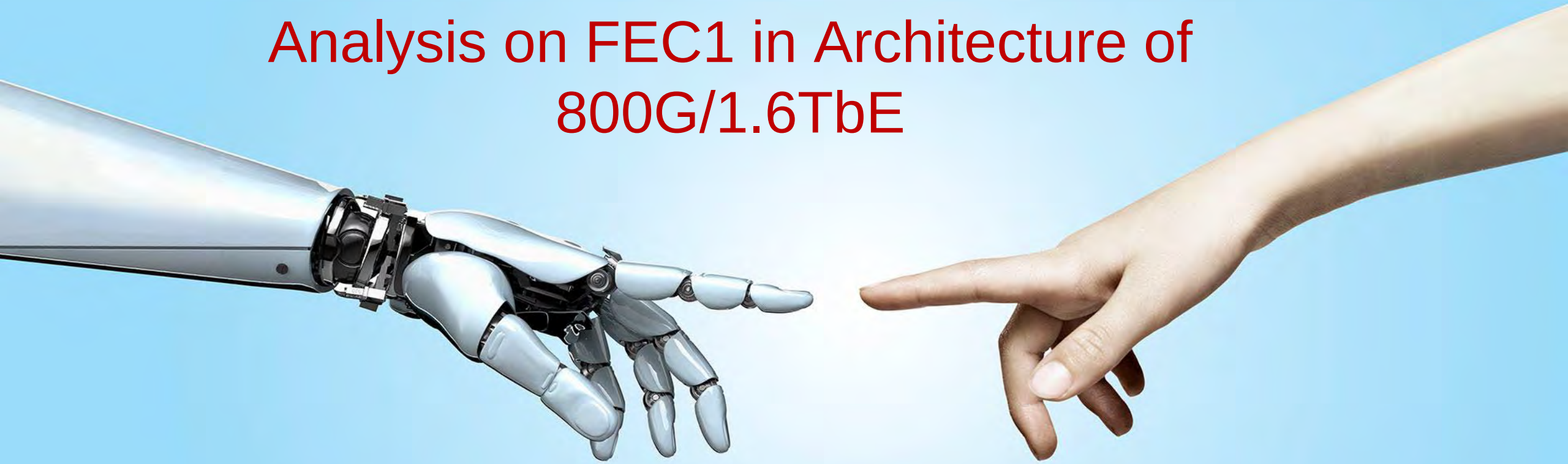
Analysis on FEC1 in Architecture of 800G/1.6TbE



