

Interleaved 100GbE FEC Sublayer

IEEE P802.3ck

November 2018 Bangkok

Mark Gustlin – Cisco

Pete Anslow – Ciena

Gary Nicholl - Cisco

Shawn Nicholl – Xilinx

Dave Ofelt – Juniper

Kapil Shrikhande – Innovium

Jeff Slavick - Broadcom

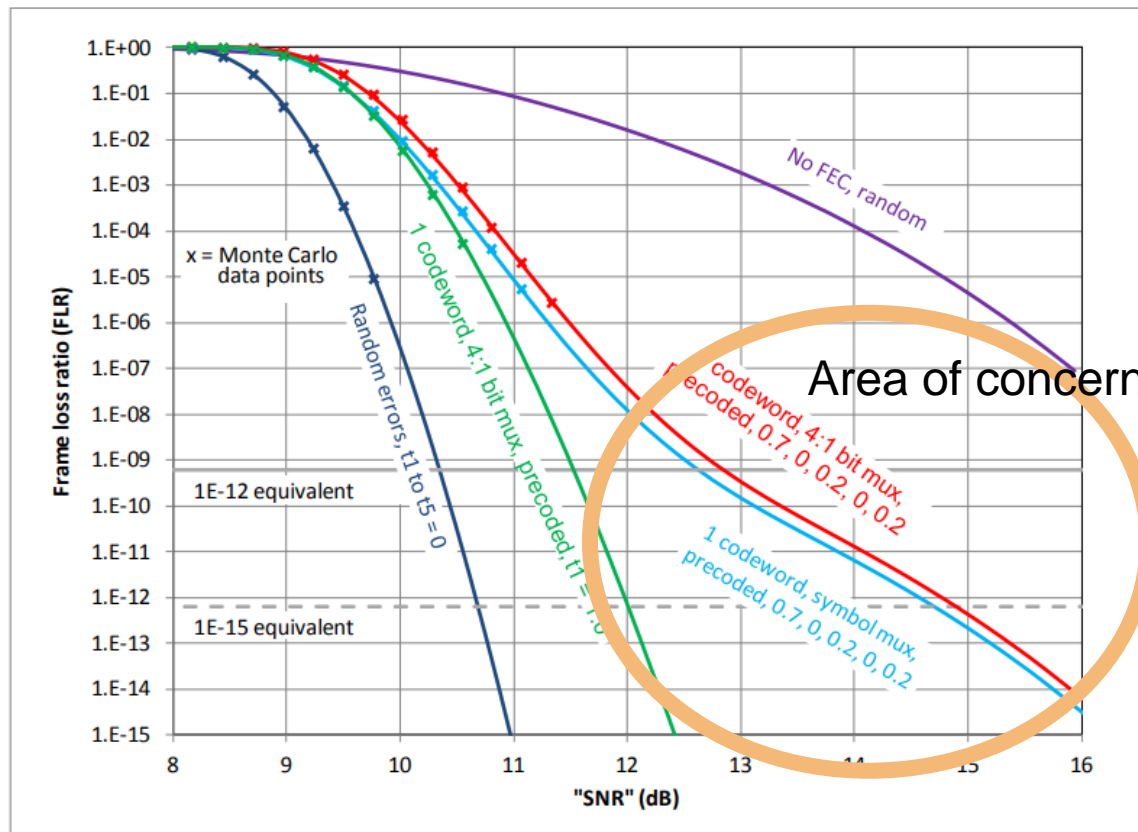
Introduction

- Pete Anslow showed in [anslow_3ck_01_0918.pdf](#) that there is some FEC performance concerns for 100GbE with multi-tap DFEs
- This presentation looks at a possible new FEC sublayer that can improve the performance for these cases

Burst Error Impact on 100G FEC Gain

- Pete Anslow showed in [anslow_3ck_01_0918.pdf](#) that there is concern with the 100G FEC performance with multi-tap DFE burst errors, even with precoding

