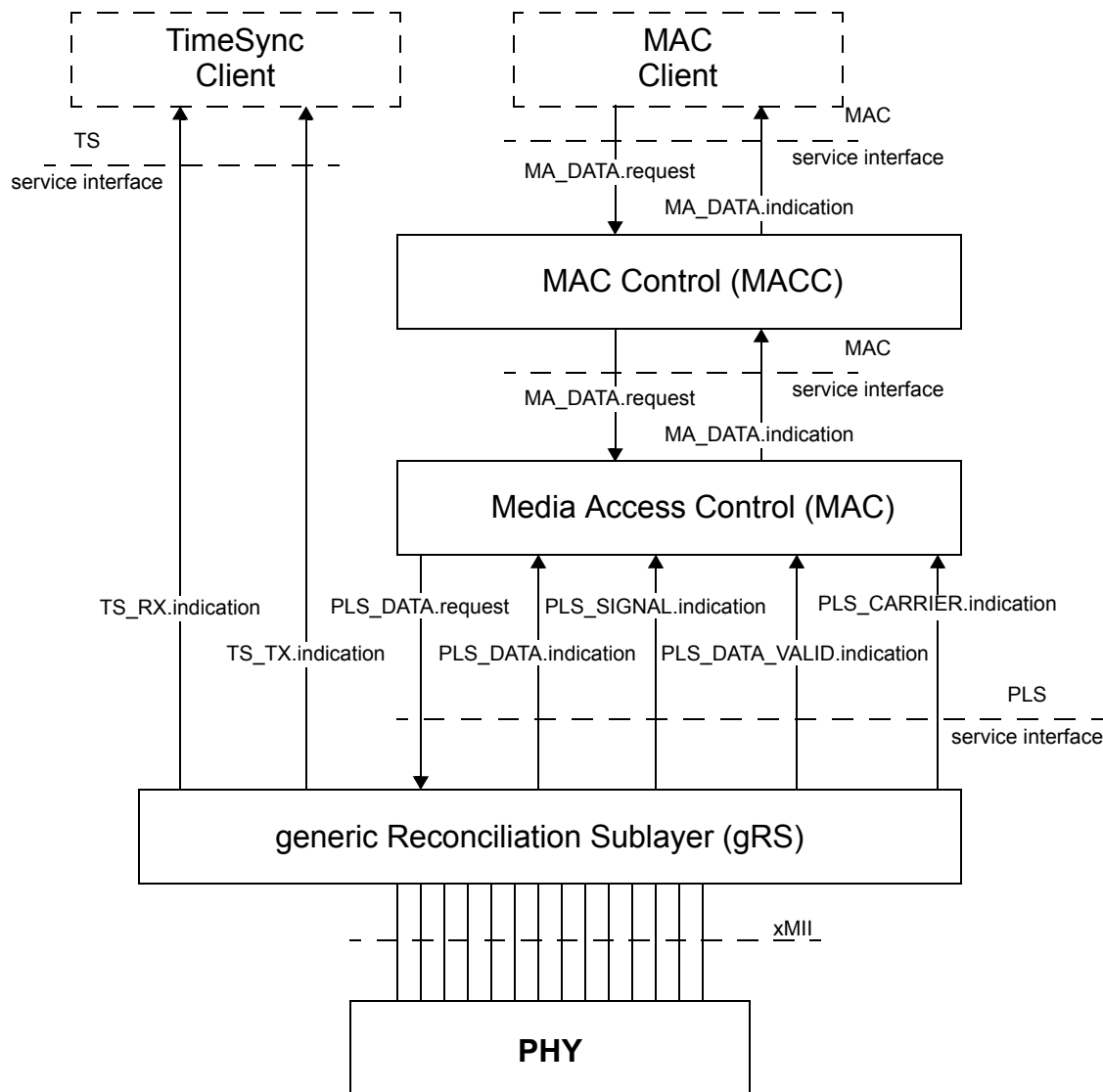


Figure 90–1—Relationship of the TimeSync Client, TSSI and gRS sublayer relative to MAC and MAC Client and associated interfaces

NOTE 1—In this figure, the xMII is used as a generic term for the Media Independent Interfaces for implementations of 100 Mb/s and above. For example: for 100 Mb/s implementations this interface is called MII; for 1 Gb/s implementations, it is called GMII; for 10 Gb/s implementations, it is called XGMII; etc.

NOTE 2—Optional Low Power Idle (LPI) Client service interface not shown.

90.4.1.2 Responsibilities of TimeSync Client

The TimeSync Client can use the indication of the event corresponding to the egress and ingress of packets provided by the TSSI, combined with knowledge of the protocol frames, to select the egress and ingress times relevant to the protocol. Which frames are of interest to any particular protocol is beyond the scope of this standard.

The TimeSync Client can use the indication of the event corresponding to the egress and ingress of packets at the xMII provided by the TSSI, together with the information provided by the TimeSync PHY transmit data delay and TimeSync PHY receive data delay (see 30.13.1).

90.4.2 TSSI

The following specifies the service interface provided by the RS to the TimeSync Client. These services are described in an abstract manner and do not imply any particular implementation. The model used in this service specification is identical to that used in 1.2.2.

The following primitives are defined:

- TS_TX.indication
- TS_RX.indication

90.4.3 Detailed service specification

90.4.3.1 TS_TX.indication primitive

This primitive defines the transfer of the indication of an event between gRS and the TimeSync Client.

90.4.3.1.1 Semantics

The semantics of the primitive are as follows:

TS_TX.indication(SFD)

The SFD parameter can take only one possible value, DETECTED. When asserted (SFD = DETECTED), the TimeSync Client is notified that a valid SFD was detected by the gRS sublayer TS_SFD_Detect_TX function (see 90.5.1) in the xMII transmit signals.

90.4.3.1.2 Condition for generation

This primitive is generated by the gRS sublayer in response to detection of a valid SFD in the data stream transmitted across the xMII transmit signals. Specific conditions for generation of this primitive are described in 90.5.1.

90.4.3.1.3 Effect of receipt

The receipt of this primitive by the TimeSync Client is undefined.

90.4.3.2 TS_RX.indication primitive

This primitive defines the transfer of the indication of an event between gRS and the TimeSync Client.

90.4.3.2.1 Semantics

The semantics of the primitive are as follows:

TS_RX.indication(SFD)

The SFD parameter can take only one possible value, DETECTED. When asserted (SFD = DETECTED), the TimeSync Client is notified that a valid SFD was detected by the gRS sublayer TS_SFD_Detect_RX function (see 90.5.2) in the xMII receive signals.

90.4.3.2.2 Condition for generation

This primitive is generated by the gRS sublayer in response to detection of a valid SFD in the data stream received across the xMII receive signals. Specific conditions for generation of this primitive are described in 90.5.2.

90.4.3.2.3 Effect of receipt

The receipt of this primitive by the TimeSync Client is undefined.

90.5 generic Reconciliation Sublayer (gRS)

Within the scope of this clause, the term generic Reconciliation Sublayer (gRS) is used to denote any IEEE 802.3 Reconciliation Sublayer (RS) used to interface a MAC with any PHY supporting the TimeSync capability through the xMII.

For the purpose of the TimeSync capabilities, two new functions are defined in this subclause, namely TS_SFD_Detect_TX (see 90.5.1) and TS_SFD_Detect_RX (see 90.5.2), which are responsible for generation of the TS_TX.indication and TS_RX.indication service primitives, as defined in 90.4. Figure 90–2 presents the TS_SFD_Detect_TX and TS_SFD_Detect_RX functions and their location within the RS sublayer.

90.5.1 TS_SFD_Detect_TX function

The TS_SFD_Detect_TX function observes the xMII transmit signals and detects the occurrence of the Start Frame Delimiter (SFD, see 3.1.1 and 3.2.2) in compliance with the specifications of the given type of the instantiated xMII.

The service primitive across the TSSI i.e. TS_TX.indication shall be generated only when the SFD is detected on the transmit signals of the xMII (SFD=DETECTED).

90.5.2 TS_SFD_Detect_RX function

The TS_SFD_Detect_RX function observes the xMII receive signals and detects the occurrence of the Start Frame (SFD, see 3.1.1 and 3.2.2) in compliance with the specifications of the given type of the instantiated xMII.

The service primitive across the TSSI i.e. TS_RX.indication shall be generated only when the SFD is detected on the receive signals of the xMII (SFD=DETECTED).

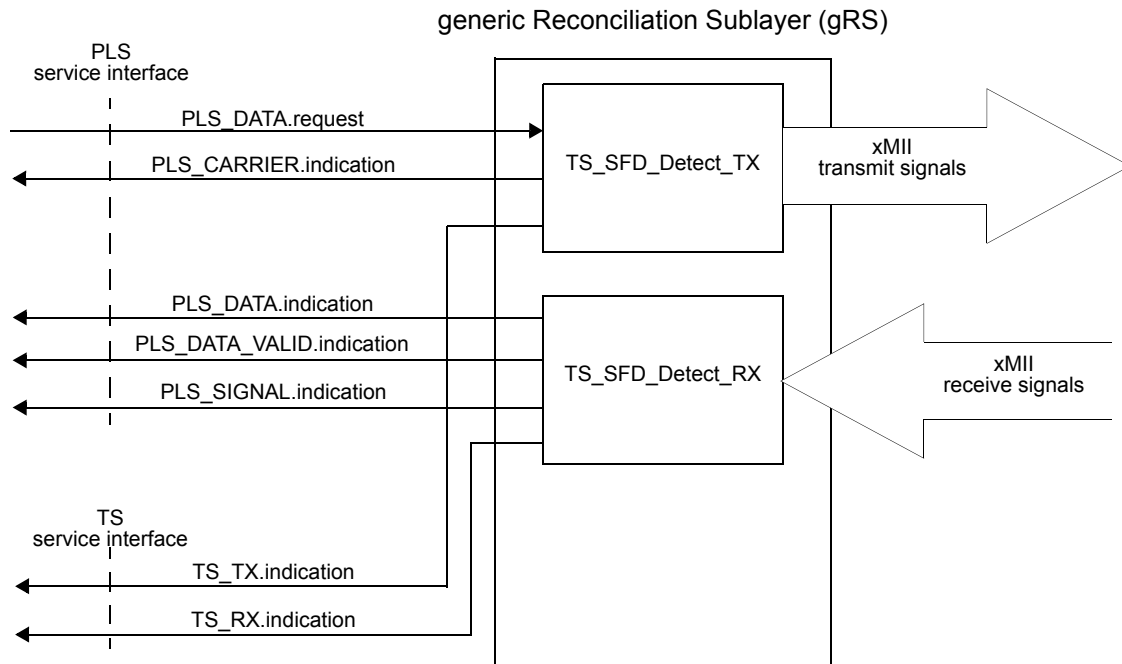


Figure 90–2—TS_SFD_Detect_TX and TS_SFD_Detect_RX functions within the generic Reconciliation Sublayer (gRS)

90.6 Overview of management features

Clause 30 describes the management functions for any of the IEEE Std 802.3 compliant PHYs. Objects defined in Clause 30 for the support of TimeSync capability are summarized below:

- oTimeSync managed object class, as defined in 30.13.1
- aTimeSyncCapabilityTX and aTimeSyncCapabilityRX managed objects, reflecting the status of a series of MDIO capability registers (1.1800, 2.1800, 3.1800, 4.1800, 5.1800, and 6.1800), as defined in 30.13.1.1 and 30.13.1.2, respectively
- aTimeSyncDelayTXmax and aTimeSyncDelayTXmin managed objects, representing the aggregate value of the series of transmit path data delay registers present in the instantiated MDIO registers (for each MMD, in case of multiple instances) 1.1801 through 1.1804, 2.1801 through 2.1804, 3.1801 through 3.1804, 4.1801 through 4.1804, 5.1801 through 5.1804, and 6.1801 through 6.1804, as defined in 30.13.1.3 and 30.13.1.4, respectively
- aTimeSyncDelayRXmax and aTimeSyncDelayRXmin managed objects, representing the aggregate value of the series of receive path data delay registers present in the instantiated MDIO registers (for each MMD, in case of multiple instances) 1.1805 through 1.1808, 2.1805 through 2.1808, 3.1805 through 3.1808, 4.1805 through 4.1808, 5.1805 through 5.1808, and 6.1805 through 6.1808, as defined in 30.13.1.5 and 30.13.1.6, respectively

The Management Data Input/Output (MDIO) capability described in Clause 45 defines several variables that provide TimeSync status information for the PMD, as shown in Table 90–1:

Table 90–1—Summary of TimeSync features in Clause 45

Register	Name	Reference
1.1800	TimeSync PMA/PMD capability register	45.2.1.104
1.1801 through 1.1804	TimeSync PMA/PMD transmit path data delay	45.2.1.105
1.1805 through 1.1808	TimeSync PMA/PMD receive path data delay	45.2.1.106
2.1800	TimeSync WIS capability register	45.2.2.20
2.1801 through 2.1804	TimeSync WIS transmit path data delay	45.2.2.21
2.1805 through 2.1808	TimeSync WIS receive path data delay	45.2.2.22
3.1800	TimeSync PCS capability register	45.2.3.48
3.1801 through 3.1804	TimeSync PCS transmit path data delay	45.2.3.49
3.1805 through 3.1808	TimeSync PCS receive path data delay	45.2.3.50
4.1800	TimeSync PHY XS capability register	45.2.4.12
4.1801 through 4.1804	TimeSync PHY XS transmit path data delay	45.2.4.13
4.1805 through 4.1808	TimeSync PHY XS receive path data delay	45.2.4.14
5.1800	TimeSync DTE XS capability register	45.2.5.12
5.1801 through 5.1804	TimeSync DTE XS transmit path data delay	45.2.5.13
5.1805 through 5.1808	TimeSync DTE XS receive path data delay	45.2.5.14
6.1800	TimeSync TC capability register	45.2.6.14
6.1801 through 6.1804	TimeSync TC transmit path data delay	45.2.6.15
6.1805 through 6.1808	TimeSync TC receive path data delay	45.2.6.16

90.7 Data delay measurement

The TimeSync capability requires measurement of data delay in the transmit and receive paths, as shown in Figure 90–3. The transmit path data delay is measured from the input of the beginning of the SFD at the xMII to its presentation by the PHY to the MDI. The receive path data delay is measured from the input of the beginning of the SFD at the MDI to its presentation by the PHY to the xMII.

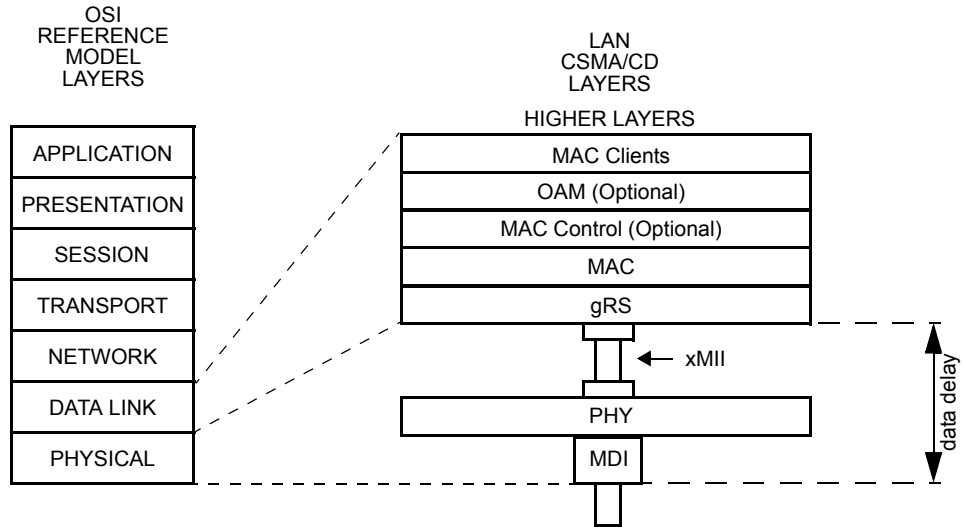


Figure 90–3—Data delay measurement

NOTE-The measured data delay in either the transmit or receive path should not include delay resulting from any length of the medium, regardless of the type of the medium used by the given PHY. It is observed that even if any medium-imposed delay is included in the measured data delay, it does not affect the operation of the TimeSync functions and only shifts the timing reference point farther from the MDI.

The obtained data delay measurement shall be reported in the form of a quartet of values; the maximum transmit data delay, the minimum transmit data delay, the maximum receive data delay, and the minimum receive data delay, as defined for the oTimeSync managed object class (30.13.1).

NOTE-The data delay values represent only the data delay in the PHY sublayers. The TimeSync Client may need to adjust for delays within the gRS. For example, the TimeSync Client may need to subtract the value of the delay associated with the TS_SFD_Detect_TX function from the sum of the minimum transmit data delay values reported by individual MMD(s). Likewise, the TimeSync Client may need to add the value of the delay associated with the TS_SFD_Detect_RX function to the sum of the maximum receive data delay values reported by individual MMD(s).

90.8 Protocol implementation conformance statement (PICS) proforma for Clause 90, Ethernet support for time synchronization protocols¹⁹

90.8.1 Introduction

The supplier of a protocol implementation that is claimed to conform to Clause 90 Ethernet support for time synchronization protocols, 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.

90.8.2 Identification

90.8.2.1 Implementation identification

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

90.8.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Clause 90, Ethernet support for time synchronization protocols
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 <input type="checkbox"/> Yes <input type="checkbox"/> (See Clause 21; the answer Yes means that the implementation does not conform to IEEE Std 802.3-2012)	
Date of Statement	

¹⁹Copyright 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.

90.8.3 TSSI indication

Item	Feature	Subclause	Value/Comment	Status	Support
TS_TX	TS_TX.indication generation	90.5.1	TS_TX.indication is generated only when the SFD is detected on the transmit signals of the xMII.	M	Yes []
TS_RX	TS_RX.indication generation	90.5.2	TS_RX.indication is generated only when the SFD is detected on the receive signals of the xMII.	M	Yes []

90.8.4 Data delay reporting

Item	Feature	Subclause	Value/Comment	Status	Support
Report	Data delay reporting format	90.7	Data delay measurement is reported in the form of a quartet of values; the maximum transmit data delay, the minimum transmit data delay, the maximum receive data delay, and the minimum receive data delay, for each MMD instantiated in the device.	M	Yes []

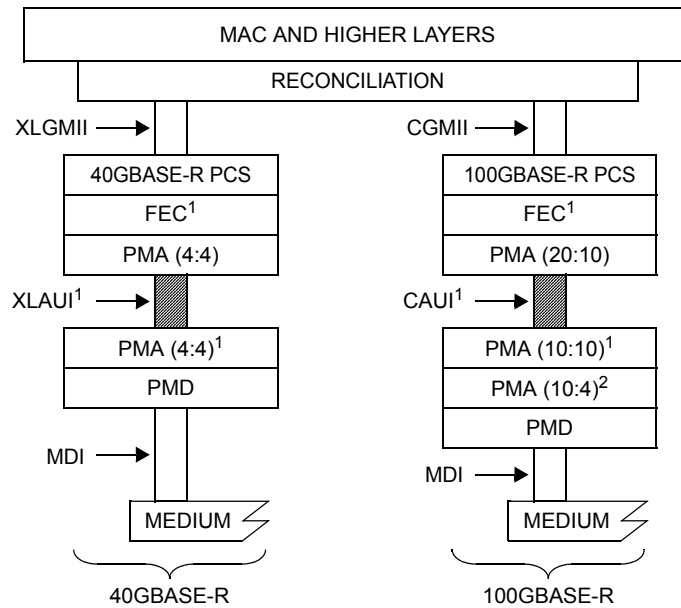
Annex 83A

(normative)

40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)

83A.1 Overview

This annex defines the functional and electrical characteristics for the optional 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI). Figure 83A–1 shows the relationships of the XLGMII, PMA, XLAUI, and PMD for 40 Gb/s and CGMII, PMA, CAUI, and PMD for 100 Gb/s.



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

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

Figure 83A–1—Example relationship of XLAUI and CAUI to IEEE 802.3 CSMA/CD LAN model

The purpose of the optional XLAUI or CAUI is to provide a flexible chip-to-chip and chip-module interconnect for 40 Gb/s or 100 Gb/s components. Annex 83A provides compliance requirements for XLAUI/CAUI transmitters and receivers while Annex 83B specifies the electrical requirements for the chip-module interconnection.

The XLAUI/CAUI allows interconnect distances of approximately 25 cm over printed circuit board including one connector, see 83A.4.

An example application of CAUI includes providing a physical connection between a ten-lane 100 Gb/s PMA and a 10:4 PMA mapping element. An example application of XLAUI is to provide lane extension for interfacing MAC and PHY components in a 40 Gb/s Ethernet system distributed across a circuit board.

