Architecture and Logic Overview

IEEE P802.3df Task Force
IEEE 802.3 Jan 2022
Interim Session

Mark Gustlin – Cisco

Goals for this presentation

- Review and highlight potential topics for the .3df Logic Track to explore
  - What does the logic architecture need to support
  - Logic architectural options
  - FEC structures options
  - Other misc. items
  - Summary/Needs

Focus of this Presentation

Organization of Project Work

Logic	Electrical	Optical
• Amendments to MAC, RS, and MAC PHY interfaces • RS and MII • Extender Sublayers? • PCS functions • PMA functions • Provide support to enable mapping over OTN	• Extender Sublayer? • C2C AUIs • C2M AUIs • Copper PMDs • Channel characteristics for electrical interfaces and PMDs	• Optical PMDs • MDIs? • Media Characteristics
FEC Architecture and Budget		
• Overall Architecture	• FEC related to electrical interfaces and PMDs	• FEC related to Optical PMDs
• Management related to Logic functions (Clauses 30, 45, etc.)	• Management related to electrical interfaces and PMDs (Clauses 30, 45, etc.)	• Management related to Optical PMDs (Clauses 30, 45, etc.)

Further insight to be provided by Track Leadership @ 18 Jan 2022 Meeting

From: dambrosia_3df_01_220111.pdf

What Does the Logic Architecture Need to Support?

- The coding needs of the electrical and optical interfaces (AUIs/PMDs)
  - Packet delineation, control code insertion, base line wander control, transition density control, error correction and false packet acceptance rejection
- The requirements of each interface can be different and vary over time and the #lanes
  - Multi-generational interoperability can be a consideration
  - For example, supporting 100Gbps/lane AUIs with 200Gbps/lane PMDs etc.
- We need to be able to add, remove or concatenate FECs as needed for a given interface
  - Details depend on the FEC structures chosen in this task force
- Desire is a single high-level architecture which gives us the flexibility to do all the above
  - Same architecture is desirable even if this project is split into different timelines

Support for Adopted Objectives

- We desire consistent Logic Architecture/structure for supporting all the various PMDs and AUIs of this project

Adopted Physical Layer Objectives

Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Technology Reuse

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Potential for either direct detect and / or coherent signaling technology