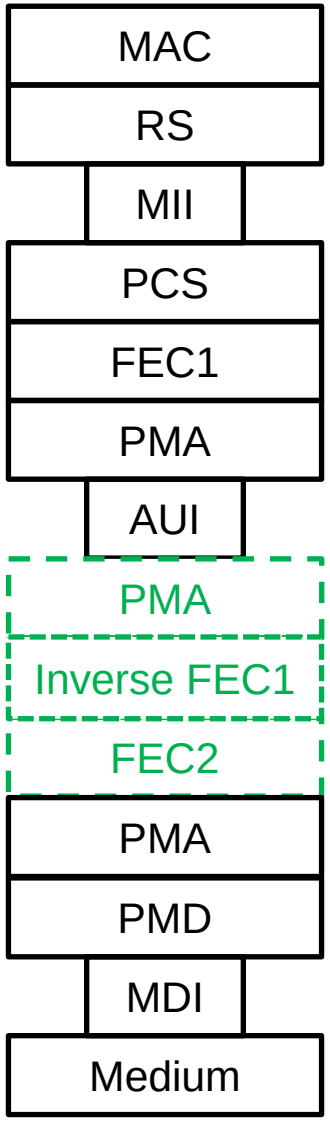
Xinyuan Wang, Xiang He

HUAWEI

Expectation for Logic Layer Architecture

- Architecture, especially PCS/FEC/PMA related, should enable diversified implementations and interconnecting scenarios. IEEE 802.3bs was a good example of successful practice in defining an architecture that:
  - Enabled diversified transceiver technologies, modulations, equalization techniques in IEEE 802.3 200/400 Gb/s Ethernet standard.
  - Supported “50 Gb/s per lane based AUI ↔ Optical PHY ↔ 100 Gb/s per lane based AUI” interoperation.
  - PCS, FEC, PMA sub-block can be implemented in either switch ASIC or oDSP, etc.
- Backward compatible to 100 Gb/s per lane optical and electrical from IEEE 802.3 400GbE standard.
- Forward looking compatibility to enable:
  - 200 Gb/s per lane AUI and optical PHY
  - 800 Gb/s or 1.6 Tb/s coherent, with 200 Gb/s and 400 Gb/s per lane (modulation channel) technology
  - Enable multiple Ethernet rates or breakout, 200GbE, 400GbE, 800GbE, 1.6TbE etc.

Assumption for Logic layer Architecture



- A hybrid architecture to enable all FEC schemes based on a separated FEC sublayer within logic layer stack
- 100 Gb/s and 200 Gb/s SerDes based AUI
  - Same pre-FEC BER @ 1E-5
- FEC Scheme: End-to End, without PMA, Inverse FEC1, FEC2
- FEC Scheme: Encapsulated or Concatenated, without Inverse FEC1
- FEC Scheme: Segment, with PMA, Inverse FEC1, FEC2