The optional XLAUI/CAUI interface has the following characteristics:

- a) Independent transmit and receive data paths
- b) Differential AC coupled signaling with low voltage swing
- c) Self-timed interface
- d) Shared technology with other 40 Gb/s or 100 Gb/s interfaces
- e) Utilization of 64B/66B coding

83A.1.1 Summary of major concepts

The following is a list of the major concepts of XLAUI and CAUI:

- a) The optional XLAUI/CAUI interface can be inserted between PMA layers in the IEEE 802.3 CSMA/CD LAN model to transparently enable chip-to-chip communication
- b) The XLAUI is organized into four lanes, the CAUI is organized into ten lanes
- c) The XLAUI/CAUI interface is a parallel electrical interface with each lane running at a nominal rate of 10.3125 Gb/s

83A.1.2 Rate of operation

The XLAUI interface supports the 40 Gb/s data rate and the CAUI interface supports the 100 Gb/s data rate. For 40 Gb/s applications, the data stream shall be presented in four lanes as described in Clause 83. For 100 Gb/s applications, the data stream shall be presented in ten lanes as described in Clause 83. The data is 64B/66B coded. The nominal signaling rate for each lane in both 40 Gb/s and 100 Gb/s applications shall be 10.3125 Gb/s.

83A.2 XLAUI/CAUI link block diagram

XLAUI/CAUI link is illustrated in Figure 83A–2. XLAUI/CAUI channel is defined from the transmit pad to the receive pad including any AC coupling in the path.

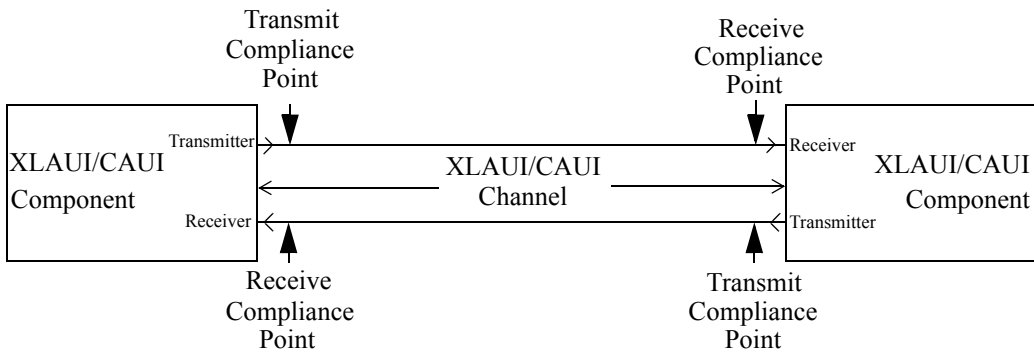


Figure 83A-2—Definition of transmit and receive compliance points

83A.2.1 Transmitter compliance points

The reference differential insertion loss, expressed in decibels, between the transmitter and the transmit compliance point is defined in Equation (83A-1) and illustrated in Figure 83A-3. The effects of differences between the actual insertion loss and the reference insertion loss are to be accounted in the measurements.

$$Insertion_loss(f) = -0.00086 + 0.2286\sqrt{f} + 0.08386f \text{ dB} \quad 0.01 \leq f \leq 11.1 \quad (83A-1)$$

where

Insertion_loss(f) is the differential insertion loss at frequency *f*
f is the frequency in GHz

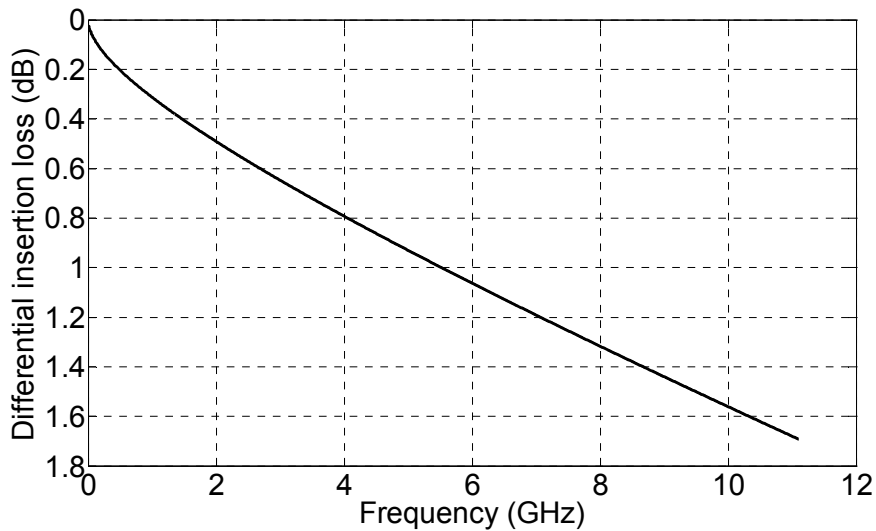


Figure 83A-3—Insertion loss between transmitter and transmit compliance point

83A.2.2 Receiver compliance points

The reference differential insertion loss, expressed in decibels, between the receiver compliance point and the receiver is defined in Equation (83A-2) and illustrated in Figure 83A-4. The effects of differences between the actual insertion loss and the reference insertion loss are to be accounted in the measurements.

$$Insertion_loss(f) = -0.00086 + 0.2286\sqrt{f} + 0.08386f \text{ dB} \quad 0.01 \leq f \leq 11.1 \quad (83A-2)$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

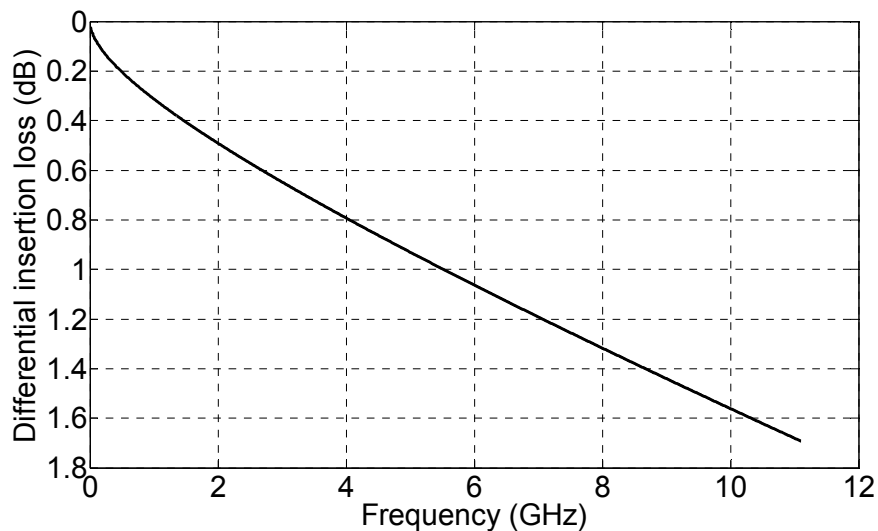


Figure 83A-4—Insertion loss between receive compliance point and receiver

83A.3 XLAUI/CAUI electrical characteristics

The electrical characteristics of the XLAUI/CAUI interface are specified such that they can be applied within a variety of 40 Gb/s Ethernet or 100 Gb/s Ethernet equipment types. The electrical characteristics for XLAUI/CAUI shall meet the specifications defined in 83A.3.1, 83A.3.2, 83A.3.3, and 83A.3.4.

83A.3.1 Signal levels

The XLAUI/CAUI is a low-swing AC coupled differential interface. AC coupling allows for interoperability between components operating from different supply voltages. Differential signal swings are defined in the following subclauses, and depend on several factors such as transmitter de-emphasis and transmission line losses.

83A.3.2 Signal paths

The XLAUI/CAUI signal paths are point-to-point connections. Each path corresponds to a XLAUI/CAUI lane, and is comprised of two complementary signals making a balanced differential pair. For XLAUI, there are four differential paths in each direction for a total of eight pairs, or sixteen connections. For CAUI, there are ten differential paths in each direction for a total of twenty pairs, or 40 connections.

83A.3.3 Transmitter characteristics

The XLAUI/CAUI transmitter characteristics measured at the transmitter compliance point are specified in Table 83A-1. The XLAUI/CAUI signaling rate shall be the signaling rate defined in Table 83A-1.

Table 83A-1—Transmitter characteristics

Parameter	Subclause reference	Value	Units
Signaling rate per lane (range)	—	10.3125 ± 100 ppm	GBd
Single-ended output voltage maximum minimum	83A.3.3.1	4 -0.4	V V
Maximum differential output voltage, peak-to-peak	83A.3.3.1	760	mV
Minimum de-emphasis	83A.3.3.1	4.4	dB
Maximum de-emphasis	83A.3.3.1	7	dB
Minimum VMA	83A.3.3.1	See Equation (83A-4)	mV
Maximum termination mismatch at 1 MHz	86A.5.3.2	5	%
Maximum output AC common-mode voltage, RMS	86A.5.3.1	15	mV
Minimum output rise and fall time (20% to 80%)	83A.3.3.2	24	ps
Differential output return loss	83A.3.3.3	See Equation (83A-5)	dB
Common-mode output return loss	83A.3.3.4	See Equation (83A-6)	dB
Maximum Total Jitter	83A.3.3.5	0.32	UI
Maximum Deterministic Jitter	83A.3.3.5	0.17	UI
Transmitter eye mask definition X1	83A.3.3.5	0.16	UI
Transmitter eye mask definition X2	83A.3.3.5	0.38	UI
Transmitter eye mask definition Y1	83A.3.3.5	200	mV
Transmitter eye mask definition Y2	83A.3.3.5	380	mV

83A.3.3.1 Output amplitude

Driver differential output amplitude shall be less than the maximum differential output voltage defined in Table 83A-1 including any transmit de-emphasis. DC-referenced logic levels are not defined since the receiver is AC coupled. Single-ended output voltage shall be within the range specified in Table 83A-1 with respect to ground.

De-emphasis shall be greater than the minimum de-emphasis and less than the maximum de-emphasis defined in table Table 83A-1. De-emphasis is defined as the ratio between the amplitude following a transition and the amplitude during a non-transition bit as seen in Equation (83A-3). VMA is defined in 86A.5.3.5 using the square wave or PRBS9 (Pattern 4) defined in Table 86-11.

See Figure 83A-5 for an illustration of absolute driver output voltage limits, definition of differential peak-to-peak amplitude, and definition of the parameters used to calculate de-emphasis. $SL_{i<P>}$ and $SL_{i<N>}$ are positive and negative sides of a differential signal pair for lane i ($i = 0, 1, 2, 3$ for XLAUI. For CAUI $i = 0:9$).

$$\text{De-emphasis (dB)} = 20\log_{10}\left(\frac{\text{Differential peak-to-peak amplitude}}{\text{VMA}}\right) \quad (83A-3)$$

$$\text{Minimum VMA (mV)} = (234.64 - 2.13x + 0.18x^2) \times 1.32(10^{-y/20}) \quad (83A-4)$$

where

- x is the rise or fall time (whichever is larger) in ps
- y is de-emphasis value in dB

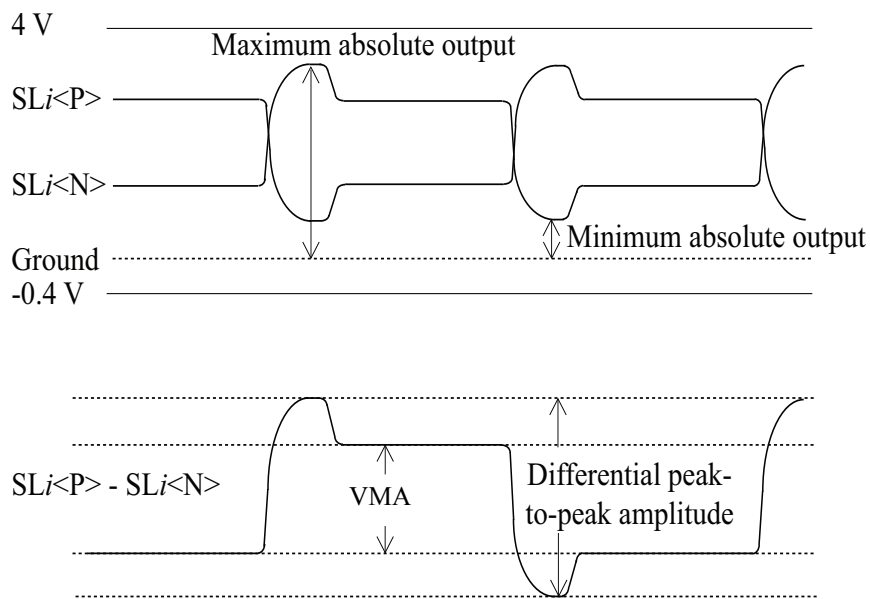


Figure 83A-5—Driver output voltage limits and definitions

83A.3.3.2 Rise/fall time

Differential rise/fall times shall be greater than the minimum output rise and fall time defined in Table 83A-1, as measured from the 20% to the 80% levels. Shorter transitions may result in excessive high-frequency components and increase EMI and crosstalk. The upper limit is defined by the transmit eye mask shown in Figure 83A-8. Rise/fall time is measured with de-emphasis off as defined in 83A.5.1. Rise/fall time measurements are taken using a square wave test pattern as defined in 83.5.10.

83A.3.3.3 Differential output return loss

Differential output return loss shall meet the requirements defined in Table 83A-1. Differential output return loss is given in Equation (83A-5) and is illustrated Figure 83A-6. Differential output return loss includes contributions from on-chip circuitry, chip packaging, and any off-chip components related to the driver. This output impedance requirement applies to all valid output levels. The reference impedance for differential return loss measurements is 100 Ω. The maximum termination mismatch at 1 MHz shall be less than the requirement defined in Table 83A-1.

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 & 0.01 \leq f < 2.125 \\ 6.5 - 13.33 \log_{10} \left(\frac{f}{5.5} \right) & 2.125 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \tag{83A-5}$$

where

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

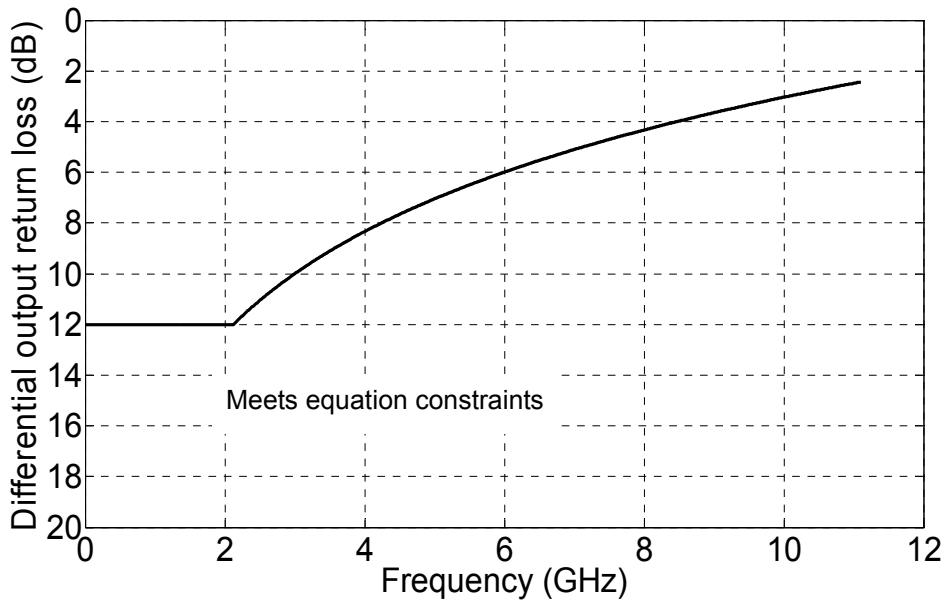


Figure 83A-6—Differential output return loss

83A.3.3.4 Common-mode output return loss

Common-mode output return loss shall meet the requirements defined in Table 83A-1. Common-mode output return loss is given in Equation (83A-6) and is illustrated in Figure 83A-7. Common-mode output return loss includes contributions from on-chip circuitry, chip packaging, and any off-chip components

related to the driver. This output impedance requirement applies to all valid output levels. The reference impedance for common-mode return loss measurements is 25 Ω

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 9 & 0.01 \leq f < 2.125 \\ 3.5 - 13.33 \log_{10} \left(\frac{f}{5.5} \right) & 2.125 \leq f < 7.1 \\ 2 & 7.1 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83A-6)$$

where

$Return_loss(f)$ is the common-mode output return loss at frequency f
 f is the frequency in GHz

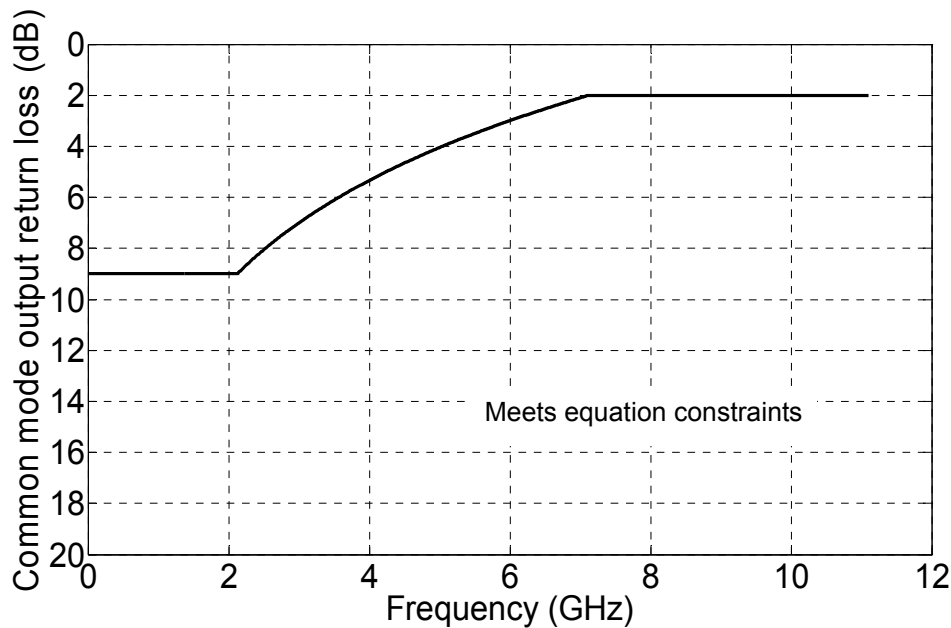


Figure 83A-7—Common-mode output return loss

83A.3.3.5 Transmitter eye mask and transmitter jitter definition

The measured transmit signal at the transmit compliance point shall meet the eye template specified in Figure 83A-8 and Table 83A-1. The template measurement requirements are specified in 83A.5.1. The measured jitter at the transmit compliance point shall be less than the maximum Total Jitter as defined in Table 83A-1 and a maximum Deterministic Jitter as defined in Table 83A-1. Jitter and eye mask measurement requirements are described in 83A.5.1 and are conducted with de-emphasis off.

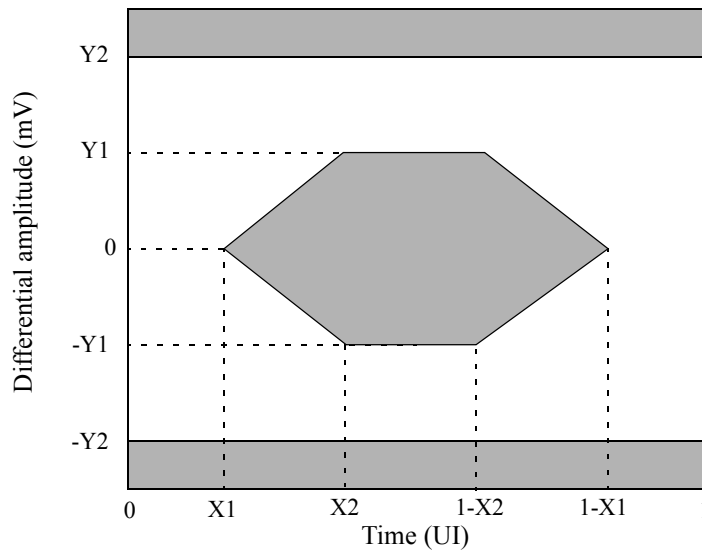


Figure 83A-8—Transmitter eye mask

83A.3.4 Receiver characteristics

Receiver characteristics at the receiver compliance point are specified in Table 83A-2 and detailed in 83A.3.4.1 through 83A.3.4.6.

