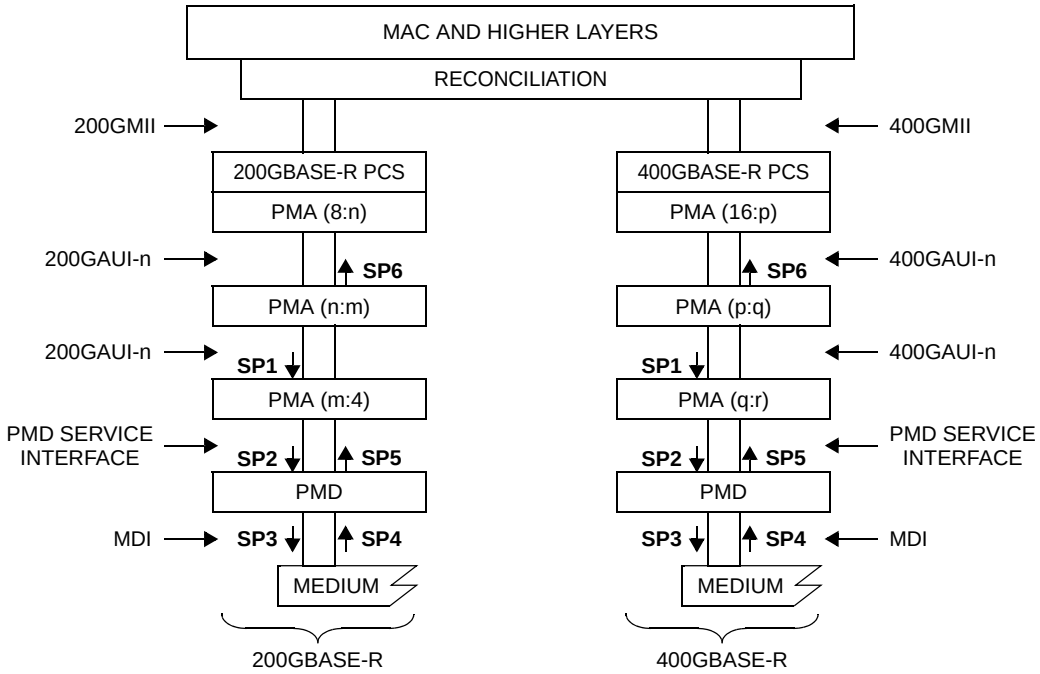


200GAUI-n = 200 Gb/s ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI-n = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM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

Figure 116–5—200GBASE-R and 400GBASE-R Skew points for multiple 200GAUI-n or 400GAUI-n

The allowable limits for Skew are shown in Table 116–7 and the allowable limits for Skew Variation are shown in Table 116–8.

The Skew requirements for the PCS, PMA and PMD sublayers are specified in the respective clauses as noted in Table 116–7 and Table 116–8.

116.6 FEC Degrade

FEC degrade is an optional feature allowing for the detection of a non-service affecting link degradation condition based on exceeding a threshold for FEC corrected errors. If there are multiple FEC decoders in a given direction of transmission between the MAC sublayers at each end of the link, a Local Degrade condition is cascaded in that direction of transmission to convey the fact that one or more FEC decoders in the path have exceeded their threshold of FEC corrected errors. If any FEC decoder in a given direction of transmission exceeds its provisioned threshold for FEC corrected errors, a Remote Degrade condition is indicated in the opposite direction of transmission from the PCS or XS closest to the MAC.

Figure 116–6 illustrates the signaling of the Remote Degraded condition in the case that there is no Clause 118 extender sublayer present between the MAC and the PCS. Note that the PCS will not initiate the

Table 116–7—Summary of Skew constraints

Skew points	Maximum Skew (ns)[a]	Maximum Skew for 200GBASE-R or 400GBASE-R PCS lane (UI)[b]	Notes[c]
SP1	29	≈ 770	See 120.5.3.1
SP2	43	≈ 1142	See 120.5.3.3, 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP3	54	≈ 1434	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP4	134	≈ 3559	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP5	145	≈ 3852	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP6	160	≈ 4250	See 120.5.3.5
At PCS receive	180	≈ 4781	See 119.2.5.1

[a]The Skew limit includes 1 ns allowance for PCB traces that are associated with the Skew points.
[b]The symbol ≈ indicates approximate equivalent of maximum Skew in UI based on 1 UI equals 37.64706 ps at PCS lane signaling rate of 26.5625 GBd.
[c]Should there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sublayer clause prevails.

