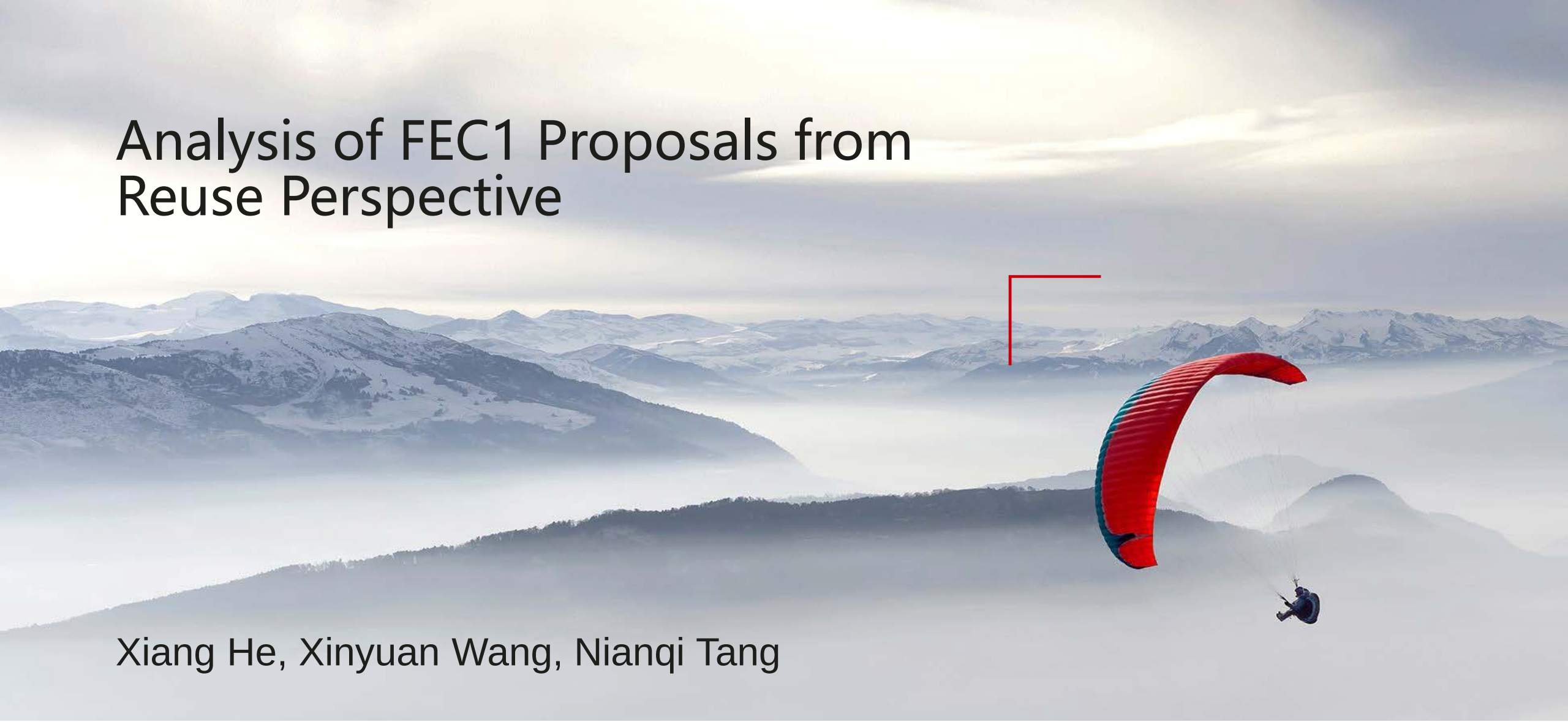
Analysis of FEC1 Proposals from Reuse Perspective



Xiang He, Xinyuan Wang, Nianqi Tang

Huawei Technologies, Co. Ltd.

HUAWEI

Candidate Options of FEC1

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)

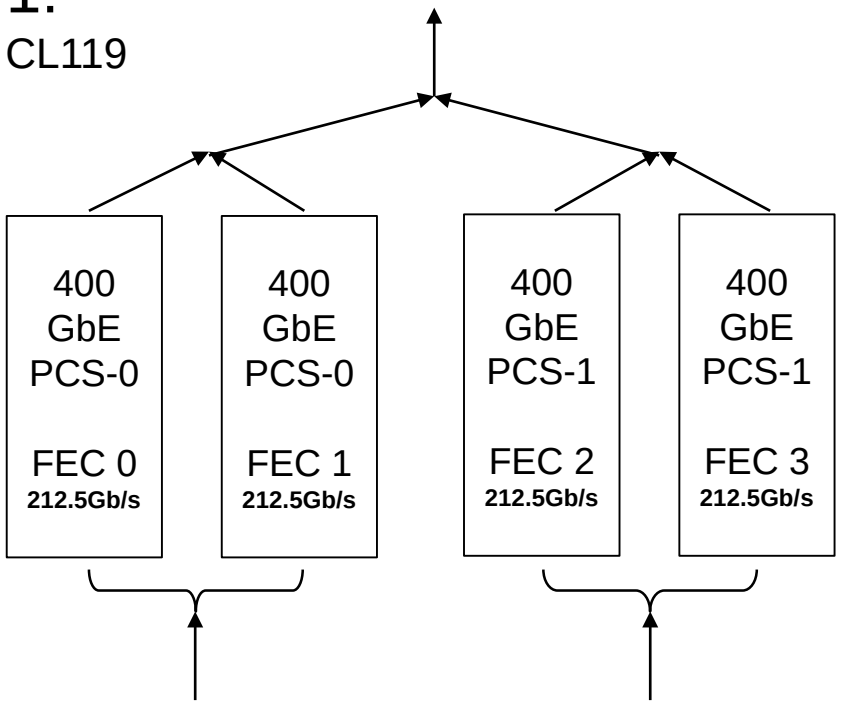
- Options:
  1. Reuse the Ethernet Technology Consortium spec
     - 66b interleaving into transcoding, 4 FEC codeword interleaving, 32 FEC lanes
     - Industry reuse
     - Fast time to market for 800GbE with 100G lanes
  2. Clause 119 like FEC
     - 2 FEC codeword interleaving
     - Shown in wang_b400g_01_210208.pdf