Table 83A-2—Receiver characteristics

Parameter	Subclause reference	Value	Units
Signaling rate per lane (range)	—	10.3125 ± 100 ppm	GBd
Minimum input AC common-mode voltage tolerance, RMS	86A.5.3.1	20	mV
Minimum input rise and fall time tolerance (20% to 80%)	83A.3.3.2	24	ps
Differential input return loss	83A.3.4.3	See Equation (83A-7)	dB
Differential to common-mode input return loss	83A.3.4.4	See Equation (83A-8)	dB
Stressed receiver tolerance			
Minimum Total Input Jitter Tolerance	83A.3.4.2	0.62	UI
Minimum Deterministic Input Jitter Tolerance	83A.3.4.2	0.42	UI
Receiver eye mask definition X1	83A.3.4.2	0.31	UI
Receiver eye mask definition X2	83A.3.4.2	0.5	UI
Receiver eye mask definition Y1	83A.3.4.2	42.5	mV
Receiver eye mask definition Y2	83A.3.4.2	425	mV

83A.3.4.1 Bit error ratio

The receiver shall operate with a BER of better than 10^{-12} in the presence of a compliant input signal as defined in 83A.3.4.2.

NOTE—A transmitter capable of operating at $TJ = 0.3$ UI and $DJ = 0.16$ UI and receiver capable of operating at stress jitter tolerance of $TJ = 0.64$ UI and $DJ = 0.41$ UI would have sufficient margin for operation at approximately BER 10^{-15} .

83A.3.4.2 Input signal definition

A compliant input signal to a XLAUI/CAUI receiver has characteristics determined by a compliant XLAUI/CAUI driver and channel. The input signal definition satisfies the eye mask defined in Table 83A-2 and Figure 83A-9. Input signal jitter does not exceed the jitter tolerance requirements specified in 83A.3.4.6. Stressed receiver measurement requirements are specified in 83A.5.2.

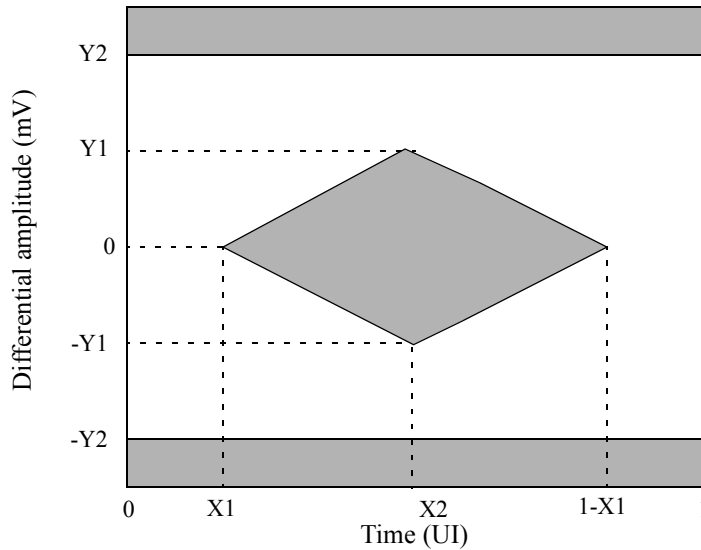


Figure 83A-9—Receiver template

83A.3.4.3 Differential input return loss

Differential input return loss shall meet the requirements defined in Table 83A-2. Differential input return loss is given in Equation (83A-7) and is illustrated in Figure 83A-10. Differential input return loss includes contributions from on-chip circuitry, chip packaging, and any off-chip components related to the receiver. The reference impedance for differential return loss measurements is 100Ω .

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 & 0.01 \leq f < 2.125 \\ 6.5 - 13.33 \log_{10} \left(\frac{f}{5.5} \right) & 2.125 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83A-7)$$

where

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

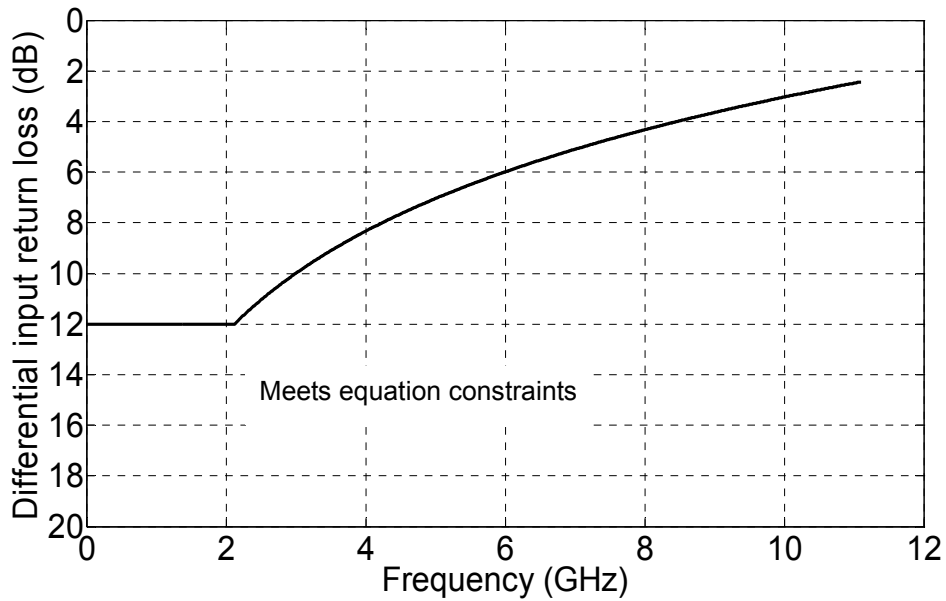


Figure 83A-10—Differential input return loss

83A.3.4.4 Differential to common-mode input return loss

Differential to common-mode input return loss shall meet the requirements defined in Table 83A-2. Differential to common-mode input return loss is given in Equation (83A-8) and is illustrated in Figure 83A-11. Differential to common-mode input return loss includes contributions from on-chip circuitry, chip packaging, and any off-chip components related to the receiver. The reference impedance for differential return loss measurements is 100 Ω. The reference impedance for common-mode return loss measurements is 25 Ω.

$$Return_loss(f) \geq 15 \quad 0.01 \leq f < 11.1 \quad (\text{dB}) \quad (83A-8)$$

where

Return_loss(f) is the differential to common-mode input return loss at frequency *f*
f is the frequency in GHz

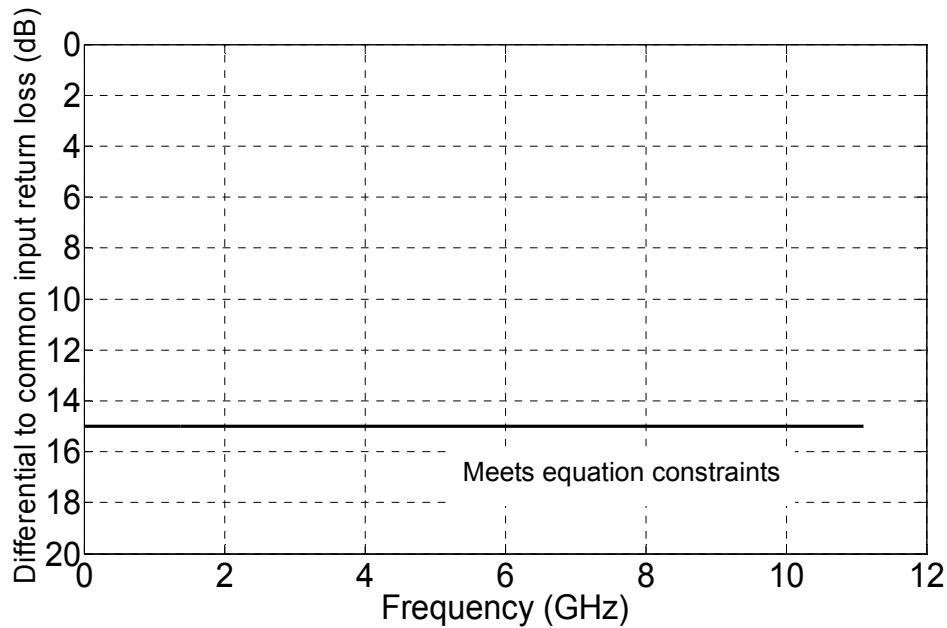


Figure 83A-11—Differential to common-mode input return loss

83A.3.4.5 AC coupling

The XLAUI/CAUI receiver shall be AC coupled to the XLAUI/CAUI transmitter to allow for maximum interoperability between various 10 Gb/s components. AC coupling is considered to be part of the receiver for the purposes of this specification except when interfacing with modules defined in Annex 83B or explicitly stated otherwise. It should be noted that there may be various methods for AC coupling in actual implementations.

83A.3.4.6 Jitter tolerance

The XLAUI/CAUI receiver shall have a peak-to-peak total jitter amplitude tolerance of at least the minimum total input jitter tolerance defined in Table 83A-2. This total jitter is composed of two components: deterministic jitter and random jitter. Deterministic jitter tolerance shall be at least the minimum deterministic input jitter tolerance defined in Table 83A-2. The XLAUI/CAUI receiver shall tolerate sinusoidal jitter with any frequency and amplitude defined by the mask of Figure 83A-12. This sub-component of deterministic jitter is intended to ensure margin for low-frequency jitter, wander, noise, crosstalk, and other variable system effects.

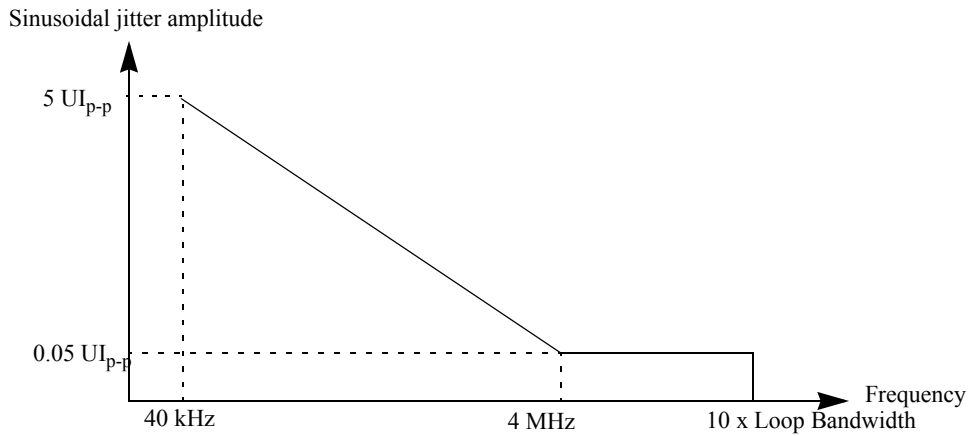


Figure 83A-12—Single-tone sinusoidal jitter mask

83A.4 Interconnect characteristics

This section describes recommended characteristics which are used to characterize a XLAUI/CAUI channel as shown in Figure 83A-2. The value for differential insertion loss is summarized in Equation (83A-9) and illustrated in Figure 83A-13. The value for minimum return loss is summarized in Equation (83A-10) and illustrated in Figure 83A-14. The channel is terminated with 100 Ω differential impedance. Other impairments such as crosstalk can have a material impact on the link performance and should be minimized.

$$Insertion_loss(f) < \left\{ \begin{array}{ll} 0.15 + 1.39\sqrt{f} + 1.4f & 0.01 \leq f < 7 \\ -15.86 + 4.2f & 7 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \tag{83A-9}$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12.5 & 0.01 \leq f < 5 \\ 12.5 - 27.5 \log_{10}\left(\frac{f}{5}\right) & 5 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \tag{83A-10}$$

where

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

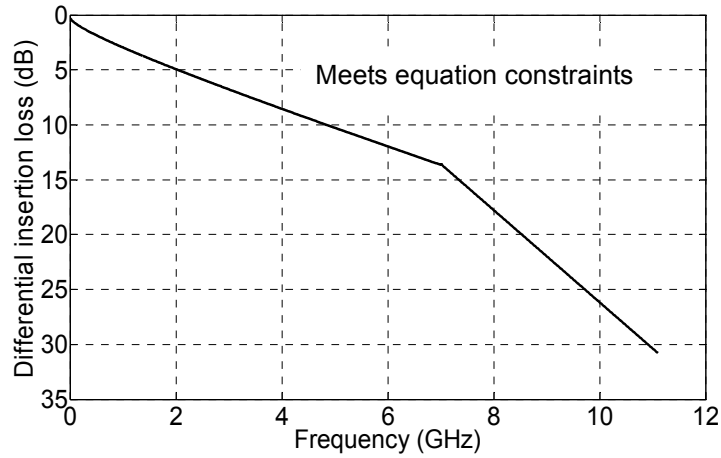


Figure 83A-13—Differential insertion loss

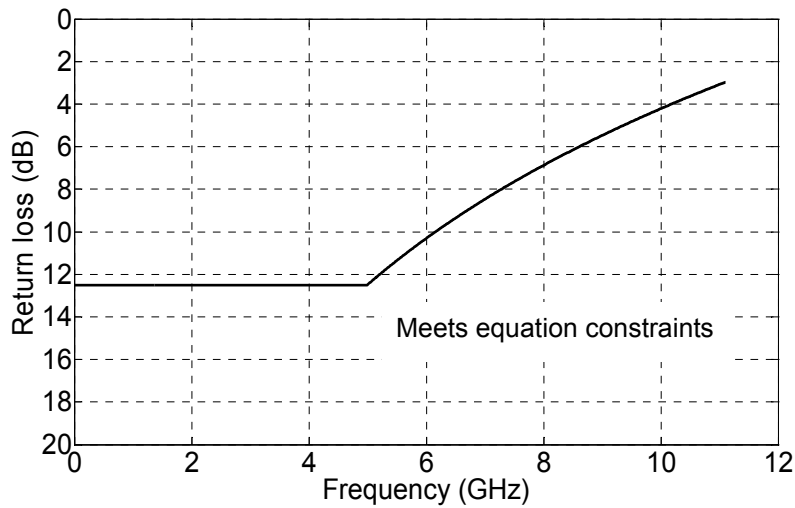


Figure 83A-14—Differential input return loss

The XLAUI/CAUI is primarily intended as a point-to-point interface of up to approximately 25 cm between integrated circuits using controlled impedance traces on low-cost printed circuit boards (PCBs). Longer reaches for the XLAUI/CAUI may be achieved by the use of better PCB materials, as the performance of an actual XLAUI/CAUI interconnect is highly dependent on the implementation.

83A.4.1 Characteristic impedance

The recommended differential characteristic impedance of circuit board trace pairs is $100 \Omega \pm 10\%$.

83A.5 Electrical parameter measurement methods

This subclause describes the measurement methodology that is to be used to verify XLAUI/CAUI compliance. Eye templates are measured with AC coupling and centered at 0 V differential. The signal waveform, eye, and jitter may be measured using a receiver with at least an equivalent 12 GHz low-pass filter response. Jitter values and eye masks are specified for BER 10^{-12} .

83A.5.1 Transmit jitter

Transmit jitter is defined with respect to a test procedure resulting in a BER bathtub curve such as that described in Annex 48B.3. For the purpose of jitter measurement, the effect of a single-pole high-pass filter with a 3 dB point at 4 MHz is applied to the jitter. The test pattern for jitter measurements shall be PRBS31 test pattern in 83.5.10 or scrambled idle in 82.2.10. Crossing times are defined with respect to the mid-point (0 V) of the AC coupled differential signal. De-emphasis shall be off during jitter testing. Transmit de-emphasis off state is defined by any setting that gives optimal performance for transmitter jitter and eye mask evaluation. All XLAUI/CAUI transmitter lanes shall be active and all XLAUI/CAUI receive lanes shall be receiving maximum amplitude and fastest rise time (as defined in Table 83A-1) during transmit jitter testing to ensure maximum lane-lane crosstalk is included in the jitter evaluation.

83A.5.2 Receiver tolerance

The XLAUI/CAUI jitter tolerance test setup in Figure 83A-15 or its equivalent shall meet the receiver eye mask defined in Table 83A-2. Applied jitter is measured using the methodology described in Annex 48B.3. Deterministic jitter is added to a clean test pattern by adding sinusoidal jitter as defined in 83A.3.4.6, along with low-pass filter stress, followed by a limiting function, and frequency-dependent attenuation. The low-pass filter stress is added until the 0.34 UI Deterministic Jitter is achieved. Frequency-dependent attenuation is then added using PCB trace or frequency-dependent attenuation which emulates PCB loss. Frequency-dependent attenuation is added until 0.42 UI Deterministic Jitter is achieved. Random jitter is added to the test signal using an interference generator which is a broadband noise source capable of producing white Gaussian noise with adjustable amplitude. The power spectral density shall be flat to ± 3 dB from 50 MHz to 6 GHz with a crest factor of no less than 5. The amplitude and output jitter of the filter stress plus limiter and random jitter injection shall meet the receiver eye mask defined in Table 83A-2. All XLAUI/CAUI lanes shall be active during jitter tolerance testing. The PRBS31 pattern defined in 83.5.10 or scrambled idle defined in 82.2.10 is used for evaluating XLAUI/CAUI jitter tolerance.

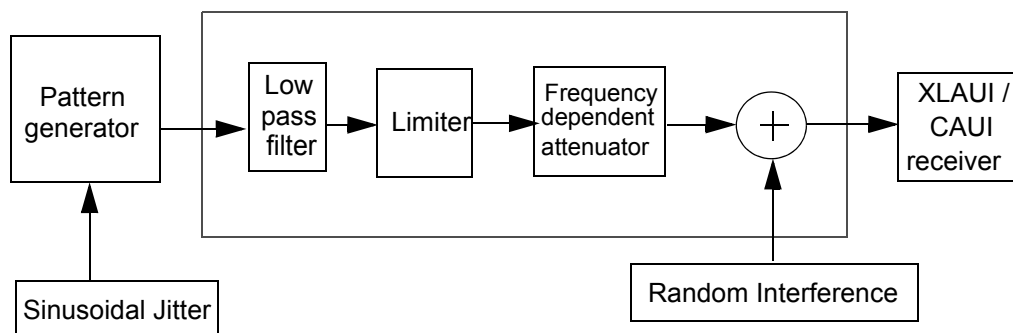


Figure 83A-15—Stressed-eye and jitter tolerance test setup

83A.6 Environmental specifications

83A.6.1 General safety

All equipment subject to this clause shall conform to applicable sections (including isolation requirements) of IEC 60950-1.

83A.6.2 Network safety

The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

83A.6.3 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.

83A.6.4 Electromagnetic compatibility

A system integrating the XLAUI/CAUI shall comply with applicable local and national codes for the limitation of electromagnetic interference.

83A.6.5 Temperature and humidity

A system integrating the XLAUI/CAUI 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.

83A.7 Protocol implementation conformance statement (PICS) proforma for Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)²⁰

83A.7.1 Introduction

The supplier of a XLAUI/CAUI implementation that is claimed to conform to Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI), 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 same, can be found in Clause 21.

83A.7.2 Identification

83A.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).	

83A.7.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Annex 83A, 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)
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 802.3-2012.)	

Date of Statement	
-------------------	--

²⁰Copyright 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.

