Table of contents

Suite 100, 2345 North First Street, San Jose, CA 95131

http://www.laranetworks.com

List of figures

Error! No table of figures entries found.

List of tables

Error! No table of figures entries found.

Change history

The following table shows the change history for this user's manual.

Version 1.0 (date)

Original version.

Version X.X (date)

Category	Description
Editorial	Description here
Technical	Description here

Note:

An “Introduction to RPR PHYs section” would be helpful.

This should provide an overview of PHYs, interfaces, etc, and needs to address mapping, since it’s not clear how to map a PHY to ringlet(s). Mapping a single PHY to a ringlet is the most obvious, but does not take advantage of PHYs that include “duplex” Layer 1 fault detection and processing (such as near-end, far-end fault reporting in 10GbE WAN PHYs).

1. Ten Gigabit Ethernet Reconciliation Sublayer (RS) and 10 Gigabit Media Independent Interface (XGMII)

Note: Text is largely borrowed from P802.3ae/D3.1. Need to insert figures. XGMII remains unchanged, RS has been modified for RPR.

1.1 Overview

This clause defines the logical and electrical characteristics for the Reconciliation Sublayer (RS) and 10 Gigabit Media Independent Interface (XGMII) between the RPR media access control (MAC) layer and various 10 Gigabit Ethernet PHYs. Figure 2 shows the relationship of the RS and XGMII to the ISO (IEEE) OSI reference model.

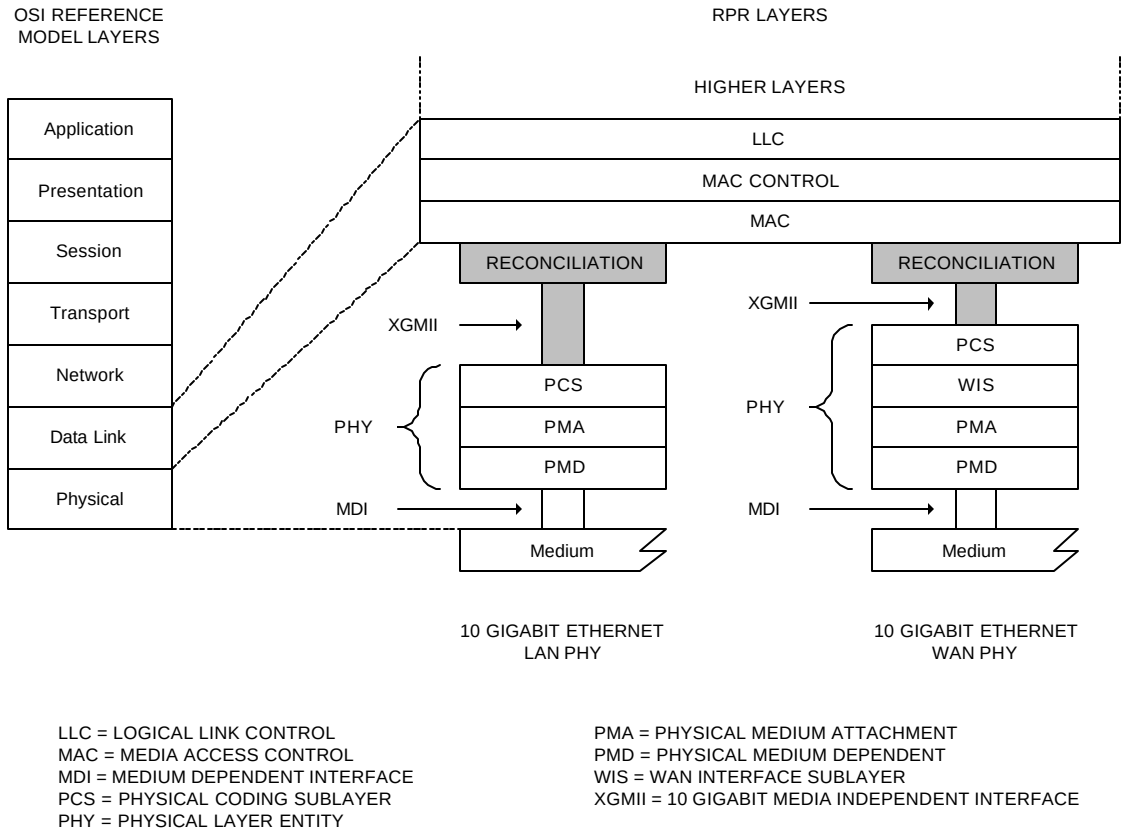

Figure 1-RS and XGMII location in the OSI protocol stack

The purpose of the RS is to convert the logical P-SAP interface of the MAC to the XGMII interface that is specified by P802.3ae as an interface to 10 Gigabit PHYs. Though the XGMII is an optional interface, it is used extensively in this standard as a basis for specification. The 10 Gb/s Physical Coding Sublayer (PCS)

is specified to the XGMII interface, so if not implemented, a conforming implementation shall behave functionally as if the RS and XGMII were implemented.

The purpose of the 10 Gigabit Media Independent Interface (XGMII) is to provide a simple, inexpensive, and easy-to-implement interconnection between the Media Access Control (MAC) sublayer and the Physical layer (PHY). The 10 Gigabit Attachment Unit Interface (XAUI) may optionally be used to extend the operational distance of the XGMII with reduced pin count.