- 106.25G per lane for AUI-8

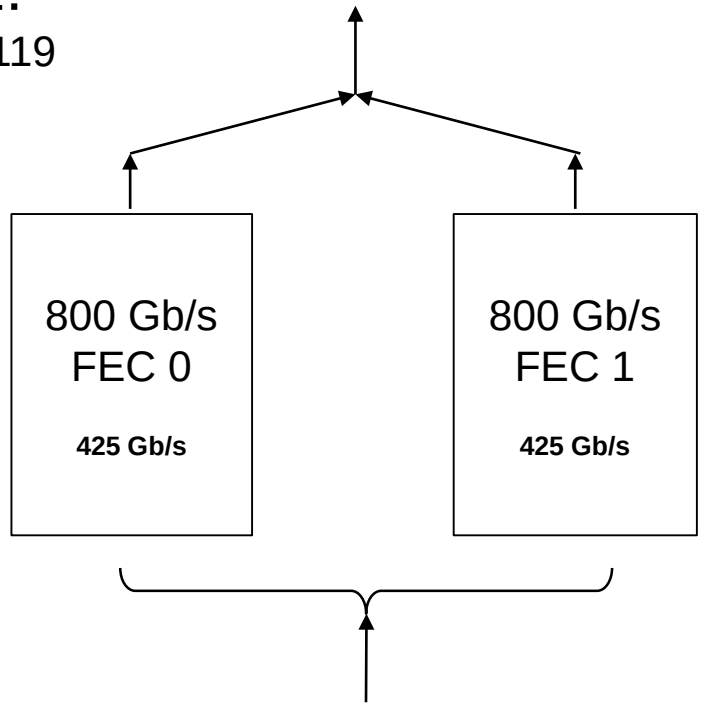
Refer to
gustlin_3df_01_220118

Proposed FEC1 Options: “2x Parallel CL119” vs “Sped Up CL 119”

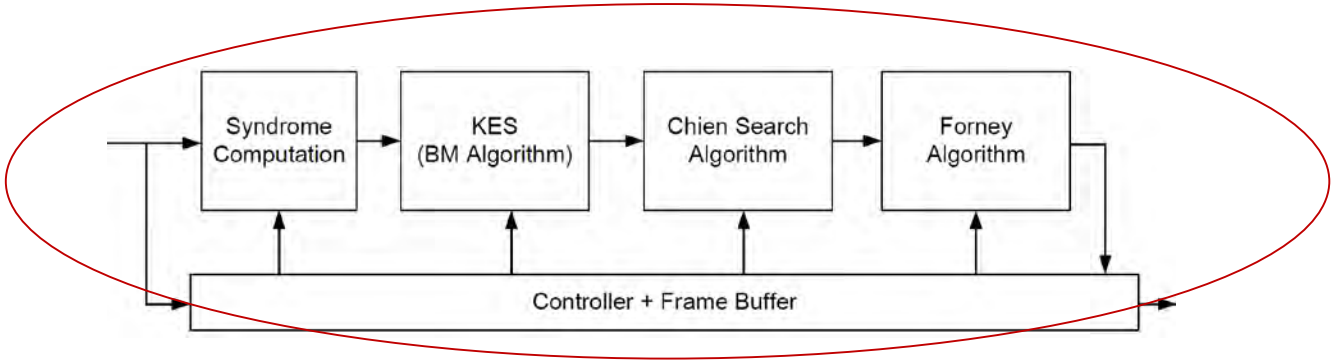
Option 1:
2x parallel CL119



Option 2:
Sped up CL119



FEC 0/1/2/3 Decoder Function Blocks

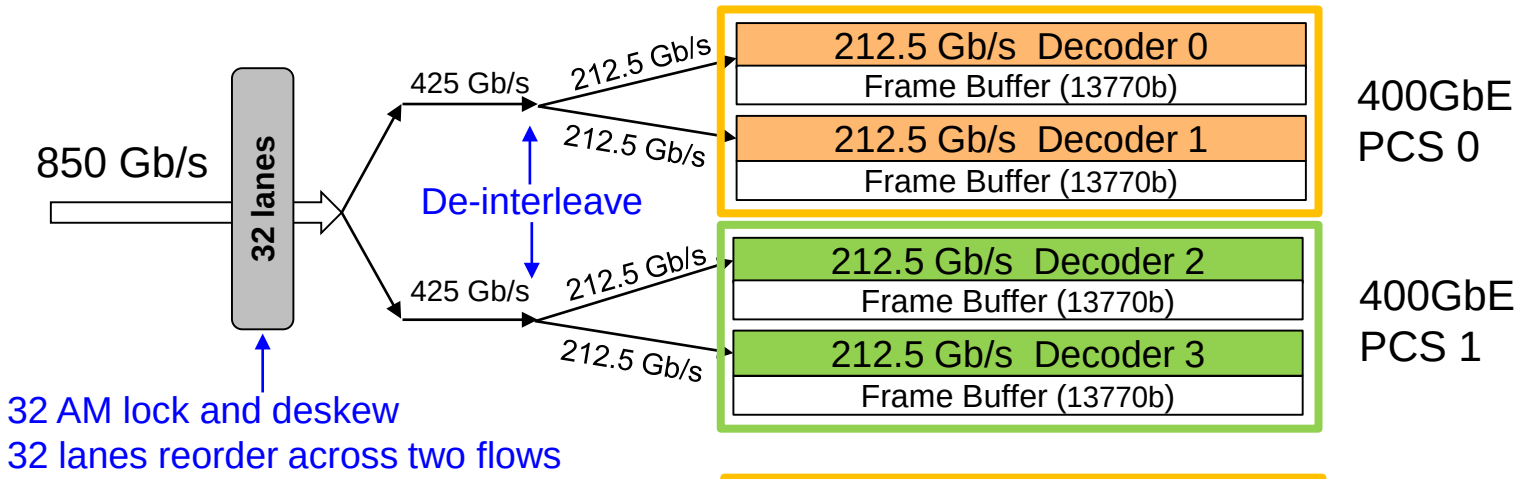


Expectations for Reuse of 802.3bs 200/400GbE

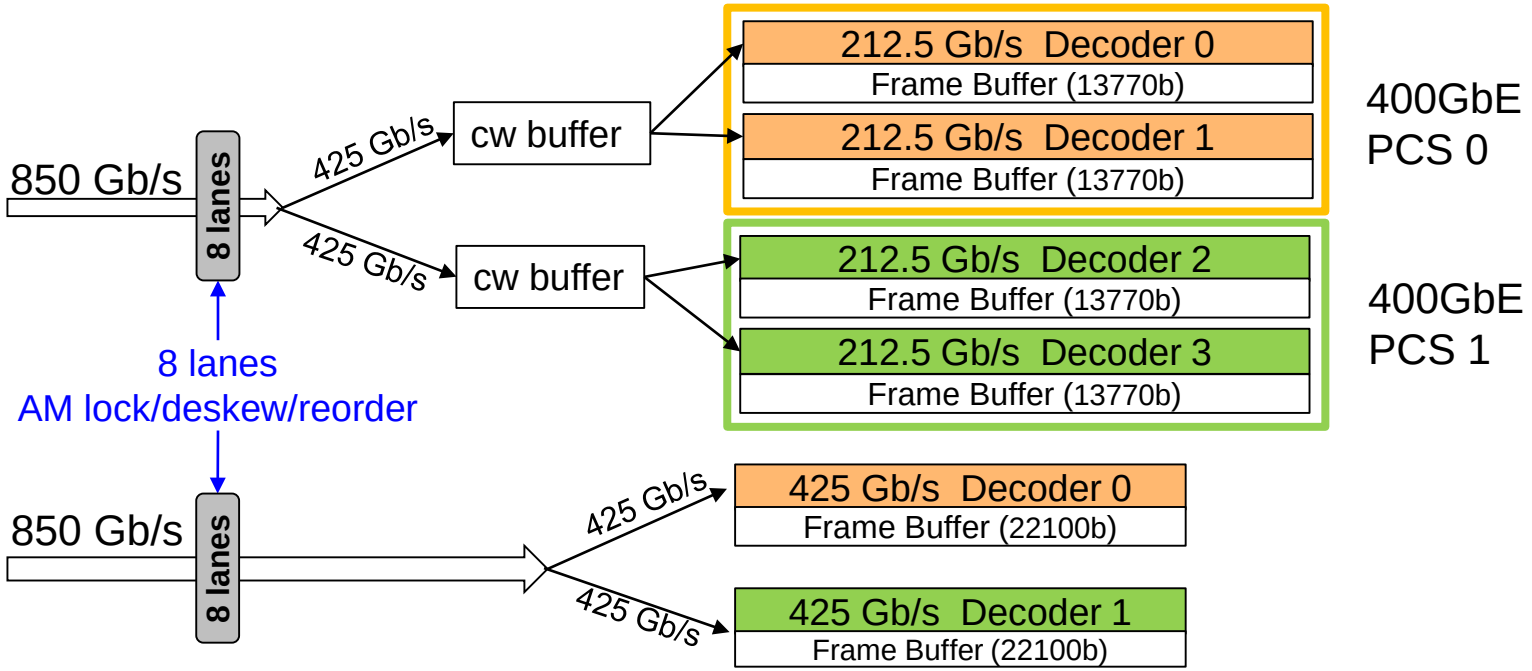
- From implementation perspective, leverage existing 400 GbE FEC sub-blocks.
  - 2 FEC decoders per 400 GbE FEC, each operating at 212.5 Gb/s throughput.
- From standard specification perspective, leverage Clause 119 of IEEE 802.3.
  - PCS encoded data distribution and recovery to decoder.
  - FEC lanes distribution, alignment and reorder mechanism.