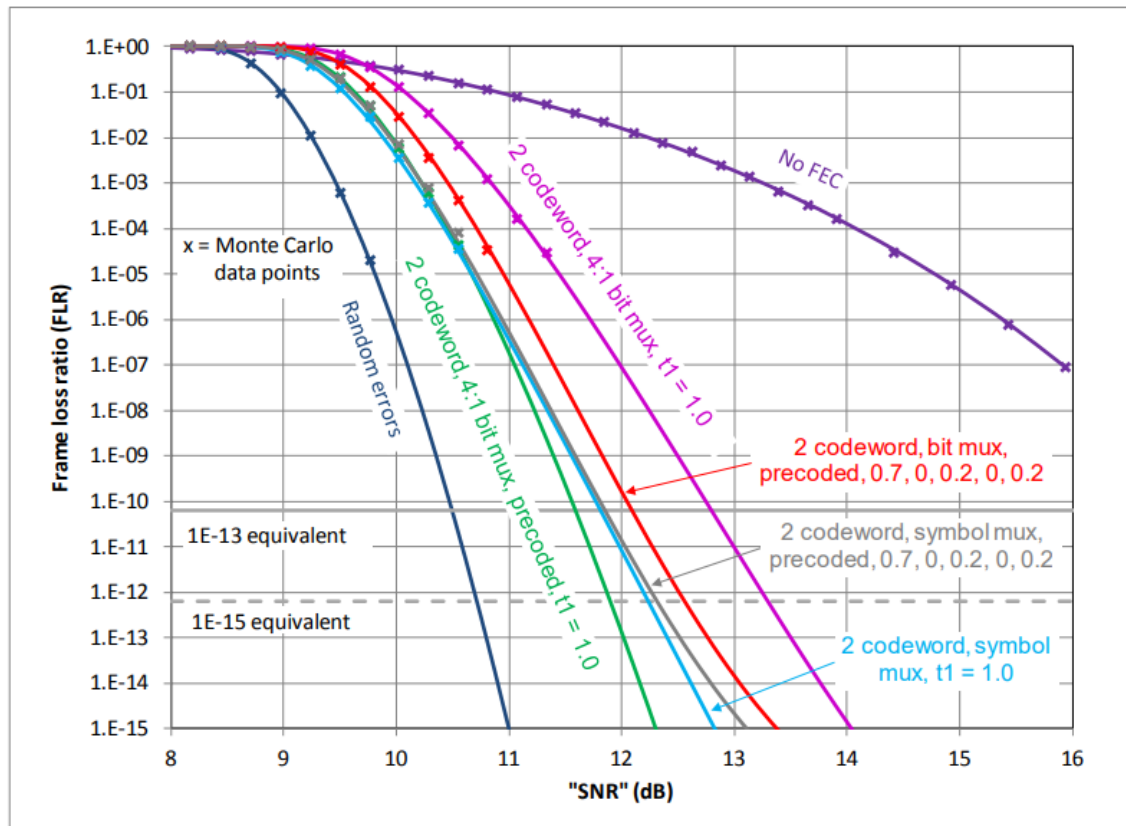
100G 5-tap DFE results (0.7, 0, 0.2, 0, 0.2) with precoding



Burst Error Impact on 400G FEC Gain

- Pete Anslow showed in [anslow_3ck_01_0918.pdf](#) that 400G does not have the same concern

400G 5-tap DFE results (0.7, 0, 0.2, 0, 0.2) with precoding



What to Do about it?

- Don't touch the 400/200G PCS/FEC
 - It doesn't have the same flaring problem
- Look at modifications to the 100G FEC architecture, but only for the longer more difficult channels
 - No changes for C2M, this preserves full link with 100GBASE-DR (and MSA optical links)
 - No changes to existing 100G per lane optical PMDs
 - Don't require PCS/FEC in optical modules
 - Look at changes for: 100GBASE-KR, 100GBASE-CR
 - C2C is a special case that needs more investigation
- Seems to be consensus that even if DFE is used for a C2M interface, the tap weights would be very low and not cause a problem with burst errors
 - Does anyone disagree with this?
 - What tap weights should we assume if we will use DFE?
- Proposed new FEC sublayer:
 - New FEC sublayer that does 2:1 FEC codeword interleaving
 - Support both 2x53G and 1x106G configurations

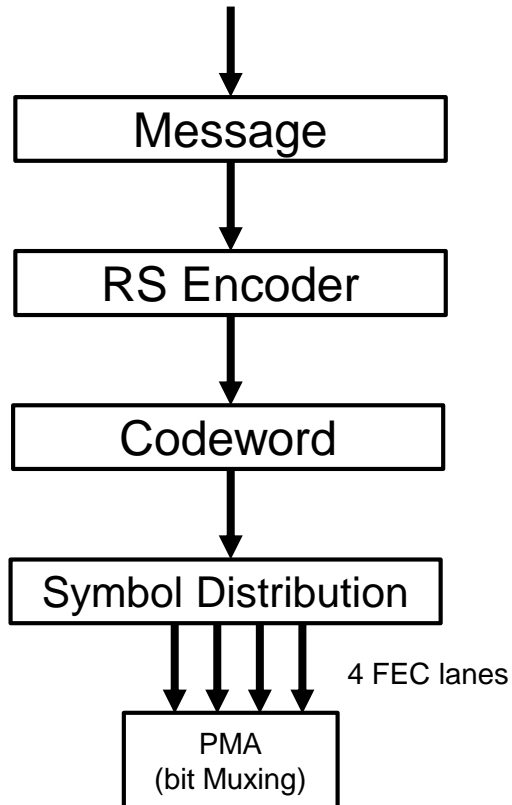
A Note About C2C Interfaces

- Still a lot of discussion around what the chip to chip interface is
- One possibility is we end up with two C2C interfaces
 - C2C-S (short) similar to what Ali Ghiasi is proposing in [ghiasi_3ck_adhoc_01_102418.pdf](#)
 - Still part of an end to end FEC budget, similar to a C2M interface
 - C2C-L (long), same/similar loss as the KR interface
 - Separate FEC domain compared to any other part of the link
 - Introduces segmented FEC
- I will use the -S and -L designations for now later in this presentation to differentiate these