83A.7.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
NOL	XLAUI is organized into four lanes, CAUI is organized into ten lanes	83A.1.2	See Clause 83	M	Yes []
RATE	Each XLAUI/CAUI lane operates at 10.3125 Gb/s	83A.1.2	10.3125 Gb/s (nominal)	M	Yes []
IO	Meets XLAUI/CAUI Electrical Characteristics	83A.3	Supports transmit and receive compliance points	M	Yes []

83A.7.4 XLAUI/CAUI transmitter requirements

Item	Feature	Subclause	Value/Comment	Status	Support
TC1	Signaling rate	83A.3.3	10.3125 GBd \pm 100 ppm	M	Yes []
TC2	Single ended output voltage range	83A.3.3.1	-0.4 V to 4 V	M	Yes []
TC3	Minimum de-emphasis	83A.3.3.1	4.4 dB	M	Yes []
TC4	Maximum de-emphasis	83A.3.3.1	7 dB	M	Yes []
TC5	Maximum termination mismatch	83A.3.3.3	5%	M	Yes []
TC6	Maximum differential output voltage	83A.3.3.1	760 mV	M	Yes []
TC7	Minimum output rise and fall time (20% to 80%)	83A.3.3.2	24 ps	M	Yes []
TC8	Differential output return loss	83A.3.3.3	See Equation (83A-5)	M	Yes []
TC9	Common-mode output return loss	83A.3.3.4	See Equation (83A-6)	M	Yes []
TC10	Transmitter eye mask	83A.3.3.5	See Figure 83A-8	M	Yes []

83A.7.5 XLAUI/CAUI receiver requirements

Item	Feature	Subclause	Value/Comment	Status	Support
RC1	Receiver BER	83A.3.4.1	Better than 10^{-12}	M	Yes []
RC2	Differential input return loss	83A.3.4.3	See Equation (83A-7)	M	Yes []
RC3	Differential to common-mode input return loss	83A.3.4.4	See Equation (83A-8)	M	Yes []
RC4	Total Jitter Tolerance	83A.3.4.6	0.62 UI	M	Yes []
RC5	Deterministic Jitter Tolerance	83A.3.4.6	0.42 UI	M	Yes []
RC6	Sinusoidal Jitter Tolerance	83A.3.4.6	See Figure 83A-12	M	Yes []
RC7	Receiver AC coupling	83A.3.4.5	Present	M	Yes []

83A.7.6 Electrical measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
EM1	De-emphasis setting during jitter measurements	83A.5.1	Off	M	Yes []
EM2	Jitter Tolerance stressed input	83A.5.2	0.42 UI DJ, 0.2 UI RJ	M	Yes []
EM3	Random jitter interference generator	83A.5.2	± 3 dB from 50 MHz to 6 GHz with a crest factor of no less than 5	M	Yes []
EM4	Meet jitter tolerance requirement with all XLAUI/CAUI channels active	83A.5.2	Yes	M	Yes []
EM5	XLAUI/CAUI Jitter Tolerance Test Pattern	83A.5.2	PRBS31 or scrambled idle	M	Yes []

83A.7.7 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
ES1	Conformance to applicable sections of IEC 60950-1	83A.6.1	Yes	M	Yes []
ES2	Compliance with applicable local and national codes for the limitation of electromagnetic interference	83A.6.4	Yes	M	Yes []

Annex 83B

(normative)

Chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)

83B.1 Overview

This annex defines the functional and electrical characteristics for the optional chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI). The purpose of this annex is to provide electrical characteristics and associated compliance points for pluggable module applications that use the XLAUI/CAUI interface and shall use the same number of lanes and signaling rate defined in Annex 83A. Figure 83B-3 and Table 83B-1 summarize an example differential insertion loss budget associated with the chip-module application. The insertion loss of Equation (83A-9), excluding a 0.5 dB connector loss at 5.15625 GHz, is linearly scaled to 7.9 dB loss at 5.15625 GHz for the host XLAUI/CAUI component, and 2.1 dB loss at 5.15625 GHz for the module as per Table 83B-1 and Equation (83B-1) for the host and Equation (83B-2) for the module. Tradeoffs can be made between the host PCB insertion loss and the connector loss. Equation (83B-1) is illustrated in Figure 83B-1 and Equation (83B-2) is illustrated in Figure 83B-2.

$$Insertion_loss(f) \leq \left\{ \begin{array}{ll} 0.111 + 1.046\sqrt{f} + 1.05f & 0.01 \leq f < 7 \\ -11.82 + 3.15f & 7 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83B-1)$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

$$Insertion_loss(f) \leq \left\{ \begin{array}{ll} 0.03 + 0.278\sqrt{f} + 0.28f & 0.01 \leq f < 7 \\ -3.155 + 0.84f & 7 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83B-2)$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

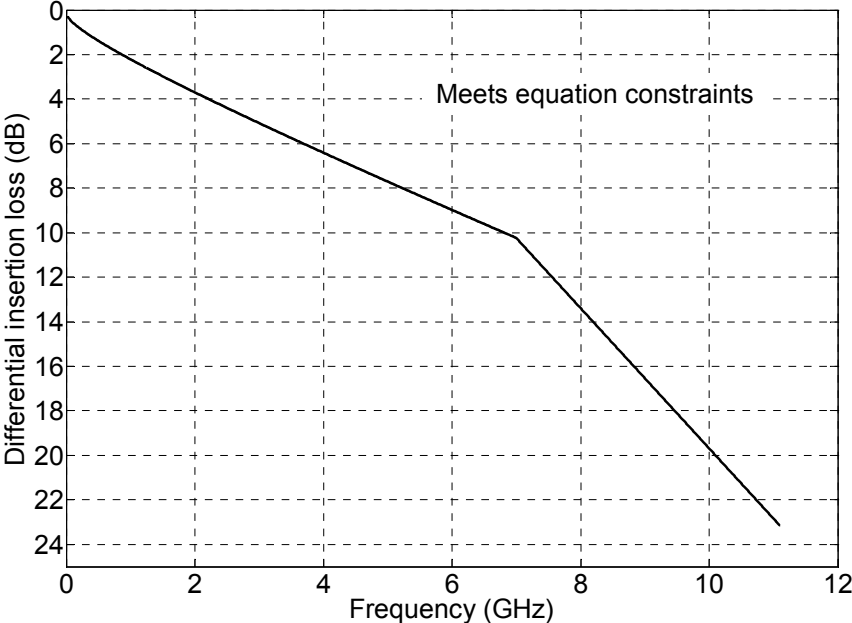


Figure 83B-1—Host insertion loss

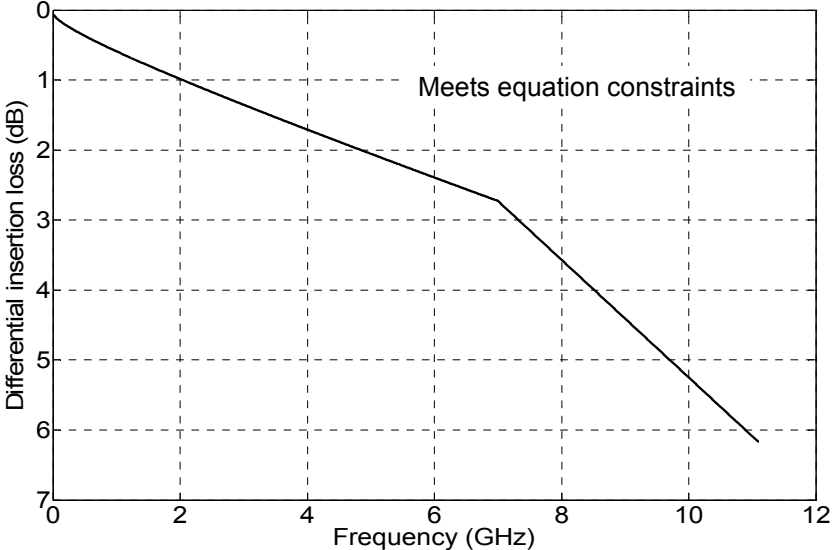


Figure 83B-2—Module insertion loss

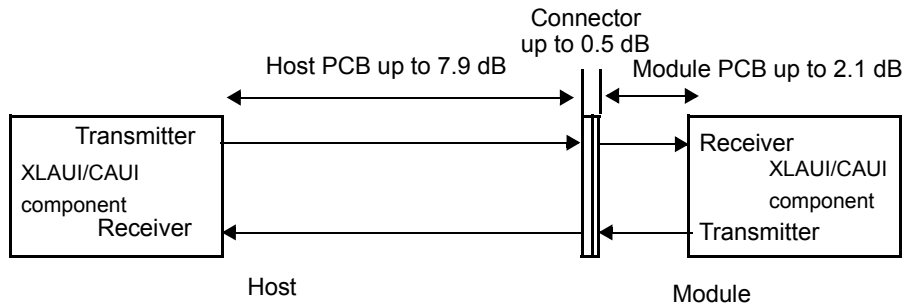


Figure 83B-3—Chip-module loss budget at 5.15625 GHz

Table 83B-1—Chip-module loss budget

Section	Differential insertion loss max. (at 5.15625 GHz)
Host XLAUI/CAUI component to connector	7.9 dB
Connector loss	0.5 dB
Connector to module XLAUI/CAUI component	2.1 dB

83B.2 Compliance point specifications for chip-module XLAUI/CAUI

The chip-module XLAUI/CAUI interface specifies compliance points around the module connector as depicted in Figure 83B-5 and Figure 83B-7. Chip-module devices shall meet the electrical characteristics defined in 83B.2, 83B.2.1, 83B.2.2, and 83B.2.3. Compliance points are defined to ensure interoperability between hosts and modules. A Module Compliance Board (MCB) is used to isolate the characteristics of the module, and a Host Compliance Board (HCB) is used to isolate the characteristics of the host. Figure 83B-5 and Figure 83B-7 include the loss associated with the HCB and MCB at 5.15625 GHz.

The reference differential insertion loss of the HCB PCB is given in Equation (83B-3) and illustrated in Figure 83B-4. The effects of differences between the insertion loss of an actual HCB and the reference insertion loss are to be accounted in the measurements.

$$Insertion_loss(f) = 0.017 + 0.5\sqrt{f} + 0.1836f \quad 0.01 \leq f \leq 11.1 \quad (83B-3)$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

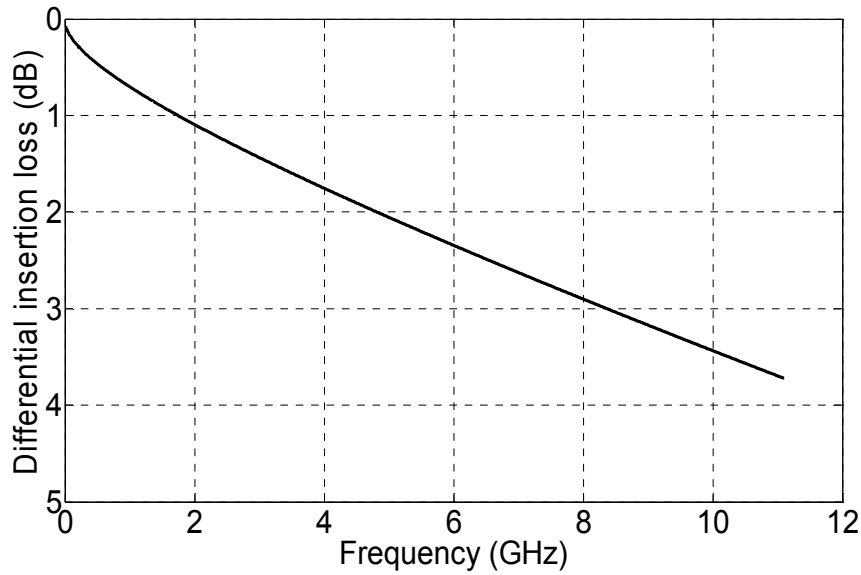


Figure 83B-4—HCB insertion loss

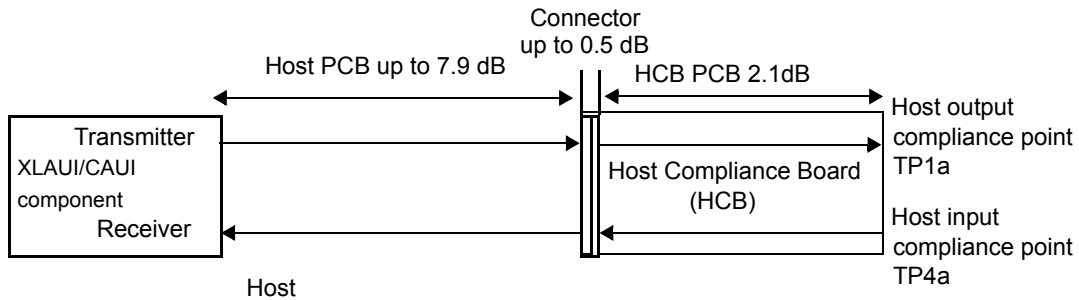


Figure 83B-5—Chip-module HCB insertion loss budget at 5.15625 GHz

The reference differential insertion loss of the MCB PCB is given in Equation (83B-4) and illustrated in Figure 83B-6. The effects of differences between the insertion loss of an actual MCB and the reference insertion loss are to be accounted in the measurements.

$$Insertion_loss(f) = -0.00086 + 0.2286\sqrt{f} + 0.08386f \text{ dB} \quad 0.01 \leq f \leq 11.1 \quad (83B-4)$$

where

$Insertion_loss(f)$ is the differential insertion loss at frequency f
 f is the frequency in GHz

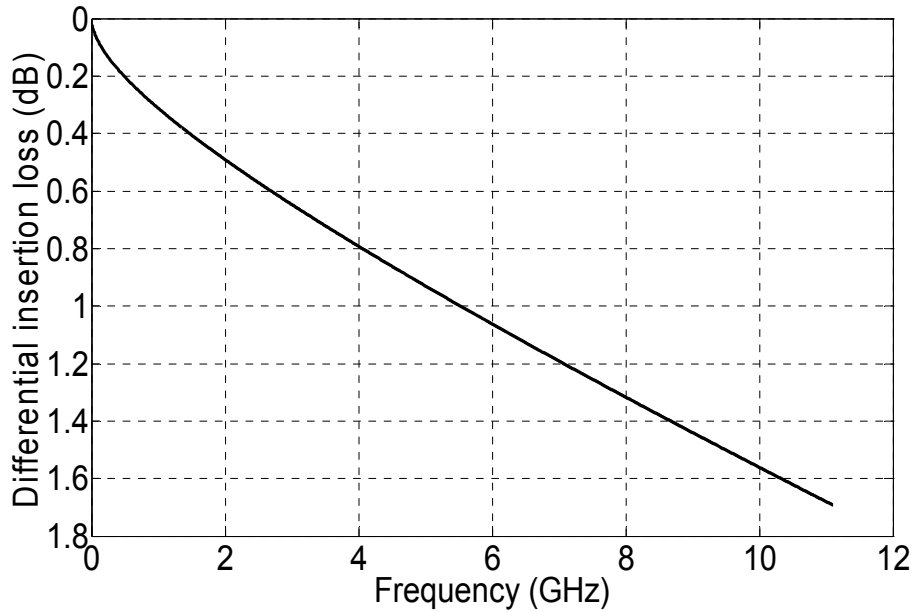


Figure 83B-6—MCB insertion loss

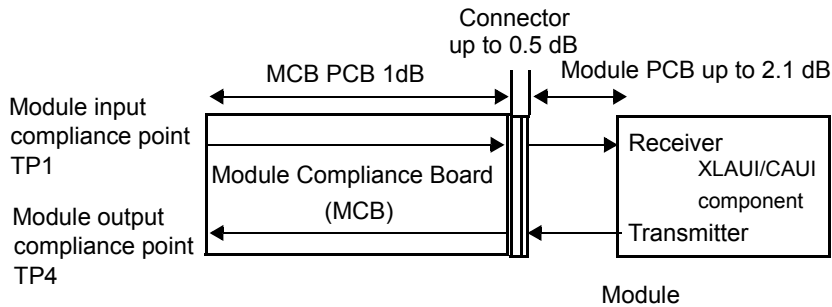


Figure 83B-7—Chip-module with MCB insertion loss budget at 5.15625 GHz

83B.2.1 Module specifications

A module that uses XLAUI/CAUI to interface with a host shall meet the characteristics outlined in Table 83B-2 and Table 83B-3 when measured using the MCB and HCB (where the HCB is used to calibrate inputs to the module). Table 83B-2 also lists the equivalent test points for the XLPPI/CPPI (see Figure 86-3).

Modules may support additional de-emphasis states, but the specification of additional states is outside the scope of this standard. De-emphasis shall be off during eye mask and jitter testing. Module electrical output de-emphasis off state is the optimal setting for module electrical output jitter and eye mask evaluation. AC coupling for both Tx and Rx paths shall be located in the module.

Table 83B-2—Specifications at module compliance points

Reference	Compliance point		Value	Unit
Minimum module differential input return loss	MCB input	TP1	See Equation (83B-5)	dB
Module input tolerance signal	HCB output	TP1a	See 83A.5.2	
Module output signal	MCB output	TP4	See Table 83B-3	
Minimum module differential output return loss	MCB output	TP4	See Equation (83B-5)	dB

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.01 \leq f < 2.19 \\ 5.56 - 8.7 \log_{10}\left(\frac{f}{5.5}\right) & 2.19 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83B-5)$$

where

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

Minimum module differential input/output return loss is illustrated in Figure 83B-8.

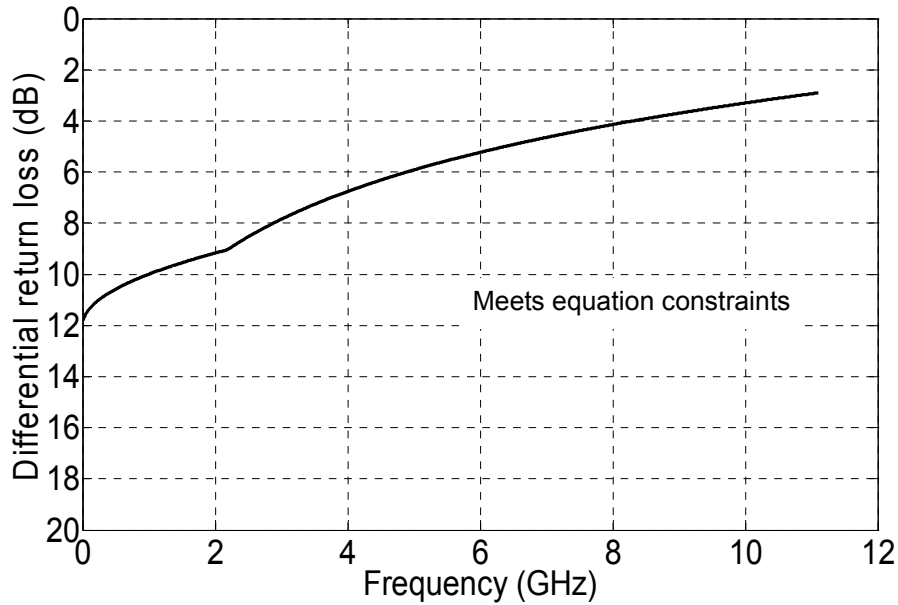


Figure 83B-8—Module input/output return loss

Table 83B-3—Module electrical output

Parameter	Subclause reference	Value	Unit
Maximum differential output voltage, peak-to-peak	83A.3.3.1	760	mV
Minimum de-emphasis	83A.3.3.1	3.5	dB
Maximum de-emphasis	83A.3.3.1	6	dB
Minimum VMA	83A.3.3.1	See Equation (83B-6)	mV
Maximum termination mismatch at 1 MHz	86A.5.3.2	5	%
Maximum output AC common-mode voltage, RMS	86A.5.3.1	15	mV
Minimum output rise and fall time (20% to 80%)	83A.3.3.2	24	ps
Maximum Total Jitter	83A.3.3.5	0.4	UI
Maximum Deterministic Jitter	83A.3.3.5	0.25	UI
Module electrical output eye mask definition X1	83A.3.3.5	0.2	UI
Module electrical output eye mask definition X2	83A.3.3.5	0.5	UI
Module electrical output eye mask definition Y1	83A.3.3.5	136	mV
Module electrical output eye mask definition Y2	83A.3.3.5	380	mV

$$\text{Minimum VMA (mV)} = (-110 - 2.13x + 0.32x^2) \times (10^{-y/20}) \quad (83B-6)$$

where

x is the rise or fall time, or 38, (whichever is larger) in ps
 y is de-emphasis value in dB

83B.2.2 Host specifications

A host that uses XLAUI/CAUI to interface with a module shall meet the characteristics outlined in Table 83B-4 and Table 83B-5 when measured using the HCB and MCB (where MCB is used to calibrate inputs to the host). Table 83B-4 also lists the equivalent test points for the XLPPI/CPPI.

Table 83B-4—Specifications at host compliance points

Reference	Compliance point		Value	Unit
Host output signal specifications	HCB output	TP1a	See Table 83B-5	
Minimum host differential output return loss	HCB output	TP1a	See Equation (83B-7)	dB
Minimum host differential input return loss	HCB input	TP4a	See Equation (83B-7)	dB
Host input tolerance signal	MCB output	TP4	See 83B.2.3	

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.01 \leq f < 2.19 \\ 5.56 - 8.71 \log_{10}\left(\frac{f}{5.5}\right) & 2.19 \leq f \leq 11.1 \end{array} \right\} \text{ (dB)} \quad (83B-7)$$

where

$Return_loss(f)$ is the differential input/output return loss at frequency f
 f is the frequency in GHz

Minimum host differential input/output return loss is illustrated in Figure 83B-9.