The RS adapts the logical P-SAP interface of the MAC to the parallel encodings of 10 Gb/s PHYs. Though the XGMII is an optional interface, it is used extensively in this standard as a basis for specification. The 10 Gb/s Physical Coding Sublayer (PCS) is specified to the XGMII interface, so if not implemented, a conforming implementation shall behave functionally as if the RS and XGMII were implemented.

The XGMII has the following characteristics:

a) It is capable of supporting 10 Gb/s operation;
b) Data and delimiters are synchronous to clock reference;
c) It provides independent 32-bit-wide transmit and receive data paths;
d) It uses signal levels compatible with common digital ASIC processes;
e) It provides for full-duplex operation only

1.1.1 Summary of major concepts

a) The RS converts the P-SAP service primitives to the parallel data paths of the XGMII;
b) The RS maps the signal set provided at the XGMII to the P-SAP service primitives provided at the MAC;
c) Each direction of data transfer is independent and serviced by data, control, and clock signals;
d) The RS generates continuous data or control characters on the transmit path and expects continuous data or control characters on the receive path;
e) The RS participates in link fault detection and reporting by monitoring the receive path for status reports that indicate an unreliable link;
f) When the XGMII is optionally extended with XAUI, two XGMII interfaces logically exist (see Figure 46–1). When extended, signals on the transmit path are from the RS to an XGXS of the XAUI interface on one XGMII and from the XGXS to the PHY on the other XGMII. When extended, signals on the receive path are from the PHY to an XGXS of the XAUI interface on one XGMII and from the XGXS to the RS on the other XGMII. The descriptions of the XGMII as between the RS and the PHY are therefore equally applicable between the RS and the XGXS or the XGXS and the PHY.

1.1.2 Application

This clause applies to the interface between the MAC and PHY. The physical implementation of the interface is primarily intended as a chip-to-chip (integrated circuit to integrated circuit) interface implemented with traces on a printed circuit board. The XGMII may also be used in other ways, for example, as a logical interface between ASIC logic modules within an integrated circuit.

This interface is used to provide media independence so that an identical media access controller may be used with all 10GBASE PHY types using either serial or wavelength division multiplexed optics.

1.1.3 Rate of operation

The XGMII supports only the 10 Gb/s data rate as defined within this clause.

1.1.4 Delay Constraints

Predictable operation of the MAC Control PAUSE operation requires that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sub-layer, 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 46–1 contains the values of maximum cumulative MAC Control, MAC and RS round-trip (sum of transmit and receive) delay in bit time as specified in 1.4 and pause quanta as specified in 31B.2.

1.1.5 Allocation of functions

The allocation of functions at the XGMII balances the need for media independence with the need for a simple and cost-effective interface. The bus width and signaling rate are applicable to short distance chip-to-chip interconnect with printed circuit board trace lengths electrically limited to approximately 7 cm. The XGMII maximizes media independence by cleanly separating the Data Link and Physical Layers of the ISO (IEEE) seven-layer reference model.

1.1.6 XGMII structure

The XGMII is composed of independent transmit and receive paths. Each direction uses 32 data signals (TXD<31:0> and RXD<31:0>), four control signals (TXC<3:0> and RXC<3:0>) and a clock (TX_CLK and RX_CLK). Figure 46–2 depicts a schematic view of the RS inputs and outputs.

The 32 TXD and four TXC signals shall be organized into four data lanes, as are the 32 RXD and four RXC signals (see Table 46–2). The four lanes in each direction share a common clock — TX_CLK for transmit and RX_CLK for receive. The four lanes are used in round-robin sequence to carry an octet stream. On transmit, each eight PHY_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, then repeating with the fifth 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.

1.1.7 Mapping of XGMII signals to P-SAP service primitives

The Reconciliation Sublayer (RS) shall map the signals provided at the XGMII to the P-SAP service primitives as shown in **Error! Reference source not found.**. The following P-SAP service primitives are defined:

PHY_DATA.request
PHY_DATA.indicate
PHY_DATA_VALID.indicate
PHY_LINK_OK.indicate
PHY_READY.indicate

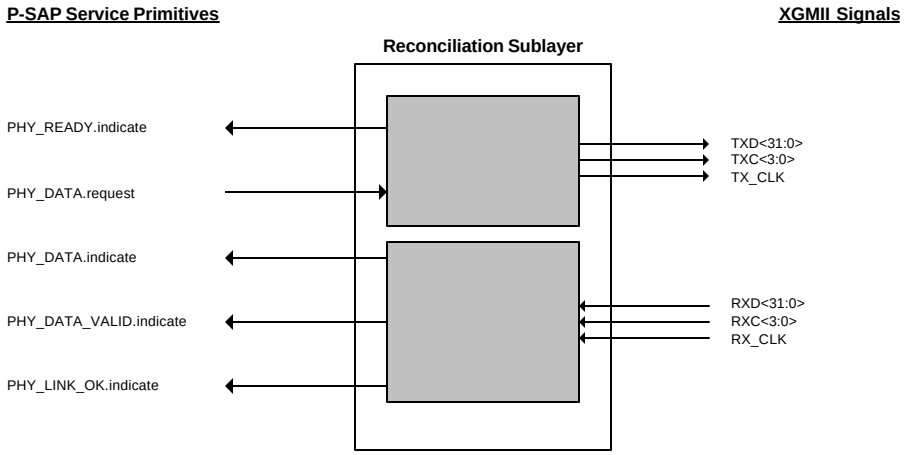

Figure 46–2—Reconciliation Sublayer (RS) inputs and outputs

1.1.7.1 Mapping of PHY_DATA.request

1.1.7.1.1 Function

Map the primitive PHY_DATA.request to the XGMII signals TXD<31:0>, TXC<3:0> and TX_CLK.