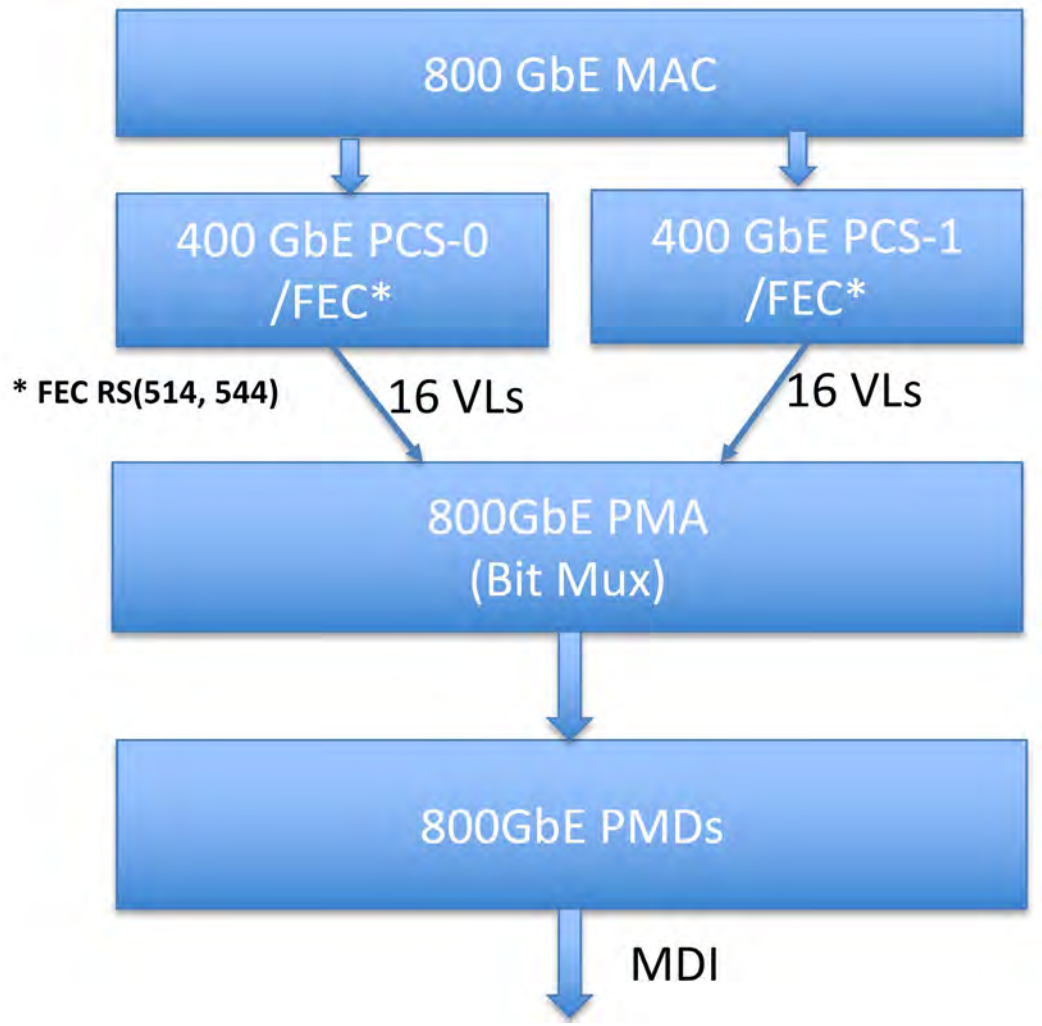
The Relationship of FEC1 and FEC2 Code

- Segmented scheme allows decoupling FEC1 and FEC2, no dependence between each other on particular FEC code selection.
  - Each FEC code have deterministic pre-FEC BER limitation for a target post-FEC BER.
- For Encapsulated or Concatenated scheme, extra consideration is needed when selecting FEC codes for FEC1 as outer code and FEC2 as inner code.
  - FEC1 can operate identically as proven technology in 802.3bs – hard decision RS(544,514).
  - FEC2 can relax raw BER requirement of the physical link, for example, for 200 Gb/s per lane IM-DD link, ~2E-3 with soft decision, ~1E-3 with hard decision as in BER Objective for Beyond 400GbE .
  - 1E-5 pre-FEC BER for 200 Gb/s per lane AUI? Then we can allocate 0.1 dBo FEC coding gain similar as in 802.3bs and some analysis example in FEC Architectures of B400GbE to Support BER Objective .
  - Interleaving FEC1 to further consider non-Poisson distributed errors from FEC2 output, with as low as possible latency.
  - For optical links, it is less challenging to have higher overhead for inner BCH/Hamming code to further lower raw BER.

Option 1 of FEC1: Ethernet Technology Consortium

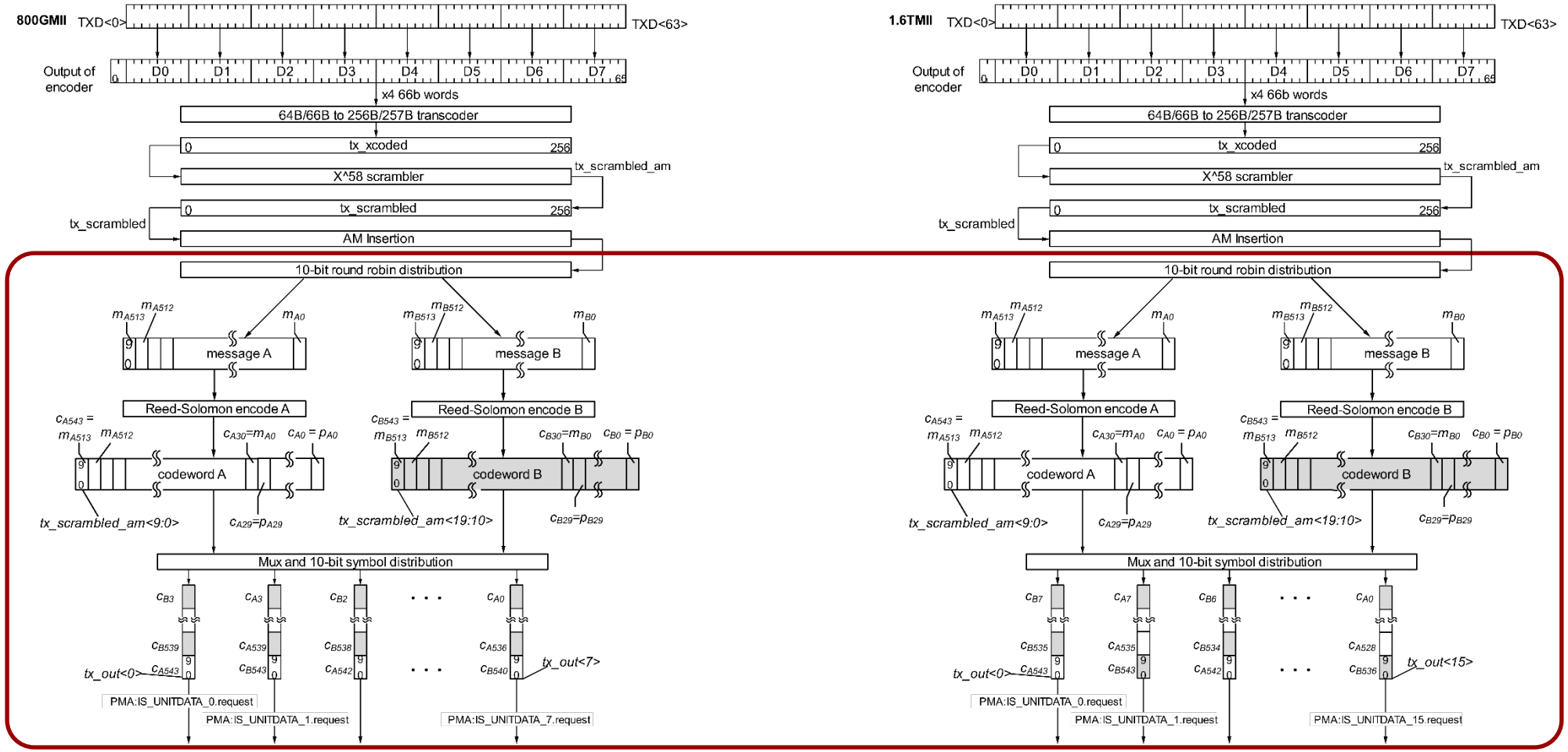
- 4 RS(544,514) FEC codeword interleaving for 800GbE, 32 FEC lanes as in Architectural Considerations and Managing PMDs Timeline.

