Google

FEC Requirements for 800GbE/1.6TbE Optics

Shuang Yin, Xiang Zhou, and Cedric Lam

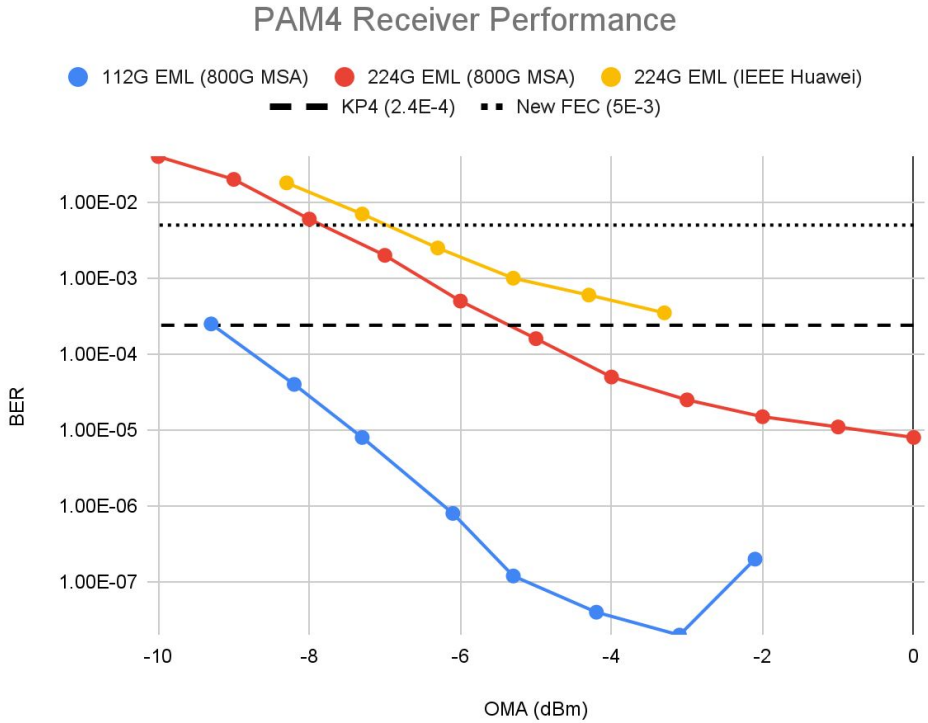
IEEE 802.3df July Plenary Meeting

Outline

- FEC requirements
  - Performance
  - Latency
  - Power
  - Backward compatibility
- Potential FEC candidates
- Other thoughts

FEC Requirements - Performance

112G EML: Enabling the next generation of cloud & AI using 800Gb/s optical modules.
224G EML: 200G per lane for future 800G and 1.6T modules.
224G EML: kuschnerov_3df_01_220222