1.1.7.1.2 Semantics of the service primitive

PHY_DATA.request (OUTPUT_UNIT)

The OUTPUT_UNIT parameter either takes the value of an octet of data or DATA_COMPLETE, and represents the transfer of an octet of data from the MAC to the RS. The value DATA_COMPLETE indicates that the MAC has no more data to transfer.

1.1.7.1.3 When generated

This primitive is generated by the MAC sublayer to request the transmission of a data octet on the physical medium, or to indicate that no more data is available for transmission.

1.1.7.1.4 Effect of receipt

The OUTPUT_UNIT values are conveyed to the PHY by the signals TXD<31:0> and TXC<3:0> on each TX_CLK edge. Each PHY_DATA.request transaction shall be mapped to a TXD signal in sequence (TXD<0:7>, ... TXD<24:31>, TXD<0:7>) as described in xxx. After four PHY_DATA.request

transactions from the MAC sublayer, the RS requests transmission of 32 data bits by the PHY. The first eight octets of the frame shall be converted as follows:

a) The first octet is converted to a Start control character and aligned to lane 0;
b) The following six octets are converted to Preamble characters;
c) The eighth character is converted to a SFD character.

The MAC does not generate interframe spacing—the RS generates IPG between MAC frames. The DATA_COMPLETE value shall initiate IPG generation as follows:

a) The DATA_COMPLETE value is mapped to a Terminate control character;
b) Interframe spacing is generated in accordance with IEEE Std 802.3 Clause 4, and mapped to IPG.

Table 46–2—Transmit and receive lane associations

1.1.7.2 Mapping of PHY_DATA.indicate

1.1.7.2.1 Function

Map the primitive PHY_DATA.indicate to the XGMII signals RXD<31:0>, RXC<3:0> and RX_CLK.

1.1.7.2.2 Semantics of the service primitive

PHY_DATA.indicate (INPUT_UNIT)

The INPUT_UNIT parameter takes the value of an octet of data, and defines the transfer of an octet of data from the RS to the MAC.

1.1.7.2.3 When generated

The INPUT_UNIT values are derived from the signals RXC<3:0> and RXD<31:0> received from the PHY on each edge of RX_CLK. Each primitive generated to the MAC sublayer entity corresponds to a PHY_DATA.request issued by the MAC at the remote end of the link connecting two DTEs. For each RXD<31:0> during frame reception, the RS shall generate four PHY_DATA.indicate transactions until the end of frame (Terminate control character), where one, two, three, or four PHY_DATA.indicate transactions will be generated from the RXD<31:0> containing the Terminate. During frame reception, each RXD signal shall be mapped in sequence into a PHY_DATA.indicate transaction (RXD<0:7>, ... RXD<24:31>, RXD<0:7>) as described in xxx.

The RS shall convert a valid Start control character and the subsequent eight octets of data to null data prior to generation of the associated PHY_DATA.indicate transactions. The RS shall not generate any PHY_DATA.indicate 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 XGMII error indications (see 46.3.3). Sequence ordered_sets are not indicated to the MAC (see 46.3.4).

1.1.7.2.4 Effect of receipt

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

1.1.7.3 Mapping of PHY_DATA_VALID.indicate

1.1.7.3.1 Function

Map the primitive PHY_DATA_VALID.indicate to the XGMII signals RXC<3:0> and RXD<31:0>.

1.1.7.3.2 Semantics of the service primitive

PHY_DATA_VALID.indicate (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 PHY_DATA.indicate primitive contains a valid data of an incoming frame. The DATA_NOT_VALID value indicates that the INPUT_UNIT parameter of the PHY_DATA.indicate primitive does not contain valid data of an incoming frame.

1.1.7.3.3 When generated

The PHY_DATA_VALID.indicate 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 PHY_DATA.indicate transaction is generated in response to reception of a Start control character on lane 0 if the prior RXC<3:0> and RXD<31:0> contained four Idle characters or a Sequence ordered set. 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 xxx).

1.1.7.3.4 Effect of receipt

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

1.1.7.4 Mapping of PHY_LINK_OK.indicate

1.1.7.4.1 Function

Map the primitive PHY_LINK_OK.indicate to the XGMII signals RXD<31:0>, RXC<3:0> and RX_CLK.

1.1.7.4.2 Semantics of the service primitive

PHY_LINK_OK.indicate (LINK_STATUS)

The LINK_STATUS parameter takes the value of OK, FAIL, or DEGRADE, and signifies the status of the PHY as indicated by Link Fault Signaling. OK indicates that no Local Fault signal has been received. FAIL indicates that a Local Fault signal has been received. DEGRADE is not generated by the RS.

1.1.7.4.3 When generated

The LINK_STATUS values are derived from the signals RXC<3:0> and RXD<31:0> received from the PHY on each edge of RX_CLK. Details TBD.

1.1.7.4.4 Effect of receipt

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

1.1.7.5 Mapping of PHY_READY.indicate

1.1.7.5.1 Function

Map the primitive PHY_READY.indicate to the XGMII signals RXD<31:0>, RXC<3:0> and RX_CLK.

