P802.3dj Draft 0.1
Architectural considerations

Matt Brown, Alphawave, P802.3dj Editor-In-Chief
Gary Nicholl, Cisco, Logic Track Lead

IEEE P802.3dj Task Force

January 2024

Introduction

- In Draft 0.1, the editorial team made some discussion-worthy decisions in favor of a concise and elegant standard.
- Define PMD service interface to be analog signal vs symbols.
- Define Inner FEC with PMA above forming 200G lanes.
- Define separate PMA for bit-muxing and symbol muxing.
- Unique nomenclature for symbol and bit multiplexing PMAs.
- Unique nomenclature for IMDD Inner FEC and Coherent Inner FEC.

PMD Service Interface

PMD service interface (SI) primitives

- During 802.3df comment resolution…
  - specified the 800GBASE-R PMA and PMD sublayer delay limits as though the equalization, clock recovery, and data recovery, etc., were part of the PMA (not PMD)
  - for symmetry amongst 100 Gb/s per lane PHYs, left the 800GBASE-R PMD SI primitives defined as though the PMD is detecting the PAM4 symbols
  - this left a bit of dissonance that might be addressed in 802.3dj for new lane rates
- In 802.3dj Draft 0.1, redefined the PMD SI primitives appropriately.

PMD SI primitives

In particular, we are discussing the primitives PMD:IS_UNITDATA_i.request PMD_IS_UNITDATA_i.indication as shown in the figure to the right.

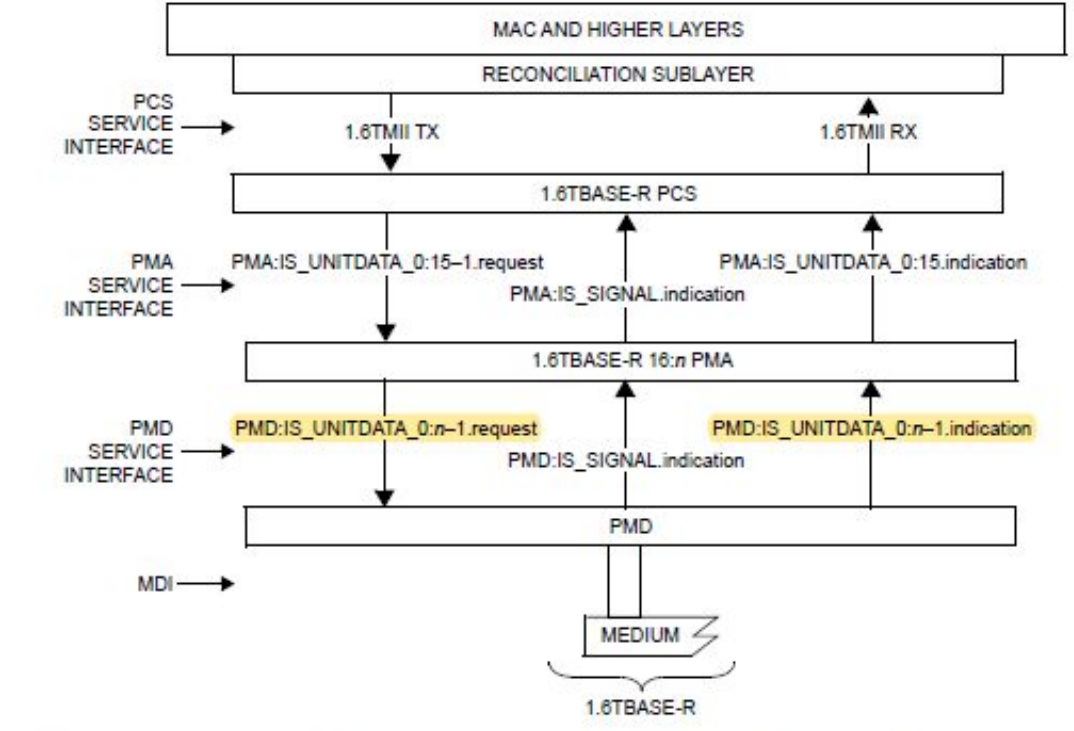


Figure 174–3—1.6TBASE-R inter-sublayer service interfaces not including 1.6TMII Extender

PMD SI primitives

In Draft 0.1, reusing all of the SI primitive definitions from Clause 116, except for the PMD SI.

The PMD SI parameters are redefined as a signals rather than symbols, leveraging work in P802.3cw.