Decoder Implementation Reuse Options

Option 1:
2x parallel CL119

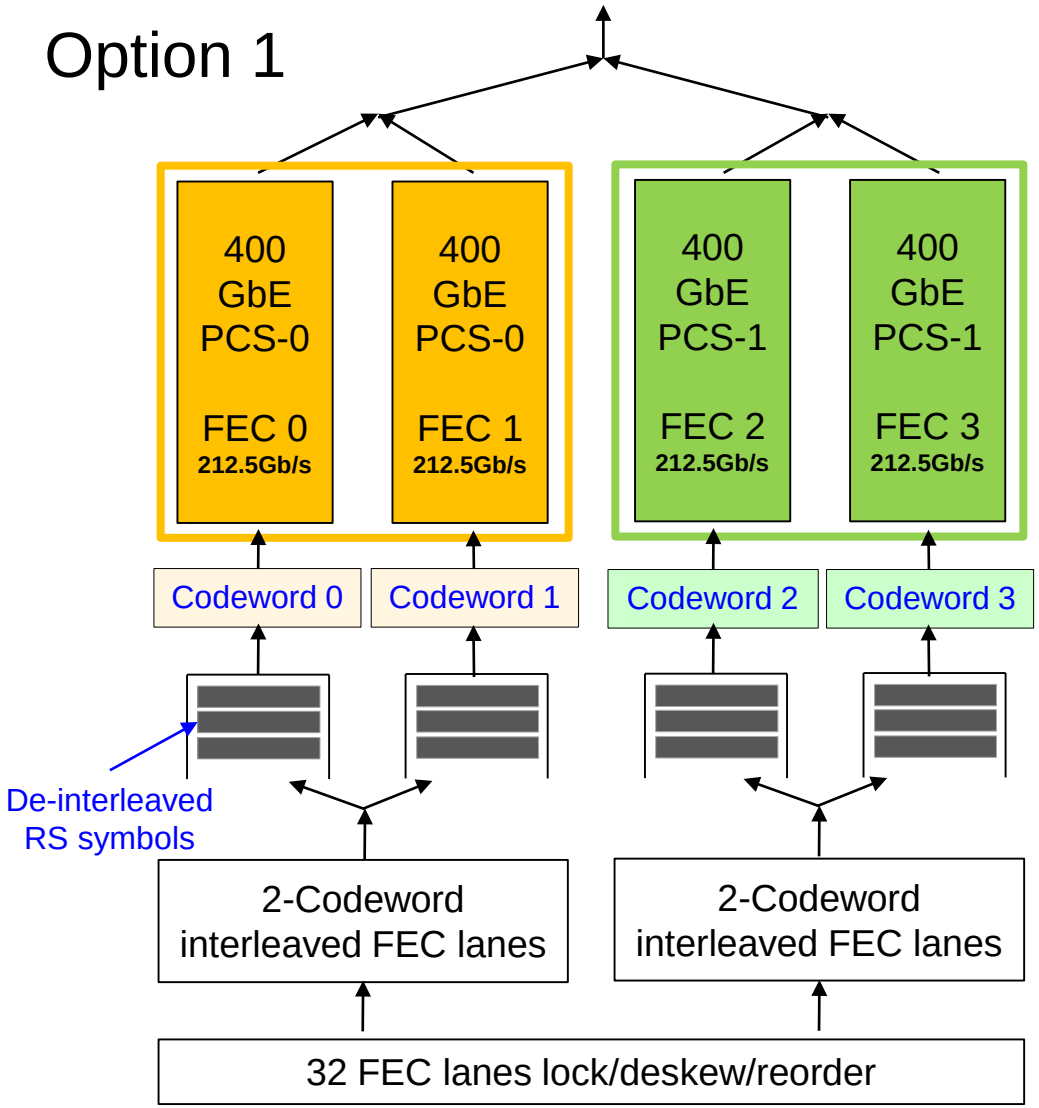


Option 2:
2 flavors of “sped up CL119” implementations



*Frame buffer sizes are evaluated based on the same Chien Search speed and same clock rate (1.25GHz) for all designs.
Actual frame buffer size for reused 212.5Gb/s decoders are generally larger due to slower Chien Search and lower clock speed.