1.1.7.5.2 Semantics of the service primitive

PHY_READY.indicate (READY_STATUS)

The READY_STATUS parameter takes the value of READY or NOT_READY, and signifies the status of a transmitted frame. If a frame is being received from the MAC and has not been transmitted, READY_STATUS indicates NOT_READY. If a frame has been received and fully transmitted, READY_STATUS indicates READY signifying that a new frame can be received.

1.1.7.5.3 When generated

The LINK_STATUS values is derived from the RS transmit state machine. Details TBD.

1.1.7.5.4 Effect of receipt

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

1.2 XGMII data stream

Data frames transmitted through the XGMII shall be transferred within the XGMII data stream. The 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 46–3.

<inter-frame><preamble><sfd><data><efd>

Figure 46–3—XGMII data stream

For the XGMII, transmission and reception of each bit and mapping of data octets to lanes shall be as shown in Figure 46–4.

1.2.1 Inter-frame <inter-frame>

The inter-frame <inter-frame> period on an XGMII transmit or receive path is an interval during which no frame data activity occurs. The interpacket gap is generated by the RS. The <inter-frame> 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 RS and receiving RS by one or more functions (e.g., RS lane alignment, PHY clock rate compensation or 10GBASE-W data rate adaptation functions). The minimum IPG at the XGMII of the receiving RS is five octets.

The signaling of link status information logically occurs in the <inter-frame> period (see xxx).

1.2.2 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 XGMII. On transmit, the RS converts the first data octet of preamble transferred from the MAC into a Start control character. On receive, the RS will convert the Start control character into a preamble data octet. The start control character is aligned to lane 0 of the XGMII 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 XGMII is unchanged from the Start Frame Delimiter (SFD) specified in 4.2.6 and is the bit sequence:

10101011

Figure 46–4—Relationship of data lanes to MAC serial bit stream

The preamble and SFD are shown above with their bits ordered for serial transmission from left to right. As shown, the left-most bit of each octet is the LSB of the octet and the right-most bit of each octet is the MSB of the octet.

The preamble and SFD are transmitted through the XGMII as octets sequentially ordered on the lanes of the XGMII. The first preamble octet is replaced with a Start control character and it is aligned to lane 0, the second octet on lane 1, the third on lane 2 and the fourth on lane 3, and the four octets are transferred on the next edge of TX_CLK. The fifth octet is assigned to lane 0 with subsequent octets sequentially assigned to the lanes with the SFD assigned to lane 3. The XGMII <preamble> and <sfd> are:

Lane 0	**Lane 1**	**Lane 2**	**Lane**	**3**
Start	10101010	10101010	10101010	
10101010		10101010	10101010	10101011

1.2.3 Data <data>

The data <data> in a well-formed frame shall consist of a set of data octets.

1.2.4 End of frame delimiter <efd>

Assertion of TXC with the appropriate Terminate control character encoding of TXD on a lane constitutes an end of frame delimiter <efd> for the transmit data stream. Similarly, assertion of RXC with the appropriate Terminate control character encoding of RXD constitutes an end of frame delimiter for the receive data stream. The XGMII shall recognize the end of frame delimiter on any of the four lanes of the XGMII.

1.2.5 Definition of Start of Packet and End of Packet Delimiters

For the purposes of Clause 30 layer management the Start of Packet delimiter is defined as the Start control character; and the End of Packet delimiter is defined as the end of the last sequential data octet preceding the Terminate control character or other control character causing a change from DATA_VALID to DATA_NOT_VALID. (See 46.1.7.5.2 and 30.3.2.1.5.)

1.3 XGMII functional specifications

The XGMII is designed to make the differences among the various media and transceiver combinations transparent to the MAC sublayer. The selection of logical control signals and the functional procedures are all designed to this end.

1.3.1 Transmit

1.3.1.1 TX_CLK (10 Gb/s transmit clock)

TX_CLK is a continuous clock used for operation at 10 Gb/s. TX_CLK provides the timing reference for the transfer of the TXC<3:0> and TXD<31:0> signals from the RS to the PHY. The values of TXC<3:0> and TXD<31:0> shall be sampled by the PHY on both the rising edge and falling edge of TX_CLK. TX_CLK is sourced by the RS.

The TX_CLK frequency shall be 156.25MHz ±0.01%, one-sixty-fourth of the MAC transmit data rate.

1.3.1.2 TXC<3:0> (transmit control)

TXC<3:0> indicate that the RS is presenting either data or control characters on the XGMII 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 deasserted 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 XGMII. TXC<3:0> are driven by the RS and shall transition synchronously with respect to both the rising and falling edges of TX_CLK. Table 46–3 specifies the permissible encodings of TXD and TXC for a XGMII 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 1.1—Permissible encodings of TXC, and TXD

TXC	TXD	Description	PHY_DATA.request parameter
0	00 through FF	Normal data transmission	ZERO, ONE (eight bits)
1	00 through 06	Reserved	--
1	07	Idle	No applicable parameter (Normal inter-frame)
1	08 through 9B	Reserved	--
1	9C	Sequence (only valid in lane 0)	No applicable parameter (Inter-frame status signal)
1	9D through FA	Reserved	--
1	FB	Start (only valid in lane 0)	No applicable parameter, replaces first eight ZERO, ONE of a frame (preamble octet)
1	FC	Reserved	--
1	FD	Terminate	DATA_COMPLETE
1	FE	Transmit error propagation	No applicable parameter
1	FF	Reserved	--
Note—Values in TXD column are in hexadecimal, most significant bit to least significant bit (ie, <7:0>).			