The TX and RX parameters are tx_signal and rx_signal.

The excerpt to the right is for 1.6TBASE-R PHYs.

174.3.3 Semantics of inter-sublayer service interface primitives

The semantics of the inter-sublayer service interface primitives for the 1.6TBASE-R sublayers are described in 116.3.3.1 through 116.3.3.3, with the exception that the PMD service interface is described in 174.3.4.

174.3.4 Semantics of the PMD service interface primitives

174.3.4.1 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 the client sublayer, where n is the number of parallel streams of data units.

Semantics of the service primitive

IS_UNITDATA_0.request(tx_signal)
IS_UNITDATA_1.request(tx_signal)
...
IS_UNITDATA_n–1.request(tx_signal)

The data conveyed by IS_UNITDATA_0:n–1.request consists of n parallel continuous streams of encoded symbols, one stream for each lane. The tx_signal parameter conveys an analog signal to be processed by the PMD.

174.3.4.2 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 to the client sublayer, where n is the number of parallel streams of data units.

Semantics of the service primitive

IS_UNITDATA_0.indication(rx_signal)
IS_UNITDATA_1.indication(rx_signal)
...
IS_UNITDATA_n–1.indication(rx_signal)

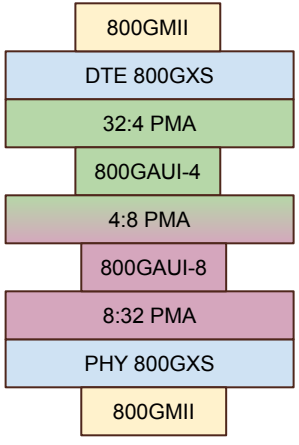
The data conveyed by IS_UNITDATA_0:n–1.indication consists of n parallel continuous streams of encoded symbols, one stream for each lane. The rx_signal parameter conveys an analog signal to be processed by the client sublayer.

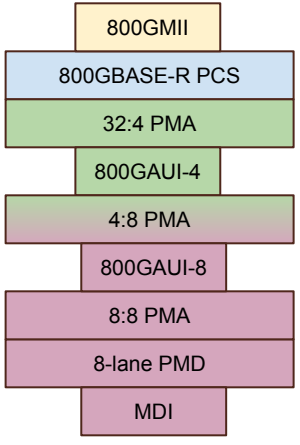
PMA symbol and bit muxing

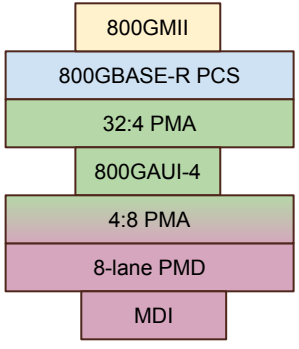
PMA, conversion between symbol-mux and bit-mux

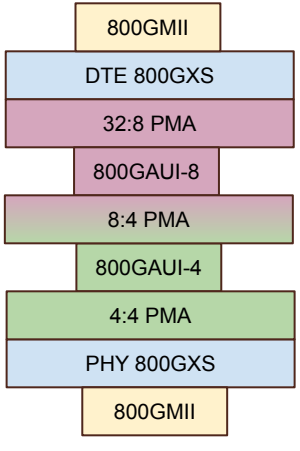
- The following is how it was portrayed in the adopted PMA baseline.
- For 200 Gb/s per lane AUI/PMD the PMA must support FEC-symbol-muxing.
- If a 100 Gb/s (or lower) AUI is used along with a 200 Gb/s AUI and/or PMD then a conversion between bit-muxing (BM) and symbol-muxing (SM) is required.
- For example, for 800GE, in addition to the 32:4, 4:4, 4:32 SM PMAs, 8:4 and 4:8 PMAs which convert between SM and BM are required.
- Some examples follow.

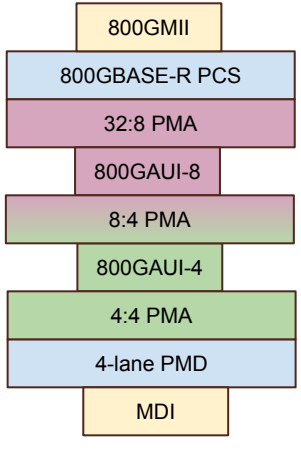
800G, with mixed interleaving per PMA (option #1)

