PCS, FEC and PMA Sublayer Baseline Proposal

IEEE P802.3ck

May 2018 Pittsburgh

Mark Gustlin - Xilinx
Gary Nicholl – Cisco
Dave Ofelt – Juniper
Jeff Slavick – Broadcom

Supporters

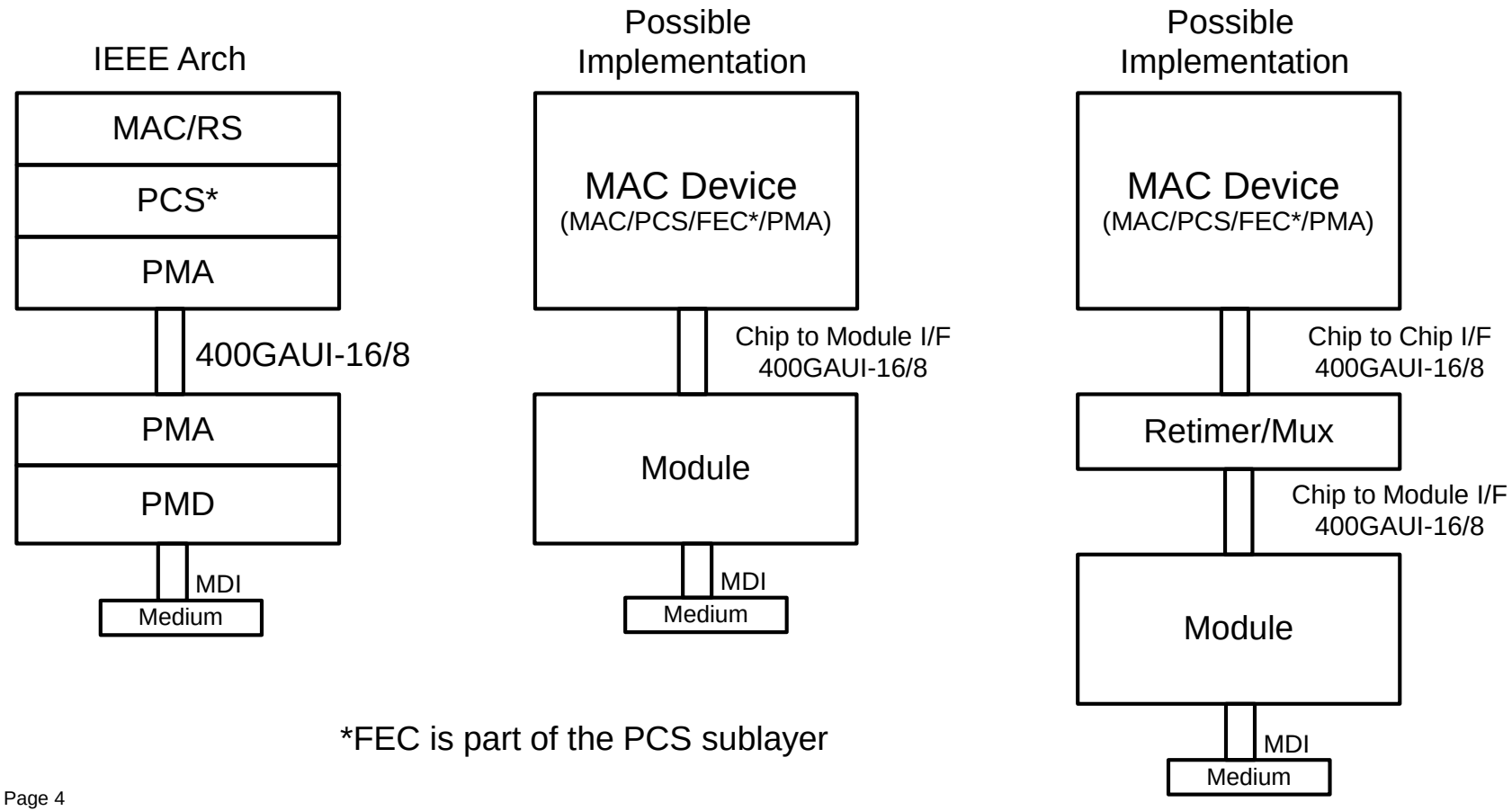
- Rob Stone – Broadcom
- Eric Baden – Broadcom
- Paul Brooks - Viavi