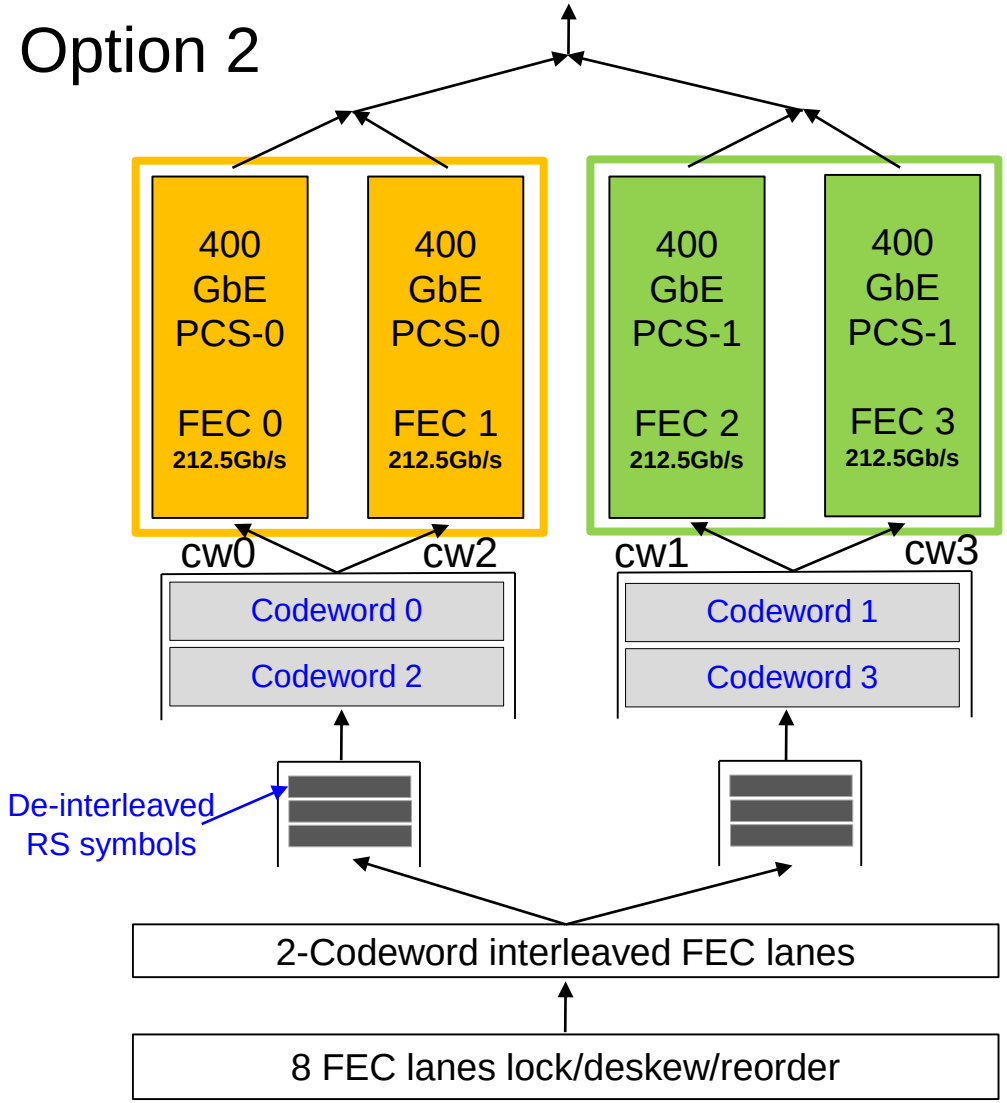
Reuse 400 GbE FEC Decoder Implementation in Option 1



- FEC Decoder 0,1, 2,3 all operate at 212.5 Gb/s throughput.
  - Friendly to implementations with lower clock rates.
- Identical implementation conforming to IEEE 802.3bs 400 GbE FEC specification above PMA.
  - FEC decoders are the most complex part, taking ~40% of Rx PCS design.
- Modification:
  - Lane reorder is performed across overall 32 lanes instead of two groups of 16 lanes.
  - Interleaving is performed over 425 Gb/s throughput, so frame time is 25.6 ns.