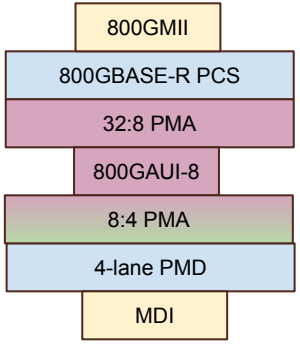








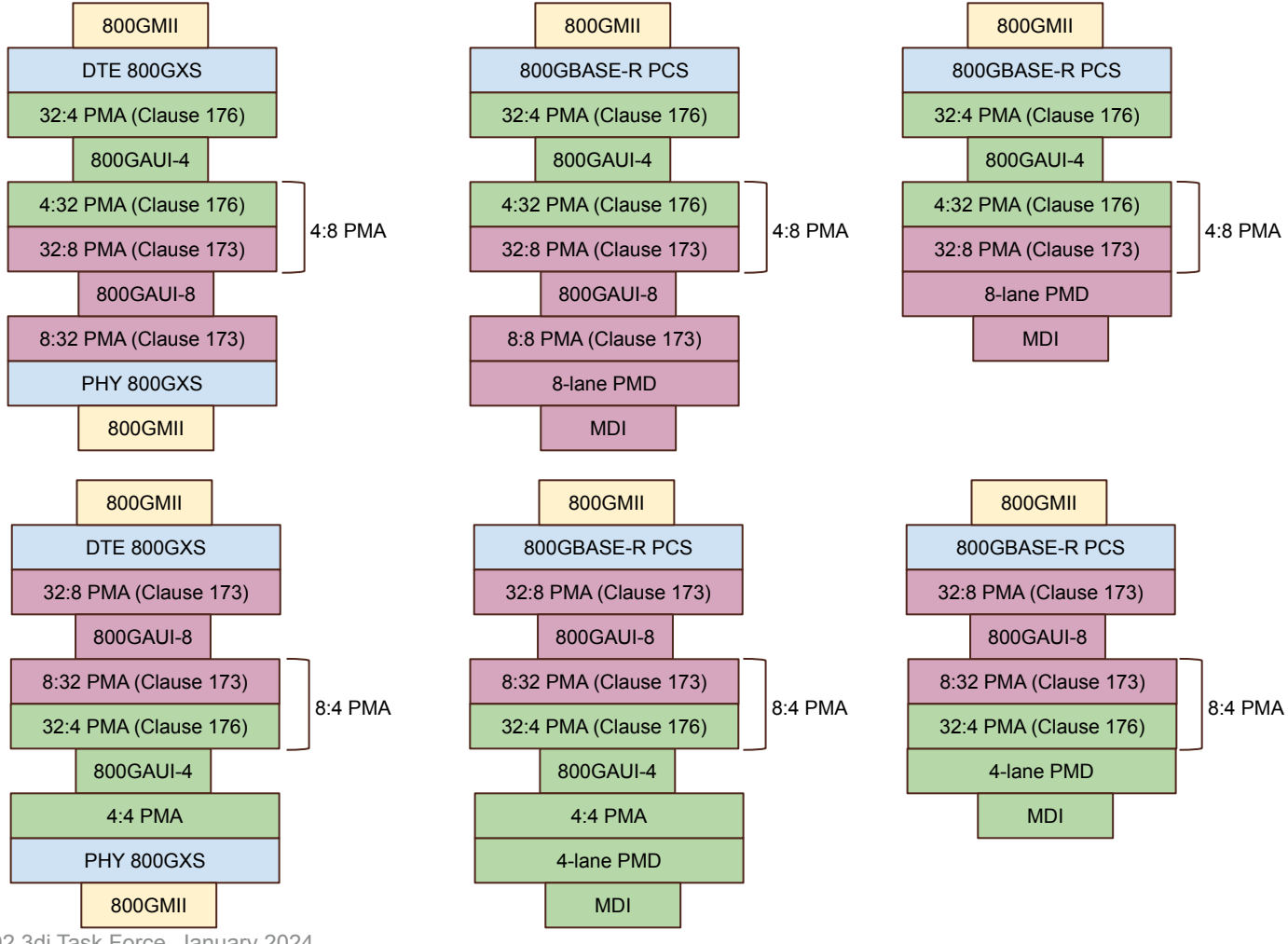




PMAs, mux conversion alternate

- Conversion between SM and BM requires that the PMA first demux to PCS lanes then mux back to physical lanes.
- So rather than defining a single PMA to convert between and have to define both SM and BM in the same clause we can use a pair of PMAs, one with BM and the other with SM.
- For examples, for 800GE, conversion from 8 lanes AUI (BM) to 4 lanes PMD (SM) we can use an 8:32 BM PMA (Clause 173) plus a 32:4 SM PMA (Clause 176).
- This was the path taken in Draft 0.1.
- Some examples follow.

