SM-PMA delay constraints

(comment # 451)

Kapil Shrikhande (Marvell), Mike Dudek (Marvell), Matt Brown (Alphawave),
Gary Nicholl (Cisco), Eugene Opsasnick (Broadcom)

IEEE 802.3dj January 2025 interim meeting

Supporters

- Jeff Slavick, Broadcom
- Andras DeKoos, Microchip

PMA delay constraints for 200G/lane (802.3dj D1.3)

***From Table 176-7 : Delay constraints** in Draft 1.3*

PMA	Type	Delay (ns)
200GE 8:1 and 1:8	m:n and n:m PMAs	Tbd
400GE 16:2 and 2:16		Tbd
800GE 32:4 and 4:32		46.08
1.6TE 16:8 and 8:16		Tbd
200GE 1:1	n:n PMAs	Tbd
400GE 2:2		Tbd
800GE 4:4		92.16
1.6TE 8:8		Tbd
1.6TE 16:16		No row in the table

176.8 Delay Constraints

There may be up to four instances of an SM-PMA within a Physical Layer, which is composed of an xBASE-R PHY and, optionally, an xMII Extender. The maximum delay (sum of transmit and receive delays) contributed by each instance of an xBASE-R PMA shall meet the values specified in Table 176–7. A description of overall system delay constraints and the definitions for bit times and pause_quanta can be found in 169.4.

In 802.3dj D1.3 :

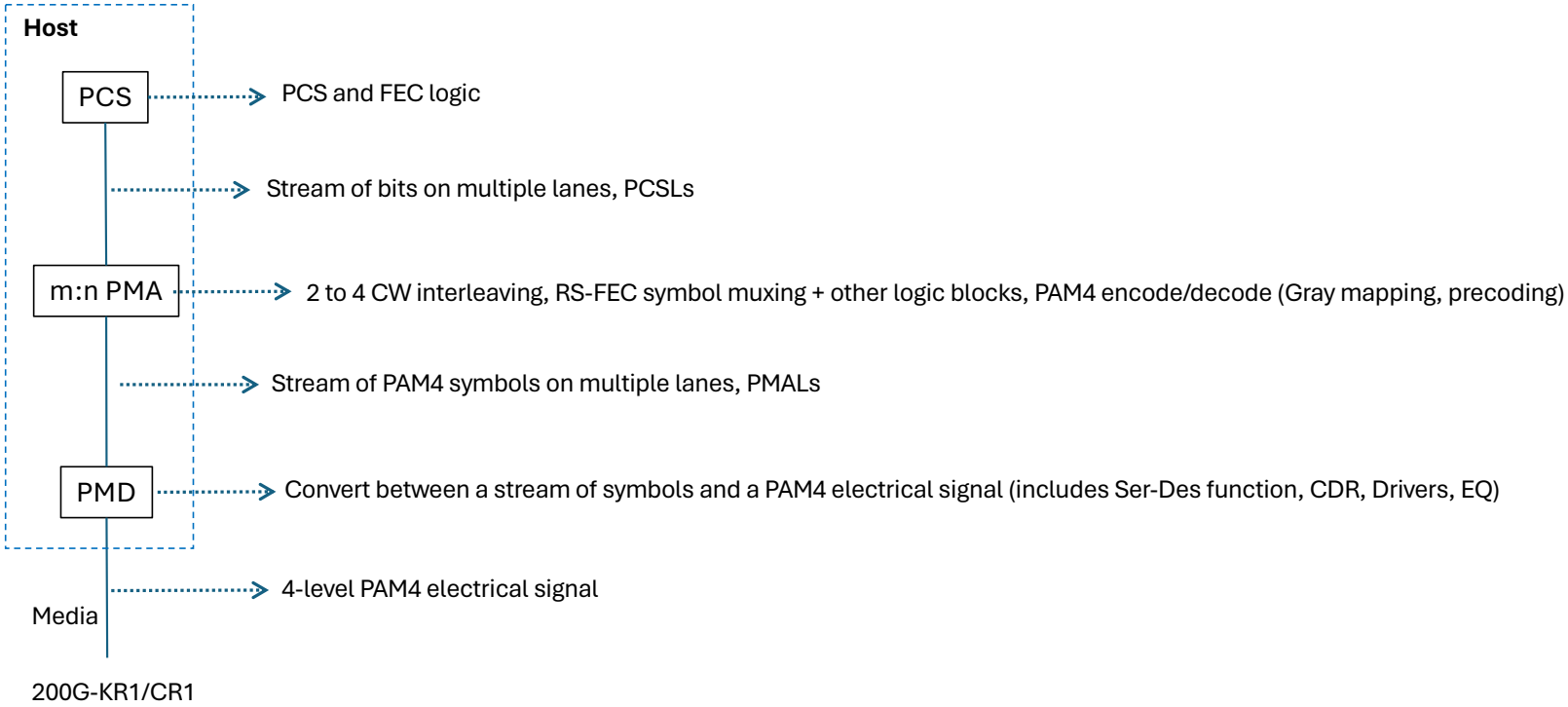
- Table contains numbers for 800GE ==> copied from 802.3df
  - 32:4 PMA delay allocation is same as 32:8 PMA from 802.3df
  - 4:4 PMA delay allocation is same as 8:8 PMA from 802.3df
  - Delay of the 4:4 PMA is 2x of 32:4 PMA
- Other PMA delays - TBD
- 1.6TbE 16:16 PMA not present in Table 176-7