Introduction

- This describes a possible PCS/FEC/PMA baseline proposal for the various Ethernet speeds covered by the 802.3ck task force
- It proposes to reuse existing PCS/FEC/PMA sublayers that have been defined in 802.3bs and 802.3cd

802.3bs Architecture – 200GbE and 400GbE

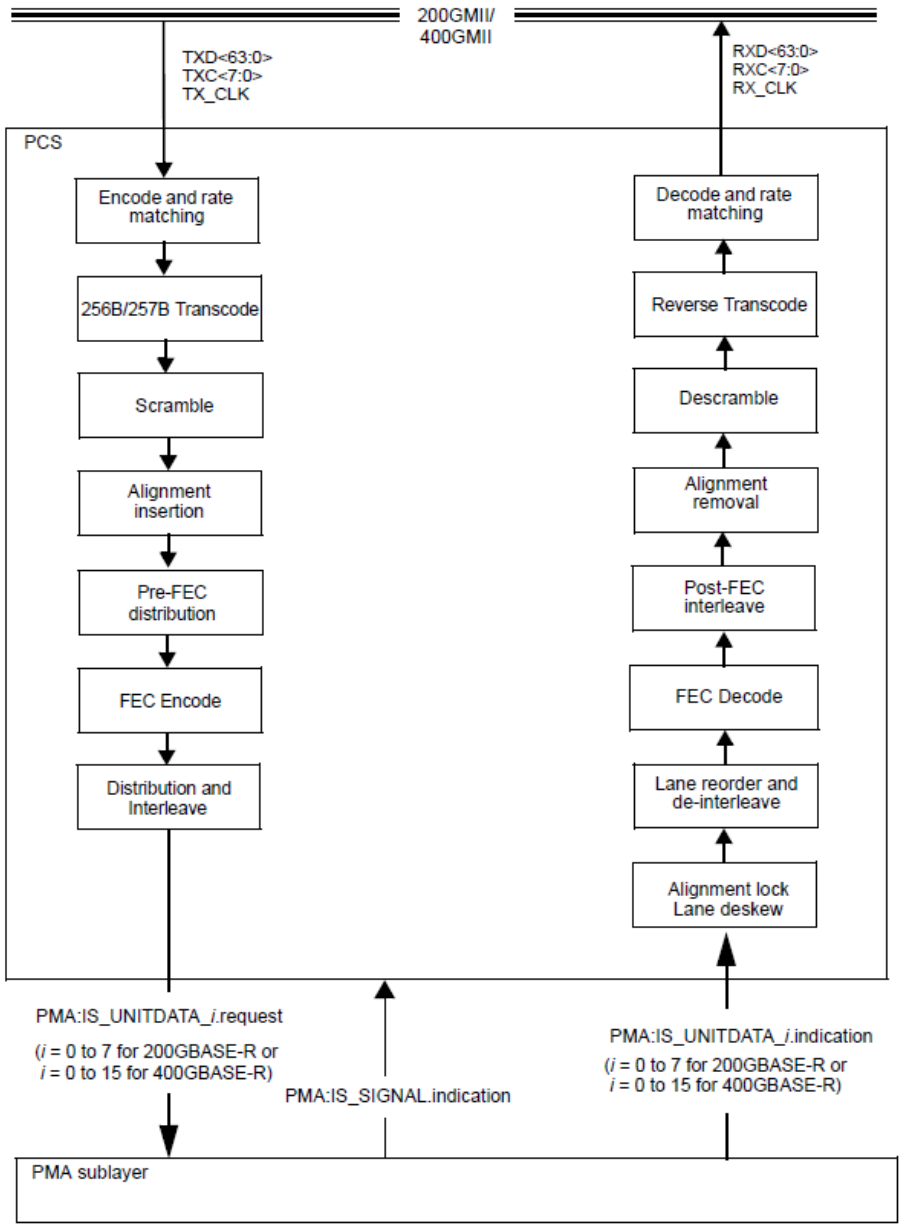
- Adopted architecture and possible implementations are shown below for 400GbE
  - 200GbE is identical except for # lanes and MAC rate
- FEC is part of the PCS sublayer utilizing the RS(544,514) aka “KP4” FEC code.
- An extender sublayer is also defined



