FEC Inner Code Bypass Options for 200G/L IMDD Optics

Brian Welch (Cisco)

Zvi Rechtman (Nvidia)

Piers Dawe (Nvidia)

Supporters

- Bill Simms (Nvidia)
- Vishnu Balan (Nvidia)
- Vasu Parthasarathy (Broadcom)
- John Calvin (Keysight)
- Rob Stone (Meta)
- Yan Zhuangyan (Huawei)
- Leon Bruckman (Huawei)
- Mark Gustlin (Cisco)
- Adee Ran (Cisco)
- Sridhar Ramesh (Maxlinear)
- Mike Li (Intel)
- Lenin Patra (Marvell)
- Jamal Riani (Marvell)
- Roberto Rodes (Coherent)†
- Vipul Bhatt (Coherent)†
- Ryan Yu (Sifotonics)
- Jamie Gaudette (Microsoft)
- Phil Sun (Credo)

† Indicates support for DR reaches only

Overview

- In the March 802.3dj task force meeting, a motion was approved to adopt an inner code as part of the FEC approach for 200G/L IMDD optics
  - https://www.ieee802.org/3/dj/public/23_03/motions_3dfdj_2303.pdf (page 19)

- This proposal even with bypassing the convolutional interleaver yields longer latency than desired for AI/ML applications.

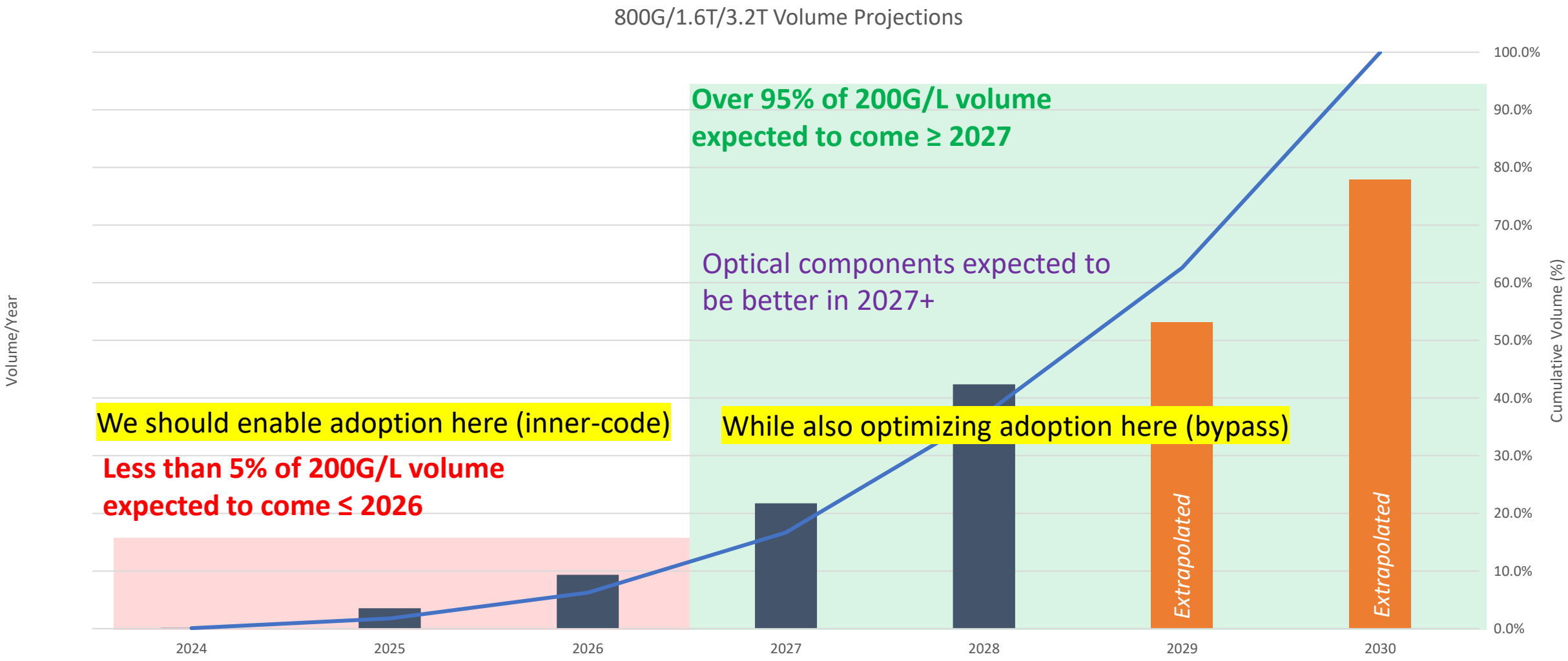
- An alternate proposal is presented to enable inner-FEC bypass as an option for latency sensitive applications.

