Consideration of Implementations and Architecture

IEEE P802.3df Task Force
Architecture & Logic Ad hoc

11 April 2022

John D'Ambrosia,
Futurewei, U.S. Subsidiary of Huawei

Introduction

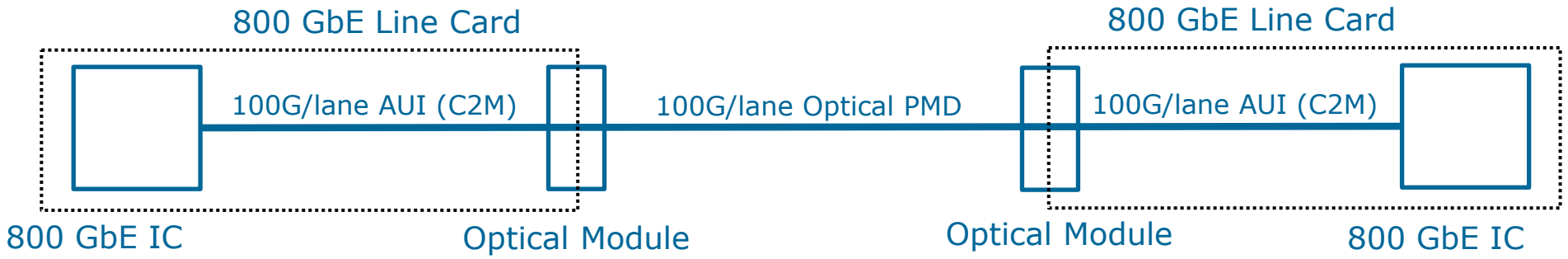
- Per ”Project Overview- IEEE P802.3df” (https://www.ieee802.org/3/B400G/public/21_1028/B400G_overview_c_211028.pdf), Slide #27
  - An initial holistic approach to start is recommended to develop an architecture that would support one or more FEC schemes
    - This requires significant analysis and technical decisions by the Task Force
- Per John D’Ambrosia – “We are writing a standard, not an implementation specification”
- Note – This presentation is conceptual and short on all the details. Let’s walk before we run.

Adopted Physical Layer Objectives & Nomenclature

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 200GAUI-1		Over 1 pair 200GBASE-CR1			Over 1 Pair TBD	Over 1 Pair TBD		
400 Gb/s	100 Gb/s							Over 4 Pair TBD		
	200 Gb/s	Over 2 lanes 400GAUI-2		Over 2 pairs 400GBASE-CR2			Over 2 Pair TBD			
800 Gb/s	100 Gb/s	Over 8 lanes 800GAUI-8	Over 8 lanes 800GBASE-KR8	Over 8 pairs 800GBASE-CR8	Over 8 pairs 800GBASE-VR8	Over 8 pairs 800GBASE-SR8	Over 8 pairs TBD	Over 8 pairs TBD		
	200 Gb/s	Over 4 lanes 800GAUI-4		Over 4 pairs 800GBASE-CR4			Over 4 pairs TBD	1) Over 4 pairs TBD 2) Over 4 λ's TBD		
	TBD								Over single SMF in each direction TBD	Over single SMF in each direction TBD
1.6 Tb/s	100 Gb/s	Over 16 lanes 1.6TAUI-16								
	200 Gb/s	Over 8 lanes 1.6TAUI-8		Over 8 pairs 1.6TBASE-CR8			Over 8 pairs TBD	Over 8 pairs TBD		

We need to approach all of this holistically

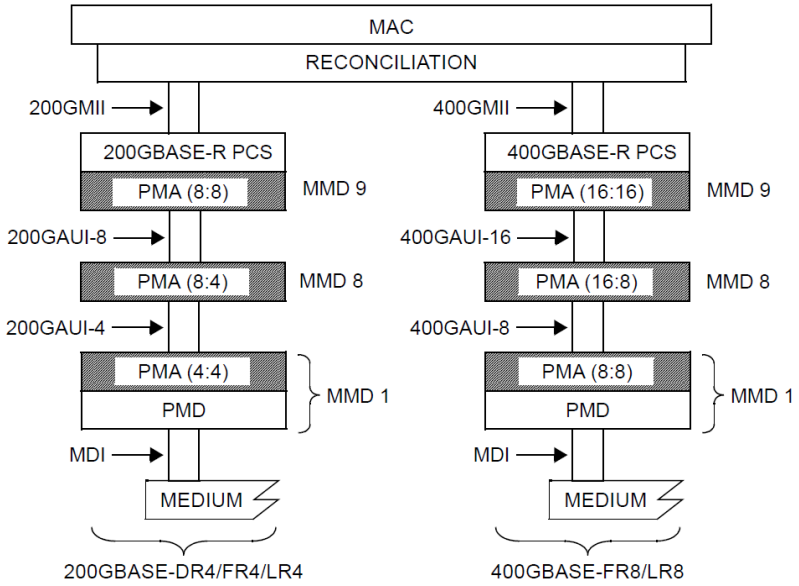
Pre-occupation with Single AUI / Line Card Implementations