1.3.1.3 TXD<31:0> (transmit data)

TXD is a bundle of 32 data signals organized into four lanes of eight signals each (TXD<7:0>, TXD<15:8>, TXD<23:16>, and TXD<31:24>) that are driven by the RS. Each lane is associated with a TXC signal as shown in Table 46–2 and shall be encoded as shown in Table 46–3. TXD<31:0> shall transition synchronously with respect to both the rising and falling edges of TX_CLK. For each high or low TX_CLK transition, data and/or control are presented on TXD<31:0> 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 and TXD<24> the least significant bit of lane 3.

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 46–5.

Figure 46–6 shows the behavior of TXD and TXC during an example transmission of a frame propagating an error.

Figure 46–5—Normal frame transmission

Figure 46–6—Transmit Error Propagation

1.3.1.4 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 two ways.

c) A MAC implementation may incorporate this RS function into its design and always insert additional idle characters to align the start of preamble on a four byte boundary. Note that this will reduce the effective data rate for certain packet sizes separated with minimum inter-frame spacing.

d) 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 three. Note that this may result in inter-frame spacing observed on the transmit XGMII that is up to three octets shorter than the minimum transmitted preamble specified in Clause 4, however the frequency of shortened inter-frame spacing is constrained by the DIC rules. The Deficit Idle Count 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 DIC as described.

1.3.2 Receive

1.3.2.1 RX_CLK (receive clock)

RX_CLK is a continuous clock which provides the timing reference for the transfer of the RXC<3:0> and RXD<31:0> signals from the PHY to the RS. RXC<3:0> and RXD<31:0> shall be sampled by the RS on both the rising and falling 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 156.25MHz ±0.01%, 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. Transitions from nominal clock to recovered clock or from recovered clock to nominal clock shall not decrease the time between adjacent edges of RX_CLK.

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

1.3.2.2 RXC<3:0> (receive control)

RXC<3:0> indicate that the PHY is presenting either recovered and decoded data or control characters on the XGMII. 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 XGMII. RXC<3:0> are driven by the PHY and shall transition synchronously with respect to both the rising and falling edges of RX_CLK. Table 46–4 specifies the permissible encodings of RXD and RXC for a XGMII 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).

Figure 46–7 shows the behavior of RXC<3:0> during an example frame reception with no errors.

Table 1.2—Permissible encodings of RXC, and RXD

TXC	TXD	Description	PHY_DATA.request parameter
0	00 through FF	Normal data transmission	ZERO, ONE (eight bits)
1	00 through 06	Reserved	--
1	07	Idle	No applicable parameter (Normal inter-frame)
1	08 through 9B	Reserved	--
1	9C	Sequence (only valid in lane 0)	No applicable parameter (Inter-frame status signal)

1	9D through FA	Reserved	--
1	FB	Start (only valid in lane 0)	No applicable parameter, replaces first eight ZERO, ONE of a frame (preamble octet)
1	FC	Reserved	--
1	FD	Terminate	No applicable parameter (Start of inter-frame)
1	FE	Receive error	No applicable parameter
1	FF	Reserved	--
Note—Values in RXD column are in hexadecimal, most significant bit to least significant bit (ie, <7:0>).			

Figure 46–7—Basic Frame Reception

1.3.2.3 RXD (receive data)

RXD is a bundle of 32 data signals (RXD<31:0>) organized into four lanes of eight signals each (RXD<7:0>, RXD<15:8>, RXD<23:16>, and RXD<31:24>) that are driven by the PHY. Each lane is associated with a RXC signal as shown in Table 46–2 and shall be encoded as shown in Table 46–4. RXD<31:0> shall transition synchronously with respect to both the rising and falling edges of RX_CLK. For each high or low RX_CLK transition, received data and/or control are presented on RXD<31:0> 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 and RXD<24> the least significant bit of lane 3. Figure 46–7 shows the behavior of RXD<31:0> during frame reception.

While the RXC of a lane is de-asserted, RXD of the lane is used by the RS to generate PHY_DATA.indicate transactions. 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 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 46.3.3.1. Figure 46–8 shows the behavior of RXC and RXD during the reception of an example frame with an error.

Figure 46–8—Reception with error

1.3.3 Error and fault handling

1.3.3.1 Response to error indications by the XGMII

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 PHY_DATA.indicate primitives for each Error control character received within a frame, and may generate eight PHY_DATA.indicate 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.

1.3.3.2 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.

1.3.3.3 Response to indication of invalid frame sequences

The 10 Gb/s PCS is required to align the Start control character to lane 0. The RS shall not indicate DATA_VALID to the MAC for a Start control character received on any other lane. Error free 10 Gb/s operation will not change the SFD alignment in lane 3. A 10 Gb/s MAC/RS implementation is not required to process a packet that has an SFD in a position other than lane 3 of the column following the column containing the Start control character.

1.3.4 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 indicates the fault status to the MAC through the PHY_LINK_OK.indicate service primitive. When the RS no longer receives fault status messages, it returns to normal operation, it signals normal status to the MAC.

Status is signaled in a four byte Sequence ordered_set as shown in Table 46–5. The PHY indicates Local Fault with a Sequence control character in lane 0 and data characters of 0x00 in lanes 1 and 2 plus a data character of 0x01 in lane 3. Remote Fault is not used. 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.