Motivations

- Inner-code FEC proposals with the 12-way convolutional interleaver add up to 280ns of latency per link/hop, with breakout applications (which tend to be shortest reach) the most impacted.
  - brown_3dj_optx_01b_230413.pdf
  - This is the equivalent of approximately 60m of single mode fiber.
  - Approximately 7% additional overhead for 1.2 decades of BER improvement
- Inner-code FEC has little gain to give up (to optimize latency), given the high overhead requirement
  - CI bypass options will degrade the coding gain, but still adds ~23 nsec latency
- Flexibility to bypass the Inner Code FEC will enable DC operators to run the 200G/L link with end-to-end KP FEC for Low Latency application.
  - Also, likely to reduce module/system power consumption
- CPO/NPO/Linear optics architectures may not have DSP as part of the optical transceiver, and as such may not have FEC inner-code available.
  - The possibility of relocating the inner code to the host may exist, however that would result in higher line rates over the AUI and likely degrade soft decision performance. Would also require customized host silicon.

2024-2028 From LightCounting March 2023 (with permission).
2029-2030 Extrapolated based on 400GE

Motivations



Technical Feasibility

- Numerous Contributions (from different contributors) have demonstrated technical feasibility of 200G links based on RS(544,514) FEC.
  - https://www.ieee802.org/3/df/public/22_10/22_1005/simms_3df_01_221005.pdf
  - https://www.ieee802.org/3/df/public/22_10/22_1011/welch_3df_01a_221011.pdf
  - https://www.ieee802.org/3/dj/public/23_01/23_0206/li_3dj_01a_230206.pdf
  - https://www.ieee802.org/3/dj/public/23_03/parthasarathy_3dj_01_2303.pdf

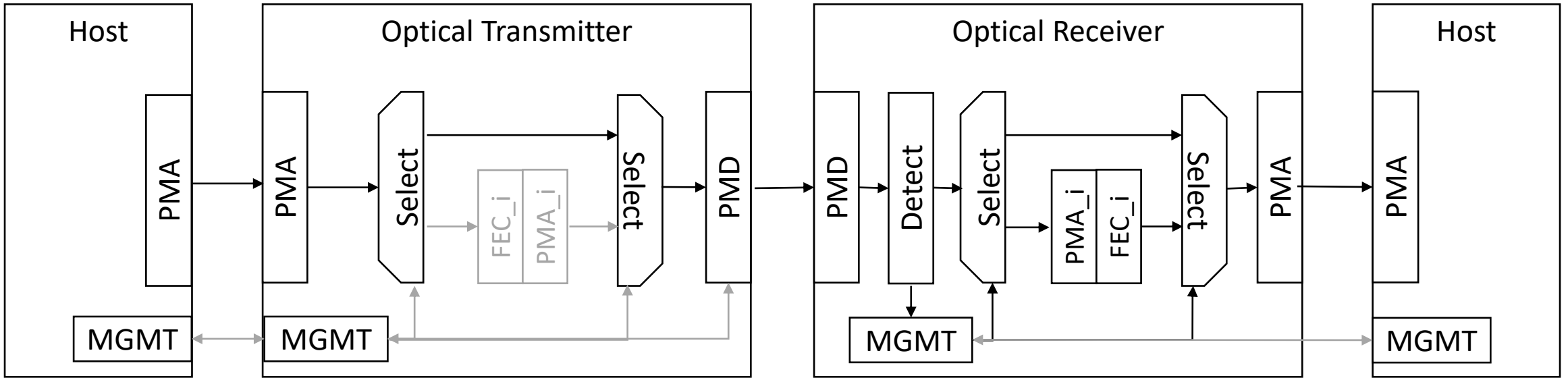
- FEC consensus presentation at last meeting proposed using RS(544,514) for all single wavelength PMD types:
  - https://www.ieee802.org/3/dj/public/23_03/patra_3dj_01b_2303.pdf

Option 1: Auto-Detect

- **Proposal**: Make the use of the inner code a transmitter optimization function, such that transmitters with better performance do not need to enable it.
  - (Minimally) Different optical specs for each mode of operation

- The receiver would need to accommodate both operating modes (with/without inner code), however many options exist to do so.
  - Rate detect: Clock recovery subsystem determines FEC mode based on what rate it locks to.
  - Data detect: Optical receiver looks for Hamming codewords and bypasses FEC if not present.
  - Receivers must support RS544 **and** RS544+Inner Code mode, with affiliated specs.