- **Simple example has been utilized in TF presentations –**
  - **Approach is not holistic**
  - **Approach is not future proof**
  - **Approach is not aligned with prior generations of Ethernet specifications**
- **Note – this presentation is not discussing PHY Sublayer details**

Partitioning Examples Per IEEE 802.3 Standards

- Clause 45 Management specifies support for 2 AUIs
- Partitioning Examples illustrate potential for 2 AUI's between single MAC and MDI
  - 25 Gb/s - see IEEE 802.3-2018 Annex 109C - Fig 109C-5
  - 40 Gb/s – see IEEE 802.3-2018 Annex 83C - Fig 83c-7
  - 50 Gb/s – See IEEE 802.3cd-2018 Annex 135A – Fig 135A-4
  - 100 Gb/s –
    - see IEEE 802.3-2018 Annex 83C - Fig 83c-7
    - See IEEE 802.3ck Annex 135A – Fig 135A-8
  - 200 Gb/s – see IEEE 802.3-2018 Annex 120A – Fig 120A-4
  - 400 Gb/s – see IEEE 802.3-2018 Annex 120A – Fig 120A-4

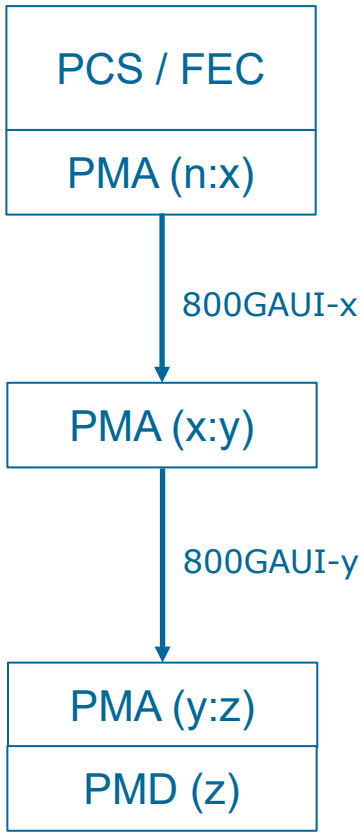


200GAUI = 200 Gb/s ATTACHMENT UNIT INTERFACE
200GMII = 200 Gb/s MEDIA INDEPENDENT INTERFACE
400GAUI = 400 Gb/s ATTACHMENT UNIT INTERFACE
400GMII = 400 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
MMD = MDIO MANAGEABLE DEVICE
PCS = PHYSICAL CODING SUBLAYER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

Figure 120A–4—Example 200GBASE-DR4/FR4/LR4 or 400GBASE-FR8/LR8 PMA layering with 200GAUI-8 or 400GAUI-16 chip-to-chip and 200GAUI-4 or 400GAUI-8 chip-to-module interfaces