Functional partitioning and accounting of PHY delay

- The 802.3dj architecture partitions the PMA, AUI, PMD functions differently from prior standards
- The functions belonging to each sublayer drive the delay constraints (for the sublayer)
- The functional partitioning is captured in a few examples in the following slides

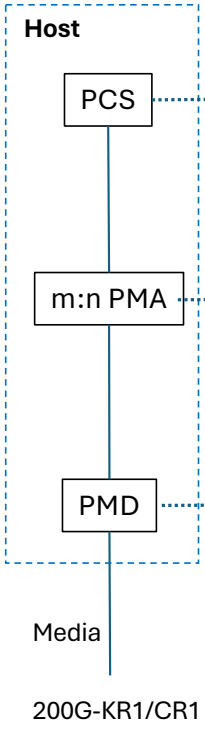
Functional partitioning and accounting of PHY delay

Example 1 of 3 : 200GE-KR1/CR1 without AUIs



Functional partitioning and accounting of PHY delay

Example 1 of 3 : 200GE-KR1/CR1 without AUIs



PCS and FEC logic.

Delay = **801.28 ns (Cl119)**

2 to 4 CW interleaving, RS-FEC symbol muxing + other logic blocks, PAM4 encode/decode (Gray mapping, precoding).

Delay = **TBD (Cl176)**

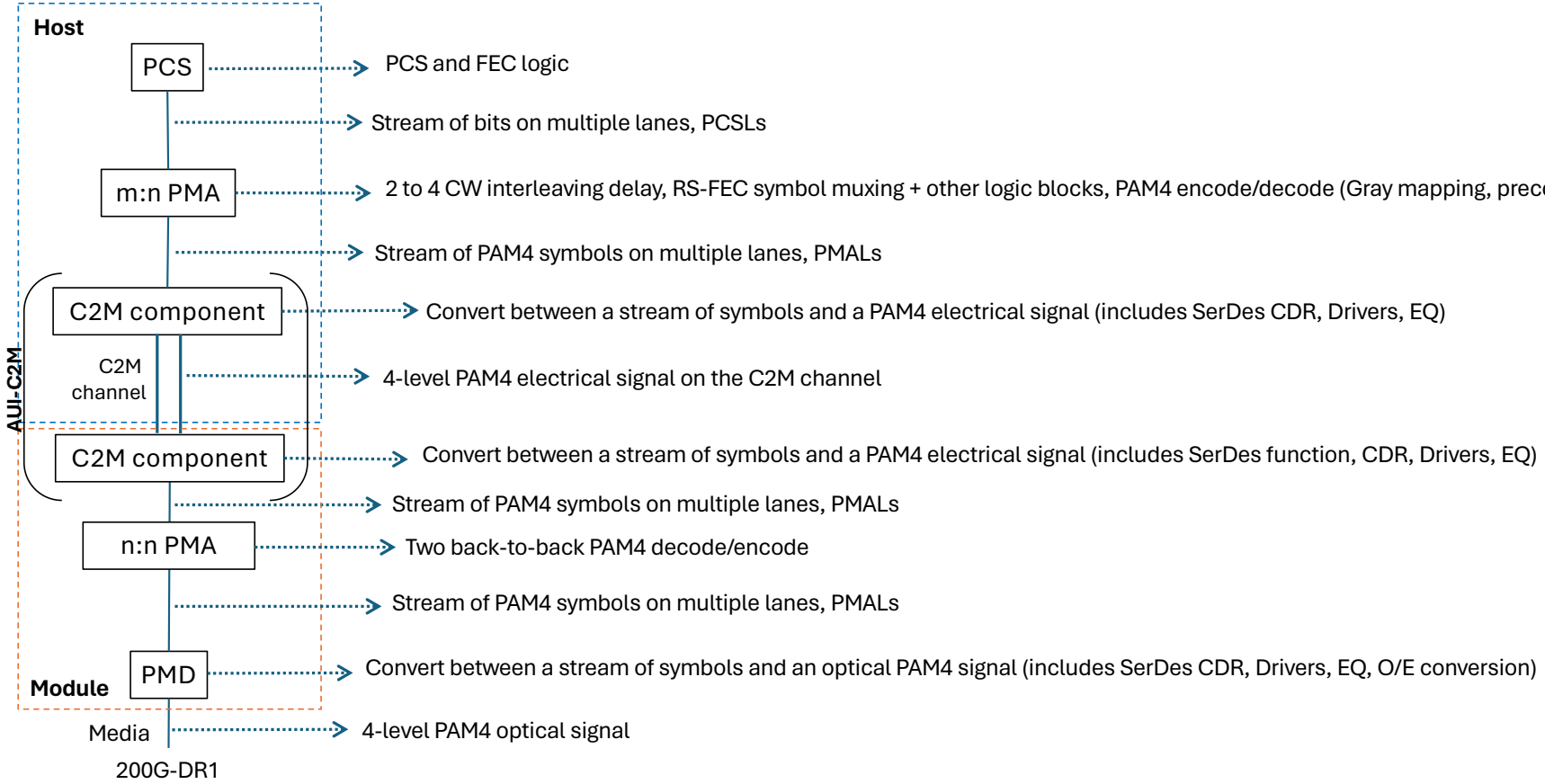
Convert between a stream of symbols and a PAM4 electrical signal (includes Ser-Des function, CDR, Drivers, EQ).