Reuse 400 GbE FEC Decoder Implementation in Option 2

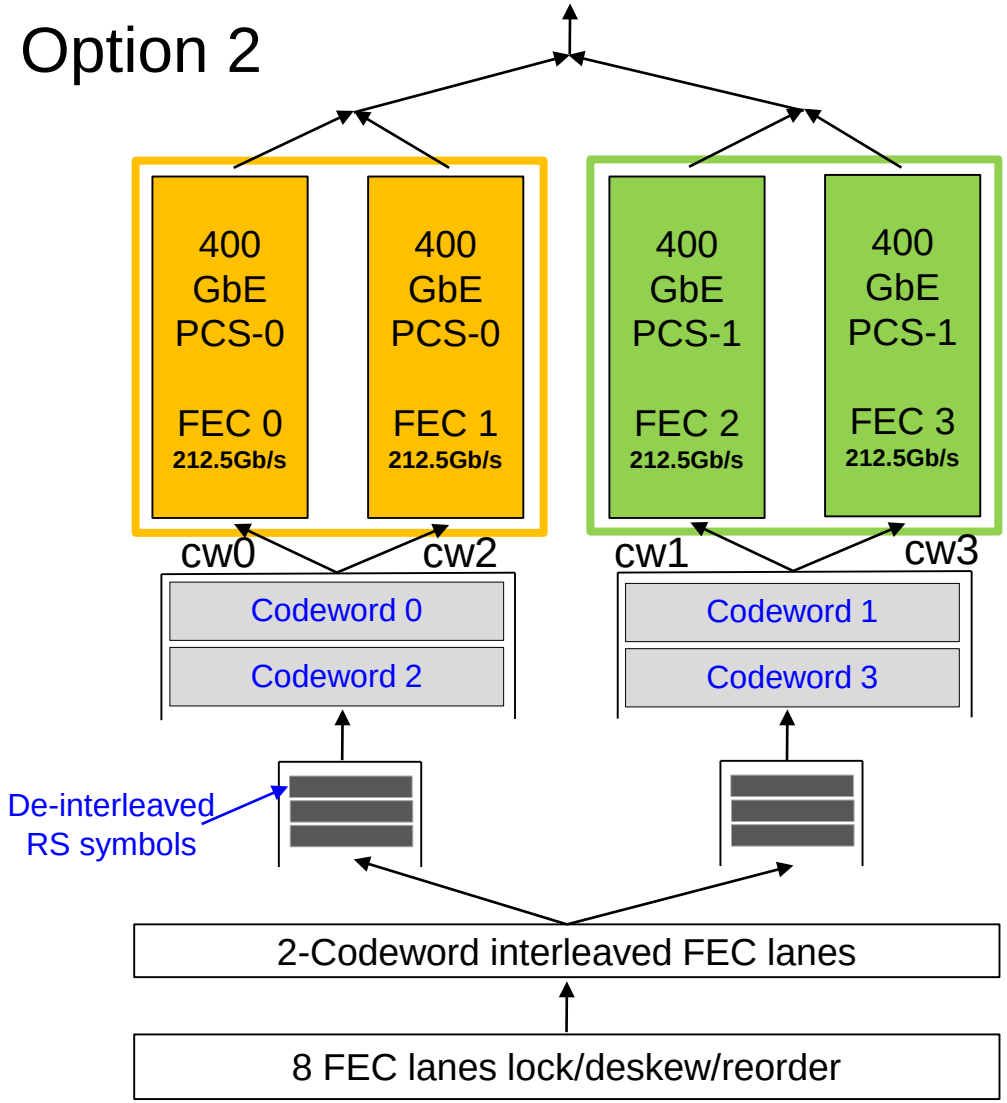
Option 2

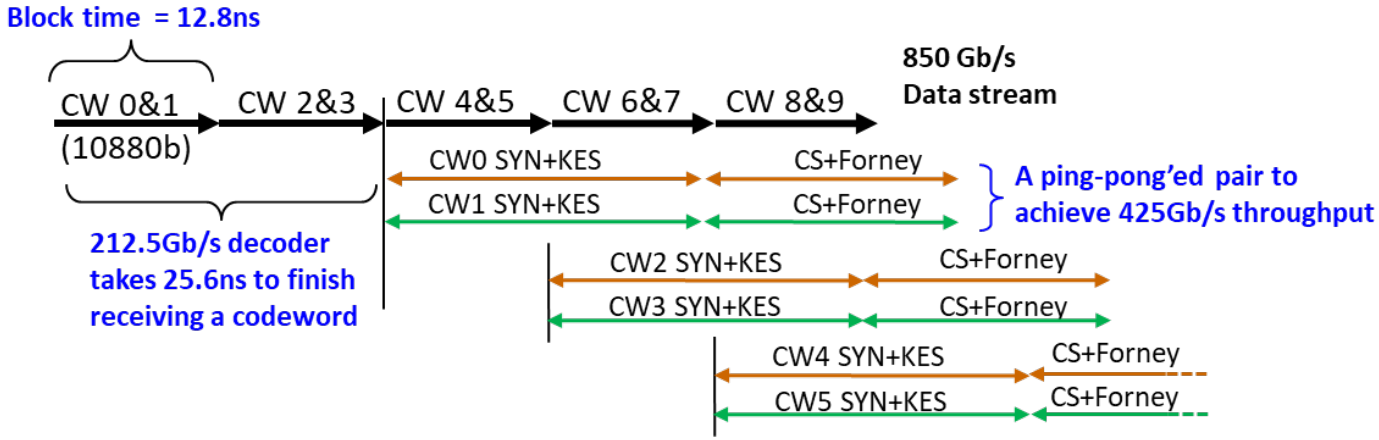


- Identical implementation conforming to IEEE 802.3bs 400 GbE FEC above the codeword buffers.
  - FEC decoders 0,1, 2,3 all operate at 212.5 Gb/s throughput.
  - More friendly to implementations with lower clock rates.
- Modifications:
  - Lane reorder is performed over 8 lanes instead of two groups of 16 lanes – similar to 200 GbE design.
  - Codeword distribution is performed across the two pairs of decoders.
- Interleaving/de-interleaving is performed over 850 Gb/s throughput.

Reuse 400 GbE FEC Decoder Implementation in Option 2, continued

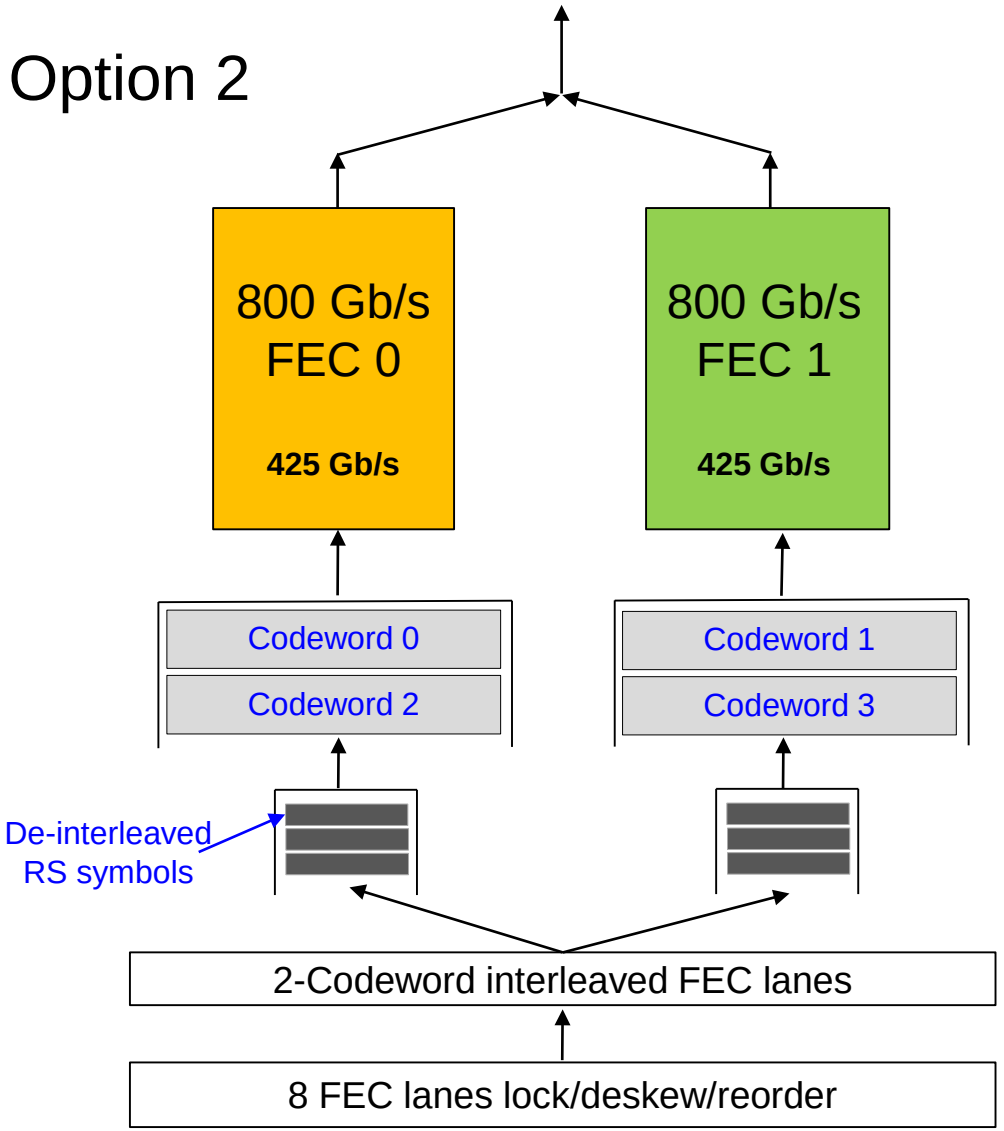
Option 2

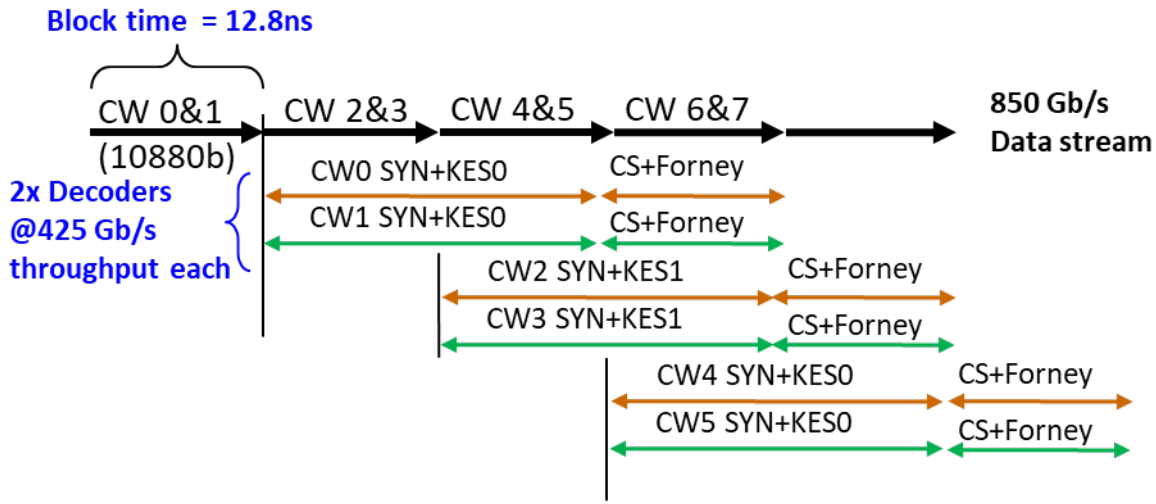




- Each interleaved pair will be fed into a pair of 212.5 Gb/s throughput decoders.
- FEC frame time is 12.8 ns (16 cylces@1.25GHz), but 212.5Gb/s throughput decoder takes 32 cycles to take in a codeword:
  - KES requires 30 clock cycles – no need to duplicate.
  - Chien Search is needs to be finished within 32 cycles.
- Total throughput is 850 Gb/s.