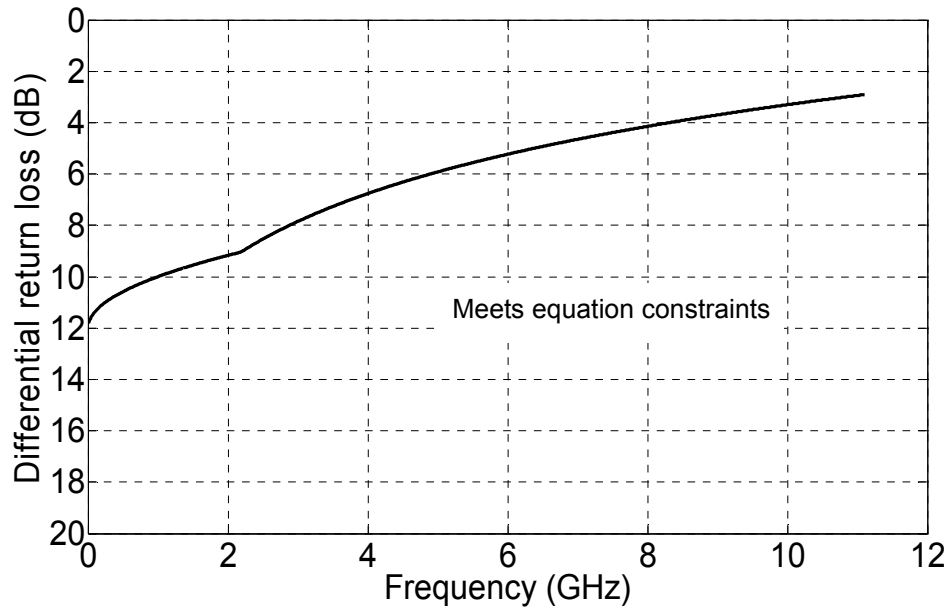


Figure 83B-9—Host input/output return loss

Table 83B-5—Host electrical output

Parameter	Subclause reference	Value	Units
Maximum output AC common-mode voltage, RMS	86A.5.3.1	20	mV
Minimum output rise and fall time (20% to 80%)	83A.3.3.2	24	ps
Maximum Total Jitter	83A.3.3.5	0.62	UI
Maximum Deterministic Jitter	83A.3.3.5	0.42	UI
Host electrical output eye mask definition X1	83A.3.3.5	0.31	UI
Host electrical output eye mask definition X2	83A.3.3.5	0.5	UI
Host electrical output eye mask definition Y1	83A.3.3.5	42.5	mV
Host electrical output eye mask definition Y2	83A.3.3.5	425	mV

83B.2.3 Host input signal tolerance

Host XLAUI/CAUI jitter tolerance evaluation shall be defined by a stressed input signal that comprises 0.25 UI Deterministic Jitter, and 0.15 UI random jitter for BER 10^{-12} . Deterministic Jitter is added to a clean test pattern as sinusoidal jitter defined in 83A.3.4.6. The limited low-pass filter stress is added until the 0.22 UI Deterministic Jitter is achieved. Frequency-dependent attenuation is then added using PCB trace or frequency-dependent attenuation that emulates PCB loss. Frequency-dependent attenuation is added until 0.25 UI Deterministic Jitter is achieved. Random jitter is added to the test signal using an interference generator, which is a broadband noise source capable of producing white Gaussian noise with adjustable

amplitude. The power spectral density shall be flat to ± 3 dB from 50 MHz to 6 GHz with a crest factor of no less than 5. Figure 83B–10 depicts a XLAUI/CAUI jitter tolerance test setup. The amplitude and output jitter of the filter stress plus limiter and random jitter injection shall meet the receiver eye mask illustrated in Figure 83A–9 with the values for X1, X2, Y1, Y2 given in Table 83B–3. All XLAUI/CAUI lanes shall be active during jitter tolerance testing. The PRBS31 pattern defined in 83.5.10 or scrambled idle defined in 82.2.10 is used for evaluating XLAUI/CAUI jitter tolerance.

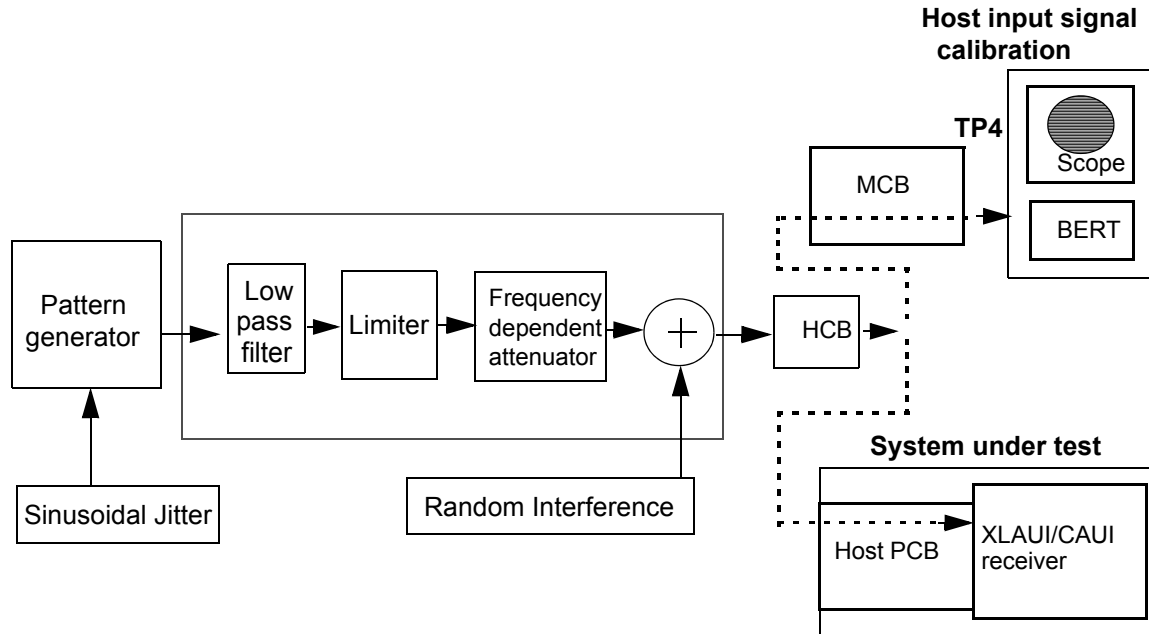


Figure 83B–10—Stressed-eye and jitter tolerance test setup

83B.3 Environmental specifications

83B.3.1 General safety

All equipment subject to this clause shall conform to applicable sections (including isolation requirements) of IEC 60950-1.

83B.3.2 Network safety

The designer is urged to consult the relevant local, national, and international safety regulations to ensure compliance with the appropriate requirements.

83B.3.3 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.

83B.3.4 Electromagnetic compatibility

A system integrating the XLAUI/CAUI shall comply with applicable local and national codes for the limitation of electromagnetic interference.

83B.3.5 Temperature and humidity

A system integrating the XLAUI/CAUI 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.

83B.4 Protocol implementation conformance statement (PICS) proforma for Annex 83B, Chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)²¹

83B.4.1 Introduction

The supplier of a chip to module XLAUI/CAUI implementation that is claimed to conform to Annex 83B, Chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI), 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 same, can be found in Clause 21.

83B.4.2 Identification

83B.4.2.1 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).	

83B.4.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Annex 83B, Chip-module 40 Gb/s Attachment Unit Interface (XLAUI) and 100 Gb/s Attachment Unit Interface (CAUI)
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 802.3-2012.)	

Date of Statement	
-------------------	--

²¹Copyright 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.

83B.4.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
NOL	XLAUI is organized into four lanes, CAUI is organized into ten lanes	83A.1.2	See Clause 83	M	Yes []
RATE	Each XLAUI/CAUI lane operates at 10.3125 Gb/s	83A.1.2	10.3125 Gb/s (nominal)	M	Yes []
IO	Meets chip-module XLAUI/CAUI Electrical Characteristics	83B.2	Supports host/module compliance points	M	Yes []

83B.4.4 Module requirements

Item	Feature	Subclause	Value/Comment	Status	Support
MC1	XLAUI/CAUI compliant module	83B.2.1	Meets requirements defined in Table 83B-2 and Table 83B-3	M	Yes []
MC2	De-emphasis setting during module jitter evaluation	83B.2.1	Off	M	Yes []
MC3	AC coupling for Tx and Rx	83B.2.1	Present in module	M	Yes []

83B.4.5 Host requirements

Item	Feature	Subclause	Value/Comment	Status	Support
HC1	XLAUI/CAUI compliant host	83B.2.2	Meets requirements defined in Table 83B-4 and Table 83B-5	M	Yes []
HC2	All XLAUI/CAUI lanes active during jitter evaluation	83B.2.3	Yes	M	Yes []
HC3	Applied jitter for host jitter tolerance evaluation	83B.2.3	0.25UI Deterministic Jitter, 0.15UI random jitter	M	Yes []

83B.4.6 Environmental specifications

Item	Feature	Subclause	Value/Comment	Status	Support
ES1	Conformance to applicable sections of IEC 60950-1	83B.3.1	Yes	M	Yes []
ES2	Compliance with applicable local and national codes for the limitation of electromagnetic interference	83B.3.4	Yes	M	Yes []

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.

83C.1 Partitioning examples with FEC

The example of FEC 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

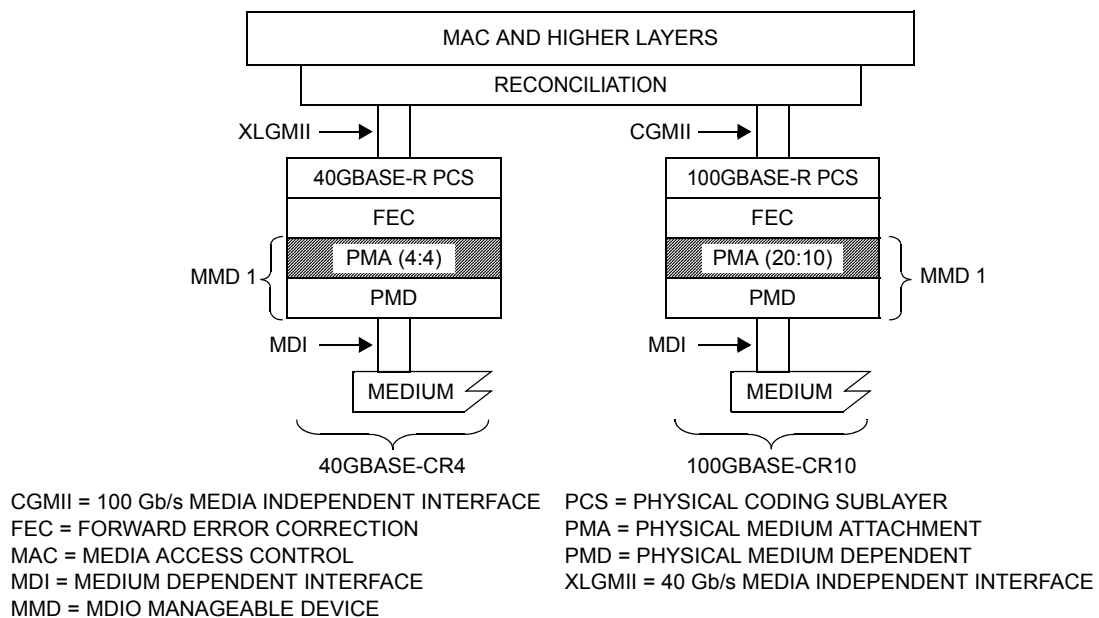
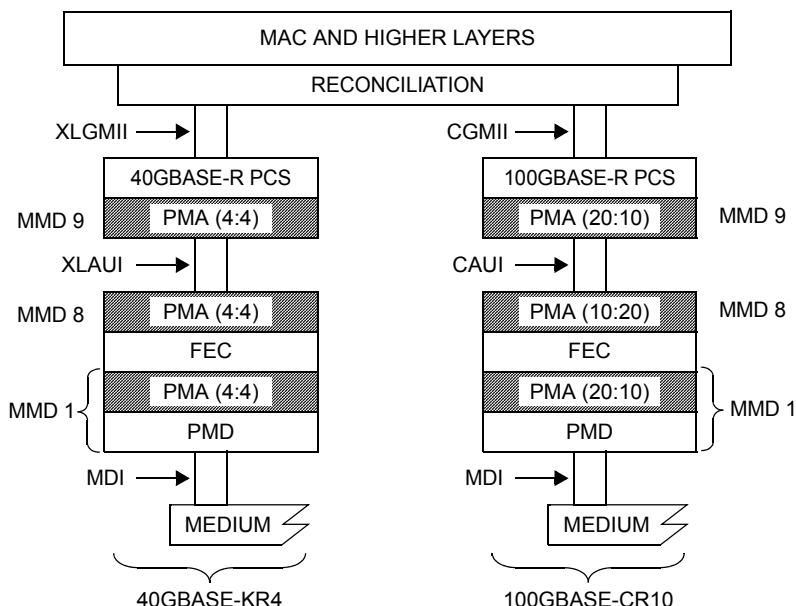


Figure 83C–1—Example FEC implemented with PCS

83C.1.2 FEC implemented with PMD



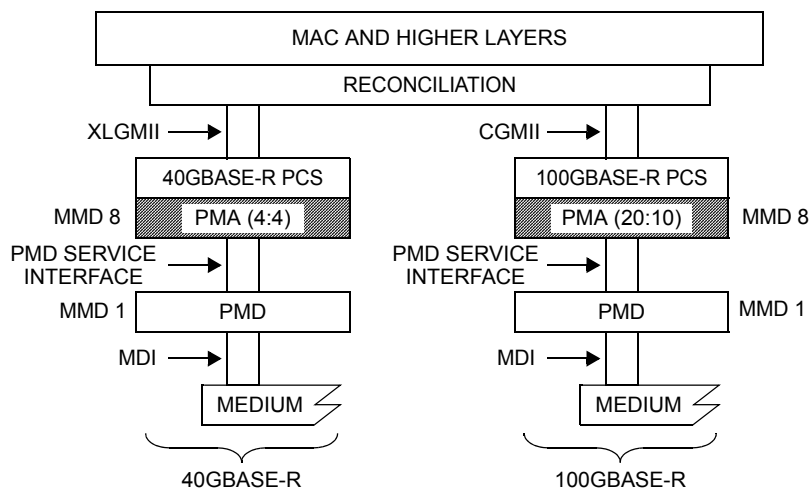
CAUI = 100 Gb/s ATTACHMENT UNIT INTERFACE
 CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
 FEC = FORWARD ERROR CORRECTION
 MAC = MEDIA ACCESS CONTROL
 MDI = MEDIUM DEPENDENT INTERFACE

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

83C.2 Partitioning examples without FEC

83C.2.1 Single PMA sublayer without 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
 XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

Figure 83C-3—Example single PMA sublayer without FEC

83C.2.2 Single XLAUI/CAUI without FEC

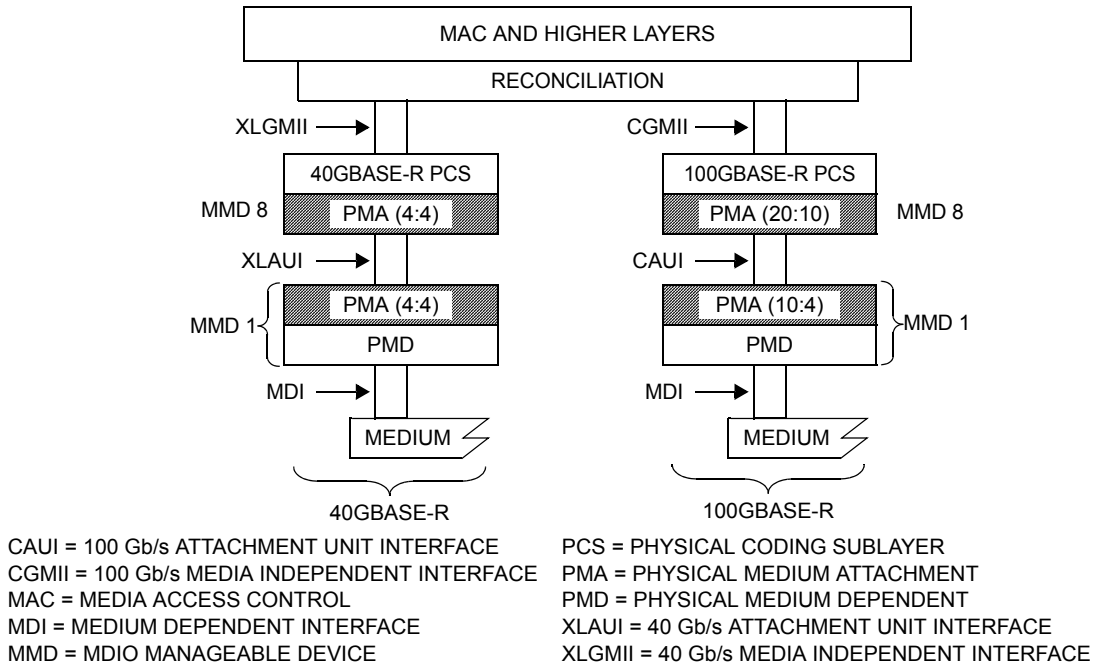


Figure 83C-4—Example single XLAUI/CAUI without FEC

83C.2.3 Separate SERDES for optical module interface

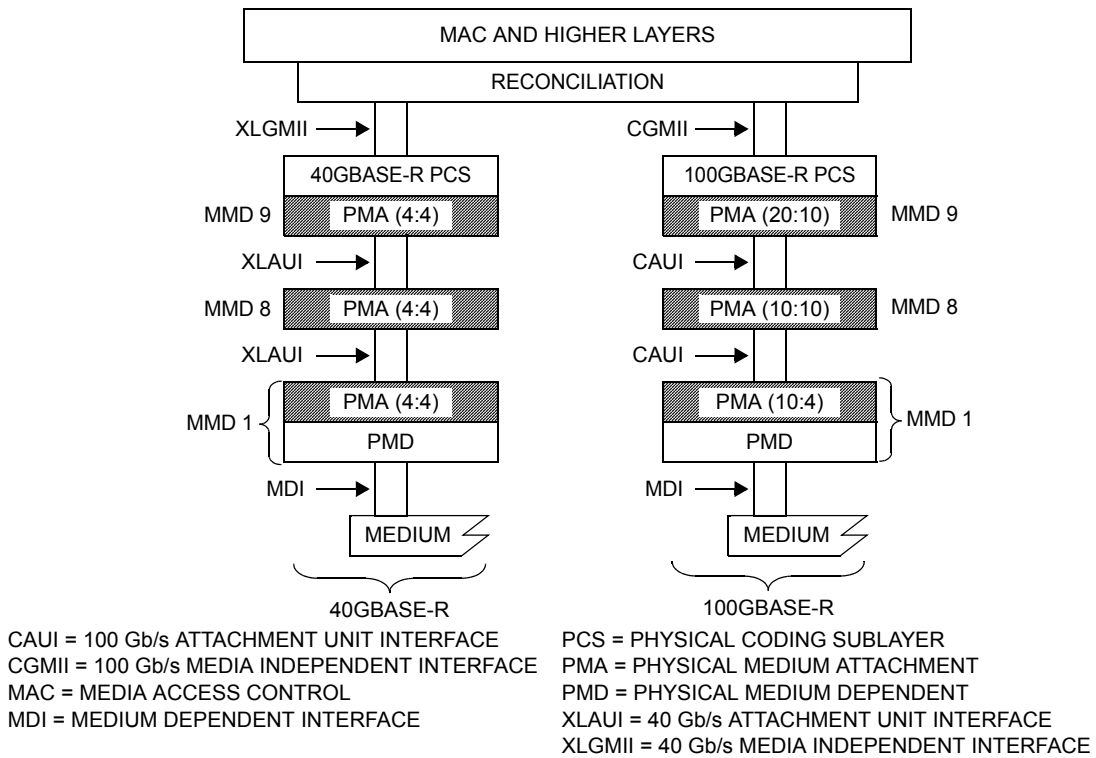


Figure 83C-5—Separate SERDES for optical module interface

Annex 85A

(informative)

40GBASE-CR4 and 100GBASE-CR10 TP0 and TP5 test point parameters

85A.1 Overview

Annex 85A 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 40GBASE-CR4 and 100GBASE-CR10 link block diagram of Figure 85–2.

85A.2 Transmitter characteristics at TP0

The specifications at TP0 are summarized in Table 85A–1

Table 85A–1—Transmitter characteristics at TP0 summary

Parameter	Subclause reference	Value	Units
Signaling rate, per lane	85.8.3.9	10.3125 ± 100 ppm	GBd
Unit interval nominal	85.8.3.9	96.969697	ps
Differential peak-to-peak output voltage (max) with Tx disabled		30	mV
Common-mode voltage limits	72.7.1.4	0 to 1.9	V
Differential output return loss (min)	72.7.1.5	See Equation (72–4) and Equation (72–5)	dB
Common-mode output return loss (min)	72.7.1.6	See Equation (72–6) and Equation (72–7)	dB
Common-mode AC output voltage (max, RMS)		30	mV
Transition time (20%–80%)	72.7.1.7	24 to 47	ps
Max output jitter (peak-to-peak) ^a	72.7.1.9		
Random jitter ^b		0.15	UI
Deterministic jitter		0.15	UI
Duty Cycle Distortion ^c		0.035	UI
Total jitter		0.28	UI

^aJitter is measured with emphasis off.

^bJitter is specified at BER 10^{-12} .

^cDuty Cycle Distortion is considered part of the deterministic jitter distribution.

85A.3 Receiver characteristics at TP5

The receiver characteristics at TP5 are summarized in Table 85A–2.