Delay = **61.44ns including 14ns for the media and 20ns budget for MLSE (Cl178/179)**

This might be a little low, we should consider increasing the allocation

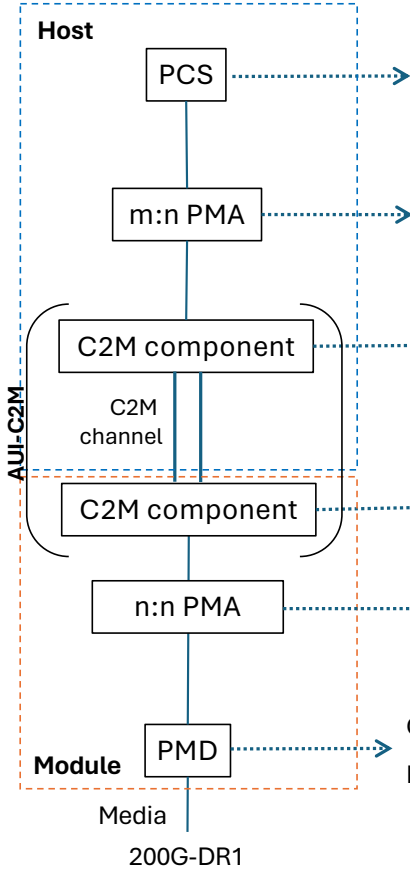
Functional partitioning and accounting of PHY delay

Example 2 of 3: 200GE-DR1 with C2M AUI interface



Functional partitioning and accounting of PHY delay

Example 2 of 3: 200GE-DR1 with C2M AUI interface



PCS and FEC logic

Delay = **801.28 ns (Cl119)**

2 to 4 CW interleaving delay, RS-FEC symbol muxing + other logic blocks, PAM4 encode/decode (Gray mapping, precoding)

Delay = **TBD (Cl176)**

Convert between a stream of symbols and a PAM4 analog signal (includes SerDes CDR, Drivers, EQ)

Delay = **None (no delay constraint section in C2M Annex 176D)**

Convert between a stream of symbols and a PAM4 analog signal (includes SerDes function, CDR, Drivers, EQ)