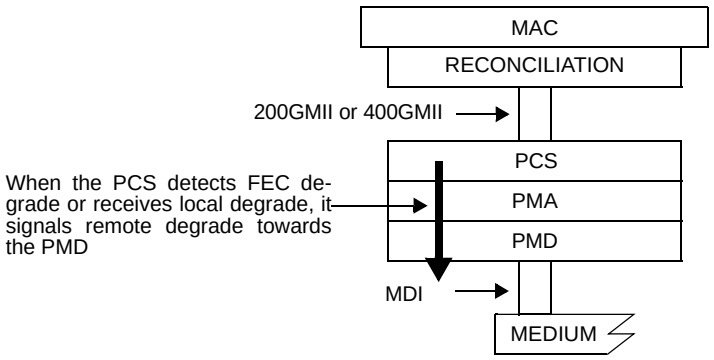
Table 116–8—Summary of Skew Variation constraints

Skew points	Maximum Skew Variation (ns)	Maximum Skew Variation for 26.5625 GBd PMD lane (UI)[a]	Maximum Skew Variation for 53.125 GBd PMD lane (UI)[b]	Notes[c]
SP1	0.2	≈ 5	N/A	See 120.5.3.1
SP2	0.4	≈11	N/A	See 120.5.3.3, 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP3	0.6	≈ 16	≈ 32	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP4	3.4	≈ 90	≈ 181	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP5	3.6	≈ 96	N/A	See 121.3.2, 122.3.2, 123.3.2, or 124.3.2
SP6	3.8	≈ 101	N/A	See 120.5.3.5
At PCS receive	4	≈ 106	N/A	See 119.2.5.1

[a]The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI based on 1 UI equals 37.64706 ps at PMD lane signaling rate of 26.5625 GBd.
[b]The symbol ≈ indicates approximate equivalent of maximum Skew Variation in UI based on 1 UI equals 18.82353 ps at PMD lane signaling rate of 53.125 GBd.
[c]Should there be a discrepancy between this table and the Skew requirements of the relevant sublayer clause, the sublayer clause prevails.

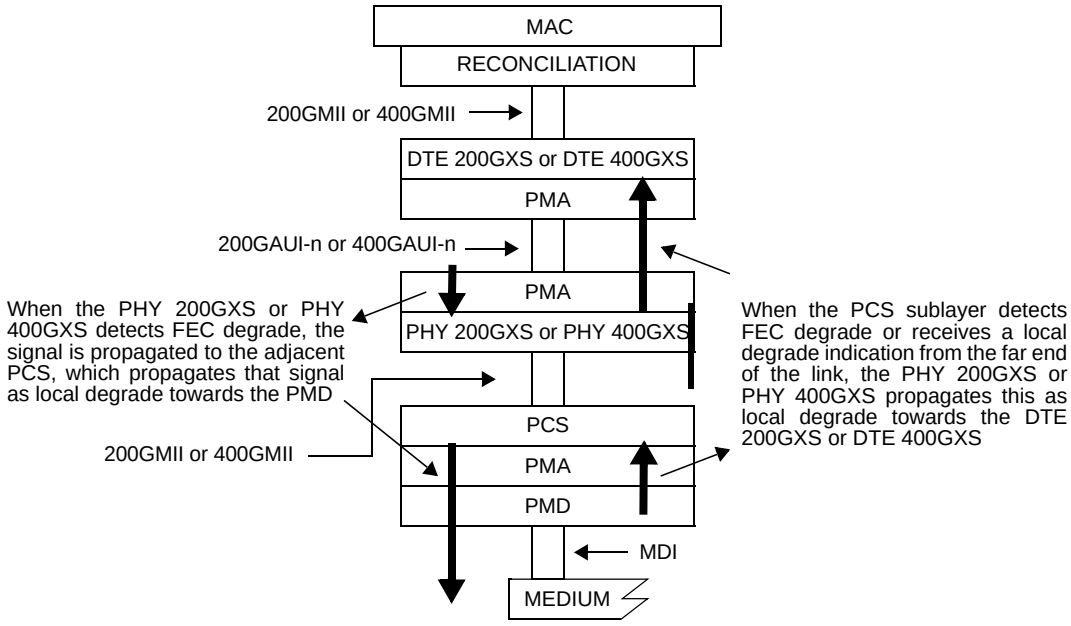
signaling for local degrade in this configuration as there are no additional FEC decoders in the receive direction between the PCS and the MAC.