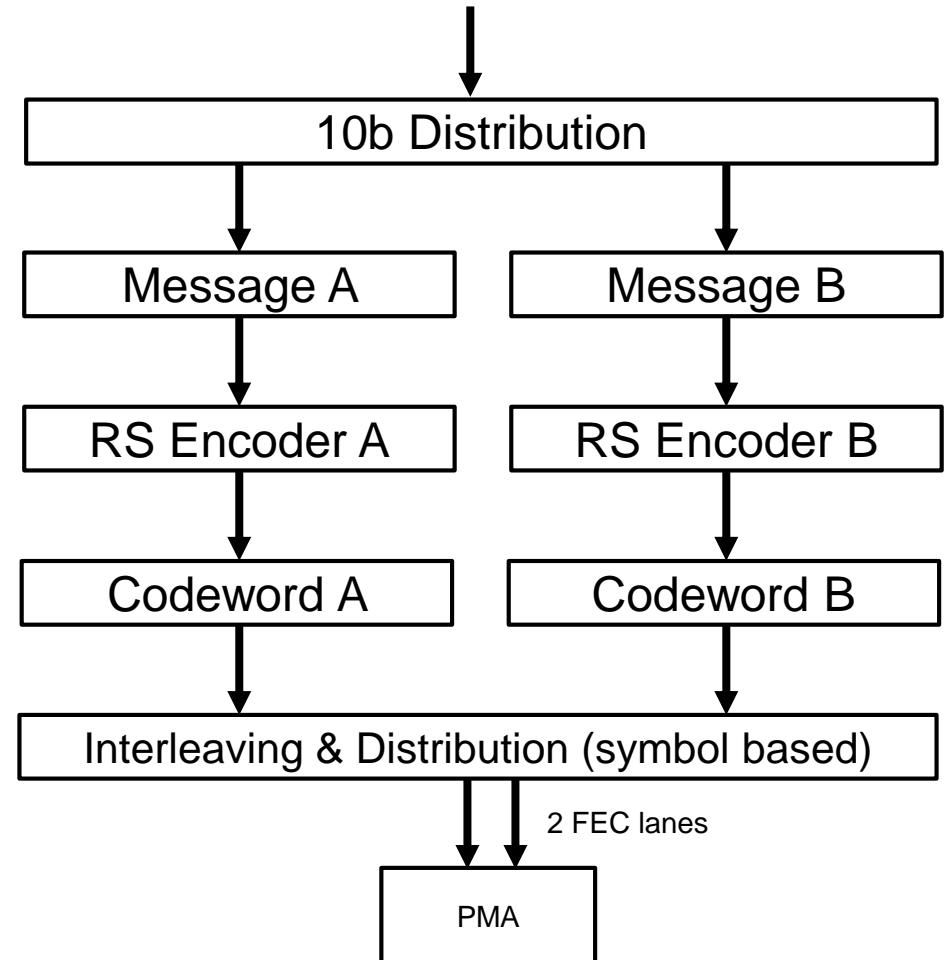
Possible New 100G FEC Sublayer

➤ Based on 2x50G RS(544,514) FEC interleaving

A portion of today's Clause 91 FEC



A portion of a possible new FEC sublayer



Latency for the 100G Interleaved FEC Sublayer

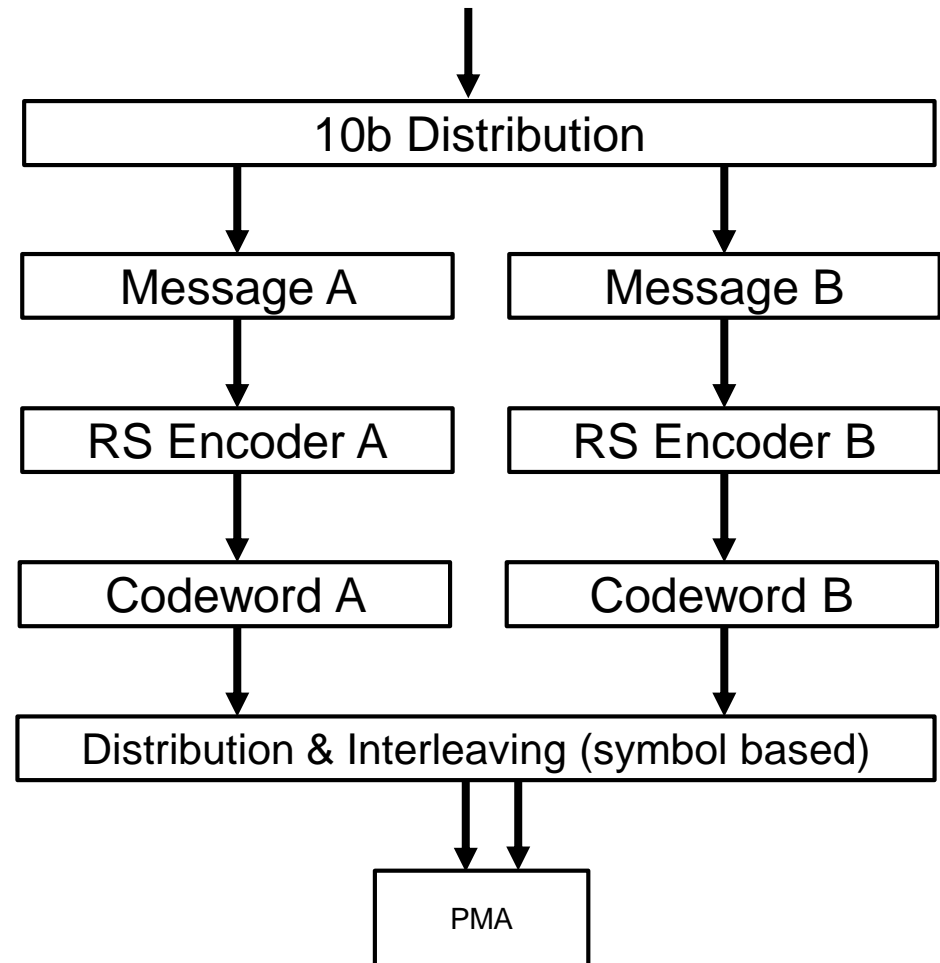
Current Clause 91 RS544