*FEC is part of the PCS sublayer

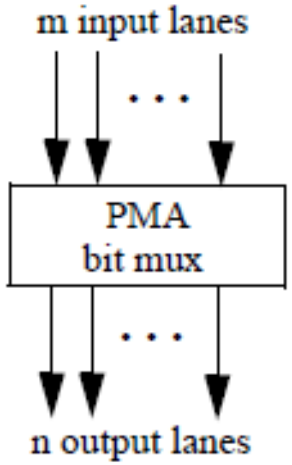
802.3bs PCS

- PCS processing flow is shown in the figure
- The PCS distributes data to 16 PCS lanes for 400GbE and 8 PCS lane for 200GbE
- Pre-FEC distribution plays the data out to two FEC codewords



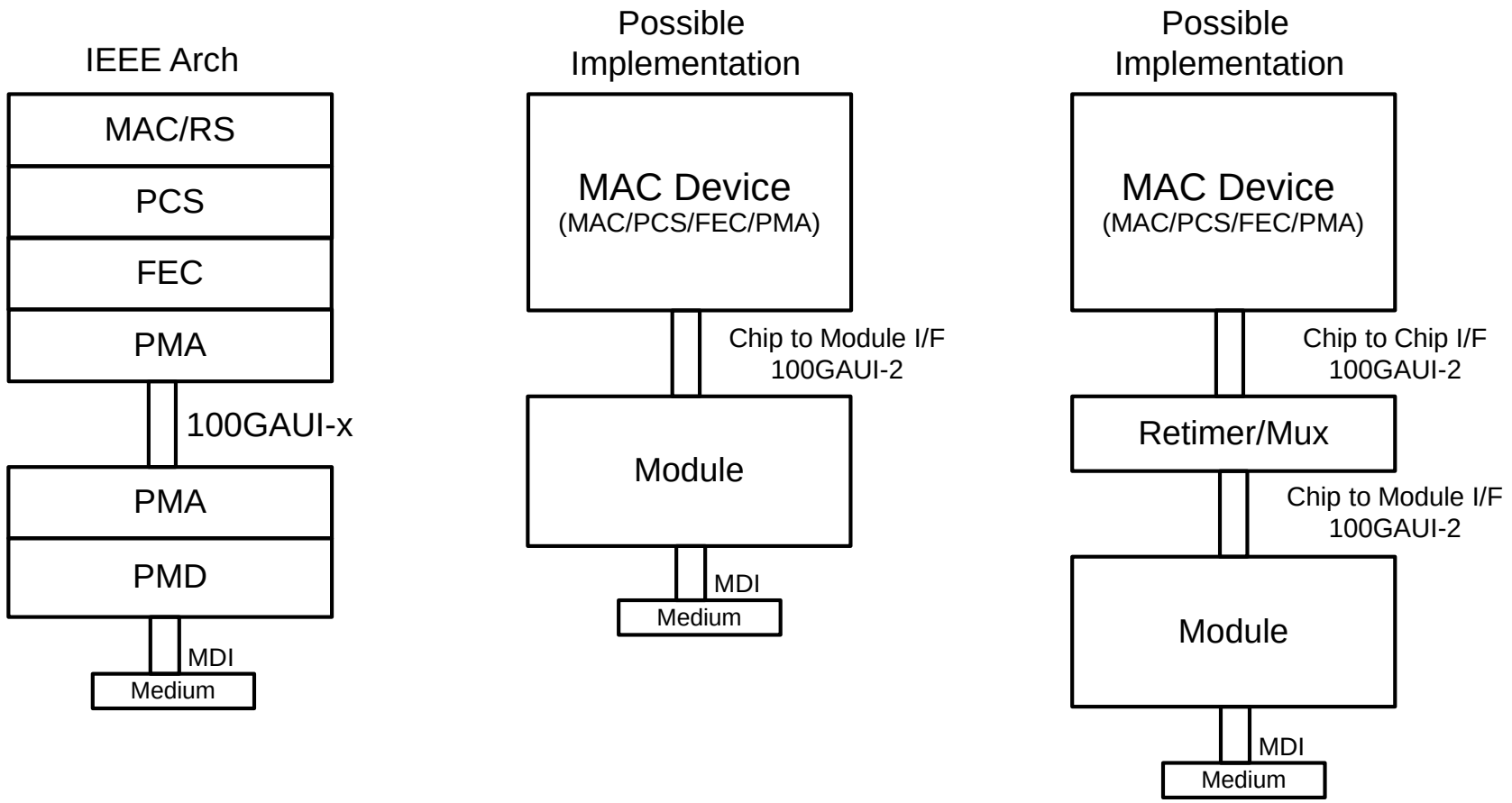
802.3bs PMA

- From a muxing point of view, the PMA is simple, m input lanes are bit muxed to n output lanes
- Bit muxing is blind, lanes can move around, the RX PCS sorts things out
- Clock content and baseline wander simulations have been performed for 1:1, 2:1 and 4:1 muxing scenarios
- Does not support precoding for 200G/400G AUIs or optical PMDs