Making it all work together

From: https://www.ieee802.org/3/B400G/public/21_1028/B400G_overview_c_211028.pdf

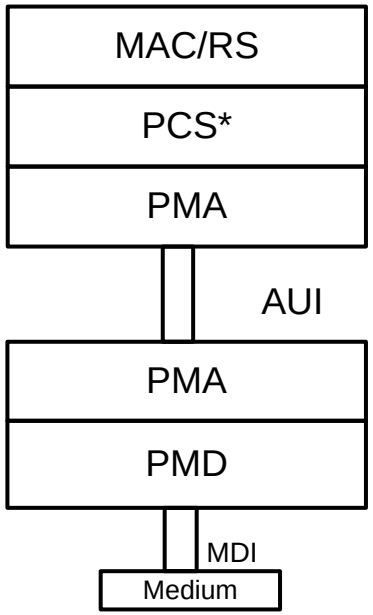
Architectural Options

Current 200GbE/400GbE Architecture

- PCS and FEC are integrated into a single block

802.3bs architecture

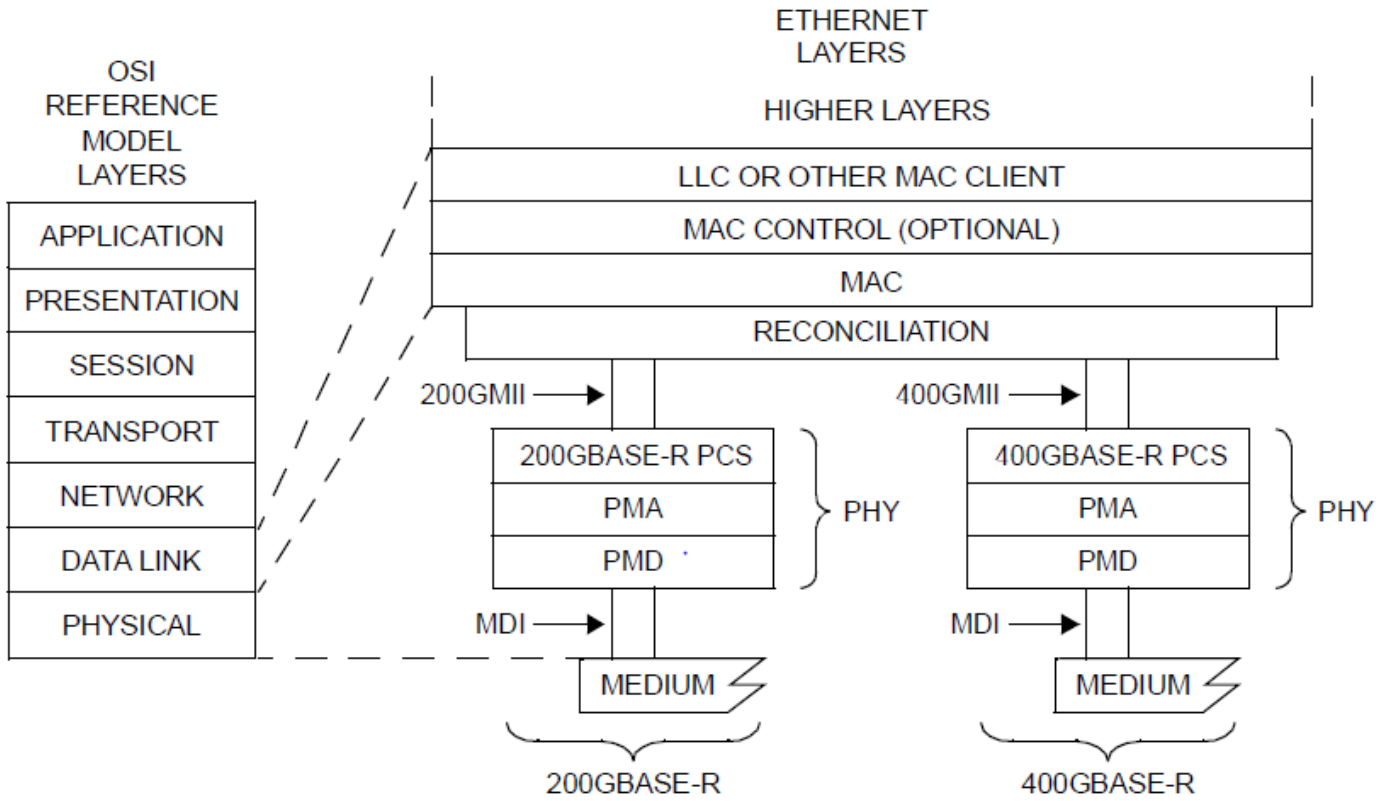
End to end FEC



*FEC is part of the PCS sublayer

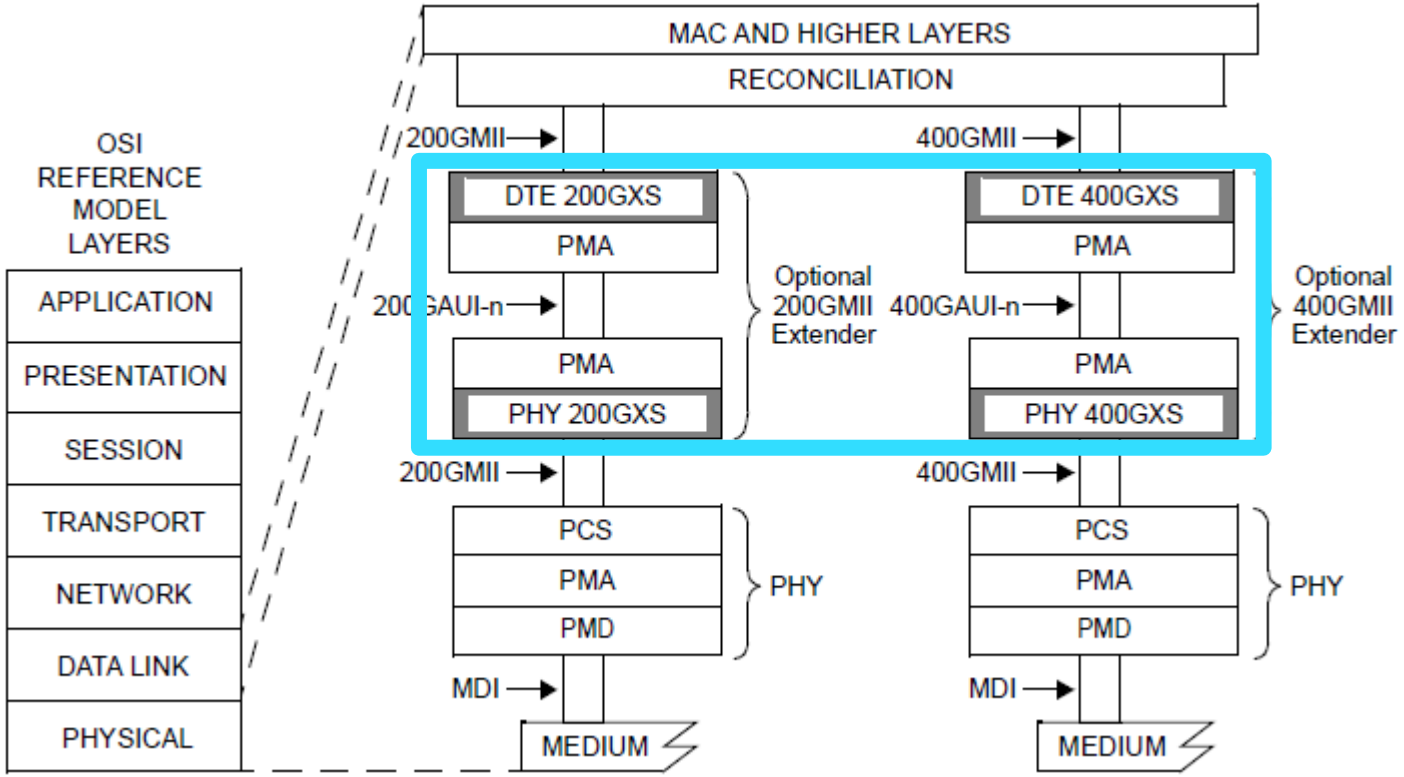
Full 200GbE/400GbE Architecture

- Architecture in the larger OSI context from the specification



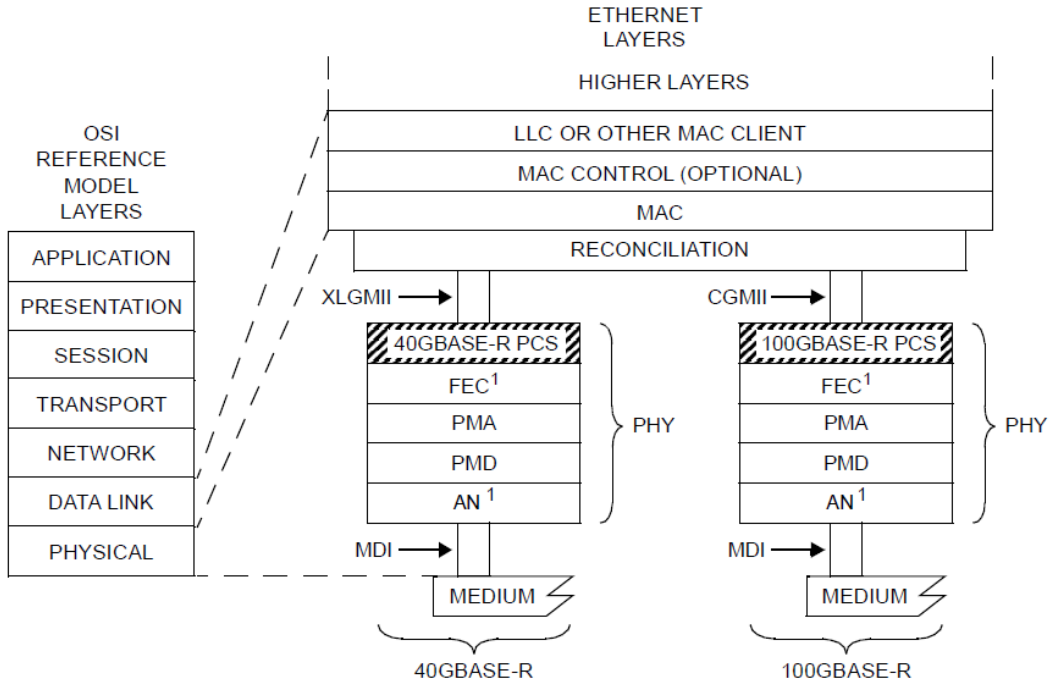
200GbE/400GbE Extender Sublayer