Latency	Contributor
51ns	Block time
50-100ns	Processing*
101-151ns	Total

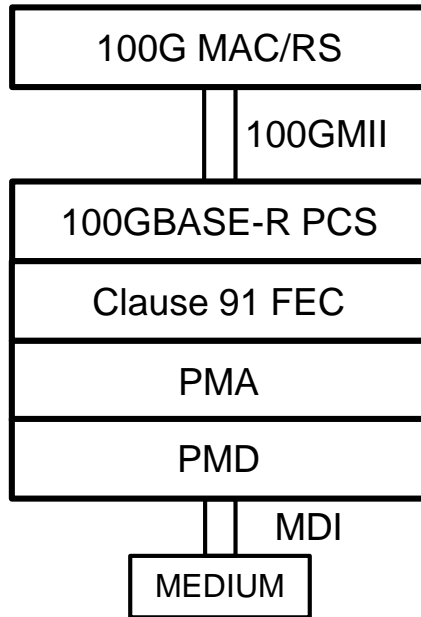
Potential RS544 Interleaved

Latency	Contributor
102ns	Block time
50-150ns	Processing*
152-252ns	Total

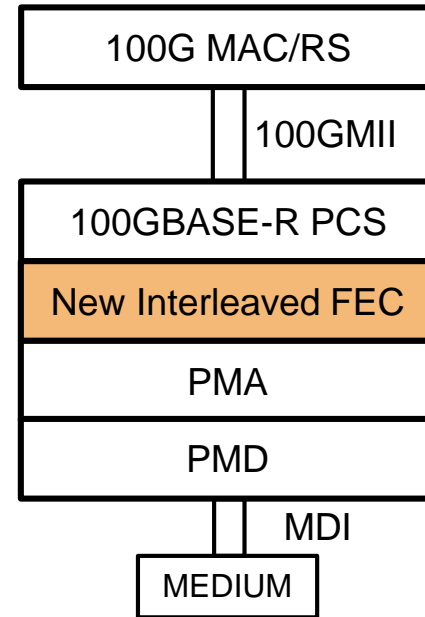
*depends on parallelism/latency tradeoffs