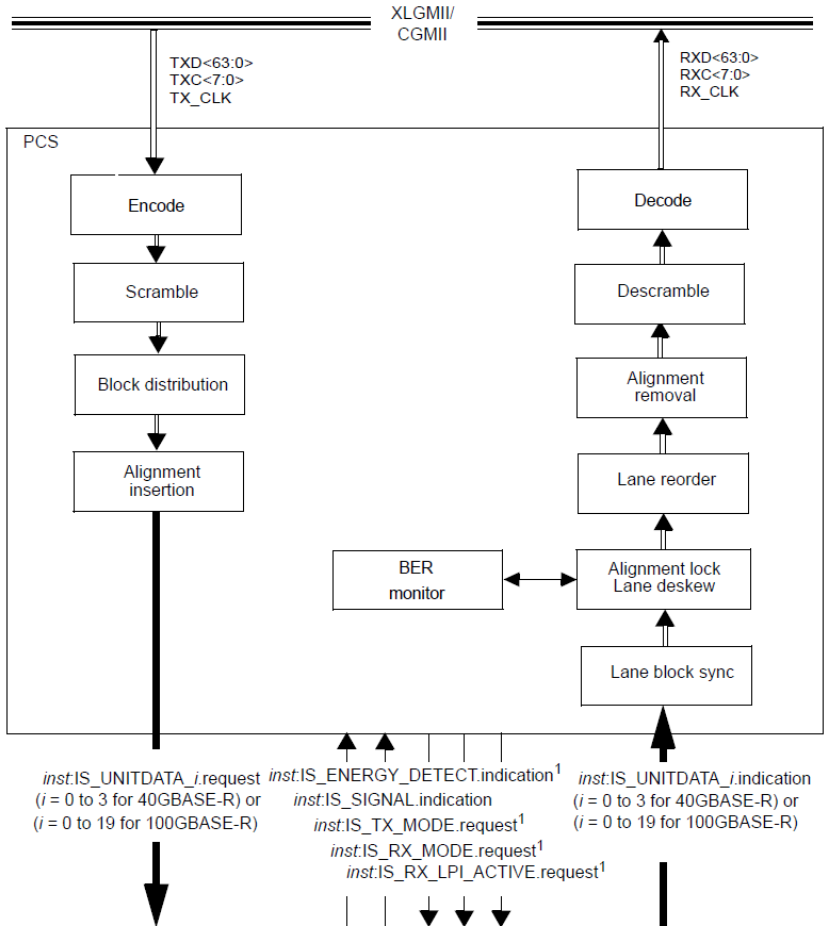
802.3cd Architecture – 100GbE

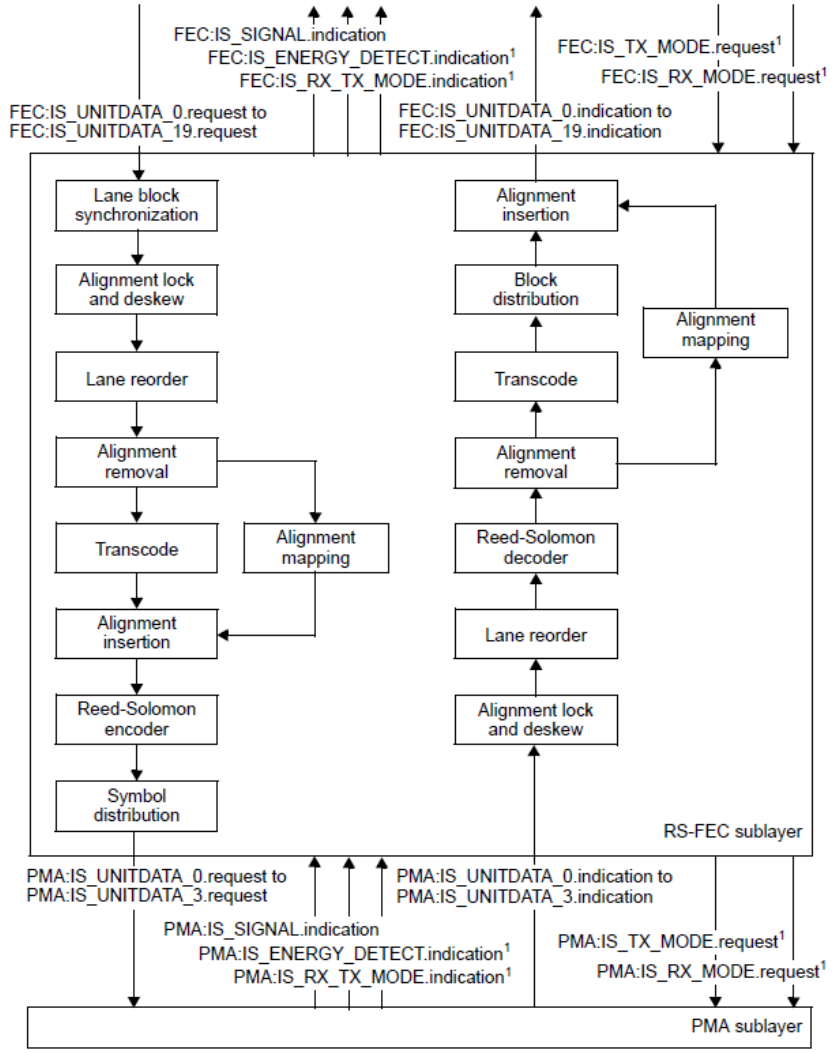
- Adopted architecture and possible implementations are shown below for 100GbE
- FEC is in the FEC sublayer, RS(544,514) aka “KP4” FEC
  - An AUI may exist between the FEC and PCS sublayers



802.3cd PCS/FEC Sublayers

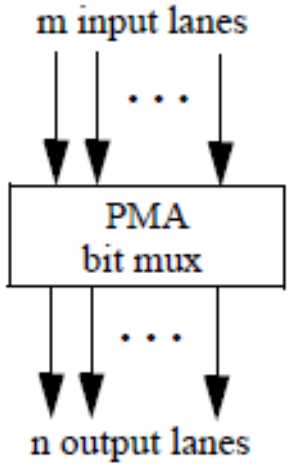
- PCS processing flow is shown in the figure to the left, FEC to the right