Further Evolution of Option 2 with More Benefits

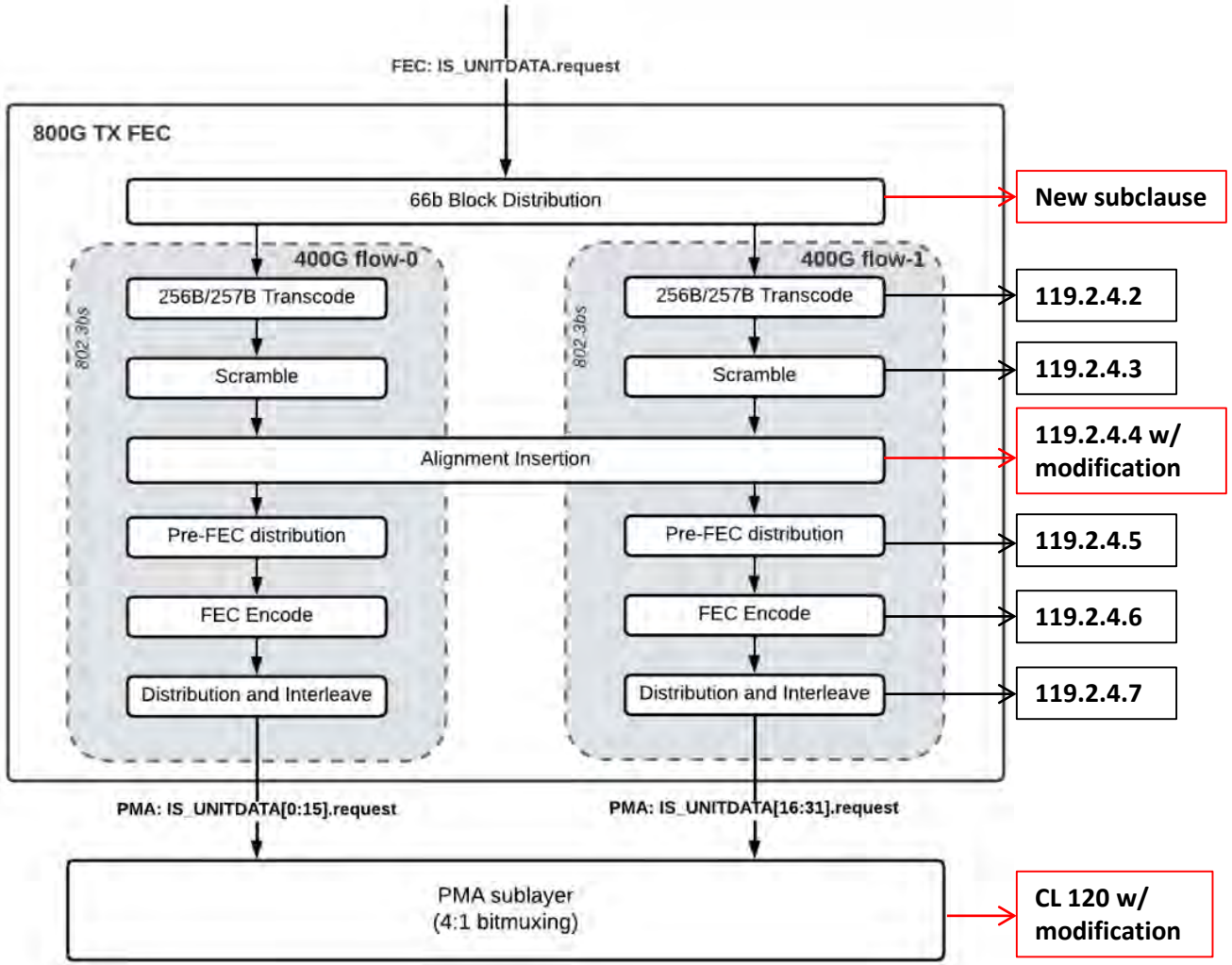




- FEC decoders 0 and 1 operate at 425 Gb/s throughput.
  - Lower latency when working with advanced process with higher clock rates.
  - Saving 10Kbits frame buffer.
- Can be viewed as a 4X/2X sped-up 200/400 GbE FEC.
- Benefit from Moore's Law – when moving to new process node with higher clock rate, lower latency, area and power can be achieved comparing to Option 1.

Option 1 Could Largely Reuse CL119 Sub-clauses of IEEE 802.3bs

- Reuse IEEE 802.3 Clause 119, with some modification needed:
  - **No change:**
    - PCS encode and decode
    - Transcode
    - Scramble and de-scramble.
    - FEC encode and decode, including interleaving.
  - **Needs modifications:**
    - New function: 64B/66B block distribution.
    - AM insertion: aligned over the two flows.
    - AM lock: 32 patterns instead of 16.
    - Lane reorder and deskew: over 32 lanes.
    - Error Marking: performed over the whole 800G flow?