Table 1.3—Sequence ordered_sets

Lane 0	Lane 1	Lane 2	Lane 3	Description
Sequence	0x00	0x00	0x00	Reserved
Sequence	0x00	0x00	0x01	Local Fault
Sequence	0x00	0x00	0x02	Remote Fault, Not Used
Sequence	≥0x00	≥0x00	≥0x03	Reserved
Note—Values in Lane 1, Lane 2, and Lane 3 columns are in hexadecimal, most significant bit to least significant bit (i.e., <7:0>). The link fault signaling state machine allows future standardization of reserved Sequence ordered sets for functions other than link fault indications.				

The RS reports the fault status of the link. Local Fault indicates a fault detected on the receive data path between the remote RS and the local RS. The RS shall implement the link fault signaling state machine (see Figure 46– 9).

1.3.4.2 Conventions

The notation used in the state diagram follows the conventions of 21.5. The notation ++ after a counter indicates it is to be incremented.

1.3.4.3 Variables and counters

The link fault signaling state machine 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 both the rising and falling clock transitions) unless reset.

fault_sequence
: A new column received on RXC<3:0> and RXD<31:0> comprising a Sequence ordered_set of four bytes and consisting of a Sequence control character in lane 0 and a seq_type in lanes 1, 2, and 3 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.
: Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3.

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.
Remote Fault; 0x00 in lane 1, 0x00 in lane 2, 0x02 in lane 3.

1.3.4.4 State Diagram

Figure 46–9—Link Fault Signaling State Machine

The link fault signaling state machine specifies the RS monitoring of RXC<3:0> and RXD<31:0> for Sequence ordered_sets. The variable link_fault is set to indicate the value of a received Sequence ordered_set when the following conditions have been met:

a) Four fault_sequences containing the same fault value have been received
b) Without receiving any fault_sequence within a period of 128 columns

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<3:0> and TXD<31:0> is controlled by the variable link_fault.

a) link_fault = OK
The RS shall send MAC frames as requested through the PHY 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.

1.4 XGMII electrical characteristics

The XGMII electrical characteristics shall meet the requirements of P802.3ae Subclause 46.4.

2. Ten Gigabit Ethernet XGMII Extender Sublayer (XGXS), and 10 Gigabit Attachment Unit Interface (XAUI)

2.1 Overview

2.1.1 Scope

This clause defines the logical and electrical characteristics for the 10 Gigabit Ethernet XGMII Extender Sublayer (XGXS) and 10 Gigabit Attachment Unit Interface (XAUI) between the RPR media access control (MAC) layer and various 10 Gigabit Ethernet PHYs.

2.1.2 Objectives

The following are the objectives of the Resilient Packet Ring RS and XGMII:

a) Support the Resilient Packet Ring MAC;
b) Provide an interface from the RPR RS to 10 Gigabit Ethernet PHYs;
c) Support a BER objective of 10^{-12} .

2.1.3 Summary of major concepts

a) The optional XGXS/XAUI is identical to that specified for 10 Gigabit Ethernet.

2.1.4 Relationship to other sublayers

Figure 2 shows the relationship of the RS and XGMII to the ISO Open System Interconnection (OSI) reference model.

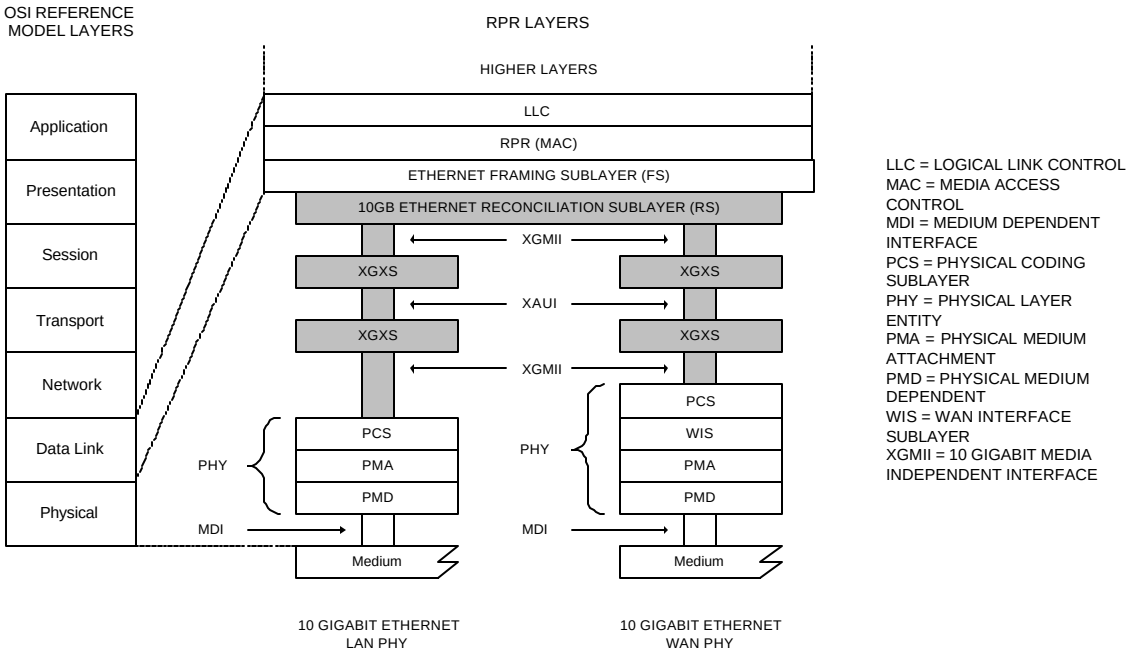

Figure 2-RS, XGMII, and XGXS/XAUI location in the OSI protocol stack

2.2 Requirements

2.2.1 XGXS and XAUI

The optional XGMII Extender Sublayer (XGXS) and optional 10 Gigabit Attachment Unit Interface (XAUI), if implemented, shall meet all requirements specified in Clause 47 of P802.3ae.