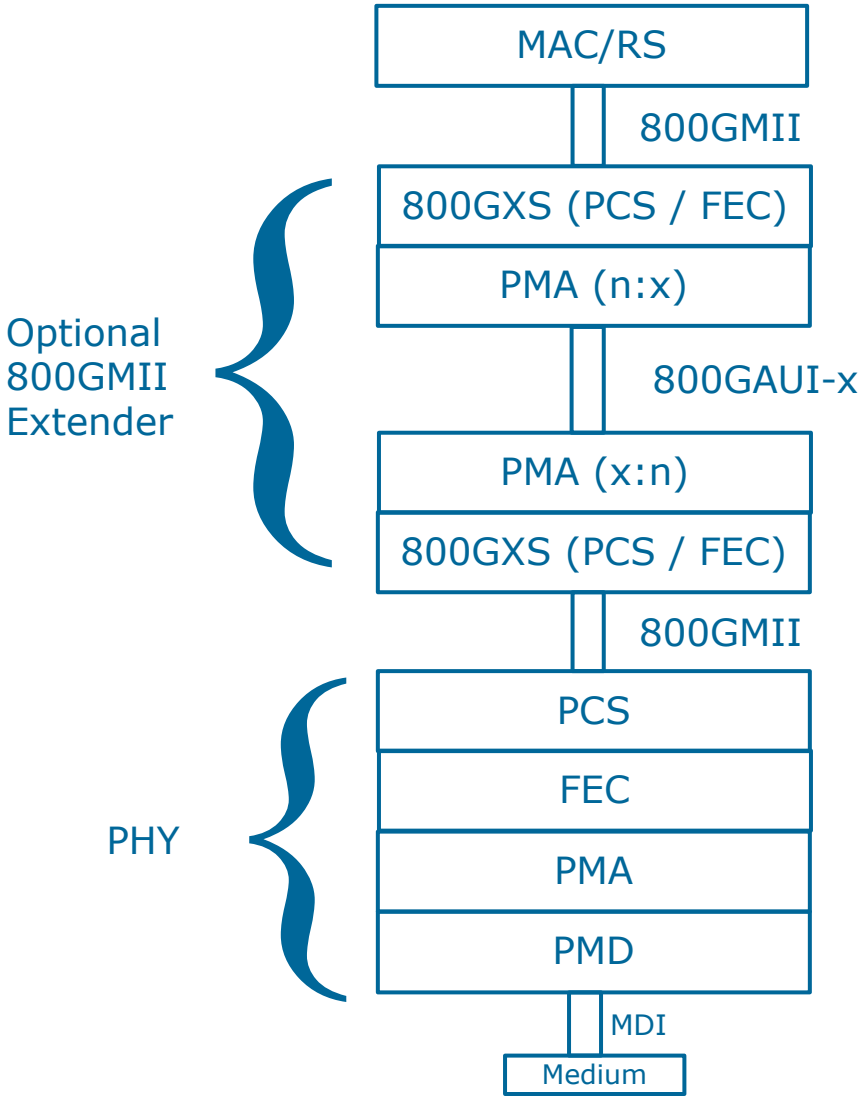
Example Diagram

Remembering the PMA / AUI



- PCS / FEC to be determined. Logical interfaces between PCS / FEC and FEC / PMA TBD.
- PMA in prior Ethernet rates provided muxing / demuxing
  - AUI's are physical instantiations of the connection between two adjacent PMA sublayers.
  - Assume that the PMA will allow transition between 800GAUI-8 and 800GAUI-4. For the purpose of this presentation, it is described as "mux / demux". Should it be general for future higher signaling rate / lane?
  - Assuming any necessary FEC is not in the PMA sublayer
  - Full functionality of PMA sublayer (or other intermediate sublayers) is not defined at this point.
- Example to left
  - Not clear how many lanes are fed into top PMA
  - Top PMA sublayer need to be able to mux / demux to 8 or 4 lanes (x) dependent upon which AUI is being supported
  - Middle PMA sublayer is either muxing /demuxing or simple retiming to whatever AUI is being supported (y)
  - Bottom PMA sublayer is either muxing /demuxing or simple retiming to whatever PMD is being supported (z)

What About an Extender Sublayer?