- No proposal for 1.6TbE MAC/PCS from ETC yet. Some assumptions were discussed during Feb 15th Meeting:
  - 4X400 GbE PCS instances, PCS-0,1,2,3
  - 4X16 FEC lanes @ 25 Gb/s per lane
  - PMA: Bit Mux



Option 2 of FEC1: Clause 119 Like

- 2 RS(544,514) FEC codeword interleaving for both 800 GbE and 1.6 TbE, with 8 FEC lanes and 16 FEC lanes respectively corresponding to 100 Gb/s per lane AUI objective.



Difference between the Two Options of FEC 1

	Option 1		Option 2	
	800GbE	**1.6TbE**	**800GbE**	**1.6TbE**
Interleave	4X FEC codeword	8X FEC codeword	2X FEC codeword	2X FEC codeword
FEC Lanes	32 FEC lanes @ 25 Gb/s	64 FEC lanes @ 25 Gb/s	8 FEC lanes @ 100 Gb/s	16 FEC lanes @ 100 Gb/s
Alignment Marker	32	64	8	16
PMA	4:1 for 100 Gb/s per lane 8:1 for 200 Gb/s per lane *16:1 for 400 Gb/s per lane*	4:1 for 100 Gb/s per lane 8:1 for 200 Gb/s per lane *16:1 for 400 Gb/s per lane*	1:1 for 100 Gb/s per lane 2:1 for 200 Gb/s per lane *4:1 for 400 Gb/s per lane*	1:1 for 100 Gb/s per lane 2:1 for 200 Gb/s per lane *4:1 for 400 Gb/s per lane*

FEC Performance and Latency

- Higher burst error tolerance when interleaving more FEC codewords
  - The coding gain benefit depends on physical link specification.
  - IEEE 802.3bs chose 2-way rather than 4-way FEC codewords interleaving, further covering IEEE 802.3cu/ck .
- Option 1 has 2X/4X latency comparing to Option 2
  - For 800 GbE: 25.6ns versus 12.8ns for block time of RS(544,514) decode
  - For 1.6 TbE: 25.6ns versus 6.4ns for block time of RS(544,514) decode

FEC Lane and Alignment Maker: Implementation Complexity

- For PCS/PMA architecture, “support for any logical lane on any physical lane” is preferred, and reorder is needed to reconstruct the original stream of the interleaved FEC codewords.
- Option 1: 64:64 or 32:32 cross-bar reorder circuit for 800 GbE and1.6 TbE respectively.
- Option 2: 16:16 or 8:8 cross-bar reorder circuit for 800 GbE and 1.6TbE respectively.
- Option 1 requires more than 16 alignment makers provided by 802.3bs.