802.3cd PMA

- From a muxing point of view, the PMA is simple, m input lanes are bit muxed to n output lanes
- Bit muxing is blind, lanes can move around, the RX PCS sorts things out
- Clock content and baseline wander simulations have been performed for 1:1, 2:1 and 4:1 muxing scenarios
- Supports precoding to reduce the impact of burst errors
  - Except for C2M and optical PMDs
  - Must implement in TX, optional to implement in RX and the feature is optional to enable



Thoughts on the Re-Use

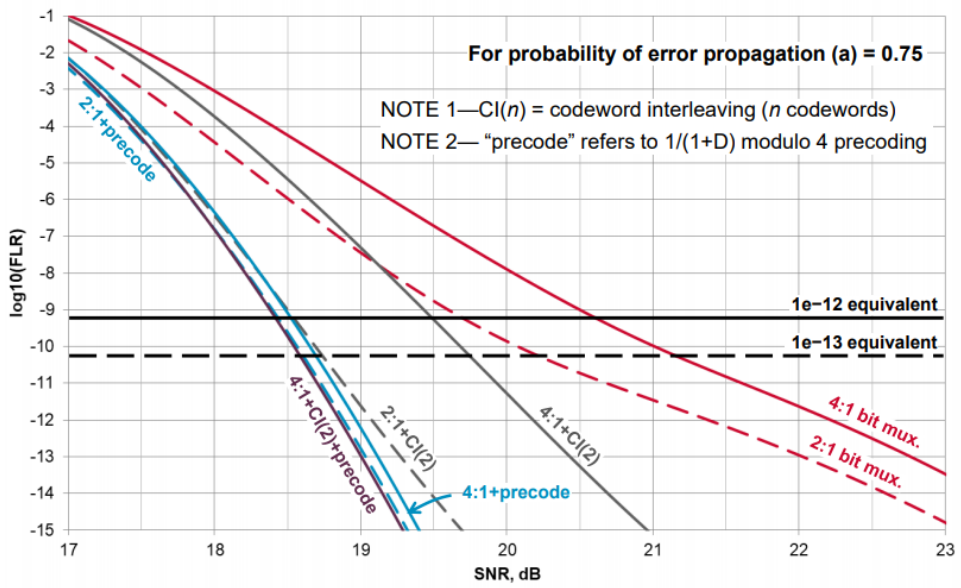
- There has been a big investment in the 802.3bs/cd architectures in the industry
- And there is a lot of desire to re-use this architecture for 100Gb/s electrical interfaces
  - They are already used for 100G per lane optical interfaces for 100GBASE-DR and 400GBASE-DR4 interfaces
  - As well as for a number of industry MSAs looking at 2km/10km 100Gb/s per lane interfaces
- There has been some work indicating that the RS(544,514) FEC might work at these electrical lane speeds
- Preserving bit muxing allows for reuse of 50Gb/s SerDes based devices with external 2:1 bit muxes when supporting new PMDs

Some Work Examples

- Adam Healey and Cathy Liu showed the following work; it indicates that utilizing Precoding in this project can help mitigate the multiplexing impacts
- Today precoding is supported as an option in 802.3cd (except for C2M and optical PMDs), but was not part of 802.3bs

SNR required for target frame loss ratio

Evaluate performance of defined error correction schemes with 4:1 bit mux.



Prob. of initial PAM-4 symbol error

$$SER_1 = \frac{3}{4}\mathrm{erfc}\left(\sqrt{\frac{SNR}{2 \times 5.4}}\right)$$ ← PAM-4 signal variance

1e−12 equivalent (100 Gb/s Ethernet)

Test case	SNR req'd, dB
4:1 bit mux.	20.6
4:1+precode	18.52
4:1+CI(2)	19.48
4:1+CI(2)+precode	18.41

1e−13 equivalent (200/400 Gb/s Ethernet)

Test case	SNR req'd, dB
4:1 bit mux.	21.13
4:1+precode	18.69
4:1+CI(2)	19.73
4:1+CI(2)+precode	18.58