- An extender sublayer is defined to allow us to remove the RS544 FEC and get back to the MII
- This is used for a 400GBASE-ZR interface, for instance
- Implementations don't really go back to the MII interface



Full 100GbE Architecture

- Architecture in the larger OSI context from the specification
- FEC is a separate sublayer from the PCS



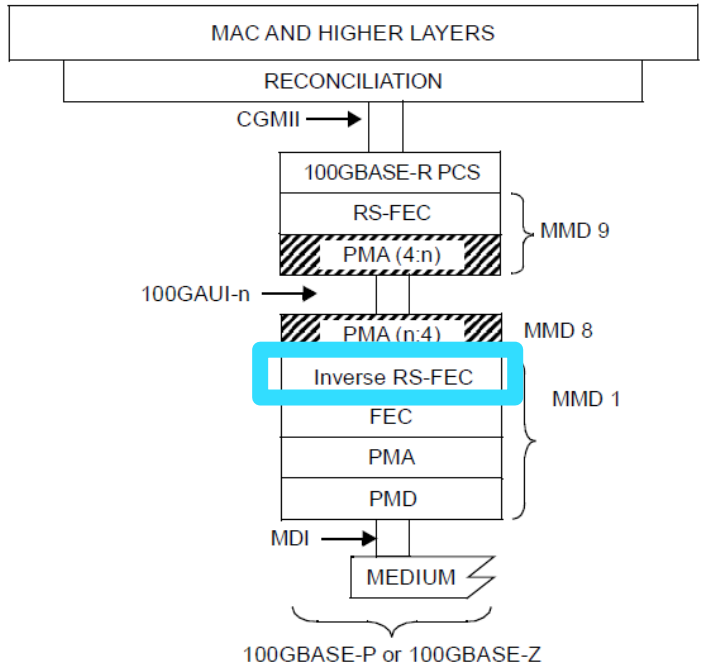
AN = AUTO-NEGOTIATION
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
LLC = LOGICAL LINK CONTROL
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PHY = PHYSICAL LAYER DEVICE
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
XLGMII = 40 Gb/s MEDIA INDEPENDENT INTERFACE