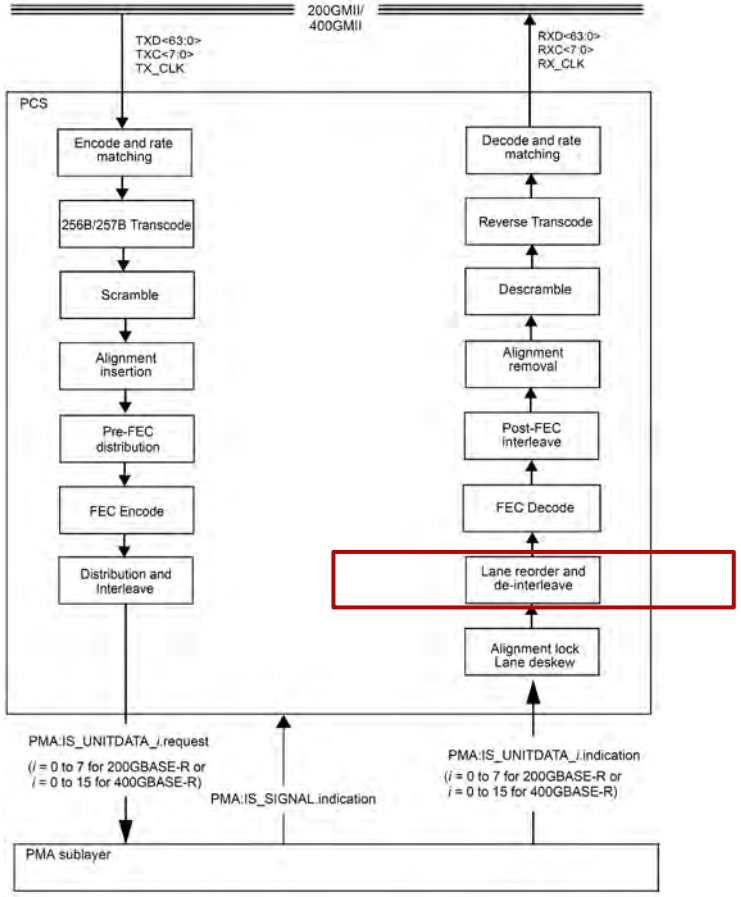


Figure 119–2—Functional block diagram

Clock Content Challenge of Bit Mux Issue

- During IEEE 802.3bs development, low transition density issue was discovered when multiplexing multiple FEC lanes into one higher rate lane with specific skews, as noted in Clause 120.5.3 for 4:1 muxing:
  - Note that since the number of input lanes and output lanes for a 200GBASE-R or 400GBASE-R PMA is always a power of two, many PMAs converting between different numbers of lanes normally simply multiplex two or four input lanes to one output lane, or demultiplex two or four output lanes from one input lane. However, any PMA implementation which produces an allowable order of bits from all PCSLs on the output lanes is valid.

    NOTE—PMA output lanes composed of some specific combinations of four PCSLs with specific skew offsets (e.g., 400GBASE-R PCSLs 0, 2, 4, and 10 with delays 0, 1, 0, and 2 bits, respectively) may have reduced transition density.
- Supporting 200 Gb/s or *400 Gb/s* per lane in Option 1 will require 8:1 or *16:1* respectively.
- Supporting 200 Gb/s or *400 Gb/s* per lane in Option 2 will require 2:1 or *4:1* respectively.
- **Further analysis on clock content issue is required if Option 1 is selected.**

Summary: FEC1 proposal

- **For backward compatibility, lower risk, and technology reuse perspective,**
  - Work towards a baseline based on Option 2 as in slide #6, same architecture as in IEEE 802.3bs with data rate quadrupled from 200 Gb/s and 400 Gb/s to 800 Gb/s and 1.6 Tb/s.

Thank you