Delay = **None (no delay constraint section in C2M Annex 176D)**

Two back-to-back PAM4 decode/encode, Delay = **TBD (Cl176)**

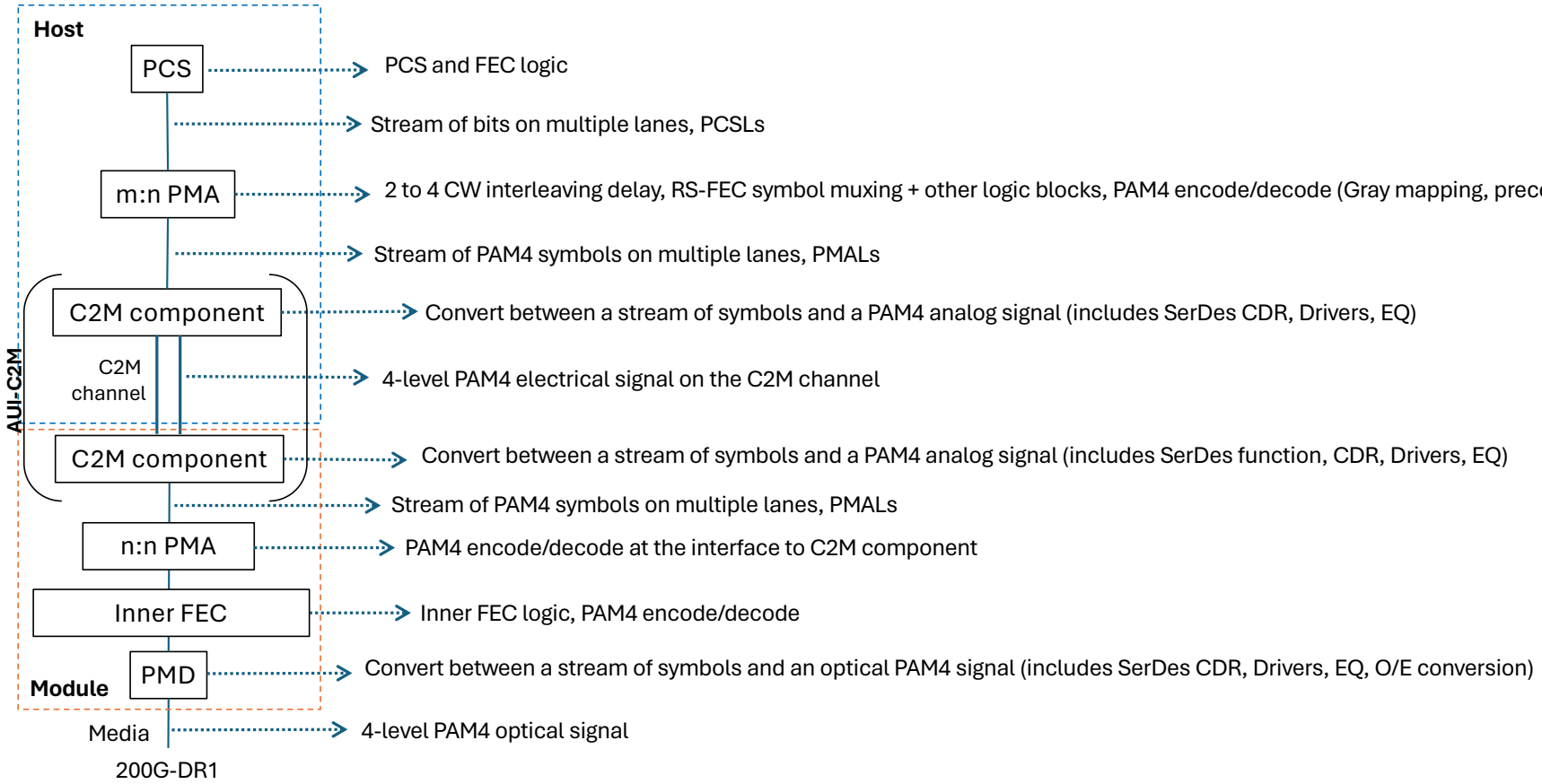
Convert between a stream of symbols and an optical PAM4 signal (includes SerDes CDR, Drivers, EQ, O/E conversion)

Delay = **20.48 ns (Cl180)**

This delay value does not include the “Serdes/DSP” functions, need to increase the allocation

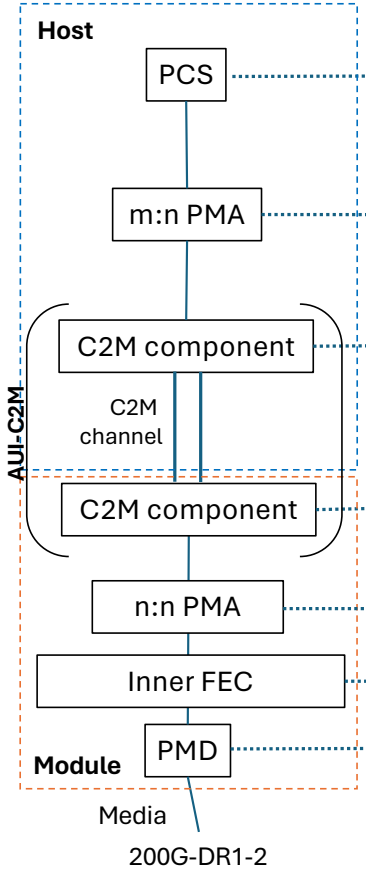
Functional partitioning and accounting of PHY delay