Architectural View

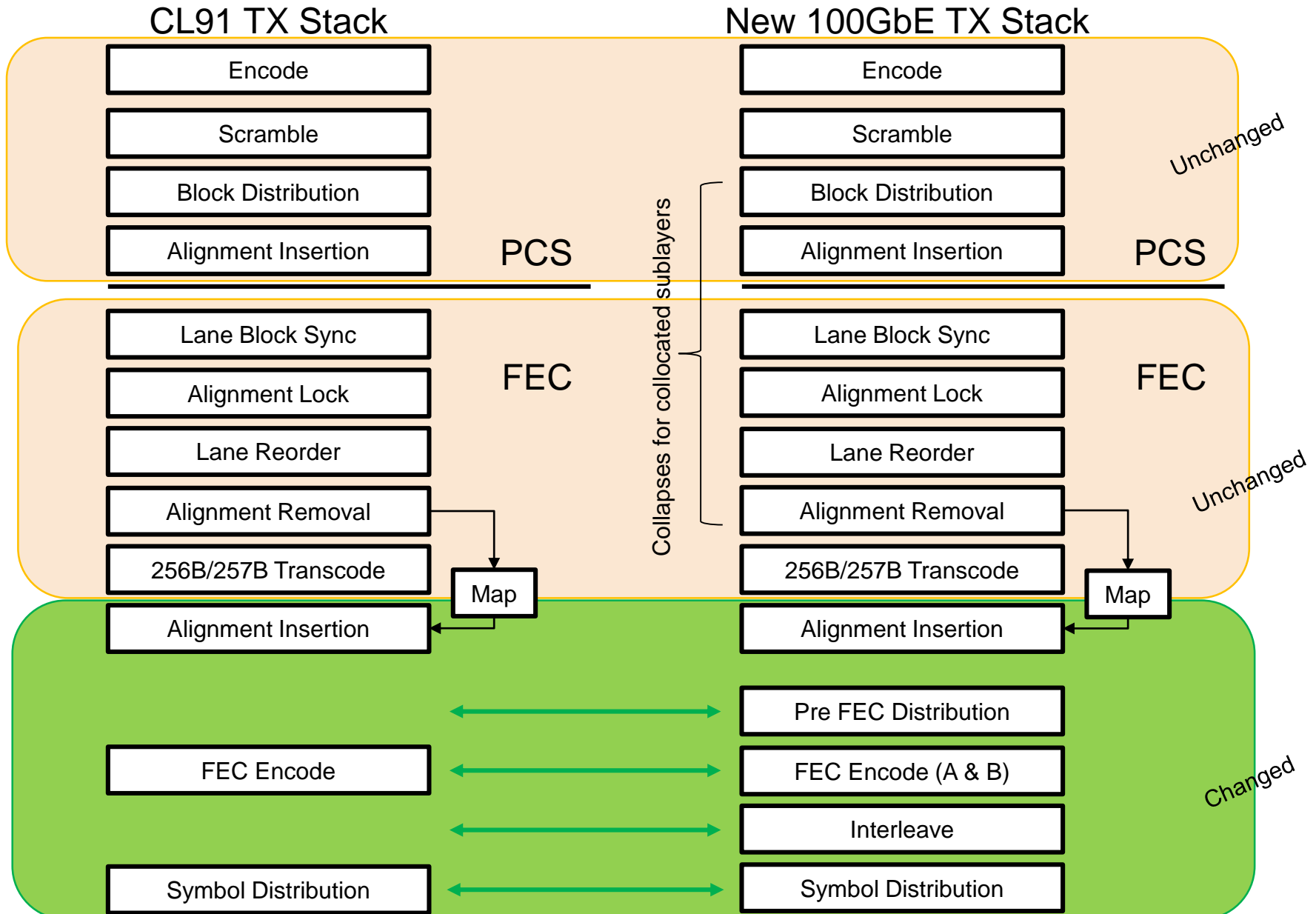


100GBASE-DR
100GAUI-1 C2M I/F

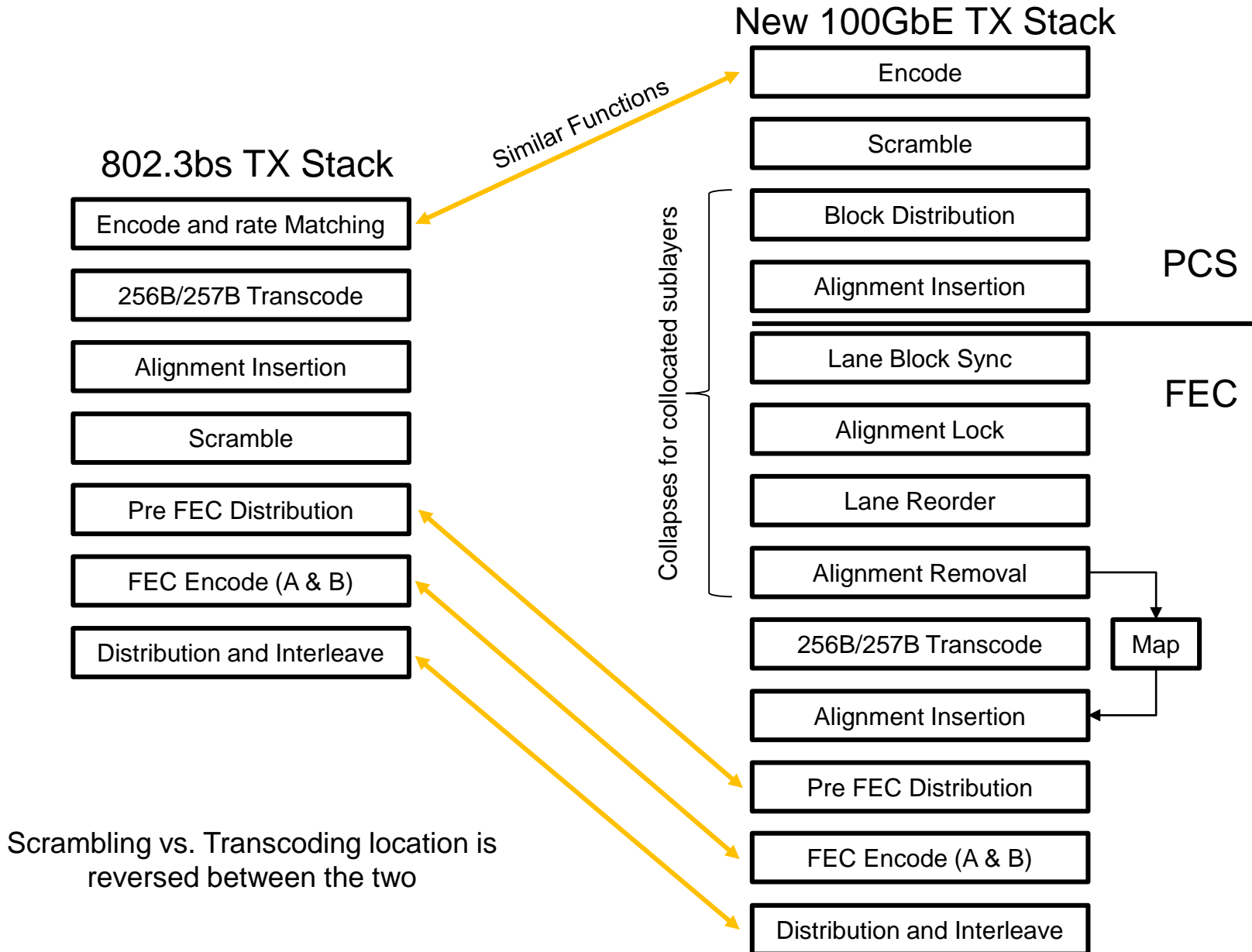


100GBASE-KR
100GBASE-CR

Interleaved Protocol Stack Comparison with CL91

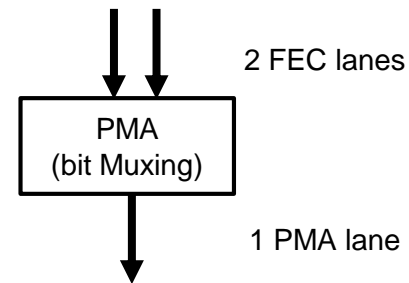
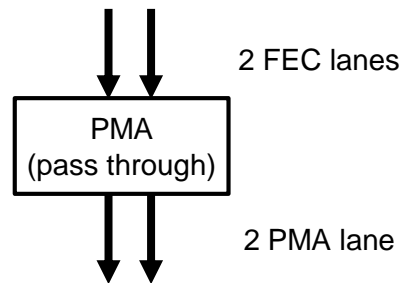


Interleaved Protocol Stack Comparison with 802.3bs



PMA for New FEC Sublayer

- PMA can be used as a pass through (2x53G)
- Or to bit mux down to a single lane (1x106G)
- Simple bit muxing



Alignment Markers

➤ Existing Clause 82/91 Marker Format

FEC lane, <i>i</i>	Reed-Solomon symbol index, <i>k</i> (10-bit symbols)																																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33		
0	0	amp_tx_0					63	0	amp_tx_4					63	0	amp_tx_8					63	0	amp_tx_12					63	0	amp_tx_16					63	