NOTE 1—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 Ethernet model

100GbE Inverse FEC Sublayer

- No Extender sublayer in this case, Inverse RS-FEC sublayer instead

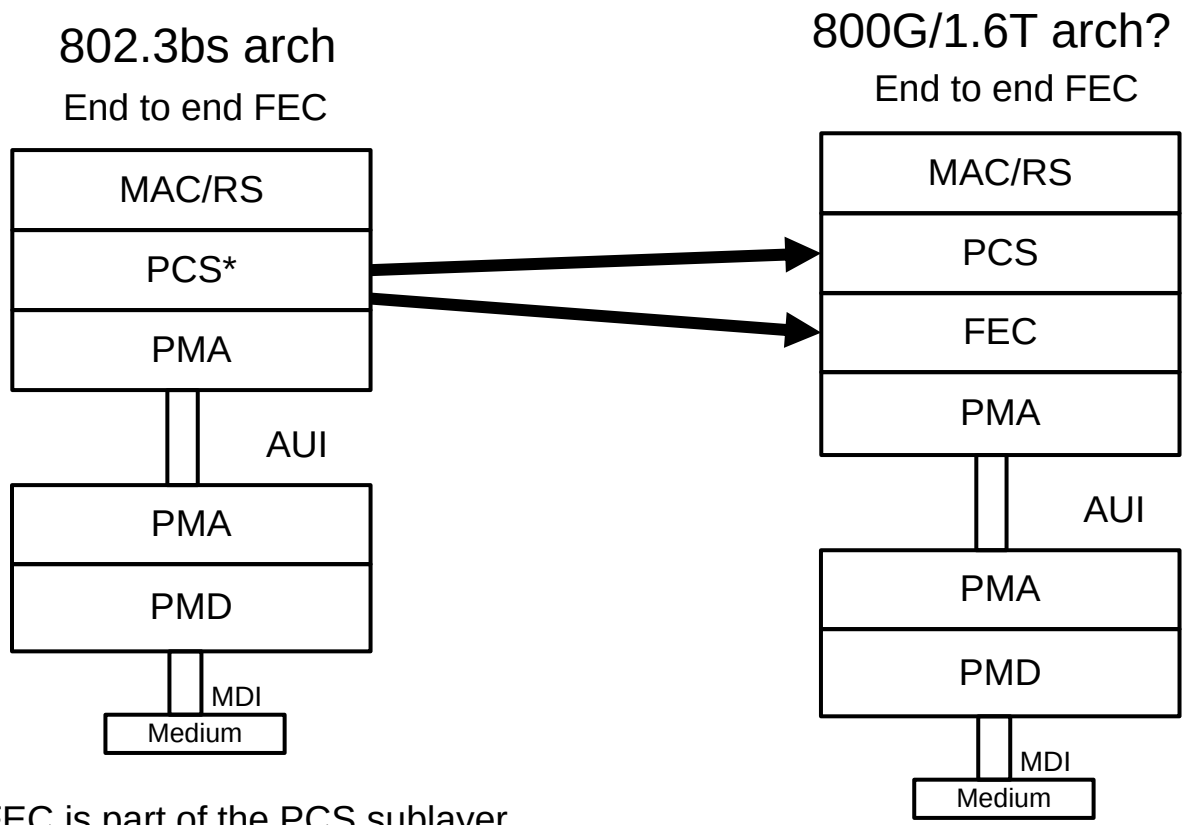


100GAUI-n = 100 Gb/s n-LANE ATTACHMENT UNIT INTERFACE
CAUI-4 = 100 Gb/s FOUR-LANE ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
FEC = FORWARD ERROR CORRECTION
MAC = MEDIA ACCESS CONTROL
MMD = MDIO MANAGEABLE DEVICE
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
RS-FEC = REED-SOLOMON FORWARD ERROR CORRECTION

Figure 135A–7—Example single 100GAUI-n with Inverse RS-FEC

Possible Architecture – 800G/1.6T

- Should we change the high-level architecture for 800G/1.6T?
  - Go back to the 100GbE architecture (though with some differences)?
- The FEC as its own sublayer can make the specification/evolution simpler
- Use Inverse FEC when needed (no extender sublayer)



*FEC is part of the PCS sublayer

Notes on Proposed 800G/1.6T PCS Architecture

- We can define a scalable PCS with a logical interface to the FEC sublayer
  - Allows us to bolt on new FEC as needed
  - To remove FEC, need to define an inverse FEC sublayer

- The logical interface is: a single 66b interface run at whatever rate is needed for a MAC rate
  - Assumes we reuse 64B/66B encoding

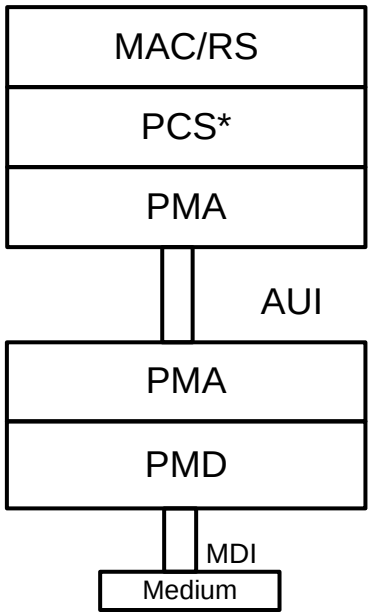
- + Some control signals

- No notion of PCS lanes (a FEC sublayer must be adjacent to the PCS sublayer)
  - This is different than 100GbE architecture where the PCS and FEC can be physically separated (with an AUI)
  - FEC sublayer does all lane mapping/reordering

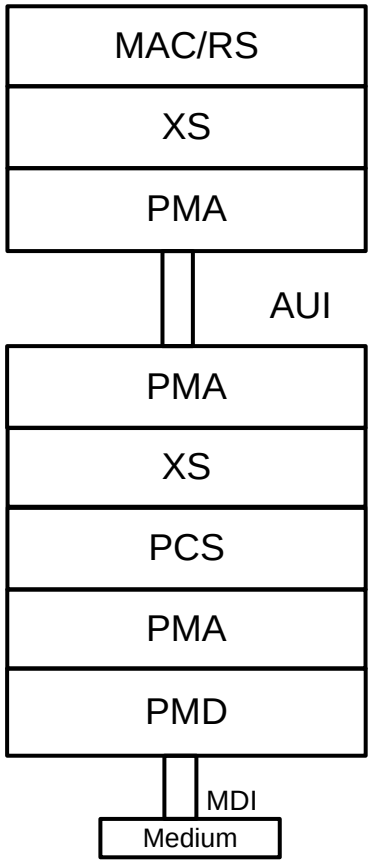
200GbE/400GbE Architecture for 200Gbps/lane