- Target **BER of ~5e-3** for 200G optics lane to lower Rx SNR requirement

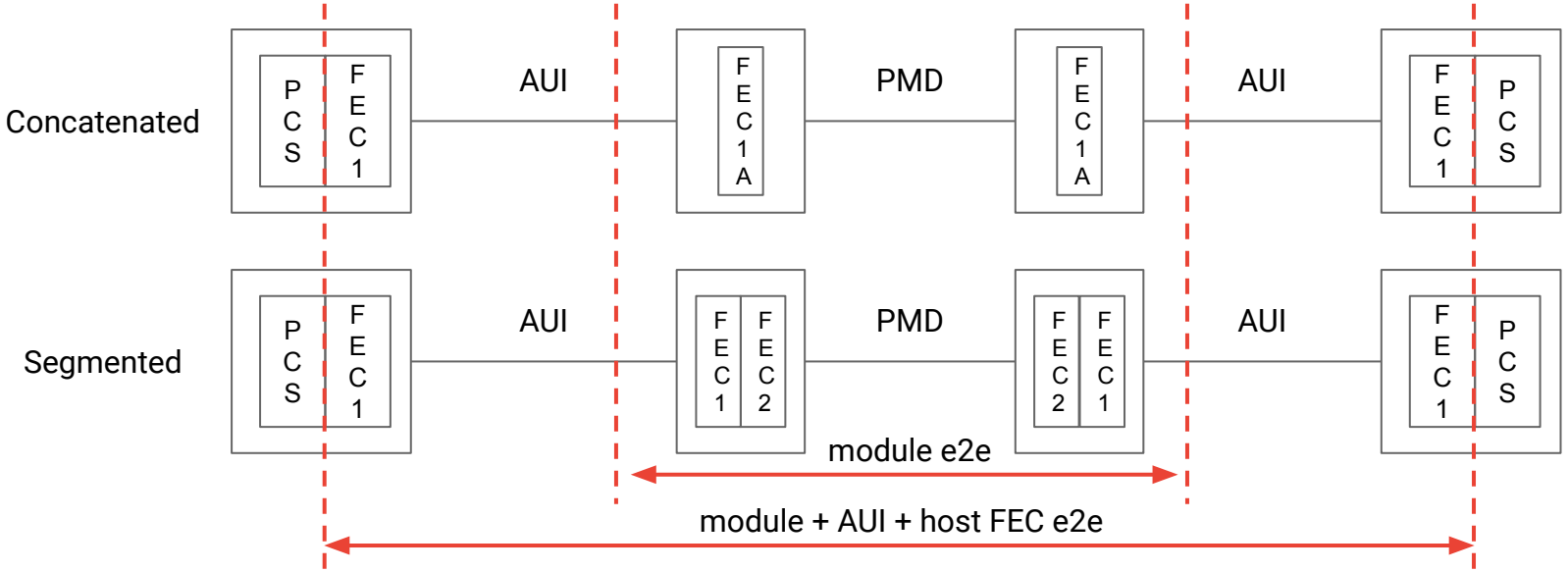
FEC Requirements - Latency

- **Latency matters**
  - AI/HPC, disaggregated architecture (RDMA/RoCE), etc.
    - High performance Ethernet NEA (zhuang_nea_01_220622)
    - Sub-10 us cluster/network fabric, i.e., **~100 ns on FEC matters**
  - Traditional network operators on private line related applications (wang_3df_01a_220609)
- Current 100G per lane 400GbE module, e.g., 400G-FR4, has ~200 ns e2e latency (~100 ns module DSP* + ~100 ns host KP4 FEC*)
- For 200G per lane 800GbE/1.6TbE module, target **< 400 ns** e2e latency
  - Allocate additional latency for a higher coding gain FEC
  - Support majority of intra data center use cases
  - Align with 800G LR1 requirement in oif2021.369.02 (300 ns module e2e + ~100 ns host KP4 FEC)

* R. Nagarajan et al., Low Power DSP-Based Transceivers for Data Center Optical Fiber Communications (Invited Tutorial), JLT 39 (16), 2021.

FEC Latency Example

FEC1: KP4
FEC1A: Hamming (128, 120)
FEC2: KP4 + Hamming (128, 120)