1	0	amp_tx_1					63	0	amp_tx_5					63	0	amp_tx_9					63	0	amp_tx_13					63	0	amp_tx_17					63	
2	0	amp_tx_2					63	0	amp_tx_6					63	0	amp_tx_10					63	0	amp_tx_14					63	0	amp_tx_18					63	
3	0	amp_tx_3					63	0	amp_tx_7					63	0	amp_tx_11					63	0	amp_tx_15					63	0	amp_tx_19					63	

= 5-bit pad

tx_scrambled

Figure 91-4—Alignment marker mapping to FEC lanes

➤ Proposed New Interleaved Marker Format (supporting two FEC lanes)

- amp_tx_0 = am0, amp_tx_1 = am0, amp_tx_2 = am2 etc.
- No repetition at the end of the AM block

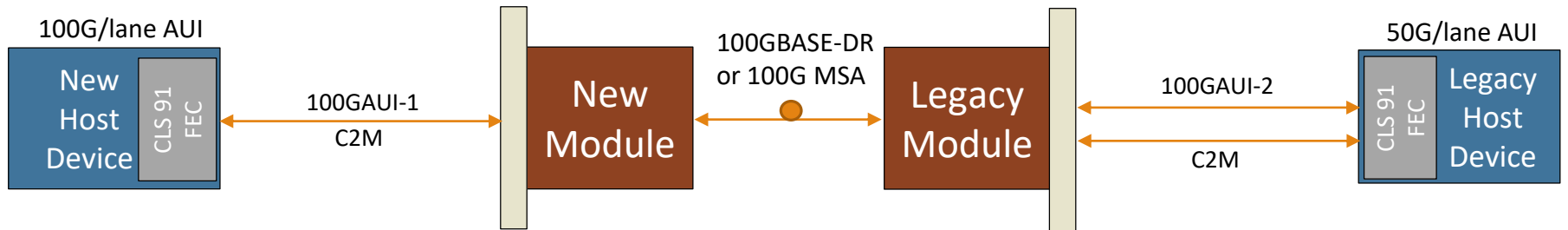
FEC Lane, i	Reed-Solomon symbol index, k (10-bit Symbols)																															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	amp_tx_0					amp_tx_2					amp_tx_4																					
1	amp_tx_1					amp_tx_3					amp_tx_5																					