- Assuming a different FEC than CL119 is needed for 200Gbps/lane PMD/AUIs
- Use the existing architecture (from 802.3bs)
- Define a **new** PCS sublayer with a new FEC as appropriate/needed
  - Refer to clause 119 for any unchanged portions
  - Use an extender sublayer when AUI and PMD require different FEC formats

802.3bs architecture



*FEC is part of the PCS sublayer



FEC Strategies

Support for Adopted Objectives – 800GbE/100G

100 Gbps/lane

- Let's looks at the FEC strategy for just the 800GbE with 100Gbps/lane throughout the link

Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Technology Reuse

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Potential for either direct detect and / or coherent signaling technology

800GbE 100Gbps/lane FEC Strategy

100 Gbps/lane

- Applies to 100G lanes
  - Compatible with 50G AUIs also

- End to end FEC reusing RS(544,514,10)
- A unique FEC sublayer (likely different from 200G/lane)

- Options:
  1. Reuse the Ethernet Technology Consortium spec
     - 66b interleaving into transcoding, 4 FEC codeword interleaving, 32 FEC lanes
     - Industry reuse
     - Fast time to market for 800GbE with 100G lanes
  2. Clause 119 like FEC
     - 2 FEC codeword interleaving
     - Shown in wang_b400g_01_210208.pdf

- 106.25G per lane for AUI-8

Support for Adopted Objectives – 200G/FEC

200 Gbps/lane

- Let's looks at the FEC strategy for 200Gbps/lane

Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Technology Reuse

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Potential for either direct detect and / or coherent signaling technology

200Gbps/lane Possible High Level FEC Strategies

200 Gbps/lane

- Big goals:
  - Support power and cost-effective PMDs/ASICs at 200Gbps/lane
  - Support 200G FEC across AUIs at 100Gbps/lane and 200Gbps/lane
    - Assuming we are not reusing the clause 119 RS-FEC directly

- Big questions:
  - What FEC gain is required by the PMDs/AUIs at 200Gbps/lane?
  - What FEC latency is tolerable (may vary by PMD/AUI)?
  - What error models are dominant (Gaussian vs. non-Gaussian)
  - These questions will likely take a long time to answer
    - Wrapped up in the modulation chosen per interface type
    - Equalization used, etc.

200Gbps/lane Possible High Level FEC Strategies - cont

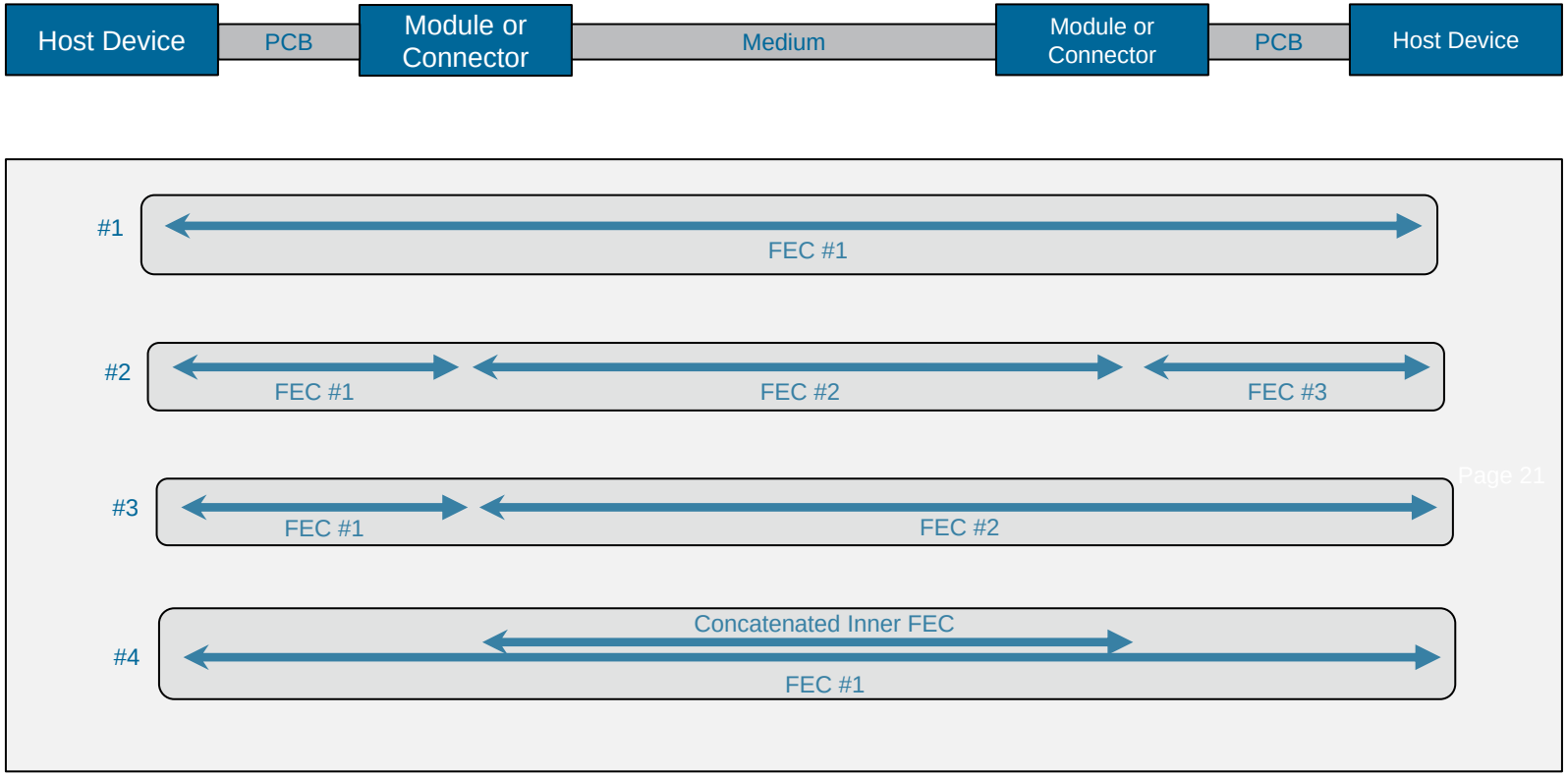
- What FEC architecture do we need/want?
  - End to end vs. segmented vs. concatenated structures (or other)
  - Is it PMD dependent, AUI dependent?
    - Almost certainly for some PMDs (40km for instance)
  - Where to optimize power vs. flexibility vs. gain vs. latency
    - Latency less critical for optical PMDs?