Table 85A–2—Receiver characteristics at TP5 summary

Parameter	Subclause reference	Value	Units
Bit error ratio	85.8.4.3	10^{-12}	
Signaling rate, per lane	85.8.4.4	10.3125 ± 100 ppm	GBd
Unit interval (UI) nominal	85.8.4.4	96.969697	ps
Receiver coupling	85.8.4.5	AC	
Differential input peak-to-peak amplitude (max)	72.7.2.4	1200 ^a	mV
Differential input return loss (min) ^b	72.7.2.5	See Equation (72–4) and Equation (72–5)	dB
Differential to common-mode input return loss	85.8.4	10 min from 10 MHz to 10 GHz	dB

^aThe receiver shall tolerate amplitudes up to 1600 mV without permanent damage.

^bRelative to 100 Ω differential.

85A.4 Transmitter and receiver differential printed circuit board trace loss

With the insertion loss TP0 to TP2 or TP3 to TP5 given in 85.8.3.4 and an assumed mated connector loss of 1.74 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 (85A–1). The maximum insertion loss allocation for the transmitter and receiver differential controlled impedance printed circuit boards is 7 dB at 5.15625 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 $IL_{PCBmax}(f)$.

$$IL_{PCB}(f) \leq IL_{PCBmax}(f) = (0.3)[20\log_{10}(e)(b_1\sqrt{f} + b_2f + b_3f^2 + b_4f^3)] \quad \text{(dB)} \quad (85A-1)$$

for $10 \text{ MHz} \leq f \leq 7500 \text{ MHz}$.

where

f is the frequency in Hz

$IL_{PCB}(f)$ is the insertion loss for the transmitter and receiver PCB

$IL_{PCBmax}(f)$ is the maximum insertion loss for the transmitter and receiver PCB

$$b_1 = 2 \times 10^{-5}$$

$$b_2 = 1.1 \times 10^{-10}$$

$$b_3 = 3.2 \times 10^{-20}$$

$$b_4 = -1.2 \times 10^{-30}$$

$$e \approx 2.71828$$

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 (85A-2) and the coefficients b_1 through b_4 are given in Equation (85A-1). The minimum insertion loss for the transmitter or the receiver differential controlled impedance printed circuit board is one half of the minimum insertion loss $IL_{PCBmin}(f)$.

$$IL_{PCB}(f) \geq IL_{PCBmin}(f) = (0.0574)[20\log_{10}(e)(b_1\sqrt{f} + b_2f + b_3f^2 + b_4f^3)] \text{ (dB)} \quad (85A-2)$$

for $10 \text{ MHz} \leq f \leq 7500 \text{ MHz}$.

where

f is the frequency in Hz

$IL_{PCB}(f)$ is the insertion loss for the transmitter and receiver PCB

$IL_{PCBmin}(f)$ is the minimum insertion loss for the transmitter and receiver PCB

85A.5 Channel insertion loss

This subclause provides information on channel insertion losses for intended topologies ranging from 0.5 m to 7 m in length. The maximum channel insertion loss associated with the 7 m topology is determined using Equation (85A-3). The channel insertion loss associated with the 0.5 m topology and a maximum host channel is determined by Equation (85A-4). The channel insertion loss budget at 5.15625 GHz is illustrated in Figure 85A-1.

The maximum channel insertion loss is determined using Equation (85A-3). The maximum channel insertion loss is 24.44 dB at 5.15625 GHz.

$$IL_{Chmax}(f) = IL_{Camax}(f) + 2IL_{Host}(f) - 2IL_{MatedTF}(f) \text{ (dB)} \quad (85A-3)$$

for $50 \text{ MHz} \leq f \leq 7500 \text{ MHz}$.

where

f is the frequency in MHz

$IL_{Chmax}(f)$ is the maximum channel insertion loss between TP0 and TP5

$IL_{Camax}(f)$ is the maximum cable assembly insertion loss using Equation (85-19)

$IL_{Host}(f)$ is the maximum insertion loss from TP0 to TP2 or TP3 to TP5 using Equation (85-14)

$IL_{MatedTF}(f)$ is the maximum insertion loss of the mated test fixture using Equation (85-36)

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 (85A-4).

$$(IL_{Ch0.5m}(f) = 0.275IL_{Camax}(f) + 2IL_{Host}(f) - 2IL_{MatedTF}(f) \text{ (dB)} \quad (85A-4)$$

for $50 \text{ MHz} \leq f \leq 7500 \text{ MHz}$.

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

$IL_{C_{max}}(f)$ is the maximum cable assembly insertion loss from TP1 to TP4 using Equation (85–19)

$IL_{Host}(f)$ is the maximum insertion loss from TP0 to TP2 or TP3 to TP5 using Equation (85–14)

$IL_{MatedTF}(f)$ is the maximum insertion loss of the mated test fixture using Equation (85–36)

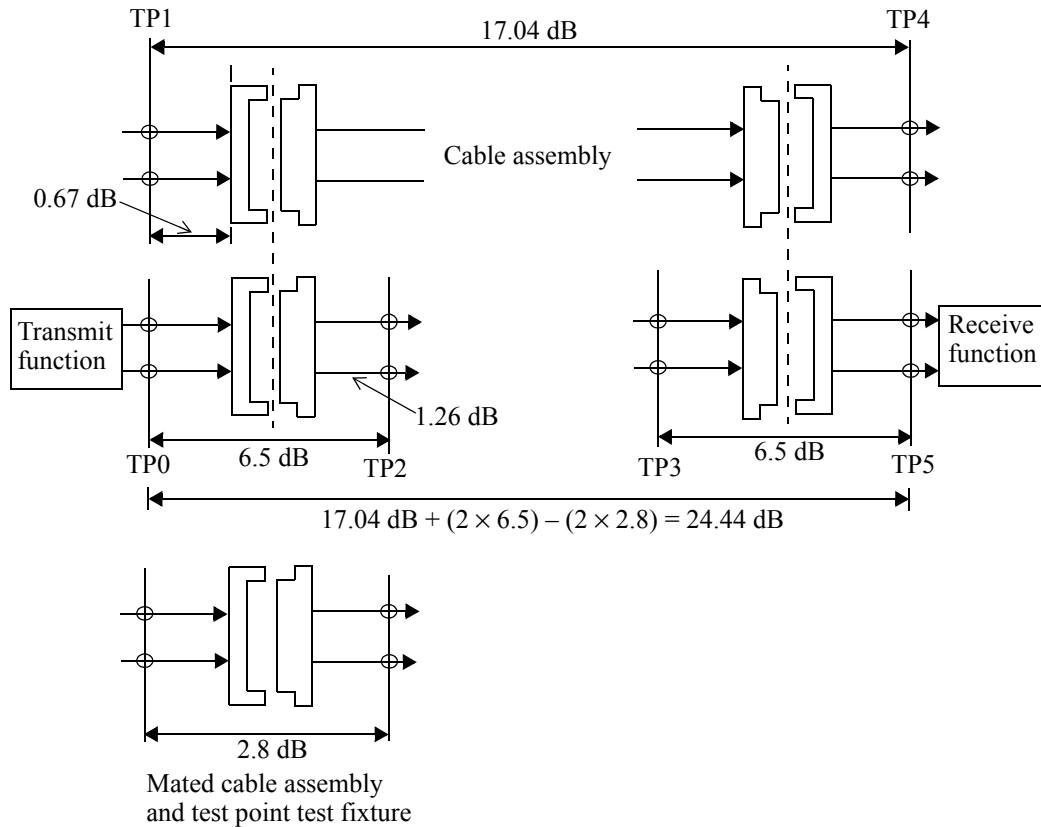


Figure 85A–1—Illustration of channel insertion loss budget at 5.15625 GHz

85A.6 Channel return loss

The return loss of each lane of the 40GBASE-CR4 or 100GBASE-CR10 channel is recommended to meet the values determined using Equation (85–25).

85A.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 (85A–5).

$$ILD(f) = IL_{Ch}(f) - IL_{fitted}(f) \text{ (dB)} \quad (85A-5)$$

where

f is the frequency in MHz
 $ILD(f)$ is the channel insertion loss deviation at frequency f
 $IL_{\text{fitted}}(f)$ is defined in Equation (85–19)

Given the channel insertion loss is at N uniformly-spaced frequencies f_n spanning the frequency range 50 MHz to 7500 MHz with a maximum frequency spacing of 10 MHz, the coefficients of the fitted channel insertion loss are determined using Equation (85–20) and Equation (85–21).

The channel insertion loss deviation is recommended to be within the region defined by Equation (85–23) and Equation (85–24) for all frequencies from 50 MHz to 7500 MHz.

85A.8 Channel integrated crosstalk noise (ICN)

Since four lanes or ten 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 or ten adjacent transmitters. The channel multiple disturber NEXT loss, $MDNEXT_loss_1(f)$, is specified using the individual NEXT losses as shown in Equation (85–26).

In addition, the far-end Crosstalk (FEXT) that is coupled into a receiver will be from the three or nine 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 (85–27).

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 10000 MHz with a maximum frequency spacing of 10 MHz, the RMS value of the integrated crosstalk noise is determined using Equation (85–28) through Equation (85–32) and the parameters shown in Table 85–11.

The total integrated crosstalk RMS noise voltage of the channel is recommended to meet the values determined using Equation (85A–6) illustrated in Figure 85A–2.

$$\sigma_{x, ch} \leq \left\{ \begin{array}{ll} 10 & 3 \leq IL \leq 7.5 \\ 13.4 - 0.45IL & 7.5 < IL \leq 24.44 \end{array} \right\} \quad (\text{mV}) \quad (85A-6)$$

where IL is the value of the channel insertion loss in dB at 5.15625 GHz.



Figure 85A-2—Channel integrated crosstalk noise

Annex 86A

(normative)

Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPPI) and 100GBASE-SR10 (CPPI)

86A.1 Overview

The Parallel Physical Interface (nPPI) is an optional instantiation of the PMD service interface for the PMDs in Clause 86, which is described in 86.2 and 80.3, or for the 40GBASE-LR4 PMD service interface described in 87.2 and 80.3. It allows the construction of compact optical transceiver modules for 40GBASE-SR4, 40GBASE-LR4, or 100GBASE-SR10 with no clock and data recovery circuits inside. The 40 Gb/s Parallel Physical Interface (XLPPPI, four lanes) is used with 40GBASE-SR4 or 40GBASE-LR4, and the 100 Gb/s Parallel Physical Interface (CPPI, ten lanes) with 100GBASE-SR10. The term “nPPI” denotes either or both.

The PMD and PMA attached to the nPPI are required to comply with the delay, Skew, and Skew Variation requirements in 86.3, 83.5.3, and 83.5.4 as appropriate. The PMD MDIO function mapping given in 86.4 may apply. The PMD functional specifications are as given in 86.5.

This annex is arranged as follows: following the overview, a reminder of the block diagram, a brief introduction to the test points, and lane assignments, 86A.4 contains the electrical specifications for nPPI from host to module (Tx side) and then module to host (Rx side). Test points, compliance boards, and electrical parameters are defined in 86A.5. The host PCB (“channel”) response recommendation is provided in 86A.6. Safety, installation, environment, and labeling is addressed in 86A.7 and PICS is provided in 86A.8. This annex is very closely related to Clause 86.

86A.2 Block diagram and test points

The PMD block diagram is shown in Figure 86–2. Figure 86–3 shows the test points.

The nPPI is standardized at the test points described in 86.8.1. The transmit side electrical signal (host electrical output and module electrical input) is defined at the output of the Host Compliance Board (TP1a), and other specifications of the module electrical input port are defined at the input of the Module Compliance Board (TP1). The receive side electrical signal (module electrical output and host electrical input) is defined at the output of the Module Compliance Board (TP4), and other specifications of the host electrical input port are defined at the input of the Host Compliance Board (TP4a). Test points and compliance boards are defined more thoroughly in 86A.5.1.

86A.3 Lane assignments

There are no lane assignments (within a group of transmit or receive lanes) for XLPPPI or CPPI. While it is expected that a PMD will map electrical lane i to optical lane i and vice versa, there is no need to define the physical ordering of the lanes, as the PCS is capable of receiving the lanes in any arrangement.

86A.4 Electrical specifications for nPPI

The signaling rate for a lane of an XLPPi or CPPI interface shall be as defined in Table 86–2. 86A.4.1 and 86A.4.2 specify the host to module (Tx side) and module to host (Rx side) respectively of the nPPI. Parameters are defined in 86A.5 and 86.8. A recommended PCB (“channel”) response for the host (PMA) is provided in 86A.6. Test points are defined in 86A.5.1.

86A.4.1 nPPI host to module electrical specifications

Each output lane and signal of the nPPI host (PMA), if measured at TP1a (see 86A.5.1) with the specified crosstalk signals applied on all input lanes, shall meet the specifications of Table 86A–1 per the definitions in 86A.5. Each lane of the nPPI module (PMD) electrical input, if measured at TP1 and TP1a with all Rx lanes (module output) operating, shall meet the specifications of Table 86A–2 per the definitions in 86A.5. The module electrical input shall be AC coupled, i.e., it shall present a high DC common-mode impedance at TP1. There may be various methods for AC coupling in actual implementations.

Table 86A–1—nPPI host electrical output specifications at TP1a

Parameter description	Min	Max	Units	Conditions
Single ended output voltage	–0.3	4	V	Referred to signal common
AC common-mode output voltage	—	15	mV	RMS
Termination mismatch at 1 MHz	—	5	%	
Differential output return loss	See 86A.4.1.1	—	dB	
Output transition time, 20% to 80%	28	—	ps	
J2 Jitter output	—	0.17	UI	
J9 Jitter output	—	0.29	UI	
Data Dependent Pulse Width Shrinkage (DDPWS)	—	0.07	UI	
Q _{sq} for XLPPi	45	—	V/V	
Q _{sq} for CPPI	43	—	V/V	
	Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	0.11, 0.31 95, 350		UI mV	Hit ratio = 5×10^{-5}
Crosstalk source VMA, each input lane	700		mV	At TP4
Crosstalk source transition times, 20% to 80%	34		ps	At TP4

86A.4.1.1 Differential return losses at TP1 and TP1a

From 10 MHz to 11.1 GHz, the magnitude in decibels of the module differential input return loss at TP1 and the host differential output return loss at TP1a (see 86A.5.1) shall not exceed the limit given in

Table 86A–2—nPPI module electrical input specifications at TP1 and TP1a

Parameter description	Test point	Min	Max	Units	Conditions
Single ended input voltage tolerance ^a	TP1a	−0.3	4	V	Referred to TP1 signal common
AC common-mode input voltage tolerance	TP1a	15	—	mV	RMS
Differential input return loss	TP1	See 86A.4.1.1	—	dB	10 MHz to 11.1 GHz
Differential to common-mode input return loss	TP1	10	—	dB	10 MHz to 11.1 GHz
J2 Jitter tolerance	TP1a	0.17	—	UI	
J9 Jitter tolerance	TP1a	0.29	—	UI	
Data Dependent Pulse Width Shrinkage (DDPWS) tolerance	TP1a	0.07	—	UI	
		Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	TP1a	0.11, 0.31 95, 350		UI mV	Hit ratio = 5×10^{-5}
Crosstalk calibration signal VMA	TP4	850		mV	While calibrating compliance signal ^b
Crosstalk calibration signal transition times, 20% to 80%	TP4	34		ps	

^a The single ended input voltage tolerance is the allowable range of the instantaneous input signals.

^b The crosstalk calibration signals are applied to the mated HCB-MCB at TP4a and measured at TP4 following the same principles as the host electrical input calibration (see 86A.5.3.8.5). They are removed before testing.

Equation (86A–1) and illustrated in Figure 86A–1.

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.01 \leq f < 4.11 \\ 6.3 - 13 \log_{10}\left(\frac{f}{5.5}\right) & 4.11 \leq f \leq 11.1 \end{array} \right\} \text{ dB} \quad (86A-1)$$

where

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

86A.4.2 nPPI module to host electrical specifications

Each electrical output lane and signal of the nPPI module (PMD), if measured at TP4 shall meet the specifications of Table 86A–3 per the definitions in 86A.5 while the specified crosstalk sources are applied to the module’s electrical input. Each lane of the nPPI host (PMA) input shall meet the specifications of Table 86A–4 at TP4 and/or TP4a per the definitions in 86A.5. The module electrical output shall be AC coupled, i.e., it shall present a high DC common-mode impedance at TP4. There may be various methods for AC coupling in actual implementations.

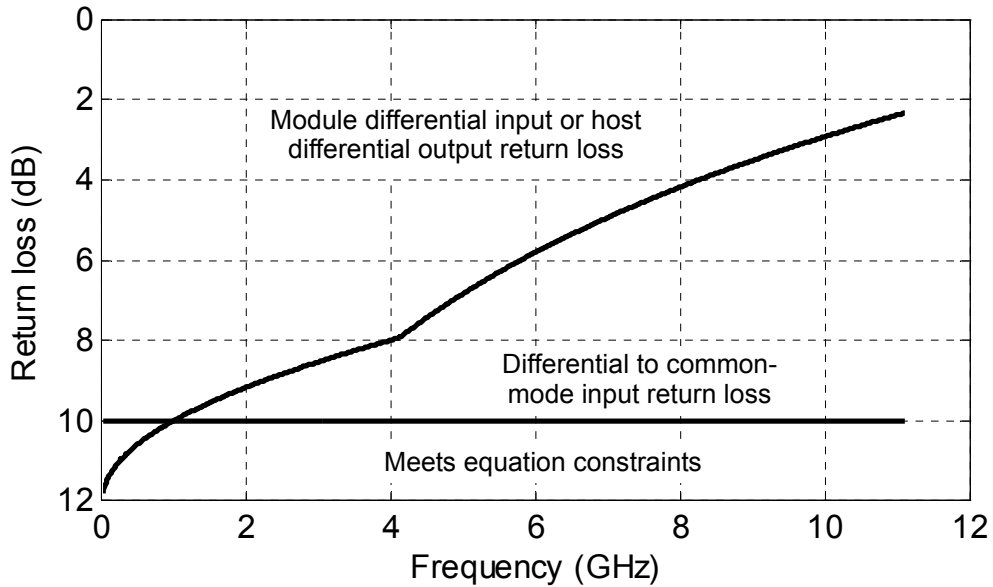


Figure 86A-1—Return loss specifications

Table 86A-3—nPPI module electrical output specifications at TP4

Parameter description	Min	Max	Units	Conditions
Single ended output voltage tolerance	-0.3	4	V	Referred to signal common
AC common-mode output voltage (RMS)	—	7.5	mV	
Termination mismatch at 1 MHz	—	5	%	
Differential output return loss	See 86A.4.2.1	—	dB	10 MHz to 11.1 GHz
Output transition time, 20% to 80%	28	—	ps	
J2 Jitter output	—	0.42	UI	
J9 Jitter output	—	0.65	UI	
Specification values				
Eye mask coordinates: X1, X2 Y1, Y2	0.29, 0.5 150, 425		UI mV	Hit ratio = 5×10^{-5}
Crosstalk source VMA, each lane	700		mV	At TP1a
Crosstalk source transition times, 20% to 80%	37		ps	At TP1a

Table 86A-4—nPPI host electrical input specifications at TP4 and TP4a

Parameter description	Test point	Min	Max	Units	Conditions
Single ended input voltage ^a	TP4	-0.3	4	V	Referred to signal common
AC common-mode input voltage tolerance	TP4	7.5	—	mV	RMS
Differential input return loss	TP4a	See 86A.4.2.1	—	dB	
Differential to common-mode input return loss	TP4a	10	—	dB	10 MHz to 11.1 GHz
Host input signal tolerance, interface BER limit		—	10^{-12}	—	
Conditions of host electrical receiver signal tolerance test: ^b					
		Specification values			
Eye mask coordinates: X1, X2 Y1, Y2	TP4	0.29, 0.5 150, 425		UI mV	Hit ratio = 5×10^{-5}
Transition time, 20% to 80%		34		ps	
J2 Jitter	TP4	0.42		UI	
J9 Jitter	TP4	0.65		UI	
Data Dependent Pulse Width Shrinkage (DDPWS)		0.34		UI	
VMA of aggressor lanes	TP4	850		mV	
Crosstalk calibration signal VMA	TP1a	700		mV	
Crosstalk calibration signal transition times, 20% to 80%	TP1a	37		ps	

^a The host is required to tolerate (work correctly with) input signals with instantaneous voltages anywhere in the specified range.

^b The specification values are test conditions for measuring signal tolerance and are not characteristics of the host (see 86A.5.3.8).

86A.4.2.1 Differential return losses at TP4 and TP4a

From 10 MHz to 11.1 GHz, the magnitude in decibels of the module differential output return loss at TP4 and the host differential input return loss at TP4a shall not exceed the limit given in Equation (86A-2) and illustrated in Figure 86A-1.