800G, without mixed interleaving per PMA (option #2)



Inner FEC and PMAs

PMA above IMDD Inner FEC

- The baselines for the IMDD Inner FEC (see Clause 177) were written with the assumption that a PMA above would multiplex PCS lanes to a 200 Gb/s physical lane.
- For 200G and 400G the PMA would also extend the RS-FEC codeword interleaving from 2 to 4 codewords.
- Traditionally, a PMA is not required between a PCS/XS and a supplementary FEC sublayer, however, in this case a PMA will always be required with or without an AUI above the Inner FEC sublayer.
- Related baselines are here:
  - https://www.ieee802.org/3/dj/public/23_03/patra_3dj_01b_2303.pdf
  - https://www.ieee802.org/3/dj/public/23_07/he_3dj_01_2307.pdf
  - https://www.ieee802.org/3/dj/public/23_07/he_3dj_02a_2307.pdf

PMA below IMDD and Coherent Inner FEC

- The baselines for both the IMDD (Clause 177) and Coherent (Clause 194) Inner FEC sublayer showed a PMA below the Inner FEC.
- However, it was not clear that a PMA is required there as some of that functionality is subsumed in the inner FEC soft decoder.
- So rather than define a PMA below the Inner FEC the draft defines the Inner FEC as including all of the (below) PMA functionality thus avoiding defining a pair of sublayers and definition of service interface between that likely would not be physically instantiated.
- Thus in Draft 0.1 the Inner FEC is defined to be immediately above the PMD, subsuming the sparse functionality of the PMA.

IMDD Inner FEC in Draft 0.1

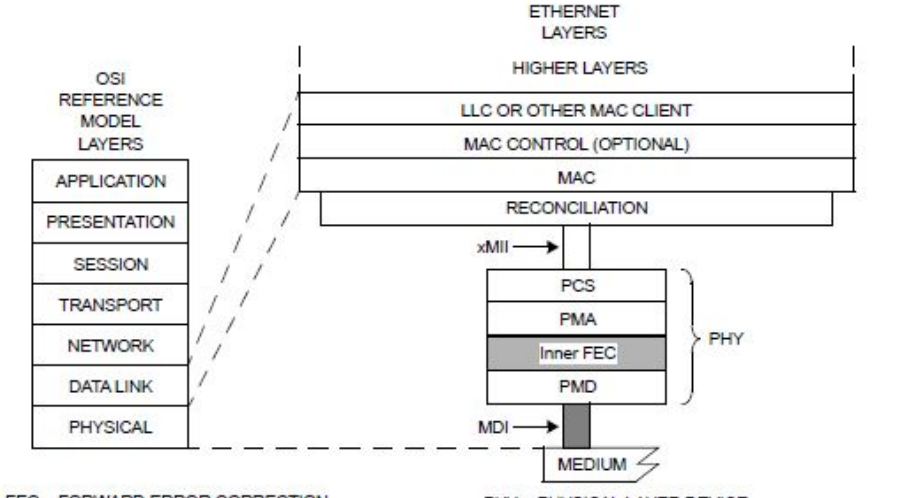


Figure 177–1—Relationship of Inner FEC and PMA to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 Ethernet model