000

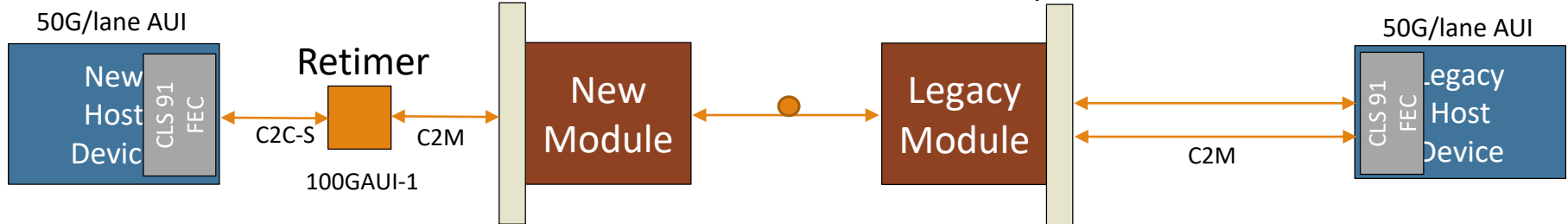
					amp_tx_16					amp_tx_18									
					amp_tx_17					amp_tx_19									

100GbE Example Use Cases - Optical

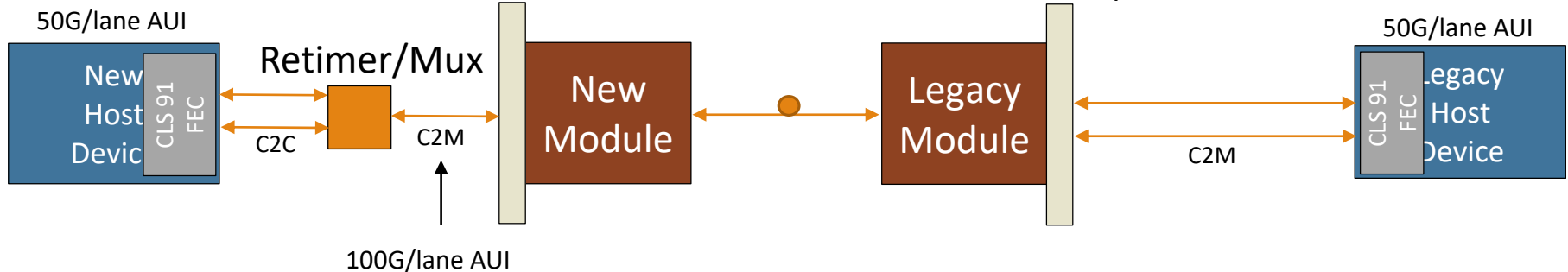
Seamless Clause 91 FEC end to end, backwards compatible



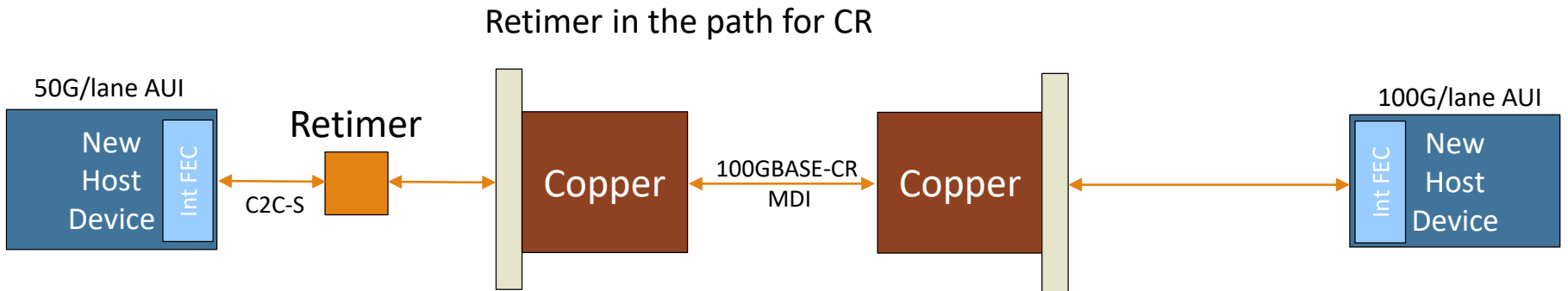
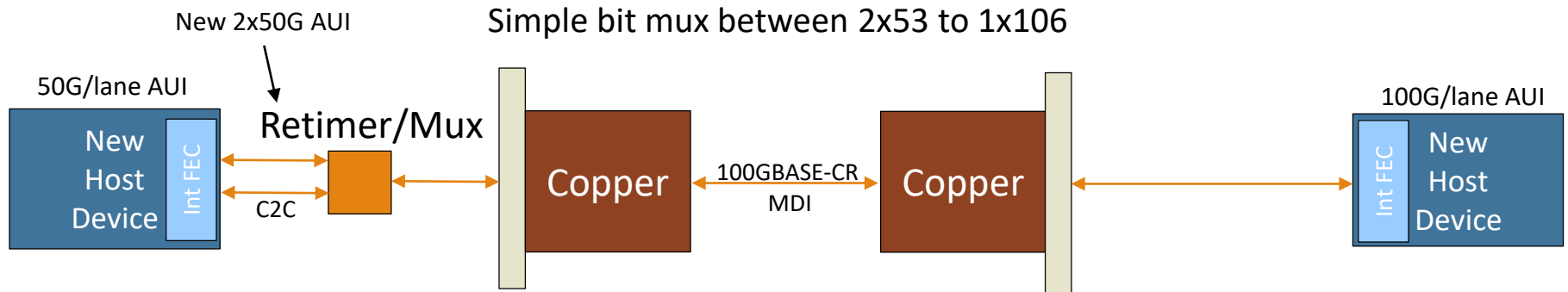
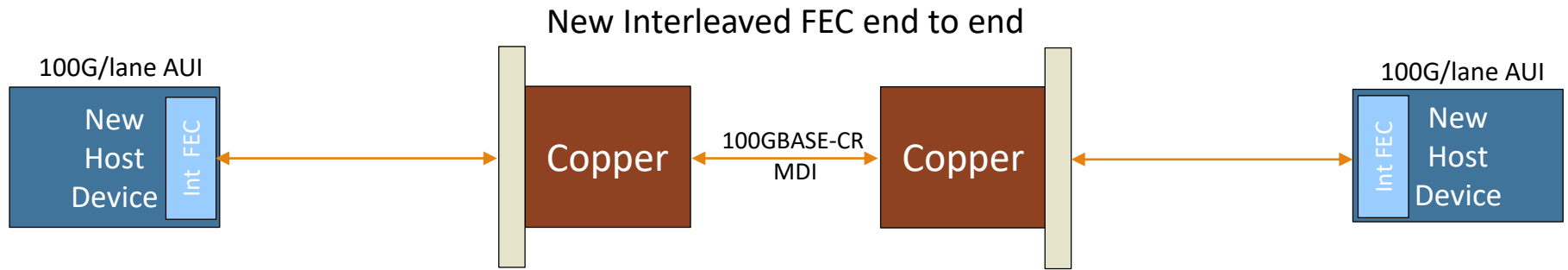
Retimer to 100G AU1-1, backwards compatible