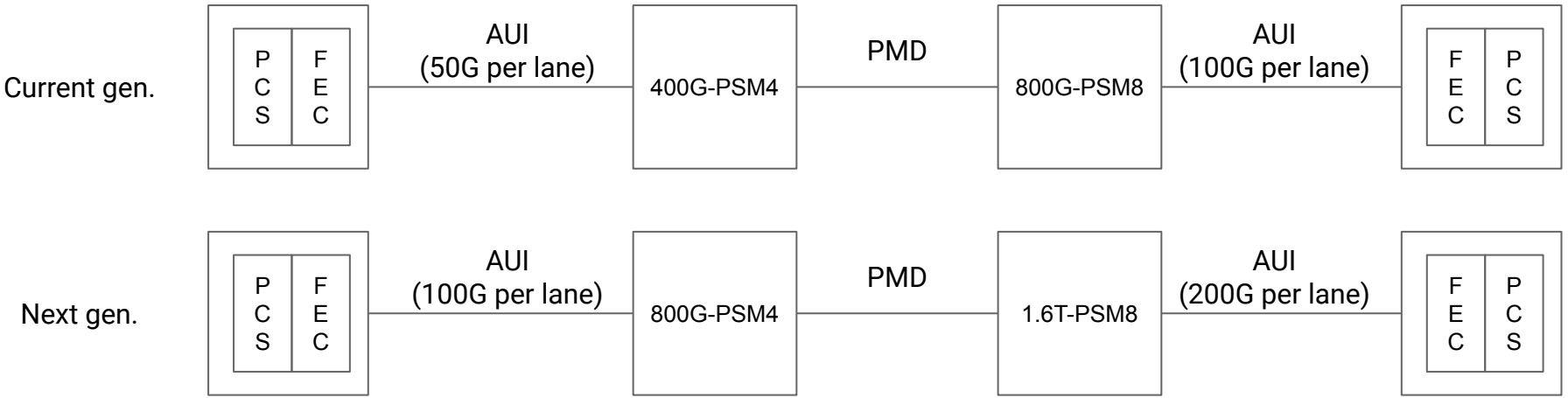
- Module e2e: ~300 ns (concatenated) vs ~500 ns (segmented)
- Module + AUI + host FEC e2e: ~400 ns (concatenated) vs ~600ns (segmented)
- Strongly prefer the concatenated FEC with ~200 ns lower latency per link

FEC Requirements - Power

- Current 100G per lane 800GbE module at ~16 W @7nm CMOS node
- Target at **< 30 W** for 200G-per-lane 1.6TbE module @5nm CMOS node
- Example study showed **module FEC power of ~0.2 W/800G** @5nm CMOS node (oif2022.274.00) is feasible in terms of performance and module power envelope

FEC Requirements - Backward Compatibility

- Host FEC backward compatibility
  - Enable smooth speed evolution without FEC adaptation in the module (simplifying module and DSP options)
  - Good to keep, don't break it unless absolutely have to!



Potential FEC Candidates

FEC Scheme	FEC Proposal	Pre-FEC BER	E2E Latency*	Module FEC Power	FEC Overhead
Concatenated	RS(544, 514) + Hamming(128, 120) patra_3df_01_220518	4.85e-3	~400 ns	Low	12.9%
End to end	RS(576, 514) he_b400g_01_210426	1.1e-3	~230 ns	0	12.1%
Segmented	RS(544, 514) + [RS(544, 514) + BCH(144, 136)] he_b400g_01_210426	2.4e-3	~420 ns	High	12.1%

* Module + AUI + host FEC e2e latency for fair comparison, assumed ~160 ns additional latency for non-FEC contributions, e.g., equalization, etc.

Other Thoughts

- FEC flexibility is needed to accommodate different PMDs (requirements)
- For concatenated FEC scheme
  - Proper interleaver design between outer and inner code is critical to overall FEC performance
  - Flexibility in interleaver (e.g., 2-way vs 12-way) is preferred to trade off between code gain/performance and latency (bliss_3df_01a_220517 and patra_3df_01_220518)
  - Appropriate inner code telemetry is needed for ease of operation
- To maintain FEC backward compatibility in 200G per lane host (KP4)
  - 200G C2M channel needs to operate better than 1e-4
  - There were early studies on this, but more detailed study is needed to evaluate its feasibility

Summary

- FEC requirements for 800GbE/1.6TbE optics (200G per lane) are elaborated in terms of performance, latency and power
- Concatenated RS(544, 514) + Hamming(128, 120) from patra_3df_01_220518 could be a good baseline solution for 200G per lane optics
- A fair comparison among different FEC approaches is needed
    - Optimizing for the lowest power and area (cost) while meeting the code gain (pre-FEC BER) and latency requirements