Example 3 of 3: 200GE-DR1-2 with Inner FEC and C2M AUI interface



Functional partitioning and accounting of PHY delay

Example 2 of 3: 200GE-DR1-2 with Inner FEC and C2M AUI interface



PCS and FEC logic

Delay = **801.28 ns (Cl119)**

2 to 4 CW interleaving delay, RS-FEC symbol muxing + other logic blocks, PAM4 encode/decode (Gray mapping, precoding)

Delay = **TBD (Cl176)**

Convert between a stream of symbols and a PAM4 analog signal (includes SerDes CDR, Drivers, EQ)

Delay = **None (no delay constraint section in C2M Annex 176D)**

Convert between a stream of symbols and a PAM4 analog signal (includes SerDes function, CDR, Drivers, EQ)

Delay = **None (no delay constraint section in C2M Annex 176D)**

PAM4 decode/encode, Delay = **TBD (Cl176)**

Inner FEC, PAM4 encode/decode, Delay = **296.96 ns**

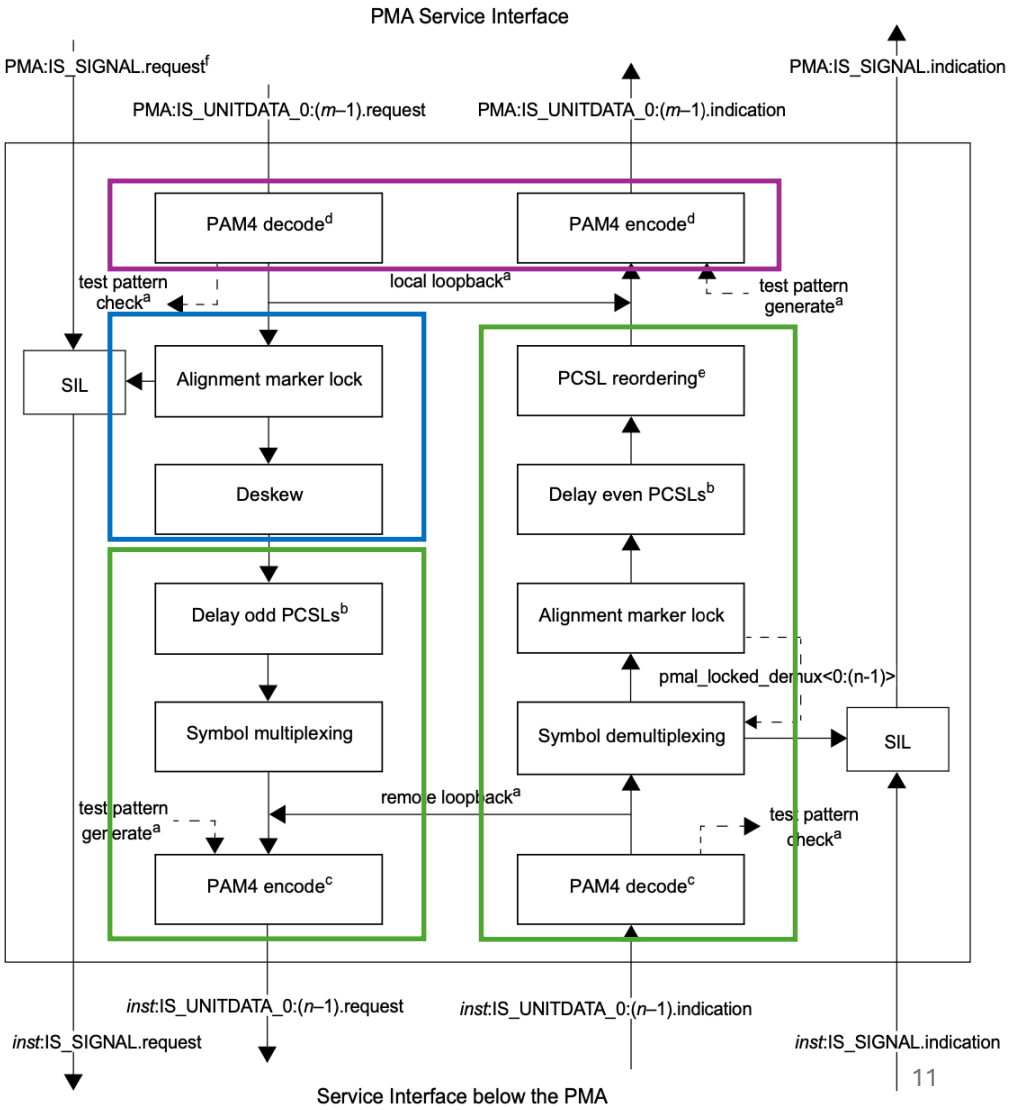
Convert between a stream of symbols and an optical PAM4 signal (includes SerDes CDR, Drivers, EQ, O/E conversion)