Retimer/mux to 100G AU1-1, backwards compatible



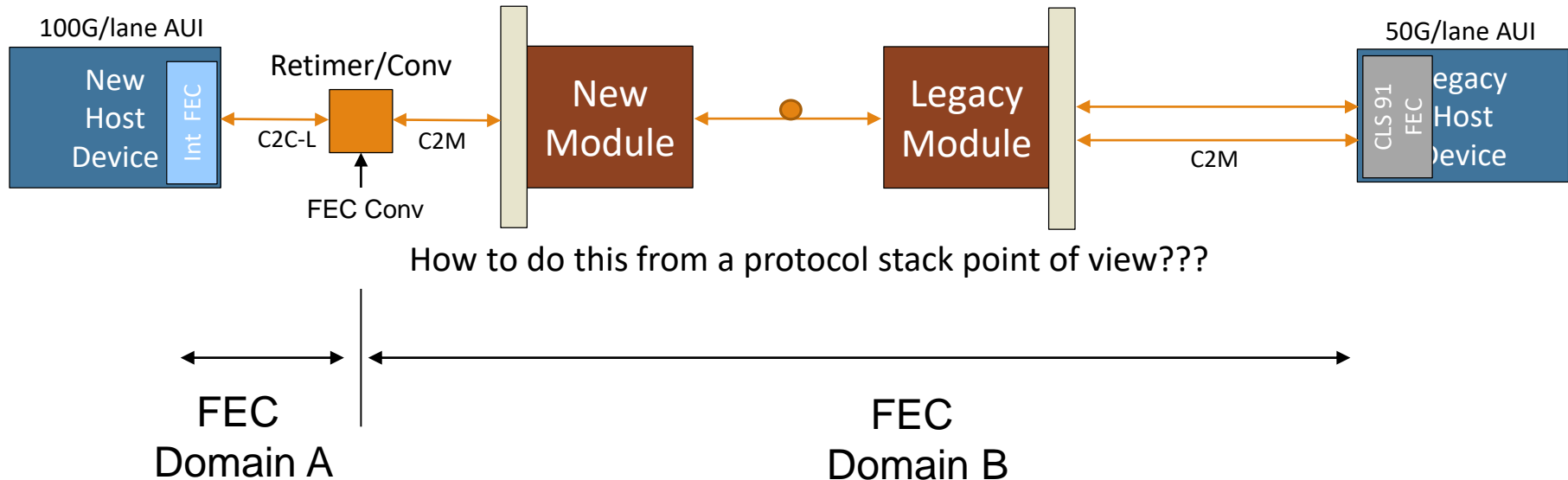
100GbE Example Use Cases – Copper



100GbE Example Use Cases – Needs More Work

- There is a separate effort to specify some sort of XS that might be usable for this scenario

“Retimer” converts from Interleaved FEC to Clause 91 FEC



More Work to be Done

- Look in detail at how to handle C2C interfaces, with bigger questions open around what does the C2C I/F look like from a loss point of view
 - Do we have two C2C interfaces, a short and a long?
 - Each with different implications on the FEC budget/partitioning?
- Agree to a model for DFE on a C2M, and run sims to see the performance
 - If DFE will be used for C2M

Conclusion

- This presentation shows a possible interleaved FEC sublayer for the harder 100GbE single lane channels

Thanks!