Observations

An Extender Sublayer enables PHY choice –

- Change PCS
- Change FEC
- Change PMD

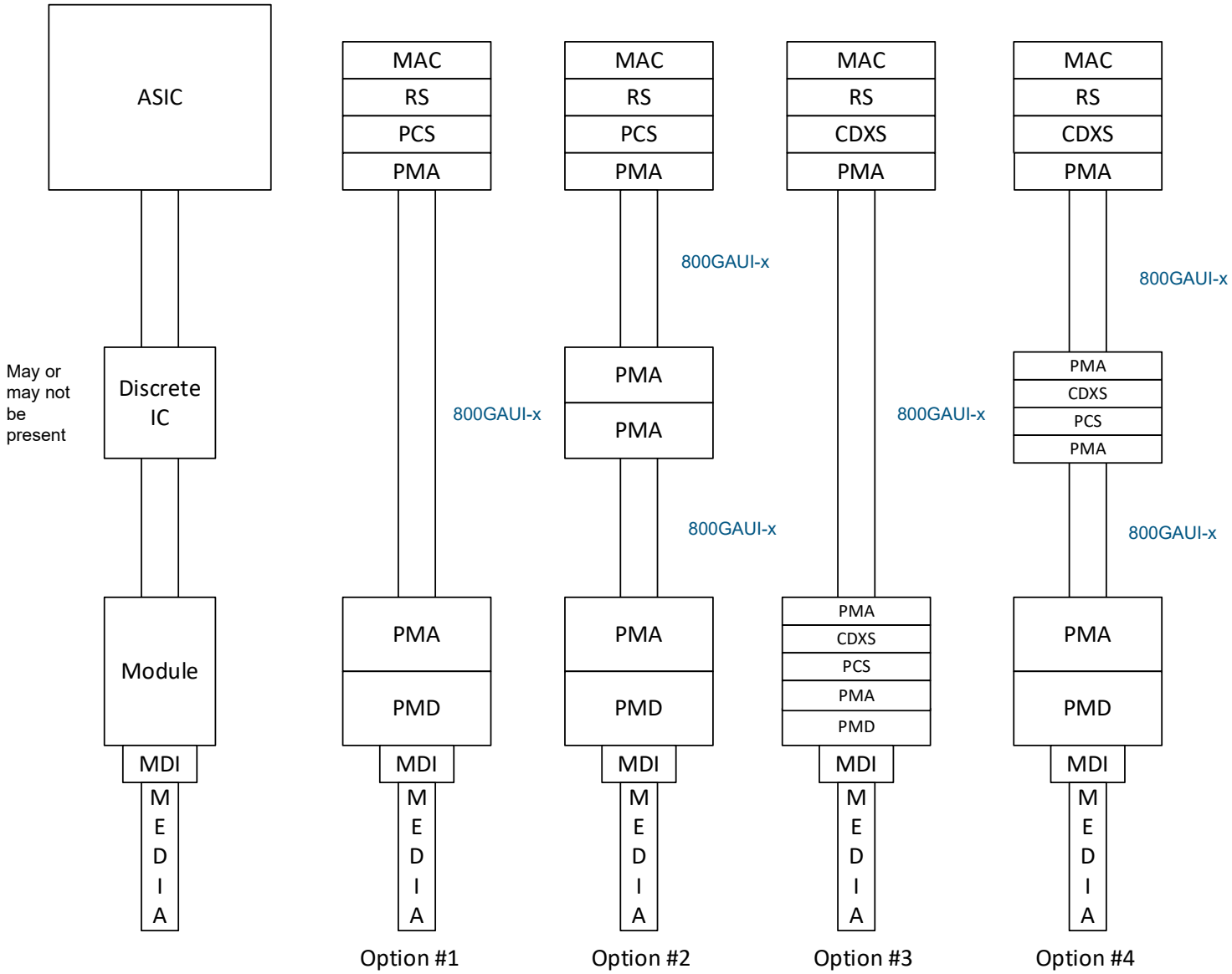
- Example – Future 800GBASE-ZR PHY
- Might be helpful for – CR PHYs, as FEC for extender sublayer could be used to cover reach on line card

In my opinion as .3ct/.3cw Chair, development of 400GBASE-ZR architecture was relatively easier / quicker than 100GBASE-ZR architecture due to presence of 400G Extender Sublayer

Cons of Extender Sublayer

- Power
- Complexity
- Latency

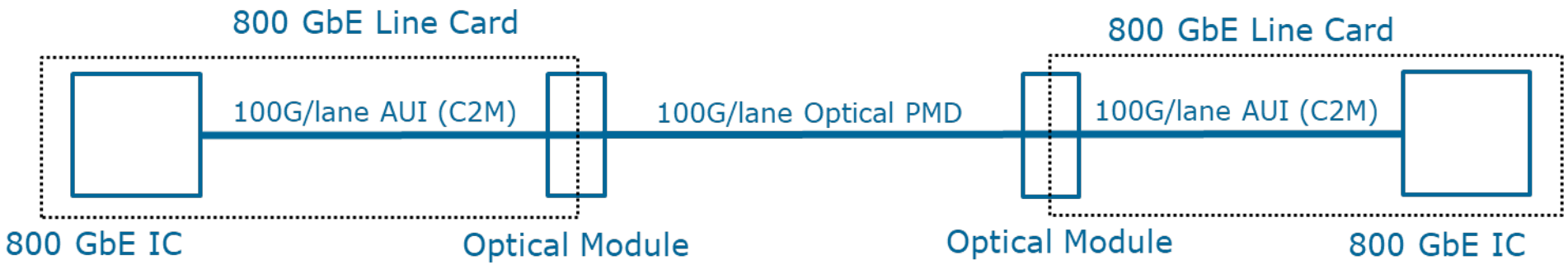
Potential 800GbE Example Implementations