Delay = **20.48 ns (Cl182)**

This delay value does not include the “Serdes/DSP” functions, need to increase the allocation

m:n SM-PMAs

- m:n SM-PMA connecting a PCS/DTE-XS to an AUI or a PMD uses the green boxes in the figure on the right
  - PAM4 encode/decode (Serdes)
  - Delay odd/even PCS lanes
  - Symbol mux/demux
  - PCSL reordering (in demux direction)
- m:n SM-PMA interfacing to a BM-PMA uses the functions in the blue box in addition to green box
  - Alignment marker lock (in muxing direction)
  - Deskew (in muxing direction)
- The functions in the Purple box are only used when there is a 1.6TAUI-16 above the PMA
- The PAM4 encode/decode is Gray mapping and precoding
  - Other functions typically associated with “Serdes” are in the C2M/C2C component or in the PMD



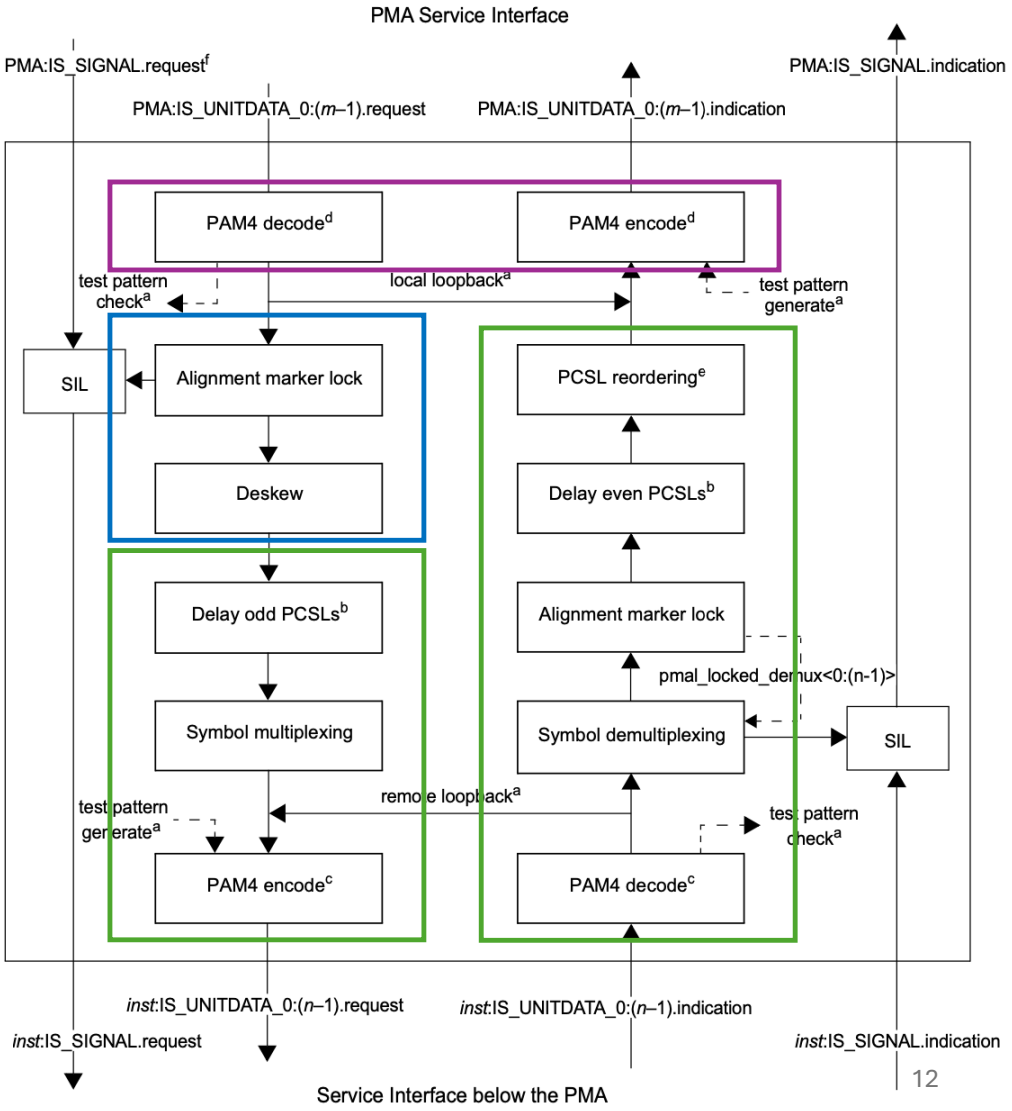
inst: PMA or PMD or FEC or AUI
SIL: Signal indication logic