3. Ten Gigabit Ethernet Physical Layer Entities (PHYs), Media Dependent Interfaces (MDIs), and Media

3.1 Overview

3.1.1 Scope

This clause defines a family of Resilient Packet Ring (RPR) Physical Layer Entities (PHYs), Media Dependent Interfaces (MDIs), and media based on 10 Gigabit Ethernet as specified in P802.3ae. The PHYs specified here are identical to those specified for 10 Gigabit Ethernet—no changes are made to the Ethernet specifications or requirements.

3.1.2 Objectives

The following are the objectives of the Resilient Packet Ring PHYs, MDIs, and media:

a) Support the Resilient Packet Ring MAC;
b) Provide Ethernet LAN PHYs operating at 10 Gb/s and Ethernet WAN PHYs operating at a SONET STS-192c/SDH VC-4-64c rate;
c) Support cable plants using optical fiber compliant with ISO/IEC 11801: 1995;
d) Allow for a nominal ring segment extent of up to 40 km between nodes;
e) Support a BER objective of 10^{-12} .

3.1.3 Summary of major concepts

a) A family of 10 Gigabit Ethernet LAN PHYs are referenced from P802.3ae, operating at 10 Gbps, summarized in ;
b) A family of 10 Gigabit Ethernet WAN PHYs are referenced from P802.3ae, operating at a data rate data rate and format compatible with SONET STS-192c and SDH VC-4-64c, summarized in ;
c) The Media Dependent Interface characteristics are referenced from P802.3ae;
d) The media requirements are referenced from P802.3ae.

Description	Reach/Fiber	10 GbE Designation	
		LAN PHY	WAN PHY
850 nm serial	~85 m/MMF	10GBASE-SR	10GBASE-SW
1310 nm serial	10 km/SMF	10GBASE-LR	10GBASE-LW
1550 nm serial	40 km/SMF	10GBASE-ER	10GBASE-EW
1310 nm WDM	10 km/SMF ~300m/MMF	10GBASE-LX4	-

Table 1 – 10 GbE PHY summary

3.1.4 Relationship to other sublayers

The relationship of the RPR Physical Layer to other RPR sublayers, and to the ISO Open System Interconnection (OSI) reference model is shown in Figure 3.

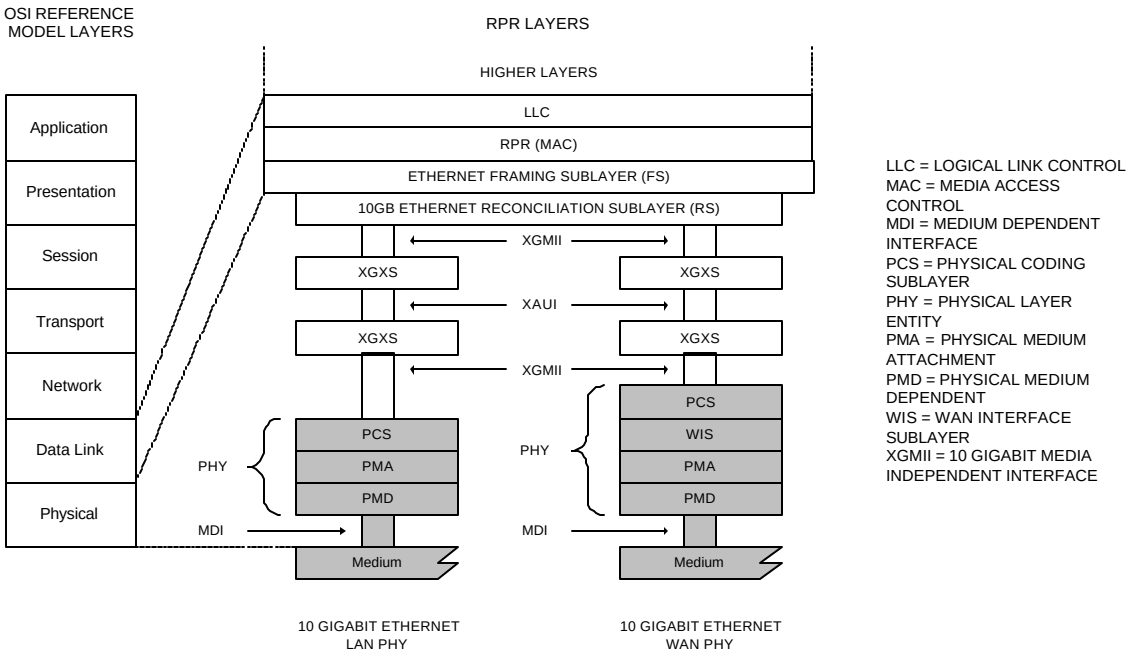

Figure 3-PHY location in the OSI protocol stack

3.1.5 PHY configurations

3.1.5.1 10 Gigabit Ethernet LAN PHYs

This clause specifies a family of Physcial Layer entities based on the P802.3ae 10 Gigabit Ethernet LAN PHY. The LAN PHY operates at a data rate of 10 Gbps. The LAN PHYs consist of the sublayers and configurations identified in Table 2.