shrikhande_3df_01_220329

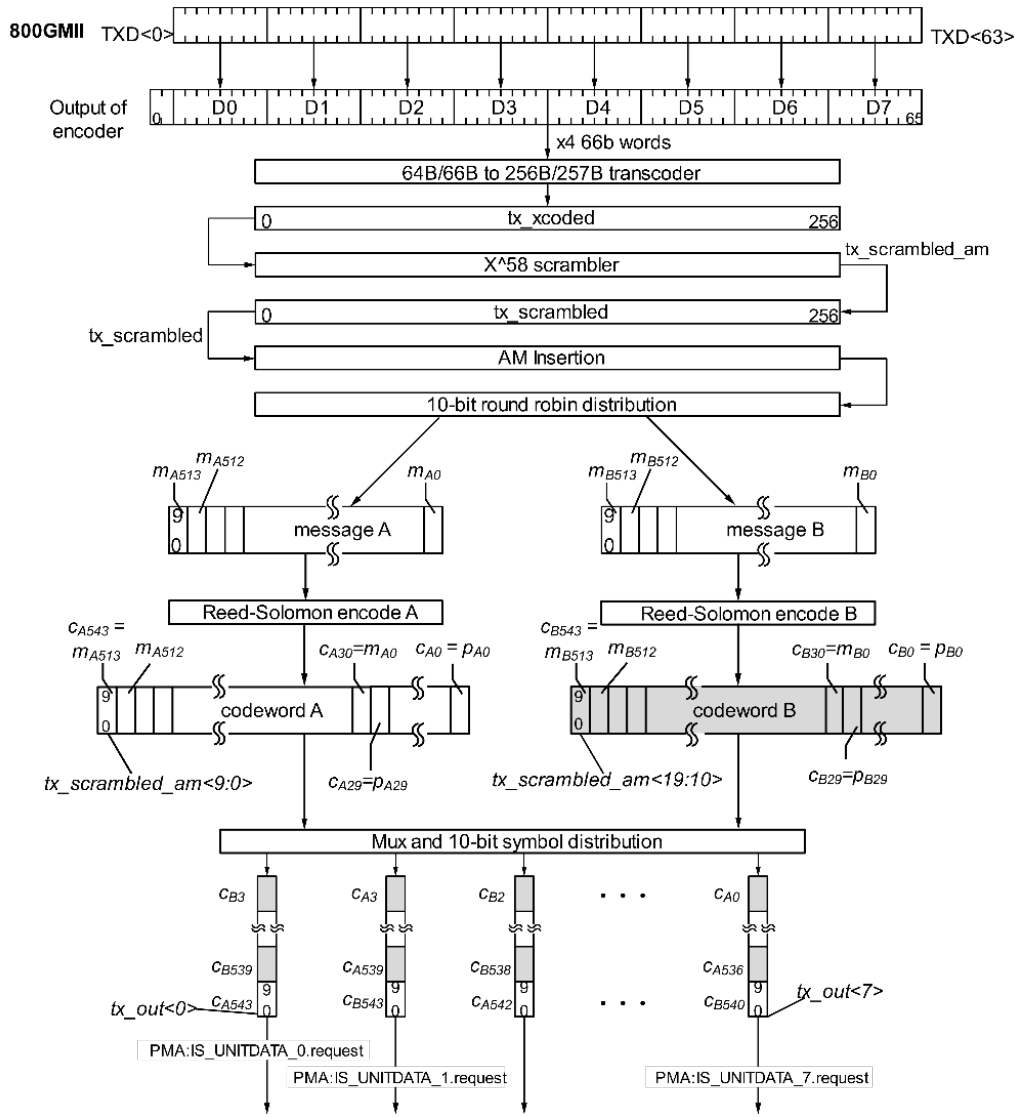
Option 2 Could Completely Reuse CL119 of IEEE 802.3bs

- **Completely reuse** IEEE 802.3 Clause 119:
  - New sub-clauses can directly refer to the existing sub-clauses in clause 119. For example, in 802.3cd-2018:

    134.5.2.7 Reed-Solomon encoder

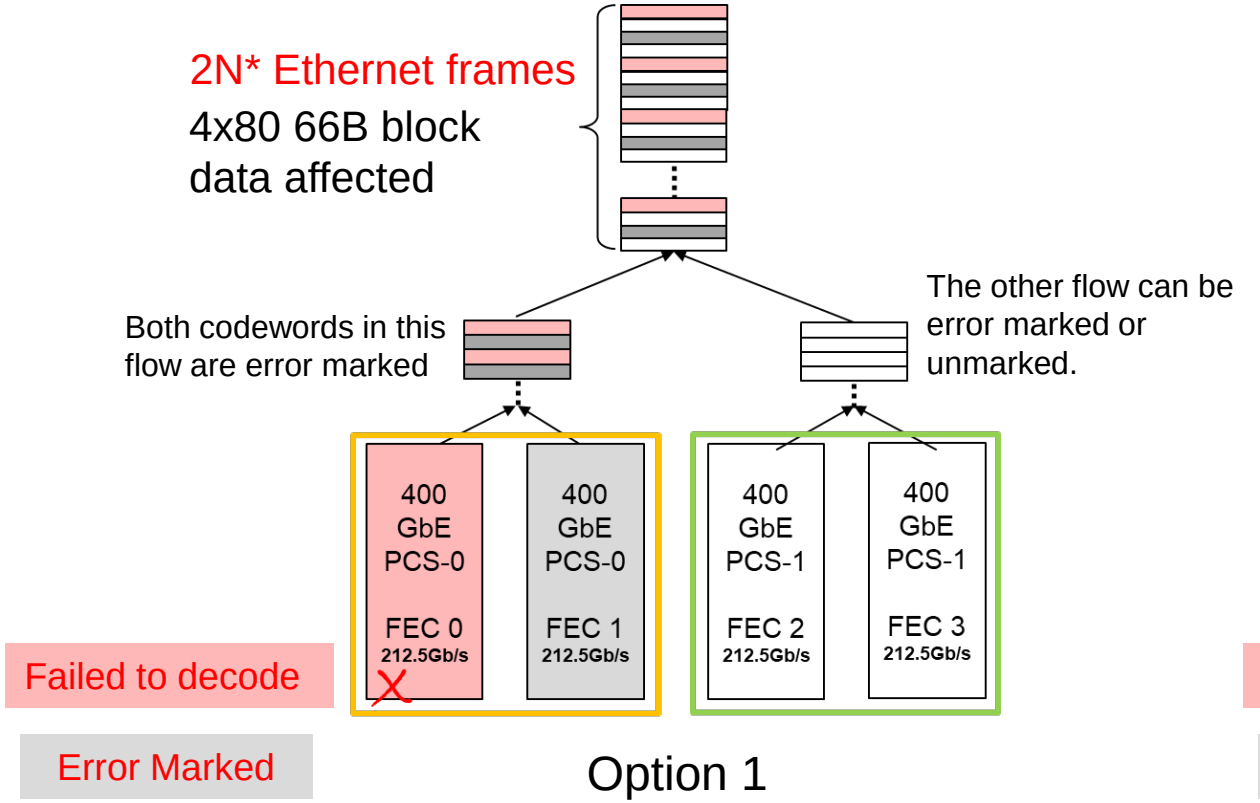
    The Reed-Solomon encoder is identical to the RS(544,514) Reed-Solomon encoder defined in 91.5.2.7.