5 | IEEE 802.3 100 Gb/s Electrical Lane Study Group

BROADCOM

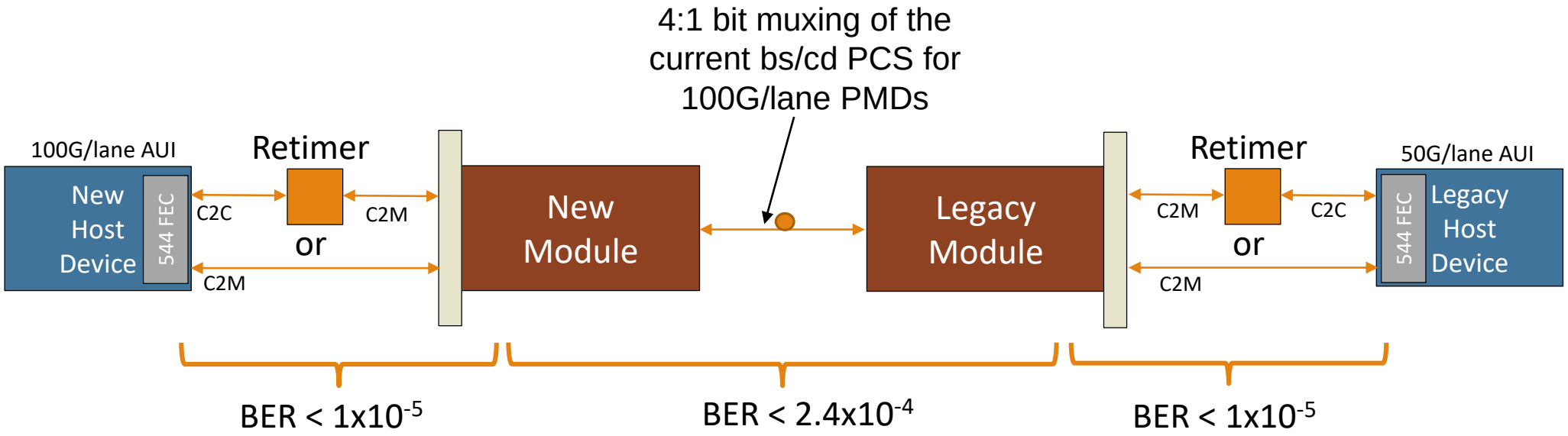
healey_100GEL_01_0318.pdf

Proposal

- Adopt the 802.3bs Extender, PCS and PMA sublayers for 200GbE and 400GbE, including the RS(544,514) FEC code for use in 802.3ck project
  - Clauses 118-120
- Adopt the 802.3cd PCS, FEC and PMA sublayers for 100GbE, including the RS(544,514) FEC code for use in 802.3ck project
  - Clauses 82, 91, 135
- Add all 802.3ck AUIs and copper/backplane PHYs to list of interfaces with precoding capability
  - As in 802.3cd Tx must support ability, Rx optional to implement, feature optional to enable

Why Supporting the Existing PCS Makes Sense

- This diagram is from ofelt_100GEL_01_0118.pdf
- 100GBASE-DR and 400GBASE-DR4 are current IEEE PMDs that use the existing PCSs
- There are a number of MSAs in flight which are extending the 100G per lane/lambda work to additional PMDs
- If we don't support the current PCSs for at least the C2M interfaces, the New Module will require more complicated circuitry to change the PCS, deskew lanes etc.
  - This will make the module more complicated, something we have always prevented by not doing things like block muxing



Conclusion

- Once sufficient simulations and other work is done to show that the RS(544,514) FEC is sufficient for the 802.3ck channels, adopt slide 12 as the baseline for the PCS/FEC/PMA sublayers
- In our opinion, we must re-use the bs/cd PCSs at least for the C2M interface due to backwards compatibility issues
- There is the most flexibility on the copper cable and backplane interfaces
  - No direct backwards compatibility issues (no defined 100G/lane cable/backplane interfaces today)
  - Though defining a different PCS/PMA for copper cable would create divergent module interfaces for copper cable vs. optical PMDs
  - Changing the PCS/PMA for these would lead to additional modes for devices

Thanks!