Figure 116–7 illustrates the signaling of the Local Degraded condition in the case that a Clause 118 extender sublayer is present between the MAC and the PCS.



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

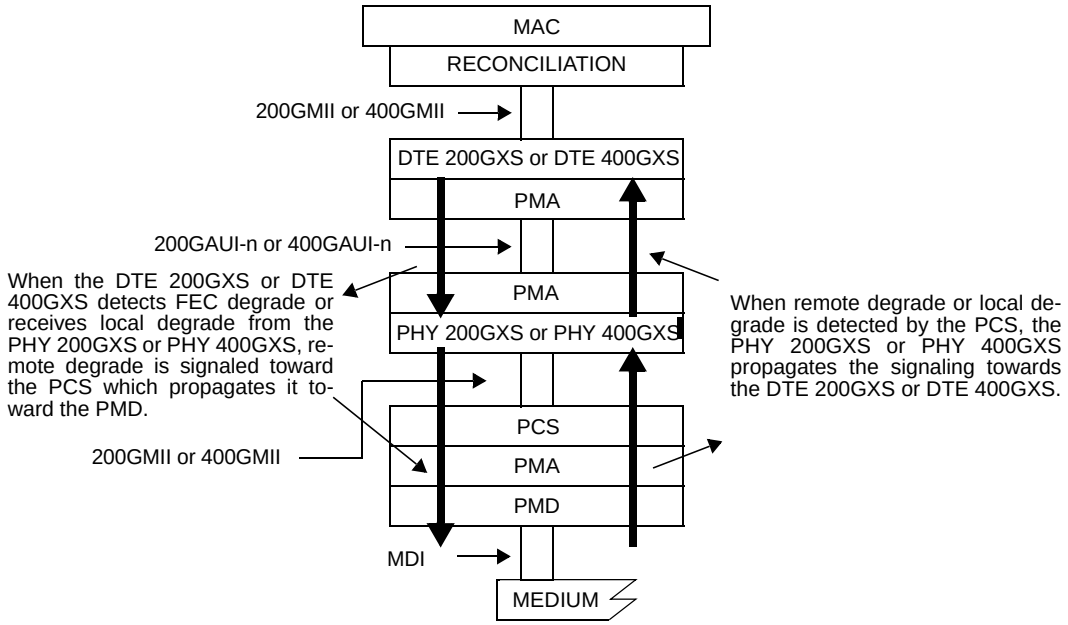
Figure 116–6—Remote Degrade signaling without Extender Sublayer



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

Figure 116–7—Local Degrade Signaling with Extender Sublayer

Figure 116–8 illustrates the signaling of the Remote Degraded condition in the case that a Clause 118 extender sublayer is present between the MAC and the PCS.



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

Figure 116–8—Remote Degrade signaling with Extender Sublayer

118.1.1 Summary of major concepts

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

a) Simple signal mapping to the 200GMII/400GMII
b) The optional 200GMII/400GMII Extender can be inserted between the Reconciliation Sublayer and the PHY to transparently extend the reach of the 200GMII/400GMII
c) Independent transmit and receive data paths
d) The 200GXS/400GXS leverages all functions in the Clause 119 PCS and supports physical instantiations of the 200GAUI-n/400GAUI-n
e) Optionally extends LPI signaling to the PHY for EEE

118.1.2 200GXS/400GXS Sublayer

The 200GXS is identical in function to the 200GBASE-R PCS in Clause 119 with the addition of the functions defined in 118.2. A single device may be configured as either a 200GXS or the 200GBASE-R PCS and may be managed through different optional management registers.