- Is the reuse of the RS544 FEC part of the strategy?
  - Benefits are obvious for reuse in large switch chips
  - Any drawbacks?

- Concerns:
  - Any increase in the AUI's data rate and that impact on net coding gain
    - Less of an issue for the optical PMDs?
  - How do the needs of copper cable impact other PMDs?

Possible FEC Schemes for 200Gbps/lane

200 Gbps/lane



#1 End to end FEC
#2 Segmented FEC
#3 Mix of #1/2
#4 Concatenated

Is the modulation consistent across the link?

From: https://www.ieee802.org/3/B400G/public/21_1028/B400G_overview_c_211028.pdf

Support for Adopted Objectives – 200GFEC/100G

100 Gbps/lane

- Let's looks at 100Gbps/lane AUIs with the 200G FEC structure

Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Technology Reuse

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

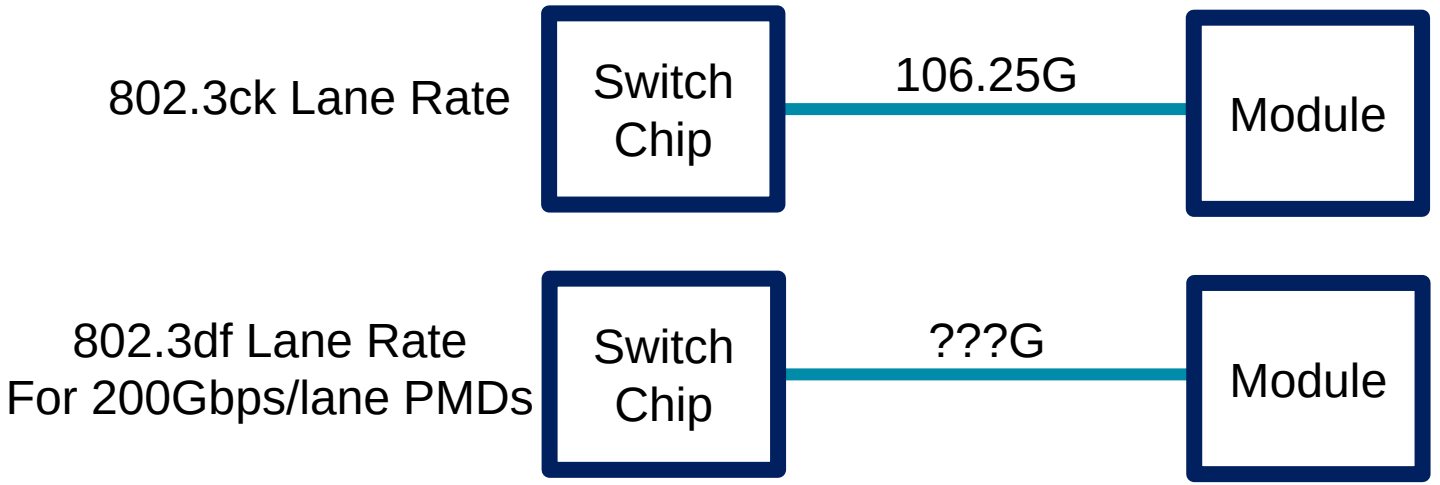
Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Potential for either direct detect and / or coherent signaling technology

100Gbps/lane AUIs with 200G FEC Structure

- Support the 200Gbps/lane FEC structures over 100Gbps/lane AUIs
- Goal is to reuse the electrical work from 802.3ck
  - This implies the lane rate is still 106.25Gb/s, is that achievable?
- If we can't achieve this goal, then the 802.3ck work must be extended for the new lane rate
  - In this task force



Support for Adopted Objectives – Longer Reach

Longer Reach

- Let's looks at Longer reach PMDs

Ethernet Rate	Assumed Signaling Rate	AUI	BP	Cu Cable	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
200 Gb/s	200 Gb/s	Over 1 lane		Over 1 pair			Over 1 Pair	Over 1 Pair		
400 Gb/s	200 Gb/s	Over 2 lanes		Over 2 pairs			Over 2 Pair			
800 Gb/s	100 Gb/s	Over 8 lanes	Over 8 lanes	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs	Over 8 pairs		
	200 Gb/s	Over 4 lanes		Over 4 pairs			Over 4 pairs	1) Over 4 pairs 2) Over 4 λ's		
	TBD								Over single SMF in each direction	Over single SMF in each direction
1.6 Tb/s	100 Gb/s	Over 16 lanes								
	200 Gb/s	Over 8 lanes		Over 8 pairs			Over 8 pairs	Over 8 pairs		