[a] Optional
[b] If 200GBASE-R, 400GBASE-R or 800GBASE-R
[c] If the sublayer below the PMA is an AUI or PMD
[d] If the sublayer above the PMA is a 1.6TAUI-16
[e] If the sublayer above the PMA is an 800GBASE-R BM-PMA
[f] PMA:IS_SIGNAL.request is not present when the sublayer above the PMA is a PCS or DTE XS

m = 8 (200GBASE-R), 16 (400GBASE-R), 32 (800GBASE-R), or 16 (1.6TBASE-R)
n = 1 (200GBASE-R), 2 (400GBASE-R), 4 (800GBASE-R), or 8 (1.6TBASE-R)

m:n SM-PMAs: Proposed delay constraints

- Delay of alternating lanes (to enable 4 CW interleaving) results in a fixed 2-CW latency for 200/400GE
  - 51.2ns for 200GE
  - 25.6ns for 400GE
  - The 2-CW delay is not applied in the 800GE/1.6TE m:n PMAs
- For the other logic blocks, assume one clock cycle delay per block (plus few extra cycles)
  - A budget of ~ 15 ns (at 1Ghz) is sufficient
  - Round up to 15.36ns = 6 PQ @ 200GE
- Propose following delay constraints:
  - 200GE 8:1 and 1:8 PMA : 15.36 + 51.2 = 66.56 ns
  - 400GE 16:2 and 2:16 PMA: 15.36 + 25.6 = 40.96 ns
  - 800GE 32:4 and 4:32 PMA : 15.36 ns
  - 1.6TE 16:8 and 8:16 PMA : 15.36 ns



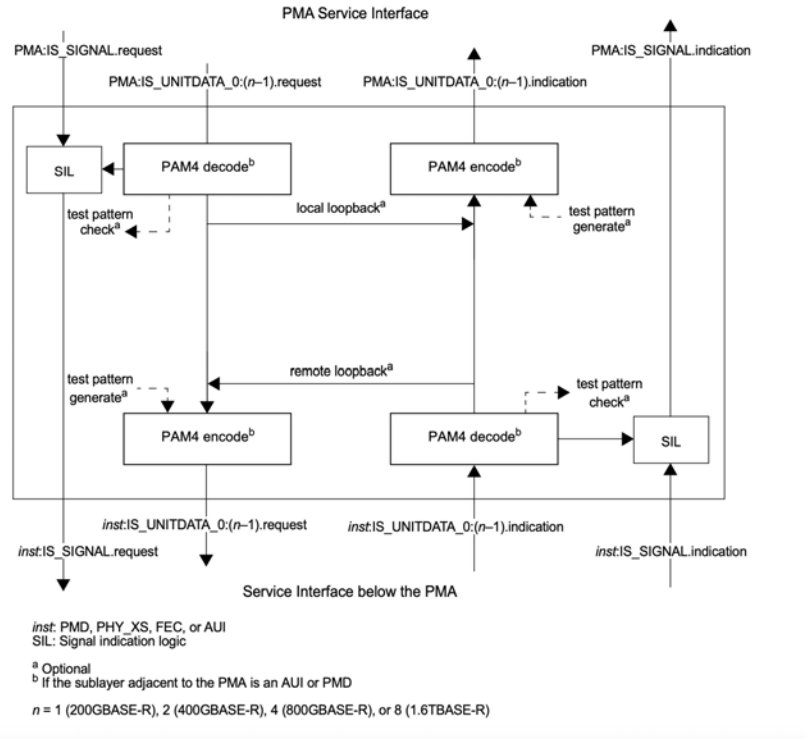
inst: PMA or PMD or FEC or AUI
SIL: Signal indication logic

[a] Optional
[b] If 200GBASE-R, 400GBASE-R or 800GBASE-R
[c] If the sublayer below the PMA is an AUI or PMD
[d] If the sublayer above the PMA is a 1.6TAUI-16
[e] If the sublayer above the PMA is an 800GBASE-R BM-PMA
[f] PMA:IS_SIGNAL.request is not present when the sublayer above the PMA is a PCS or DTE XS