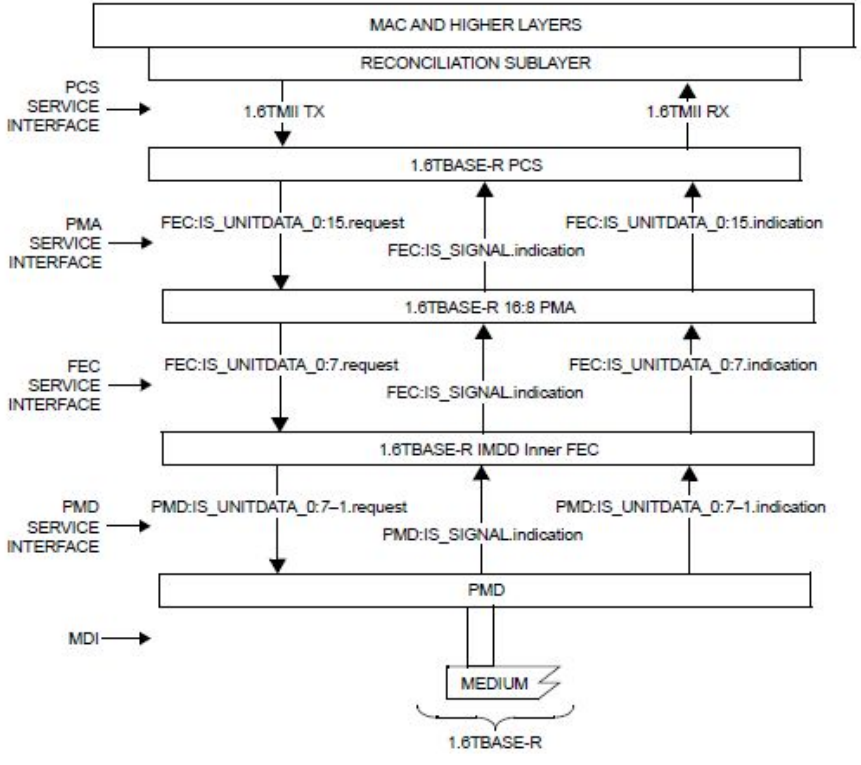
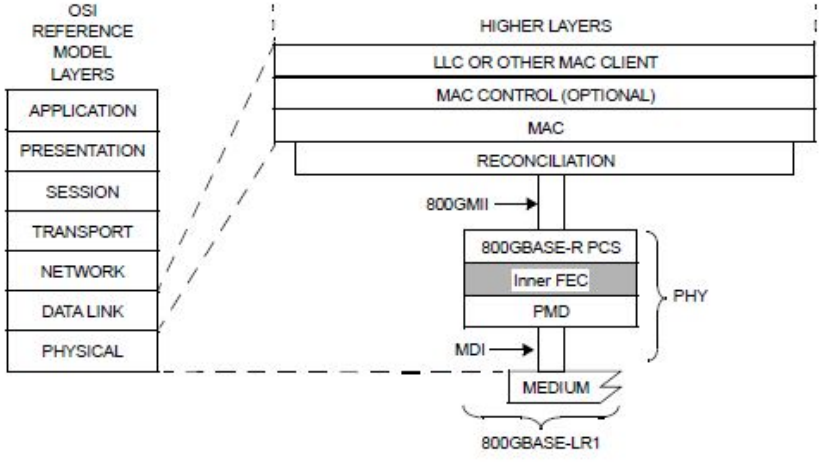


Figure 174–4—1.6TBASE-R inter-sublayer service interfaces including Clause 177 Inner FEC and not including 1.6TMII Extender

Coherent Inner FEC in Draft 0.1



800GMII = 800 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
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 184–1—Inner FEC relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 Ethernet model