Sublayer	P802.3ae Clause Reference	P802.3ae Designation			
		10GBASE-SR	10GBASE-LX4	10GBASE-LR	10GBASE-ER
8B/10B PCA & PMA	48		M[1]		
64B/66B PCS	49	M		M	M
Serial PMA	51	M		M	M
850nm Serial PMD	52	M			
1310nm Serial PMD	52			M	
1550nm Serial PMD	52				M
1300nm WWDM PMD	53		M		

[1]M = Mandatory

Table 2 - 10 Gigabit Ethernet LAN PHY configurations

3.1.5.2 10 Gigabit Ethernet WAN PHYs

This clause specifies a family of Physcial Layer entities based on the 802.3ae 10 Gigabit Ethernet WAN PHY. The WAN PHY operates at a data rate and format compatible with SONET STS-192c and SDH VC-4-64c. The WAN PHY is not intended to interoperate directly with interfaces that comply with SONET or SDH standards.

The WAN PHYs consist of the sublayers and configurations identified in Table 3.

Sublayer	P802.3ae Clause Reference	P802.3ae Designation		
		10GBASE-SW	10GBASE-LW	10GBASE-EW
64B/66B PCS	49	M[1]	M	M
WIS	50	M	M	M
Serial PMA	51	M	M	M
850nm Serial PMD	52	M		
1310nm Serial PMD	52		M	
1550nm Serial PMD	52			M

[1]M = Mandatory

Table 3 - 10 Gigabit Ethernet LAN PHY configurations

3.2 10 Gigabit Ethernet sublayers

3.2.1 8B/10B Physical Coding Sublayer (PCS) & Physical Medium Attachment (PMA) sublayer

The 10 Gigabit Ethernet 8B/10B PCS and PMA sublayer are common to the WWDM PHY implementations specified in this clause. The PCS and PMA sublayers map the interface characteristics of the PMD sublayer to the services expected by the RS and the logical and electrical characteristics of the XGMII.

An 8B/10B PCS and PMA sublayer shall comply with the requirements of Clause 48 of P802.3ae.

3.2.2 64B/66B Physical Coding Sublayer (PCS)

The 10 Gigabit Ethernet 64B/66B PCS sublayer is common to the serial PHY implementations specified in this clause. For LAN PHY implementations, the PCS sublayer maps the interface characteristics of the serial PMA sublayer to the services expected by the RS and the logical and electrical characteristics of the XGMII. Alternatively, for WAN PHY implementations, the PCS sublayer maps the interface characteristics of the serial WIS to the services expected by the RS and the logical and electrical characteristics of the XGMII.

A 64B/66B PCS shall comply with the requirements of Clause 49 of P802.3ae.

3.2.3 WAN Interface Sublayer (WIS)

The 10 Gigabit Ethernet WIS is common to the WAN PHY implementations specified in this clause. The WIS provides mapping of Ethernet data streams to STS-192c or VC-4-64c data streams at the PHY level.

A WIS shall comply with the requirements of Clause 50 of P802.3ae1.

3.2.4 Serial Physical Medium Attachment (PMA) sublayer

The 10 Gigabit Ethernet serial PMA sublayer is common to the serial PHY implementations specified in this clause. The PMA sublayer maps the interface characteristics of the PMD to those of the serial PCS.

A serial PMA sublayer shall comply with the requirements of Clause 51 of P802.3ae.

3.2.5 Serial Physical Medium Dependent (PMD) sublayer, Medium Dependent Interface (MDI), and Media

The 10 Gigabit Ethernet serial PMD sublayer is common to the serial PHY implementations specified in this clause. The PMD provides the interface from the PMA to the transmission medium. The MDI is the interface between the PMD and the fiber optic media.

A serial PMD sublayer shall comply with the requirements of Clause 52 of P802.3ae. The MDI for serial PMDs is not specified, but shall meet the interface performance requirements specified in Clause 52 of P802.3ae. The fiber optic media used with serial PMDs shall comply with the requirements of Clause 52 of P802.3ae.

3.2.6 Wide Wavelength Division Multiplexed (WWDM) Physical Medium Dependent (PMD) sublayer

The 10 Gigabit Ethernet WWDM PMD sublayer is common to the WWDM PHY implementations specified in this clause. The PMD provides the interface from the PMA to the transmission medium. The MDI is the interface between the PMD and the fiber optic media.

A WWDM PMD sublayer shall comply with the requirements of Clause 53 of P802.3ae. The MDI is not specified, but shall meet the interface performance requirements specified in Clause 53 of P802.3ae. The fiber optic media shall comply with the requirements of Clause 53 of P802.3ae.

4. Gigabit Ethernet Reconciliation Sublayer (RS) and Gigabit Media Independent Interface (GMII)

4.1 Overview

4.1.1 Scope

TBD.

5. Gigabit Ethernet PHYs

5.1 Overview

5.1.1 Scope

Reference GE PHYs specified in 802.3 with no changes. TBD which PHYs are to be supported.