The 400GXS is identical in function to the 400GBASE-R PCS in Clause 119 with the addition of the functions defined in 118.2. A single device may be configured as either a 400GXS or the 400GBASE-R PCS and may be managed through different optional management registers.

118.1.3 200GAUI-n/400GAUI-n

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

— 200GAUI-8 chip-to-chip (Annex 120B)
— 200GAUI-8 chip-to-module (Annex 120C)
— 200GAUI-4 chip-to-chip (Annex 120D)
— 200GAUI-4 chip-to-module (Annex 120E)

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

— 400GAUI-16 chip-to-chip (Annex 120B)
— 400GAUI-16 chip-to-module (Annex 120C)
— 400GAUI-8 chip-to-chip (Annex 120D)
— 400GAUI-8 chip-to-module (Annex 120E)

118.2 FEC Degrade

The propagation of FEC degrade signaling across PCS and XS sublayers is described in 116.6 and is based on the optional FEC degrade signaling described in Clause 119 with the changes described for the DTE XS in 118.2.1 and for the PHY XS in 118.2.2.

118.2.1 DTE XS FEC Degrade signaling

The variable tx_am_sf is set as follows:

tx_am_sf<2:0> = {FEC_degraded_SER + rx_local_degraded,0,0}

118.2.2 PHY XS FEC Degrade signaling

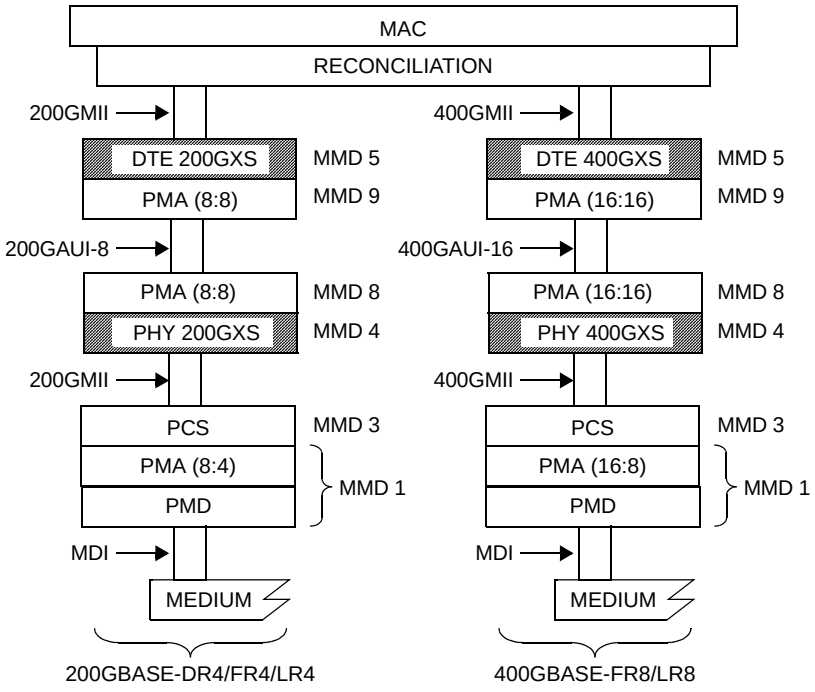
The variable tx_am_sf is set as follows:

tx_am_sf<2:0> = {PCS:rx_rm_degraded, PCS:FEC_degraded_SER + PCS:rx_local_degraded, 0}

Where PCS:rx_rm_degraded, PCS:FEC_degraded_SER, and PCS:rx_local_degraded are the rx_rm_degraded, FEC_degraded_SER, and rx_local_degraded variables from the adjacent PCS.