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 12 - 2\sqrt{f} & 0.01 \leq f < 4.11 \\ 6.3 - 13 \log_{10}\left(\frac{f}{5.5}\right) & 4.11 \leq f \leq 11.1 \end{array} \right\} \text{ dB} \quad (86A-2)$$

where

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

86A.5 Definitions of electrical parameters and measurement methods

Test points are defined in 86A.5.1, compliance boards in 86A.5.1.1, test patterns in 86A.5.2 and parameters in 86A.5.3 and 86.8. Multi-lane testing considerations are given in 86.8.2.1.

86A.5.1 Test points and compliance boards

Figure 86–3 shows the six test points for 40GBASE-SR4 and 100GBASE-SR10. These are TP1, TP1a, TP2, TP3, TP4, and TP4a; four of these are Skew points SP2, SP3, SP4, and SP5 as shown. For 40GBASE-LR4, points TP1, TP1a, TP4, TP4a, SP2, and SP5 are in equivalent positions. Figure 86–3 also shows the substitution of compliance boards for module (PMD) or host (PMA). These compliance boards are defined to connect generic test equipment to the module and host for test purposes. The module can be plugged into the Module Compliance Board (MCB), which has specified electrical parameters. The Host Compliance Board (HCB), which also has specified electrical parameters, can plug into the host. The MCB and the HCB can be plugged together for calibration of compliance signals and to check the electrical parameters of the boards. Table 86A–5 shows the parameters or signals measured at each point. Also, TP0 and TP5 define the host ends of the electrical channel, at the PMA IC.

All electrical measurements and parameters are defined as through HCB and/or MCB as appropriate, with corrections if necessary.

CAUTION

A PMD with an nPPI interface is AC coupled; however, an HCB is not. The user should take care that the test equipment does not improperly load or damage a host under test.

Table 86A–5—Parameters defined at each test point

Test point	Direction	Parameter
TP1	Looking downstream into module transmitter input	Module transmitter input return loss
TP1a	Looking upstream into host transmitter output	Host transmitter output signal and output return loss, module transmitter compliance signal calibration, host receiver compliance crosstalk signal calibration
TP4	Looking upstream into module receiver output	Module receiver output signal and output return loss, host receiver compliance signal calibration
TP4a	Looking downstream into host receiver input	Host receiver input return loss

86A.5.1.1 Compliance board parameters

The electrical characteristics of the HCB and MCB are given in 86A.5.1.1.1 and 86A.5.1.1.2. If boards are used that do not match the specifications given, the measurement results for nPPI shall be corrected for the differences between the actual HCB or MCB’s properties and the reference differential insertion losses given in 86A.5.1.1.1. As it may be impractical to correct eye measurements for a board with differential insertion loss outside the limits given in 86A.5.1.1.2, such boards shall not be used. Boards that do not meet the specifications for mated HCB-MCB in 86A.5.1.1.2 shall not be used.

86A.5.1.1.1 Reference insertion losses of HCB and MCB

The reference differential insertion loss in decibels of the HCB and of the MCB, excluding the module connector, are given in Equation (86A–3) and Equation (86A–4) respectively, and illustrated in Figure 86A–2.

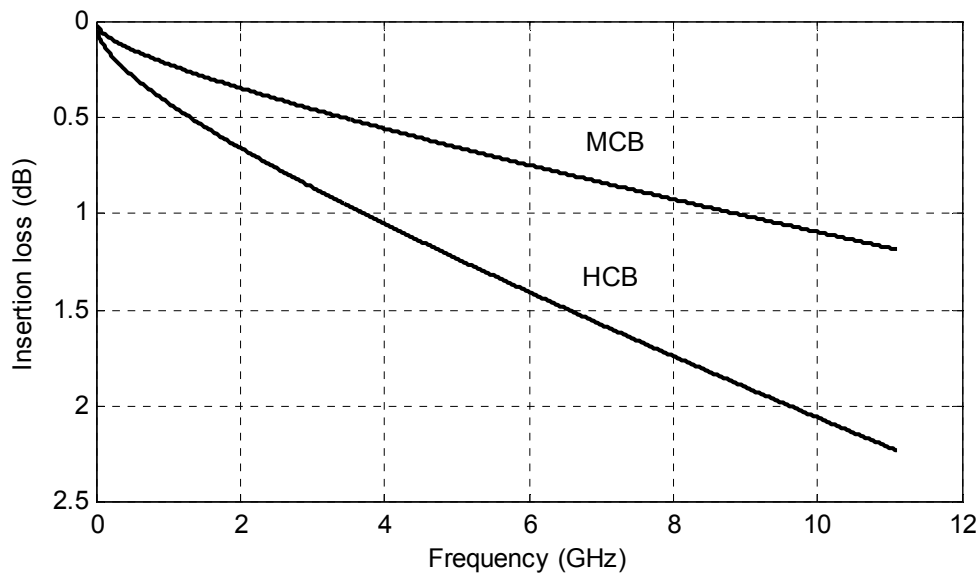


Figure 86A–2—Reference differential insertion losses of HCB, MCB excluding connector

For the HCB,

$$Insertion_loss(f) = 0.01 + 0.3\sqrt{f} + 0.11f \text{ dB} \quad 0.01 \leq f \leq 11.1 \quad (86A-3)$$

where

Insertion_loss(f) is the insertion loss at frequency *f*
f is the frequency in GHz

For the MCB,

$$Insertion_loss(f) = 0.0006 + 0.16\sqrt{f} + 0.0587f \text{ dB} \quad 0.01 \leq f \leq 11.1 \quad (86A-4)$$

where

$Insertion_loss(f)$ is the insertion loss at frequency f
 f is the frequency in GHz

86A.5.1.1.2 Electrical specifications of mated HCB and MCB

The limits on the differential insertion loss in decibels of the mated HCB and MCB (in either direction) are given in Equation (86A-5) (which defines the minimum insertion loss) and Equation (86A-6) (which defines the maximum insertion loss). These limits are illustrated in Figure 86A-3.

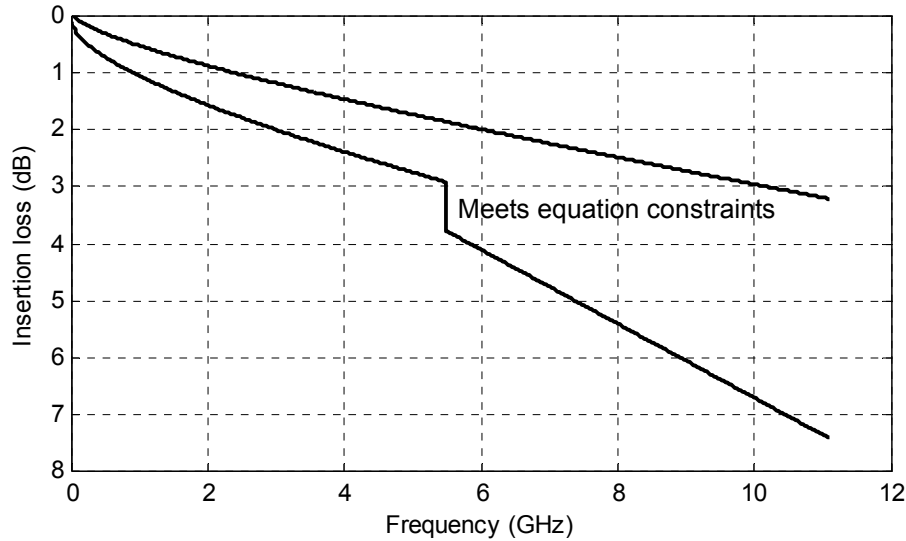


Figure 86A-3—Differential insertion loss limits of mated HCB-MCB

$$Insertion_loss(f) \geq -0.11 + 0.46\sqrt{f} + 0.16f \quad \text{dB} \quad 0.01 \leq f \leq 11.1 \quad (86A-5)$$

$$Insertion_loss(f) \leq \left\{ \begin{array}{ll} 0.029 + 0.861\sqrt{f} + 0.158f & 0.01 \leq f < 5.5 \\ 0.2 + 0.65f & 5.5 \leq f \leq 11.1 \end{array} \right\} \quad \text{dB} \quad (86A-6)$$

where

$Insertion_loss(f)$ is the insertion loss at frequency f
 f is the frequency in GHz

The limit on the differential return loss in decibels of the mated HCB and MCB, looking into the HCB or looking into the MCB, is given in Equation (86A-7), and illustrated in Figure 86A-4.

$$Return_loss(f) \geq \left\{ \begin{array}{ll} 20 - 2f & 0.01 \leq f < 2.5 \\ 15 & 2.5 \leq f < 5 \\ 13.8 - 28.85 \log_{10}\left(\frac{f}{5.5}\right) & 5 \leq f \leq 11.1 \end{array} \right\} \quad \text{dB} \quad (86A-7)$$

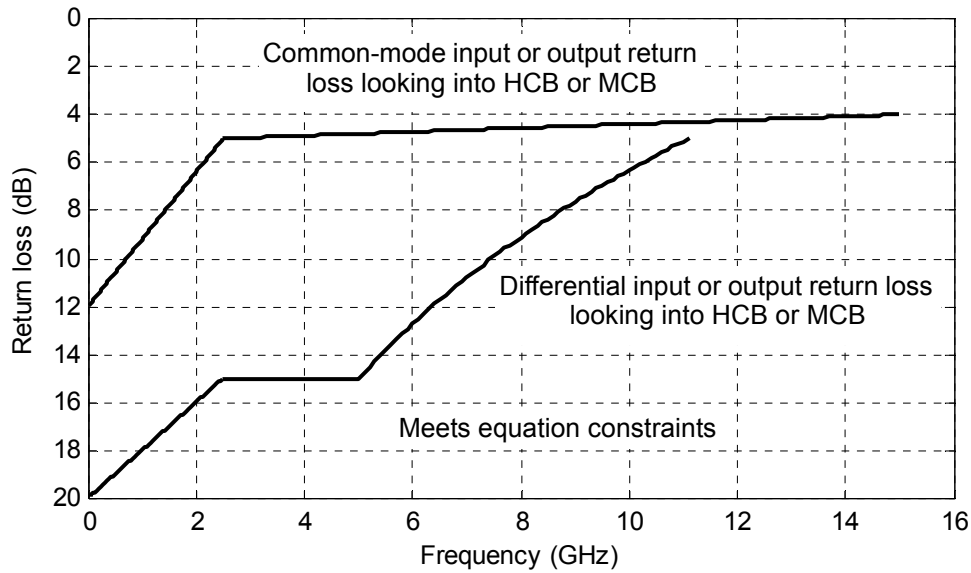


Figure 86A-4—Return loss limits of mated HCB-MCB

where

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

The limit on the common-mode return loss in decibels of the mated HCB and MCB is given in Equation (86A-8) and illustrated in Figure 86A-4.

$$Return_loss(f) \geq \begin{cases} 12 - 2.8f & 0.01 \leq f < 2.5 \\ 5.2 - 0.08f & 2.5 \leq f \leq 15 \end{cases} \text{ dB} \tag{86A-8}$$

where

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

The limit on the differential to common-mode conversion loss in decibels of the mated HCB and MCB, for input to HCB and output from MCB, or input to MCB and output from HCB, is given in Equation (86A-9) and illustrated in Figure 86A-5.

$$Mode_conversion_loss(f) \geq \begin{cases} 30 - 2.91f & 0.01 \leq f < 5.5 \\ 14 & 5.5 \leq f \leq 15 \end{cases} \text{ dB} \tag{86A-9}$$

where

Mode_conversion_loss(f) is the mode conversion loss at frequency *f*
f is the frequency in GHz

The limits on integrated crosstalk noise of the mated HCB and MCB for XLPP1 and CPPI are as specified in 85.10.9.4 for 40GBASE-CR4 and 100GBASE-CR10 respectively, with the exception that the frequency range is 0.01 GHz to 12 GHz.

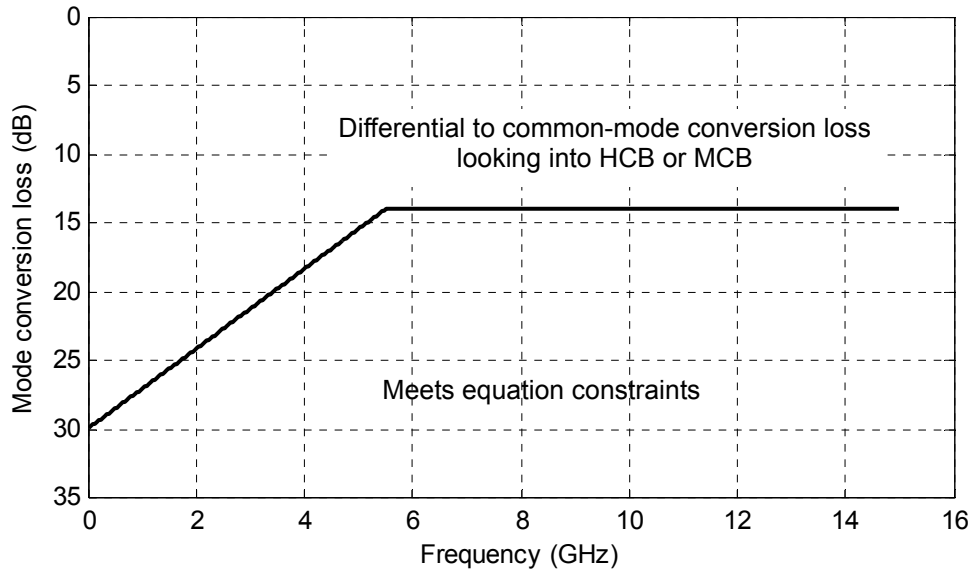


Figure 86A-5—Differential to common-mode conversion loss limit of mated HCB-MCB

86A.5.2 Test patterns and related subclauses

While compliance is to be achieved in normal operation, specific test patterns are defined for measurement consistency and to enable measurement of some parameters. Table 86-11 lists the defined test patterns, and Table 86A-6 gives the test patterns to be used in each measurement, unless otherwise specified, and also lists references to the subclauses in which each parameter is defined. Multi-lane testing considerations are

Table 86A-6—Test patterns and related subclauses

Parameter	Pattern	Related subclause
J2 Jitter	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.3.3.1
J9 Jitter	3 or 5	86.8.3.3.2
Data Dependent Pulse Width Shrinkage (DDPWS)	4	86A.5.3.4
AC common-mode voltage	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86A.5.3.1
Transition time	Square wave or 4	86A.5.3.3
Electrical waveform (eye mask)	3, 5, or valid 40GBASE-SR4 or 100GBASE-SR10 signal	86.8.3.2, 86A.5.3.6
Q_{sq}	Square wave or 4	86A.5.3.5
Host electrical receiver signal tolerance	3 or 5	86A.5.3.8

given in 86.8.2.1. As Pattern 3 is more demanding than Pattern 5 (which itself is the same or more demanding than other 40GBASE-R or 100GBASE-R bit streams) an item which is compliant using Pattern 5 is considered compliant even if it does not meet the required limit using Pattern 3.

86A.5.3 Parameter definitions

In addition to the parameter definitions in the following subclause, some definitions with dual use (both optical and electrical) are given in 86.8.3.

86A.5.3.1 AC common-mode voltage

The common-mode voltage of a differential signal at any time is the average of signal+ and signal– at that time. RMS AC common-mode voltage may be calculated by applying the histogram function over 1 UI to the common-mode signal. As AC common-mode generation is very sensitive to the cable or oscilloscope delay mismatch, it is recommended to delay match the oscilloscope inputs for any measurements.

86A.5.3.2 Termination mismatch

Termination mismatch is the percentage difference between the two low-frequency impedances to common of a differential electrical port. Termination mismatch is defined as shown in Equation (86A–10).

$$\Delta Z_M = 2 \times \frac{|Z_p - Z_n|}{Z_p + Z_n} \times 100 \% \quad (86A-10)$$

Termination mismatch can be measured by applying a low-frequency test tone to the differential inputs as shown in Figure 86A–6. The test frequency must be high enough to overcome the high pass effects of any AC coupling capacitors. The measured differential output or input impedance is designated by Z_{diff} .

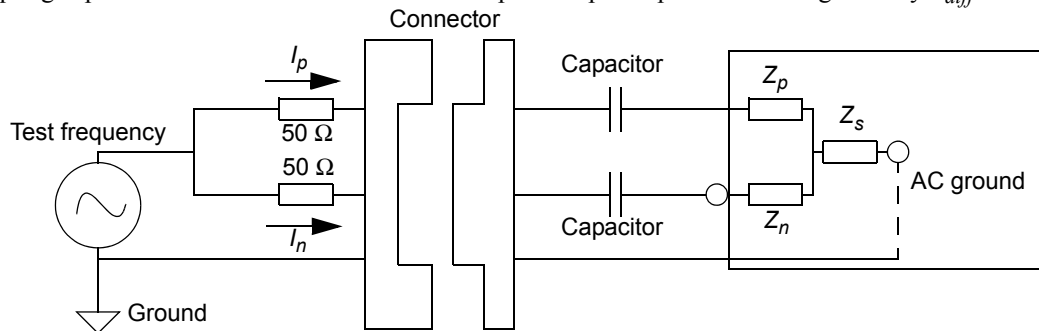


Figure 86A–6—Measurement of AC termination mismatch

For a 100 Ω port, low-frequency termination mismatch is then given by Equation (86A–11).

$$\Delta Z_M = 2 \times \frac{|I_p - I_n|}{I_p + I_n} \times \frac{Z_{diff} + 100}{Z_{diff}} \times 100 \% \quad (86A-11)$$

where I_p and I_n are the currents flowing into the port as shown in Figure 86A–6.

Z_s is the effective series impedance between the terminations Z_p and Z_n and the AC ground.

86A.5.3.3 Transition time

In this annex, transition times (rise and fall times) are defined as the time between the 20% and 80% times, or 80% and 20% times, respectively, of isolated edges.

If the test pattern is the square wave with eight ones and eight zeros, the 0% level and the 100% level are as defined by the OMA measurement procedure (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 by the average signal within windows from -3 UI to -2 UI and from 2 UI to 3 UI relative to the edge.

For electrical signals, the waveform is observed through a 12 GHz low-pass filter response (such as a Bessel-Thomson response).

NOTE—This definition is not the same as the rise and fall times typically reported by an oscilloscope from an eye diagram, which take all the edges into account.

86A.5.3.4 Data Dependent Pulse Width Shrinkage (DDPWS)

An oscilloscope, time interval analyzer, or other instrument with equivalent capability may be used to measure DDPWS. A clock recovery unit (CRU) as defined in 86.8.3.2 is used to trigger the oscilloscope. A repeating PRBS9 pseudo-random test pattern, 511 bits long, is used. For electrical jitter measurements, the measurement bandwidth is 12 GHz (such as a Bessel-Thomson response). If the measurement bandwidth affects the result, it can be corrected for by post-processing. However, a bandwidth above 12 GHz is expected to have little effect on the results.

The crossing level is the average value of the entire waveform being measured. The instrument is synchronized to the pattern repetition frequency and the waveforms or the crossing times are averaged sufficiently to remove the effects of random jitter and noise in the system. The PRBS9 pattern has 128 positive-going transitions and 128 negative-going transitions. The crossing times t_1 to t_{256} of each transition of the averaged waveform (when the averaged waveform crosses the crossing level) are found as shown in Figure 86A-7.

The DDPWS is the difference between one symbol period and the minimum of all the differences between pairs of adjacent transitions as follows in Equation (86A-12).

$$DDPWS = T - \min(t_2 - t_1, t_3 - t_2, \dots, t_{256+1} - t_{256}) \quad (86A-12)$$

where T is one symbol period.

Note that the difference from the next edge in the repeating sequence, t_{256+1} , is also considered.

86A.5.3.5 Signal to noise ratio Q_{sq}

Q_{sq} is a measure of signal to noise ratio. For an electrical signal, it is analogous to Q_{sq} defined for an optical signal in 68.6.7, and it relates the low-frequency signal amplitude to the noise in an electrical -3 dB bandwidth of 12 GHz. It is defined with all co-propagating and counter-propagating crosstalk sources active, using one of patterns 3, 5, or a valid 40GBASE-SR4 or 100GBASE-SR10 signal. The input lanes of the item under test are receiving signals that are asynchronous to those being output.

Q_{sq} may be measured using an oscilloscope as follows:

86A.5.3.8.1 Introduction

This subclause provides guidance for jitter tolerance testing at the host input (PMA) compliance point TP4/TP4a. Compliance is required with input jitter, vertical eye closure (Y1), and vertical peak level (Y2) as specified in Table 86A-4. Compliance is defined at an interface BER (the average of the BERs of all the lanes when stressed) of 10^{-12} . There are two test conditions: one each for the sensitivity and overload vertical eye parameters conditions. The reference test procedure is described in detail for a single stressed lane. Each Rx lane is tested in turn while all lanes are operated. Aggressor lanes are operated with the VMA specified in Table 86A-4.