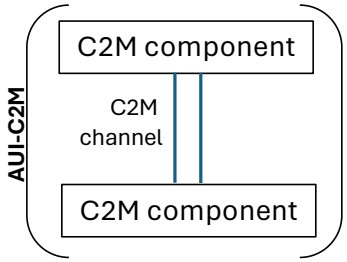
m = 8 (200GBASE-R), 16 (400GBASE-R), 32 (800GBASE-R), or 16 (1.6TBASE-R)
n = 1 (200GBASE-R), 2 (400GBASE-R), 4 (800GBASE-R), or 8 (1.6TBASE-R)

n:n SM-PMAs:
Proposed delay constraints

- n:n PMA has back-to-back PAM4 encode/decode functions
  - Other functions typically associated with “Serdes” are in the C2M/C2C component or in the PMD
- Assume one clock cycle for each of the PAM4 encode/decode logic block plus a few additional cycles (6 cycles) between blocks per direction (4 + 6 + 6 = 16 total)
  - Assume budget of ~16ns delay round-trip (1 Ghz clock)
  - Round down to 15.36 ns = 6 PQ @ 200GE
- Propose using 15.36 ns for n:n SM-PMAs



C2M component and channel delay



- C2M component delay needs to account for functions typically in a “Serdes” to drive across a C2M AUI channel
- To be consistent with the delay accounting for the PMD (which includes channel delay in one direction), add the C2M channel delay in one direction to the C2M component
- For the C2M channel – assume delay equivalent to 1m cable per direction ~ 5 ns
- For C2M component propose allocating ~ 45ns consisting of ~5ns of channel delay and ~ 40ns for the C2M component function
  - Round to 46.08 ns = 18 PQ @ 200GE
- Use 46.08 ns as the delay constraint for the C2M component and use the same value for the C2C component
  - Use same value for all Rates (200/400/800GE and 1.6TbE)

Optical PMD delay

- Current value in the D1.3 = 20.48 ns (includes 2m of fiber)
  - Same as in prior standards

- Does not account for functional partitioning in 802.3dj where the PMD now includes the “Serdes” or “DSP” functions to convert the PAM4 signal to a symbol stream

- The optics PMD delay should be increased

- Recommend ~ 74 ns for the optics PMD delay, composed of ~10ns for 2m fiber and ~64ns for the PMD function
  - Round to 74.24 ns = 29 PQ @ 200GE

- Use 74.24 ns for the Optics PMD delay (for DR4, DR4-2, FR4-500, FR4, LR4)

Proposed delay constraints values

Function	200GE delay (ns)	400GE delay (ns)	800GE delay (ns)	1.6TE delay (ns)
PCS, XS*	801.28	800	800	400
m:n SM-PMA (same for n:m)	66.56	40.96	15.36	15.36
C2M component (same for C2C)	46.08	46.08	46.08	46.08
n:n SM-PMA	15.36	15.36	15.36	15.36
Inner FEC*	296.96	176.64	116.48	86.4
CR/KR PMD	61.44 (74.24 ?)	61.44 (74.24 ?)	61.44 (74.24 ?)	61.44 (74.24 ?)
DR PMD	74.24	74.24	74.24	74.24
DR-2 PMD	74.24	74.24	74.24	74.24
FR4-500 PMD	74.24	74.24	74.24	74.24
FR4 PMD	74.24	74.24	74.24	74.24
LR4 PMD	74.24	74.24	74.24	74.24

1

Propose using the delay numbers in the table above for m:n SM-PMA, n:n SM-PMA in Clause 176
Propose using the same delay value for both C2M and C2C component to add as delay constraints in Annex 176C and 176D

2

Consider replace the 20.48ns optics PMD delay in the current draft with the 74.24 ns in the table
Consider increasing allocation for CR/KR PMD to be greater than ~62ns in the current draft. One suggestion is 74.24 ns (same as Optics PMD)

* Not proposing to make changes to the PCS, XS and Inner FEC delay numbers

Recommendation

- Propose using the delay numbers in the table from Slide 16 for m:n, n:m, and n:n SM-PMAs in Clause 176
- Propose using the C2M component delay numbers from Slide 16 for both C2M and C2C component of the AUIs to add as delay constraints in Annex 176C and 176D
- Propose using the Optics PMD delay from Slide 16 for the IM-DD PMDs
- Propose changing the CR/KR PMD delay to match the Optics PMD delay in Slide 16