118.3 200GXS and 400GXS partitioning example

A partitioning example and MMD numbering using the 200GXS and 400GXS is shown in Figure 118–2.



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

Figure 118–2—Example 200GBASE-DR4/FR4/LR4 and 400GBASE-FR8/LR8 PMA layering with 200GXS and 400GXS

118.4 200GXS and 400GXS MDIO function mapping

The optional MDIO capability described in Clause 45 defines several registers that provide control and status information for and about the 200GXS or 400GXS. If MDIO is implemented, it shall map MDIO PHY XS and DTE XS control bits to Clause 119 control variables as shown in Table 118–1 and Table 118–3, respectively. Similarly, if MDIO is implemented, it shall map MDIO PHY XS and DTE XS status bits to Clause 119 status variables as shown in Table 118–2 and Table 118–4, respectively.

119.2.4.3 Scrambler

The transcoded 257-bit block, tx_xcoded<256:0>, shall be scrambled with a self-synchronizing scrambler to generate tx_scrambled<256:0>. The scrambler polynomial is identical to that in Clause 49, see Equation (49-1) for the definition of the polynomial.

119.2.4.4 Alignment marker mapping and insertion

In order to support deskew and reordering of the individual PCS lanes at the receive PCS, alignment markers corresponding to PCS lanes are periodically inserted after being processed by the alignment marker mapping function. The alignment marker mapping function compensates for the operation of the symbol distribution function and rearranges the alignment marker bits so that they appear on the PCS lanes intact and in the desired sequence. This preserves the properties of the alignment markers (e.g. DC balance, transition density) and provides a deterministic pattern for the purpose of synchronization.

For the 200GBASE-R PCS, an alignment marker group is composed of the alignment markers for all 8 PCS lanes plus an additional 65-bit pad and a 3-bit status field to yield the equivalent of four 257-bit blocks. For the 400GBASE-R PCS, an alignment marker group is composed of the alignment markers for all 16 PCS lanes plus an additional 133-bit pad and a 3-bit status field to yield the equivalent of eight 257-bit blocks. The alignment marker group is aligned to the beginning of two FEC messages, and interrupts any data transfer that is already in progress. The pad bits at the end of the alignment marker group shall be set to a free running PRBS9 pattern, defined by the polynomial $x^9 + x^5 + 1$. The initial value of the PRBS9 pattern generator may be any pattern other than all zeros. The fixed pad within the alignment markers and the PRBS9 pad at the end of the alignment maker group are ignored on receive.

Room for the alignment marker group is created by the transmit PCS (see 119.2.4.1). Special properties of the alignment marker group are that it is not scrambled, does not conform to the encoding rules as outlined in Figure 82-5 and is not transcoded. This is possible because the alignment marker group is added after encoding, transcoding, and scrambling, and removed before descrambling, transcoding, and 64B/66B decoding. The alignment marker group is not scrambled, which allows the receiver to directly search for and find the individual alignment markers, deskew the PCS lanes, and reassemble the aggregate stream before descrambling is performed. The alignment markers are formed from a known pattern that is defined to be balanced and with many transitions and therefore scrambling is not necessary.

The format of each PCS lane's alignment marker is shown in Figure 119–4. There is a portion that is common across all alignment markers (designated as CM_0 to CM_5), a unique portion per PCS lane (designated as UM_0 to UM_5), and finally a unique pad per PCS lane (designated as UP_0 to UP_2). Common synchronization logic independent of the received PCS lane number can be used with the common portion of the alignment marker.

The content of the alignment markers shall be as shown in Table 119–1 for the 200GBASE-R PCS and as shown in Table 119–2 for the 400GBASE-R PCS. The contents depend on the PCS lane number and the octet number, with CM_0 through CM_5 being identical across all alignment markers to allow for common synchronization across lanes. The format shown in Table 119–1 defines how the alignment markers appear on a given PCS lane. In the FEC codewords, they appear in a permuted format due to the codeword interleaving that occurs before FEC codewords are distributed to PCS lanes.