Technology Reuse

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3cu, 3ck, 3db) to higher lane counts

Develop 200 Gb/s per lane electrical signaling for 1/2/4/8 lane variants of AUIs and electrical PMDs

Develop 200 Gb/s per optical fiber for 1/2/4/8 fiber based optical PMDs and 4 lambda WDM optical PMD

Potential for either direct detect and / or coherent signaling technology

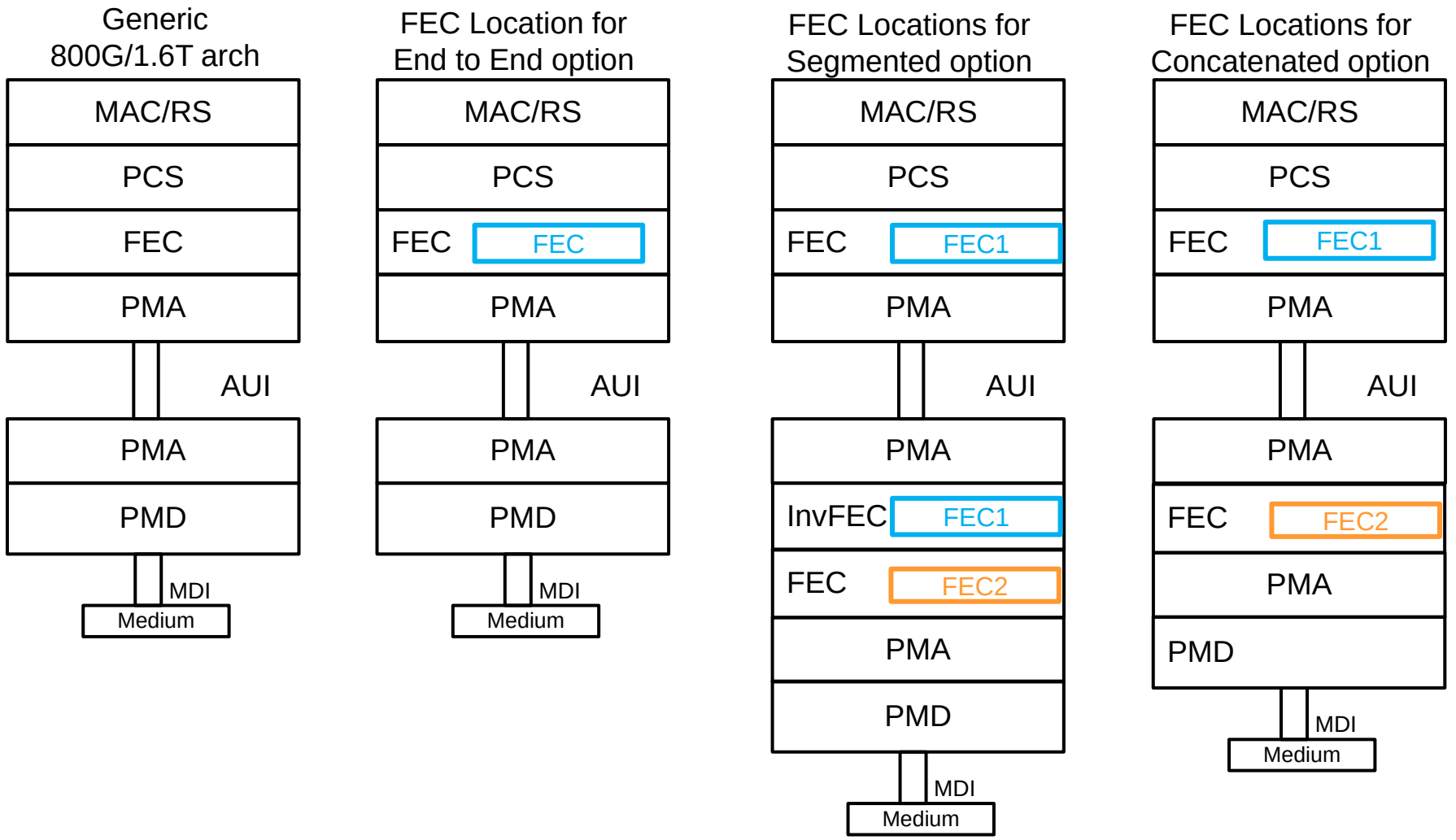
Longer Reach PMDs

- Number of lanes and technology is TBD
- Assumptions is that these PMDs will have a separate FEC strategy (higher gain) than the baseline FEC strategy
- These PMDs will need to reverse any FEC added for the AUI and apply their own FEC structures as needed
- We would use either an extender sublayer or reverse/inverse FEC architectural strategy to address these needs