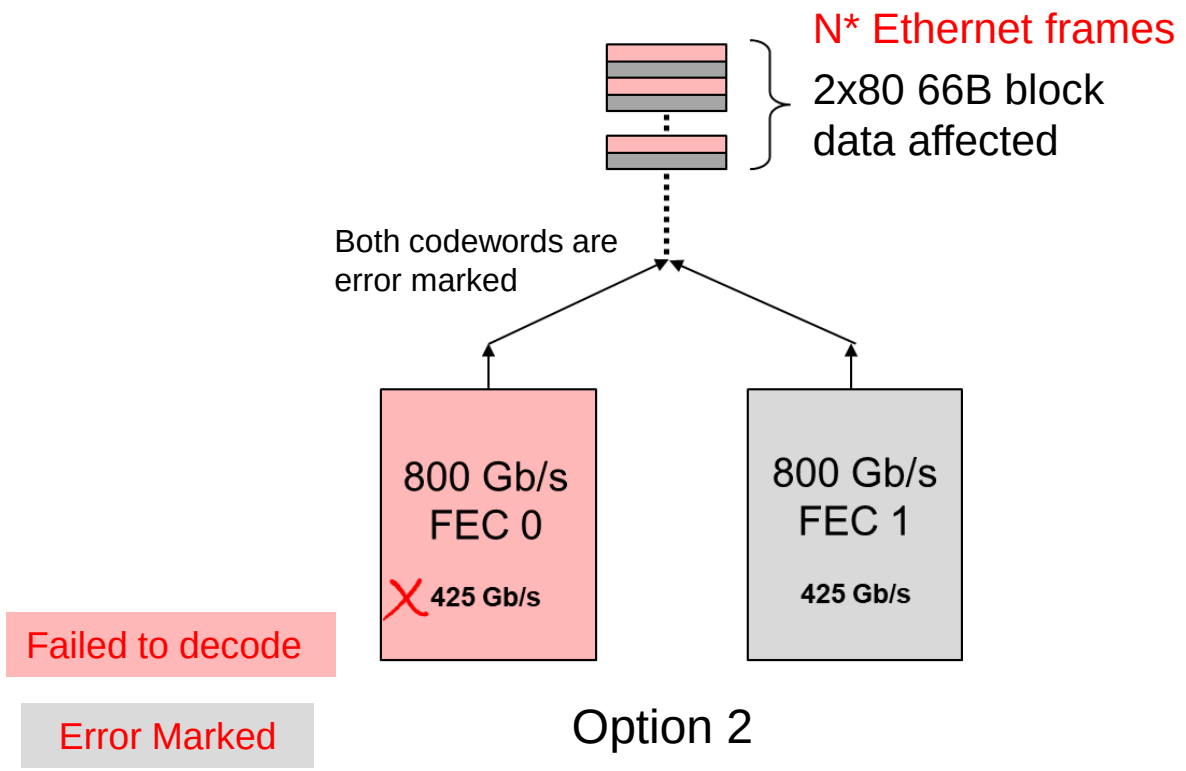
  - If FEC and PCS are two separated sublayers, text may be rearranged.



Error Marking and FLR Comparison of the Two Options

- The length of Ethernet frames affected by an uncorrectable codeword in option 2 is half of option 1.
  - Option 1 may have higher FLR than option 2 at the same post-FEC BER level.

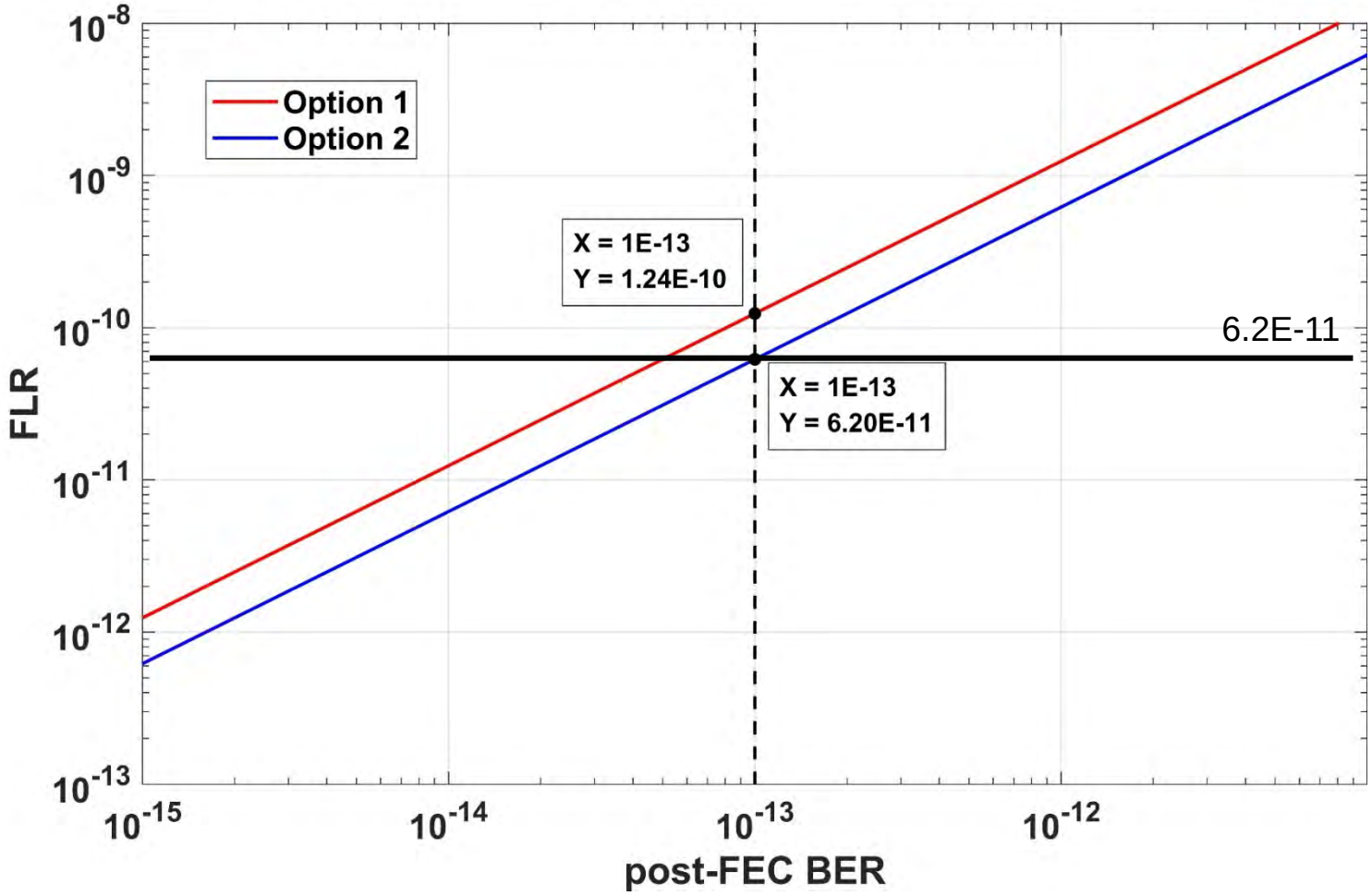




119.2.5.3:
"when the Reed-Solomon decoder determines that a codeword contains errors that were not corrected, it shall cause the PCS receive function to set every 66-bit block **within the two associated codewords** to an error block (set to EBLOCK_R)."

Post FEC BER vs FLR

- Option 1 will suffer twice as many frame loss as option 2 at the same post FEC BER @1E-13.



- The FLR for option 1 will be 1.24E-10, rather than 6.2E-11, which is the specified level for 200/400 GbE.
  - Clause 122.1.1:

    "For a complete Physical Layer, the frame loss ratio may be degraded to 6.2×10^{-11} for 64-octet frames with minimum interpacket gap due to additional errors from the electrical interfaces."

- We can either lower the BER objective, or **choose option 2 to keep the same BER and FLR**.

Summary:

- This contribution investigates the two proposed options of FEC1 with reuse of RS(544,514) from both implementation and standard specification perspectives.
  - Option 1 could largely reuse the existing 400 GbE implementation and standards with modifications.
  - Option 2 could largely reuse the existing 400 GbE implementation and completely reuse the existing standards.
- Option 2 has further advantages in lower FLR, latency, area and power.

Thank you