86A.5.3.8.2 Test equipment and setup

A jitter tolerance test configuration is shown in Figure 86A-8. A test source is used to continuously generate the test signal. The test signal is conditioned within the guidelines outlined in 86A.5.3.8.3 to exhibit the appropriate jitter stress.

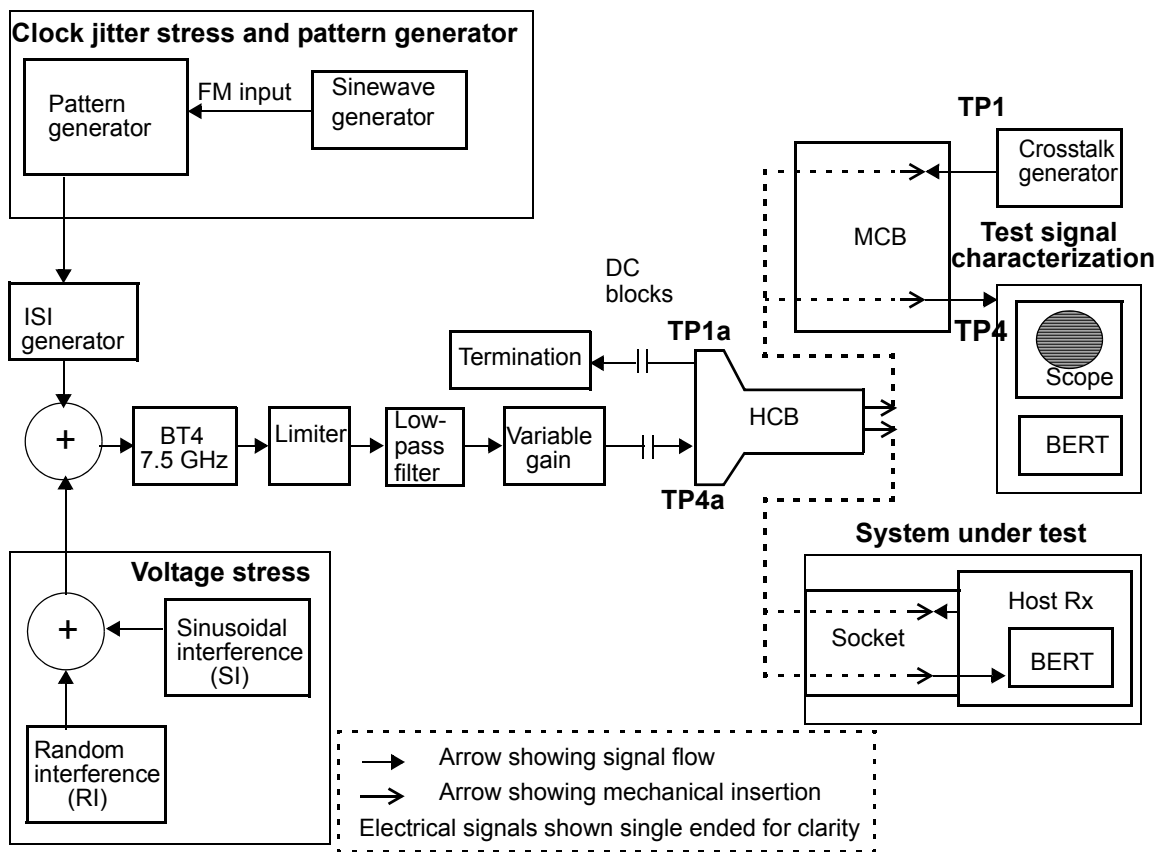


Figure 86A-8—Example jitter tolerance test configuration

An RF attenuator or other output amplitude control of the test source may be required to set the vertical eye opening of the stressed eye.

The test equipment measured at TP4a looking towards the test signal source has better than 20 dB output return loss up to 12 GHz. The output return loss of the test system when measured at TP4 with the Module Compliance Board (looking towards the test signal source) is 2 dB better than the specifications of Table 86A-3 up to 8 GHz and 1 dB better up to 11 GHz.

86A.5.3.8.3 Stressed eye jitter characteristics

This subclause describes required test signal characteristics along with considerations and suggested approaches for test signal generation. The test signal is generated by the functions shown in Figure 86A–8 or by equivalent means. Figure 86A–9 illustrates how the jitter parameters in Table 86A–4 map to the jitter components in the stressed-eye test signal.

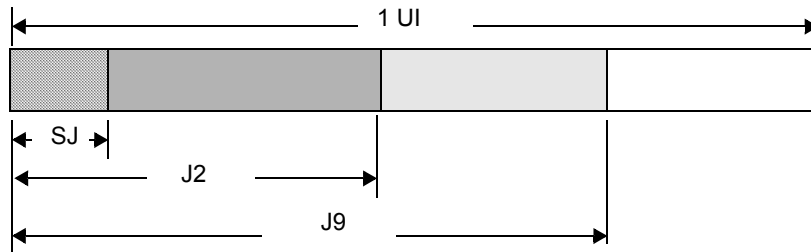


Figure 86A–9—Stressed eye jitter components

The 0.05 UI Sinusoidal Jitter (SJ) component of J2 Jitter is defined for frequencies much higher than the CDR bandwidth (e.g., ~20 MHz). At lower frequencies, the CDR must track additional applied SJ as detailed in the relevant specifications (see Figure 86A–10 and 52.8.1).

The balance of the J2 Jitter is created by the following mechanisms: an ISI generator, sinusoidal interference (SI), and random interference (RI), all passed through a limiting function.

The test signal at TP4 has DDPWS as defined by Table 86A–4. Any duty cycle distortion (DCD) in the test signal does not exceed 0.02 UI. DCD is the difference between the mean position of all falling edges and the mean position of all rising edges of a waveform. It is measured at the average level of the waveform in a 12 GHz bandwidth, using Pattern 3 (PRBS31), Pattern 4 (PRBS9), Pattern 5, or a valid 40GBASE-SR4 or 100GBASE-SR10 signal.

The ISI generator may be implemented by a low-pass filter, length of FR4 trace, length of coax cable, or other equivalent method. It is required that the resulting signal be passed through a limiter function. A suitable limiter function may be implemented using a discrete limiting amplifier followed by a low-pass filter and an attenuator. The low-pass filter emulates the bandwidth and/or slew rate of a practical limiter. The attenuator is used to set the output amplitude to minimum and maximum values allowed by the eye mask coordinates of Table 86A–4.

A voltage stress is to be applied before the limiter function. This stress is composed of a single tone sinusoidal interferer (SI) in the frequency range 100 MHz to 2 GHz and a broadband noise source (RI) with a minimum power spectrum –3 dB point of 6 GHz and minimum crest factor of 7. It is the intent that this combination of voltage stress and limiting function introduce pulse-shrinkage jitter behavior. However, no more than 20% of the J2 Jitter is created by the sinusoidal interferer.

Jitter generation mechanisms for the pattern generator are typically based on phase modulation of the clock source, edge modulation of a variable delay line or a combination thereof.

Any approach that modulates or creates the appropriate levels and frequencies of the jitter components is acceptable.

86A.5.3.8.4 Calibration

Calibration of the test signal is performed using the guidelines for test setup in 86A.5.3.8.2 and illustrated in Figure 86A–8. The aim of the calibration is to achieve a test signal exhibiting jitter stress in accordance with Table 86A–4.

The test signal is calibrated differentially into standard instrumentation loads. If complementary single-ended signals are used, they are carefully matched in both amplitude and phase.

For accurate calibration, it is imperative that all elements in the signal path (cables, DC blocks, etc.) have wide and smooth frequency response as well as linear phase response throughout the spectrum of interest. Baseline wander and overshoot/undershoot are minimized.

An AC coupling –3 dB corner frequency of 20 kHz is expected to be adequate to eliminate baseline wander effects; however, high-frequency performance is critical and must not be sacrificed by the AC coupling.

Given random jitter and the nature of the long test patterns, low probability jitter events will likely be present. It is recommended for jitter calibration that a technique that can accurately measure low probability events be used to avoid overly stressful test conditions.

It is recommended that the actual compliance test pattern be used during calibration. For jitter stress calibration it is permissible, however, to use any appropriate test pattern that still results in the creation of a compliance test pattern with the appropriate jitter stress.

86A.5.3.8.5 Calibration procedure

The vertical eye opening and peak level are set approximately to the levels specified in Table 86A–4.

With an applied calibration test pattern and no additional jitter stress applied, the intrinsic J2 Jitter and J9 Jitter of the test source due to intrinsic noise and finite bandwidth effects are measured. At this stage, J2 Jitter is less than 0.15 UI and J9 Jitter less than 0.25 UI.

SJ is added until the J2 Jitter increases by 0.05 UI above this intrinsic J2 Jitter level. The SJ jitter frequency is well above the CDR bandwidth and asynchronous to the characteristic frequencies of the signal.

Next, additional high probability jitter as specified in 86A.5.3.8.3 is added by the ISI generator until at least 80% of the J2 Jitter has been created. The sinusoidal interferer amplitude is then turned on and adjusted until the required level of J2 Jitter is achieved. The frequency of any sinusoidal interferer is asynchronous to the characteristic frequencies of the signal.

A compliant test signal exhibits Data Dependent Pulse Width Shrinkage (defined in 86A.5.3.4) as specified in Table 86A–4. This is measured with noise and clock-jitter sources turned off.

Once the required level of J2 Jitter has been achieved, turn on the crosstalk source that is set such that at the output of the Host Compliance Board, the amplitude and the transition time are as given in Table 86A–4. The crosstalk pattern is Pattern 3 (PRBS31), Pattern 5, or a valid 40GBASE-SR4 or 100GBASE-SR10 signal, and is asynchronous with the test signal. Then the RI (random interference) voltage stress is added until the specified value of J9 Jitter is achieved.

If necessary, the sine interferer is readjusted to obtain the required level of J2 Jitter and if the sinusoidal interferer is changed then the random interferer is readjusted to obtain the required level of J9 Jitter. Iterative adjustments of the sinusoidal interferer and random interferer are made until the required values of both J2 Jitter and J9 Jitter are achieved.

If necessary, the vertical eye opening is readjusted to required levels.

The vertical eye opening and peak level specifications are verified.

Care must be taken when characterizing the signal used to make receiver tolerance measurements. The intrinsic noise and jitter introduced by the calibration measurement equipment (e.g., filters, oscilloscope and BERT) must be accounted for and controlled. If equipment imperfections affect the results materially, corrections such as root-sum-of-squares deconvolution of Gaussian noise and jitter are used.

86A.5.3.8.6 Test procedure

Testing is performed differentially through a Host Compliance Board (see 86A.5.1).

Using a test signal arranged according to 86A.5.3.8.2 and calibrated according to 86A.5.3.8.5, operate the system with the test pattern specified in Table 86A-6. Each lane is tested in turn while all are operated. Aggressor lanes are operated with the VMA specified in Table 86A-4. The BERs of all the lanes when stressed are averaged to form the interface BER. See 86.8.2.1.

All signals and reference clocks that operate during normal operation are active during the test including all the other host lanes in both directions. The test signal and the host’s transmitted signals are asynchronous. The host transmits Pattern 3 (PRBS31), Pattern 5, or a valid 40GBASE-SR4 or 100GBASE-SR10 signal. The sinusoidal jitter is stepped across the frequency and amplitude range according to Table 86A-7 and illustrated in Figure 86A-10, while monitoring the BER of the lane(s). The interface BER of a compliant host receiver remains below 10^{-12} .

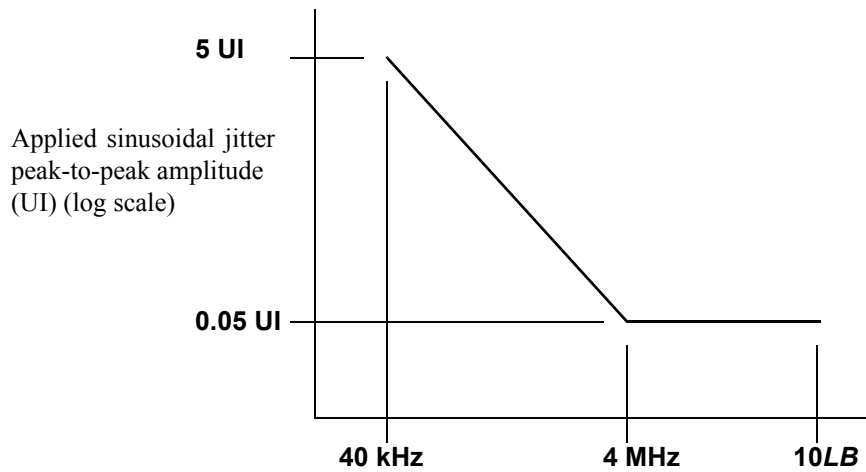


Figure 86A-10—Mask of the sinusoidal component of jitter tolerance

Table 86A-7—Applied sinusoidal jitter

Frequency range	Sinusoidal jitter, peak-to-peak, (UI)
$f < 40 \text{ kHz}$	Not specified
$40 \text{ kHz} \leq f < 4 \text{ MHz}$	$2 \times 10^5 / f$
$4 \text{ MHz} \leq f \leq 10 \text{ LB}^a$	0.05

^aLB = loop bandwidth; upper frequency bound for added sine jitter should be at least 10 times the loop bandwidth of the receiver being tested.

86A.6 Recommended electrical channel

The recommended limits for the differential insertion loss in decibels of the host PCB and connector mated to the HCB, between the PMA IC (TP0 or TP5) and TP1a or TP4a, are given in Equation (86A-14) and Equation (86A-15), and illustrated in Figure 86A-11. It is recommended that

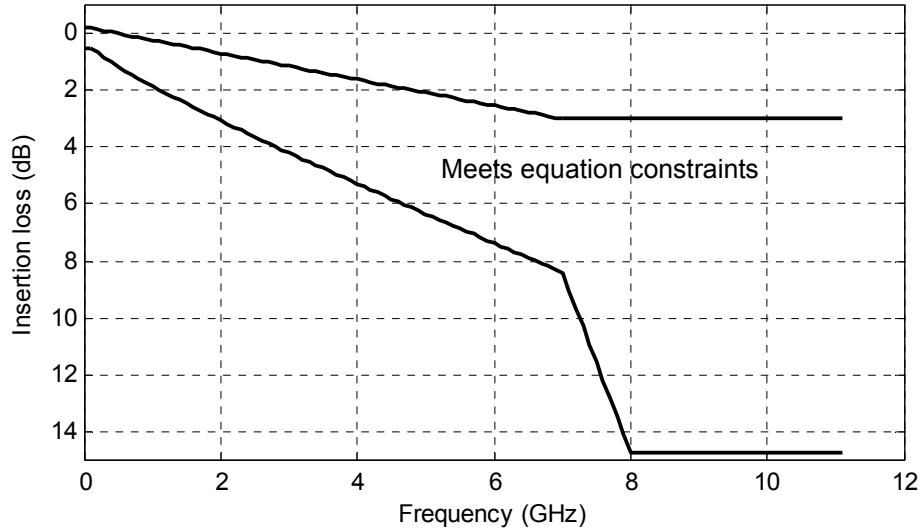


Figure 86A-11—Recommended insertion loss limits of host PCB, connector and HCB

$$Insertion_loss(f) \leq \left\{ \begin{array}{ll} 0.5 & 0.01 \leq f < 0.11 \\ 0.114 + 0.8914\sqrt{f} + 0.846f & 0.11 \leq f < 7 \\ -35.91 + 6.3291f & 7 \leq f < 8 \\ 14.72 & 8 \leq f \leq 11.1 \end{array} \right\} \text{ dB} \quad (86A-14)$$

and

$$Insertion_loss(f) \geq \left\{ \begin{array}{ll} -0.22 + 0.46f & 0.01 \leq f < 7 \\ 3 & 7 \leq f \leq 11.1 \end{array} \right\} \text{ dB} \quad (86A-15)$$

where

$Insertion_loss(f)$ is the insertion loss at frequency f
 f is the frequency in GHz

The recommended maximum loss of the host PCB only (without connector or HCB) at 5.15625 GHz is 4.4 dB.

86A.7 Safety, installation, environment, and labeling

86A.7.1 General safety

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

86A.7.2 Installation

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

86A.7.3 Environment

The 40GBASE-SR4 and 100GBASE-SR10 operating environment specifications are as defined in 52.11, as defined in 52.11.1 for electromagnetic emission, and as defined in 52.11.2 for temperature, humidity, and handling.

86A.7.4 PMD labeling

The 40GBASE-SR4 and 100GBASE-SR10 labeling recommendations and requirements are as defined in 52.12.

86A.8 Protocol implementation conformance statement (PICS) proforma for Annex 86A, Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPi) and 100GBASE-SR10 (CPPI)²²

86A.8.1 Introduction

The supplier of an XLPPi/CPPI implementation that is claimed to conform to Annex 86A, Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPi) and 100GBASE-SR10 (CPPI), 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 same, can be found in Clause 21.

86A.8.2 Identification

86A.8.2.1 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).	

86A.8.2.2 Protocol summary

Identification of protocol standard	IEEE Std 802.3-2012, Annex 86A, Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPi) and 100GBASE-SR10 (CPPI)
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 802.3-2012.)	

Date of Statement	
-------------------	--

²²Copyright 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.

86A.8.3 Major capabilities/options

Item	Feature	Subclause	Value/Comment	Status	Support
*R4	40GBASE-R4	86A.4	Can operate as part of 40GBASE-SR4 or 40GBASE-LR4 PMA or PMD	O.1	Yes [] No []
*R10	100GBASE-SR10	86A.4	Can operate as part of 100GBASE-SR10 PMA or PMD	O.1	Yes [] No []
*MO	Module		Items marked with MO are applicable to a module (PMD)	O.2	Yes [] No []
*HO	Host		Items marked with HO are applicable to a host (PMA)	O.2	Yes [] No []
*MD	MDIO capability	86A.1	Registers and interface supported	O	Yes [] No []

86A.8.4 PICS proforma tables for Parallel Physical Interface (nPPI) for 40GBASE-SR4 and 40GBASE-LR4 (XLPPI) and 100GBASE-SR10 (CPPI)**86A.8.4.1 PMD functional specifications**

Item	Feature	Subclause	Value/Comment	Status	Support
SF1	Compatible with 40GBASE-R or 100GBASE-R PCS and PMA and 40GBASE-SR4, 40GBASE-LR4 or 100GBASE-SR10 PMD	86A.1		M	Yes []

86A.8.4.2 Electrical specifications for nPPI

Item	Feature	Subclause	Value/Comment	Status	Support
S1	Signaling rate per Table 86–2	86A.4	10.3125 GBd ±100 ppm	M	Yes []
S2	Host output signal per Table 86A–1	86A.4.1	Per definitions in 86A.5	HO:M	Yes [] N/A []
S3	Module electrical input per Table 86A–2	86A.4.1	Per definitions in 86A.5	MO:M	Yes [] N/A []
S4	Module electrical input AC coupled	86A.4.1		MO:M	Yes [] N/A []
S5	Module electrical differential input return loss at TP1	86A.4.1.1	Per Equation (86A–1)	MO:M	Yes [] N/A []
S6	Host differential output return loss at TP1a	86A.4.1.1	Per Equation (86A–1)	HO:M	Yes [] N/A []
S7	Module electrical output per Table 86A–3	86A.4.2	Per definitions in 86A.5	MO:M	Yes [] N/A []
S8	Host input per Table 86A–4	86A.4.2	Per definitions in 86A.5	HO:M	Yes [] N/A []
S9	Module electrical output AC coupled	86A.4.2		MO:M	Yes [] N/A []
S10	Module electrical output return loss at TP4	86A.4.2.1	Per Equation (86A–2)	MO:M	Yes [] N/A []
S11	Host differential input return loss at TP4a	86A.4.2.1	Per Equation (86A–2)	HO:M	Yes [] N/A []

86A.8.4.3 Definitions of parameters and measurement methods

Item	Feature	Subclause	Value/Comment	Status	Support
SEM1	Compliance boards	86A.5.1	Use compliance boards, correct as necessary	M	Yes []
SEM2	Compliance boards	86A.5.1.1	Results corrected	M	Yes []
SEM3	Compliance boards	86A.5.1.1	Individual insertion losses per 86A.5.1.1.2	M	Yes []
SEM4	Compliance boards	86A.5.1.1	Mated HCB-MCB per 86A.5.1.1.2	M	Yes []
SEM5	Host input signal tolerance	86A.5.3.8	As specified	HO:M	Yes [] N/A []

86A.8.4.4 Environmental and safety specifications

Item	Feature	Subclause	Value/Comment	Status	Support
SES1	General safety	86A.7.1	Conforms to IEC 60950-1	M	Yes []
SES3	Electromagnetic interference	86A.7.3	Complies with applicable local and national codes for the limitation of electromagnetic interference	M	Yes []