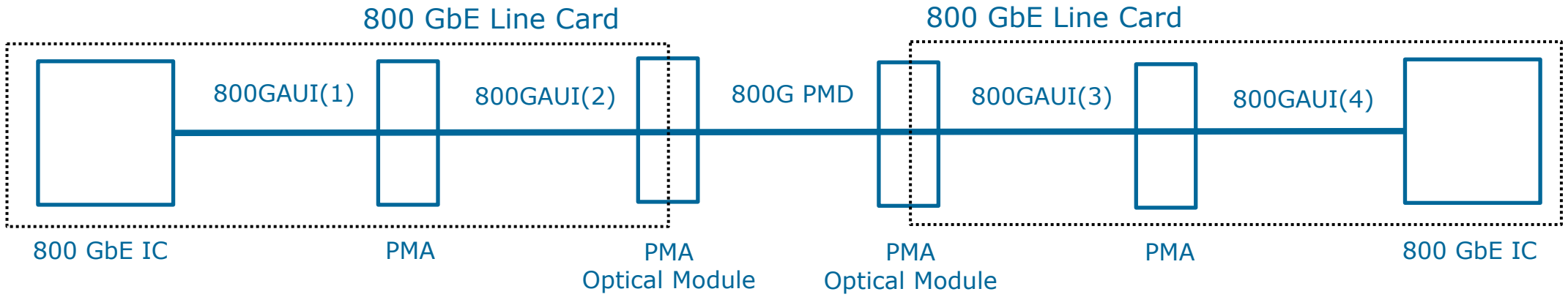
- FEC sublayers have been left off on purpose. Subsequent analysis needs to consider these different options
- "x" may be different for various interfaces cited for 800GAUI
- 800G Extender Sublayer has been shown for illustrative purposes

800 Gb/s Ethernet Base Scenario

Implementation Perspective



Standards Perspective



- **Next page looks at potential scenarios that we need to consider in the development of the IEEE P802.3df standard**
- **Does not include extender sublayer approach**

Potential Scenarios for 800 GbE (Others are possible)

	800GAUI (1)	800GAUI (2)	800G PMD	800GAUI (3)	800GAUI (4)	Notes
1	800GAUI-8	N/A	8 x 100G	N/A	800GAUI-8	
2	800GAUI-8	800GAUI-8	8 x 100G	800GAUI-8	800GAUI-8	
3	800GAUI-8	N/A	4 X 200G	N/A	800GAUI-8	
4	800GAUI-8	N/A	4 X 200G	800GAUI-8	800GAUI-8	
5	800GAUI-8	N/A	4 X 200G	N/A	800GAUI-4	
6	800GAUI-8	N/A	4 X 200G	800GAUI-4	800GAUI-4	
7	800GAUI-8	N/A	4 X 200G	800GAUI-8	800GAUI-4	
8	800GAUI-8	N/A	4 X 200G	800GAUI-4	800GAUI-8	
9	800GAUI-8	800GAUI-8	4 X 200G	N/A	800GAUI-8	
10	800GAUI-8	800GAUI-8	4 X 200G	800GAUI-8	800GAUI-8	
11	800GAUI-8	800GAUI-8	4 X 200G	N/A	800GAUI-4	
12	800GAUI-8	800GAUI-8	4 X 200G	800GAUI-4	800GAUI-4	
13	800GAUI-8	800GAUI-8	4 X 200G	800GAUI-8	800GAUI-4	
14	800GAUI-8	800GAUI-8	4 X 200G	800GAUI-4	800GAUI-8	
15	800GAUI-8	800GAUI-8	1 X 800G	N/A	800GAUI-8	
16	800GAUI-8	800GAUI-8	1 X 800G	800GAUI-8	800GAUI-8	
17	800GAUI-8	800GAUI-8	1 X 800G	N/A	800GAUI-4	
18	800GAUI-8	800GAUI-8	1 X 800G	800GAUI-4	800GAUI-4	
19	800GAUI-8	800GAUI-8	1 X 800G	800GAUI-8	800GAUI-4	
20	800GAUI-8	800GAUI-8	1 X 800G	800GAUI-4	800GAUI-8	

Summary

- The Task Force needs to consider scenarios that include up to two AUIs per line card
  - The implications of this need to be considered when developing the overall FEC architecture.
- Decision about extender sublayer?
- It is clearly complex, but we need to start considering these scenarios.