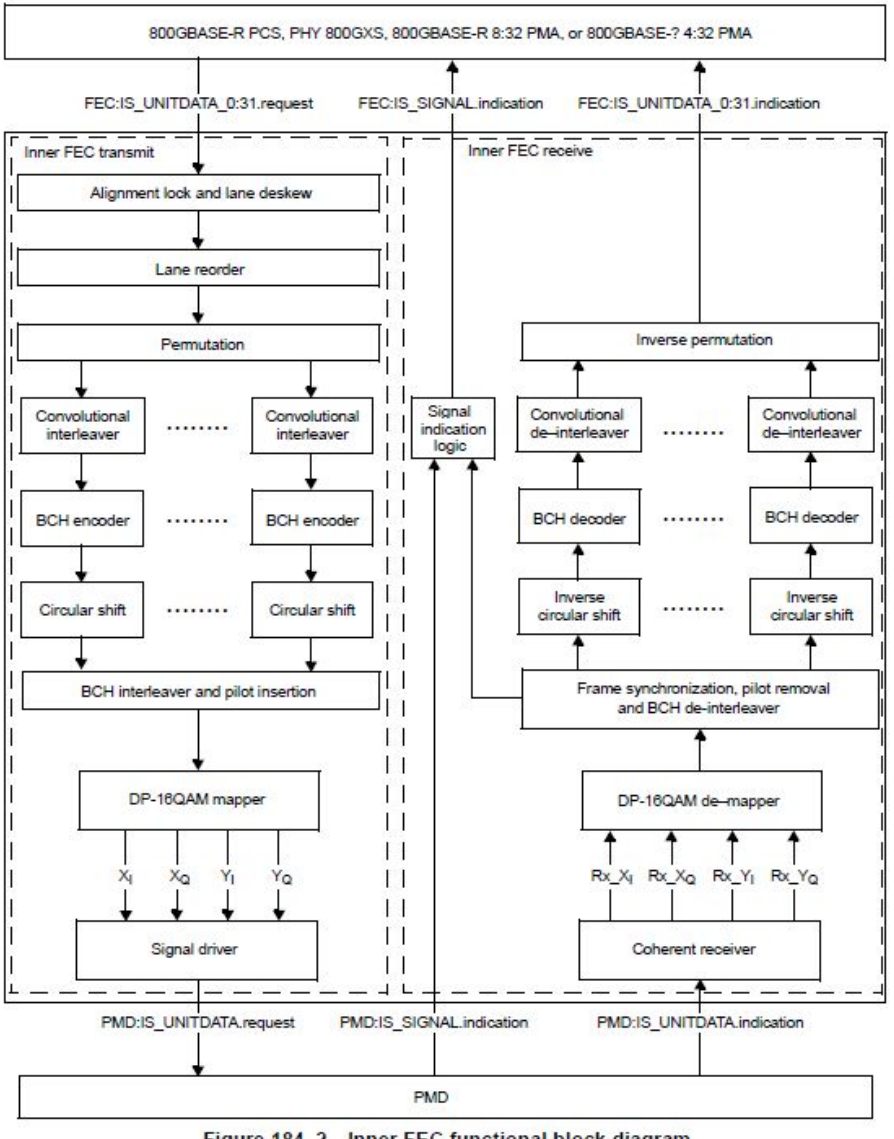


Figure 184–2—Inner FEC functional block diagram

Terminology

Terminology: Inner FEC for IMDD and Coherent PHYs

- Two Inner FEC sublayers are specified
  - “IMDD” Inner FEC (Clause 177)
  - “Coherent” Inner FEC (Clause 184)
- We have not adopted a short form to use for these and the long forms above are rather verbose to include in text, tables, etc.
- More compact, yet distinguishable, terms would be helpful in various parts of the draft.
- For the time being in Draft 0.1…
  - Using just “Inner FEC” within the Inner FEC clauses (177 and 184).
  - Elsewhere, if necessary to distinguish, using “IMDD Inner FEC” and “Coherent Inner FEC”.

cont…

169.2.4a Inner FEC sublayer

The Inner FEC sublayer provides error correction, in addition to that provided by the 800GBASE-R PCS, for the PMD.

For 800GBASE-DR4-2, 800GBASE-FR4, and 800GBASE-LR4 the Inner FEC is specified in Clause 177.

For 800GBASE-LR1 the Inner FEC is specified in Clause 184.

Terminology: Symbol/Bit muxing PMAs

- For PMDs or AUIs with 200 Gb/s per lane the PMA provides FEC-symbol multiplexing, rather than bit multiplexing.
- Terms are required to distinguish between the SM and BM PMAs especially when both are described within the same text, tables, and figures.
- Where necessary in Draft 0.1, using the the term “SM PMA” for the symbol multiplexing PMA and “BM PMA” for the bit multiplexing PMA.

cont…

Table 116–3—PHY type and clause correlation (200GBASE copper)

PHY type	Clause[a]																	
	73	78	117		118	119	120	120B	120D	120F	136	137	162	163	176	176B	178	179
	Auto-Negotiation	EEE	RS	200GMII	200GMII Extender	200GBASE-R PCS	200GBASE-R BM PMA	200GAUI-8 C2C	200GAUI-4 C2C	200GAUI-2 C2C	200GBASE-CR4 PMD	200GBASE-KR4 PMD	200GAUI-CR2 PMD	200GAUI-KR2 PMD	200GBASE-R SM PMA	200GAUI-1 C2C	200GAUI-KR1 PMD	200GAUI-CR1 PMD
200GBASE-KR1	M		M	O	O	M	O	O	O	O					M	O	M	
200GBASE-KR2	M		M	O	O	M	M	O	O	O				M	O	O		
200GBASE-KR4	M	O	M	O	O	M	M	O	O			M			O	O		
200GBASE-CR1	M		M	O	O	M	O	O	O	O					M	O		M
200GBASE-CR2	M		M	O	O	M	M	O	O	O			M		O	O		
200GBASE-CR4	M	O	M	O	O	M	M	O	O		M				O	O		

[a]O = Optional, M = Mandatory.

Thanks!