The transmit alignment marker status field allows the local PCS to communicate the status of the optional FEC degraded feature to the remote PCS. If there is no extender sublayer between the PCS and the MAC, it is set as follows:

tx_am_sf<2:0> = {FEC_degraded_SER + rx_local_degraded,0,0}

If there is a Clause 118 extender sublayer between the PCS and the MAC, it is set as follows:

tx_am_sf<2:0> = {PHY_XS:rx_rm_degraded, PHY_XS:FEC_degraded_SER, 0}

Where PHY_XS:rx_rm_degraded and PHY_XS:FEC_degraded_SER are the rx_rm_degraded and FEC_degraded_SER variables from the adjacent PHY_XS sublayer.

See 119.2.5.3 for more information on the optional FEC degrade feature.

119.2.4.4.1 AM creation for the 200GBASE-R PCS

For the 200GBASE-R PCS, the alignment marker mapping function creates a set of 8 alignment markers, and in combination with an additional 65-bit PRBS9 pad and a 3-bit status field, the PCS generates an alignment marker group. Let am_*x*<119:0> be the alignment marker for PCS lane *x*, *x*=0 to 7, where bit 0 is the first bit transmitted. The alignment markers shall be mapped to am_mapped<959:0> in a manner that yields the same result as the following process.

For *x*=0 to 7, am_*x*<119:0> is constructed as follows:

am_*x*<119:0> is set to CM_0, CM_1, CM_2, UP_0, CM_3, CM_4, CM_5, UP_1, UM_0, UM_1, UM_2, UP_2, UM_3, UM_4 and UM_5, as shown in Figure 119–4 (bits 119:0) using the values in Table 119–1 for PCS lane number *x*.

As an example, the variable am_0 is sent as (left most bit sent first, showing the first 32 bits transmitted of am_0):
01011001 01010010 01100100 10100000

The variable am_mapped is then derived from 10-bit interleaving the group of 8 alignment markers am_*x* per the following procedure:

For all *k*=0 to 11
  For all *j*=0 to 3
    if even(*k*)
      am_mapped<80*k*+20*j*+9:80*k*+20*j*> = am_{2*j*}<10*k*+9:10*k*>
      am_mapped<80*k*+20*j*+19:80*k*+20*j*+10> = am_{2*j*+1}<10*k*+9:10*k*>
    else
      am_mapped<80*k*+20*j*+9:80*k*+20*j*> = am_{2*j*+1}<10*k*+9:10*k*>
      am_mapped<80*k*+20*j*+19:80*k*+20*j*+10> = am_{2*j*}<10*k*+9:10*k*>

The additional 65-bit pad is appended to variable am_mapped as follows:
  am_mapped<1024:960> = PRBS9<64:0>
In this expression, PRBS9<0> is the first PRBS9 bit output of the 65-bit pad.

The 3-bit transmit alignment marker status field is then appended to the variable am_mapped as follows:
am_mapped<1027:1025> = tx_am_sf<2:0>

Alignment marker mapping is shown in Figure 119–5.

The alignment marker group am_mapped<1027:0> shall be inserted so it appears in the output stream every 81 920 × 257-bit blocks. The variable tx_scrambled_am<10279:0> is constructed in one of two ways. Let the set of vectors tx_scrambled_*i*<256:0> represent consecutive values of tx_scrambled<256:0>.

For a 10280-bit block with an alignment marker group inserted:
tx_scrambled_am<1027:0> = am_mapped<1027:0>
For all *i*=0 to 35
  tx_scrambled_am<257*i*+1284:257*i*+1028> = tx_scrambled_*i*<256:0>

For a 10280-bit block without an alignment marker group:
For all *i*=0 to 39

rx_local_degraded

Boolean variable that is asserted true when the receiver detects rx_am_sf<1> asserted true for two consecutive alignent marker periods. It is deasserted when rx_am_sf<1> is deasserted for two consecutive alignment marker periods. If a Clause 45 MDIO is implemented, the status of this variable is reflected in bit 3.801.6.

rx_raw<71:0>

Vector containing one 200GMII/400GMII 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.

rx_rm_degraded

Boolean variable that is asserted true when the receiver detects rx_am_sf<2> asserted true for two consecutive alignment marker periods. It is deasserted when rx_am_sf<2> is deasserted for two consecutive alignment marker periods. If a Clause 45 MDIO is implemented, the status of this variable is reflected in bit 3.801.5.

signal_ok

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

slip_done

Boolean variable that is set to true when the SLIP requested by the alignment marker lock state diagram has been completed indicating that the next candidate 120-bit block position can be tested.

test_amp

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

test_cw

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

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 200GMII/400GMII 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.