More on the Architecture with FEC in the PMD

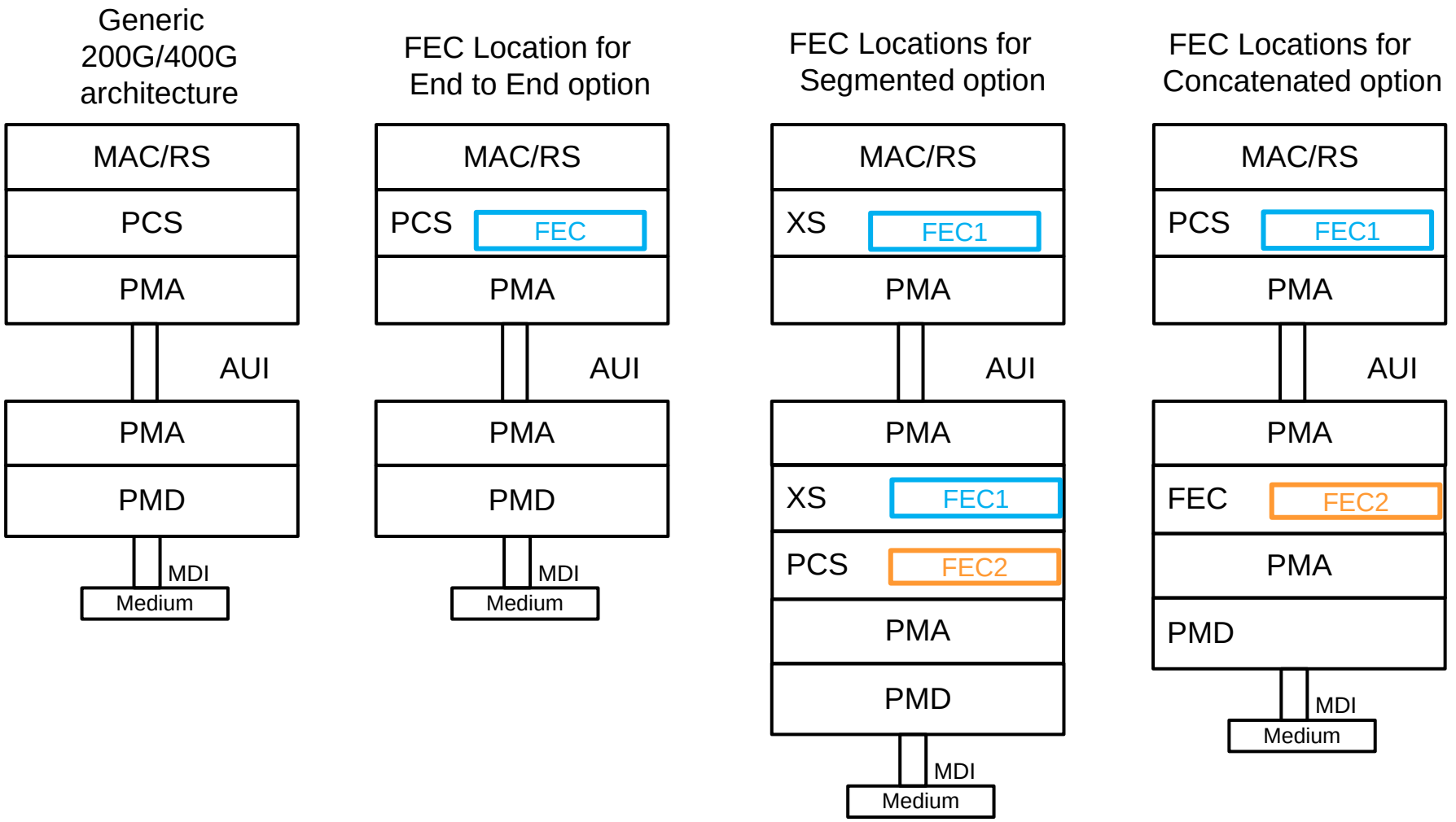
FEC in a Possible 800GbE/1.6TbE Architecture

- How 200G/Lane FEC options fit into the architecture



FEC in a the 200GbE/400GbE Architecture

- How 200G/Lane FEC options fit into the architecture



Other Thoughts..

Possible 200Gbps/lane strategy for Pre-FEC BER

- Define a pre FEC BER target for the PMDs/AUIs that allows for the difficult task we have in front of us
  - Requires higher FEC gain compared to 100Gbps/lane
  - Strikes a balance between power, complexity, cost, latency and encoded lane rates
- Then the task force can define FEC structures to enable this raw BER
- The optics and electrical groups then validate that they can live with those targets over time
  - We can always change the target if the task forces deems that desirable
- Possible targets:
  - PMDs: $Ax10^{-3}$, where A = 1, 2, 3, 4 or something else?
    - 100G/lane is $2.4x10^{-4}$
    - Not including coherent, which I assume has its own FEC strategy
  - AUIs: $Bx10^{-4}$, where B = 1, 2, 3, 4 or something else?
    - 100G/lane is $1x10^{-5}$
  - Need to discuss how the randomness of errors (or not) impacts this

Other Items

- AN/LT
  - Do we need to make any changes?
  - Is this the time to apply AN to optics?
- What is the OTN reference point?
- PMA assumptions
  - Support blind bit muxing where feasible
    - Some generations are incompatible and don’t support this
  - Any lane anywhere flexibility
  - Appropriate coding for technology choices
  - Modulation format conversion?

Summary/Needs

- Separation of the PCS/FEC sublayers seems to make sense for 800GbE/1.6TbE
  - A simple architecture flows from this
  - Inverse FEC sublayers are a little simpler than extender sublayers
- A few choices for the 100Gbps/lane FEC structure
- 200Gbps/lane FEC structure, codes etc. are wide open at this point
- What the Logic group needs from Optics/Electrical groups:
  - What FEC gain is required by the PMDs/AUIs at 200Gbps/lane?
  - What error models are dominant (Gaussian vs. non-Gaussian)
  - What raw BER is reasonable?
    - And provides reasonable tradeoffs (power vs. flexibility vs. gain vs. latency)
- The logic group needs contributions for (and ultimately to develop baseline proposals for):
  - The high-level architecture
  - The 100Gbps/lane PCS/FEC sublayers
  - The 200Gbps/lane PCS/FEC sublayers
  - The PMA sublayers
  - AN/LT

Thanks!