119.2.6.2.3 Functions

AMP_COMPARE

This function compares the values of first_pcsl and current_pcsl to determine if a valid alignment marker payload sequence has been detected and returns the result of the comparison using the variable amp_match. If current_pcsl and first_pcsl are 0, amp_match is set to true.

DECODE(rx_coded<65:0>)

Decodes the 66-bit vector returning rx_raw<71:0>, which is sent to the 200GMII/400GMII. The DECODE function shall decode the block as specified in 119.2.3.

ENCODE(tx_raw<71:0>)

Encodes the 72-bit vector returning tx_coded<65:0>. The ENCODE function shall encode the block as specified in 119.2.3.

R_TYPE(rx_coded<65:0>)

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

Values: C; The vector contains a sync header of 10 and one of the following:

a) A block type field of 0x1E and eight valid control characters other than /E/ or /LI/;

b) A block type field of 0x4B.

Table 119–5—MDIO/PCS status variable mapping

MDIO status variable	PCS register name	Register/ bit number	PCS status variable
BASE-R and MultiGBASE-T receive link status	BASE-R and MultiGBASE-T PCS status 1 register	3.32.12	PCS_status
Lane *x* aligned	Multi-lane BASE-R PCS alignment status 3 and 4 registers	3.52.7:0 3.53.7:0	am_lock<*x*>
PCS lane alignment status	Multi-lane BASE-R PCS alignment status 1 register	3.50.12	align_status
Lane *x* mapping	Lane *x* mapping register	3.400 through 3.415	pcs_lane_mapping<*x*>
PCS FEC bypass indication ability	PCS FEC status register	3.801.1	FEC_bypass_indication_ability
PCS FEC corrected codewords	PCS FEC corrected codewords counter register	3.802, 3.803	FEC_corrected_cw_counter
PCS FEC uncorrected codewords	PCS FEC uncorrected codewords counter register	3.804, 3.805	FEC_uncorrected_cw_counter
PCS FEC symbol errors, PCS lanes 0 to *x*	PCS FEC symbol error counter register, lanes 0 to *x*	3.600 to 3.631	FEC_symbol_error_counter_*i*
Tx LPI indication	PCS status 1	3.1.9	Tx LPI indication
Tx LPI received	PCS status 1	3.1.11	Tx LPI received
Rx LPI indication	PCS status 1	3.1.8	Rx LPI indication
Rx LPI received	PCS status 1	3.1.10	Rx LPI received
EEE wake error counter	EEE wake error counter	3.22	Wake_error_counter
PCS FEC degraded SER ability	PCS FEC status register	3.801.3	FEC_degraded_SER_ability
PCS FEC degraded SER	PCS FEC status register	3.801.4	FEC_degraded_SER
Local degraded SER received	PCS FEC status register	3.801.6	rx_local_degraded
Remote degraded SER received	PCS FEC status register	3.801.5	rx_rm_degraded
PCS FEC high SER	PCS FEC status register	3.801.2	hi_ser

119.4 Loopback

When the PCS is in loopback, the PCS shall accept data on the transmit path from the 200GMII/400GMII and return it on the receive path to the 200GMII/400GMII. In addition, the PCS shall transmit what it receives from the 200GMII/400GMII to the PMA sublayer, and shall ignore all data presented to it by the PMA sublayer. If a Clause 45 MDIO is implemented, then the PCS is placed in the loopback when the loopback bit from the PCS control 1 register (bit 3.0.14) is set to a one.