Option 1: Auto-Detect

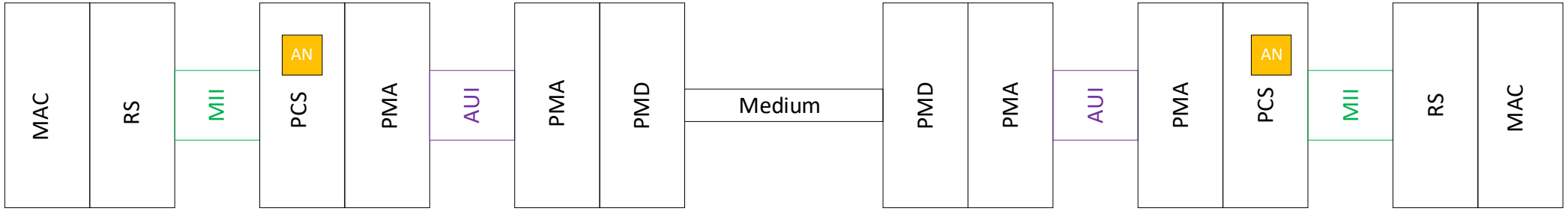


Optical transmitter chooses inner-FEC mode (with host override function)

Optical receiver detects inner-FEC mode (with host override function)

Option 2: Auto-Negotiation

- Leverage the BASE-R Negotiating concept to bypass the inner-FEC.
- **Optimize host2host link performance including both AUI(s) and PMD errors.**
- A path to negotiate additional capabilities (e.g. 12-way Convolutional Interleaver bypass/enable)
- There are several options for AN, this contribution focus on performing AN within PCSs
- **Note 1: Margin determination for auto-negotiation can be determined at the host discretion (ie, if host has specific requirements due to link aging)**
- **Note 2: Forced host control via management interface also an option for networks where both sides of the link are under common management.**



Precedent

- 802.3 has a long history of allowing performance dependent tradeoffs for optical transmitters:
  - Ex: OMA vs. TDECQ

- 802.3 also has precedent for allowing different FEC types for the same PMD spec:
  - Ex: 10GBASE-KR, 25GBASE-KR/CR

Summary

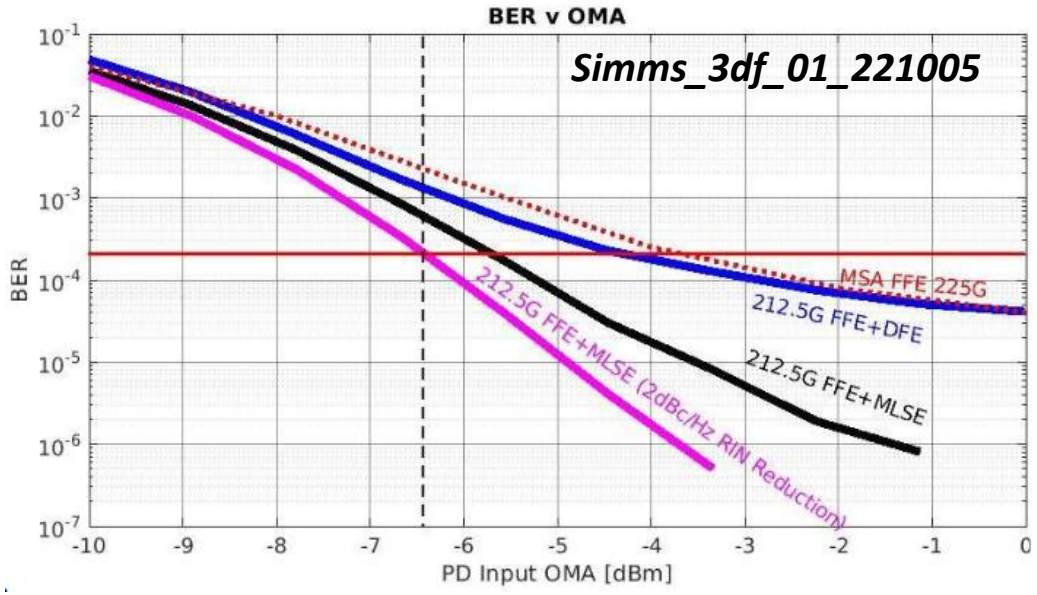
- The use of an inner-code for 200G/L adds considerable latency to an optical link/network

- Inner-code bypass can be easily supported using existing Ethernet techniques (auto-detect and/or auto-negotiation)
  - **The proposal is not to remove the inner-code, but to supplement with a bypass mode.**
  - Technical feasibility has been demonstrated through prior task force contributions.

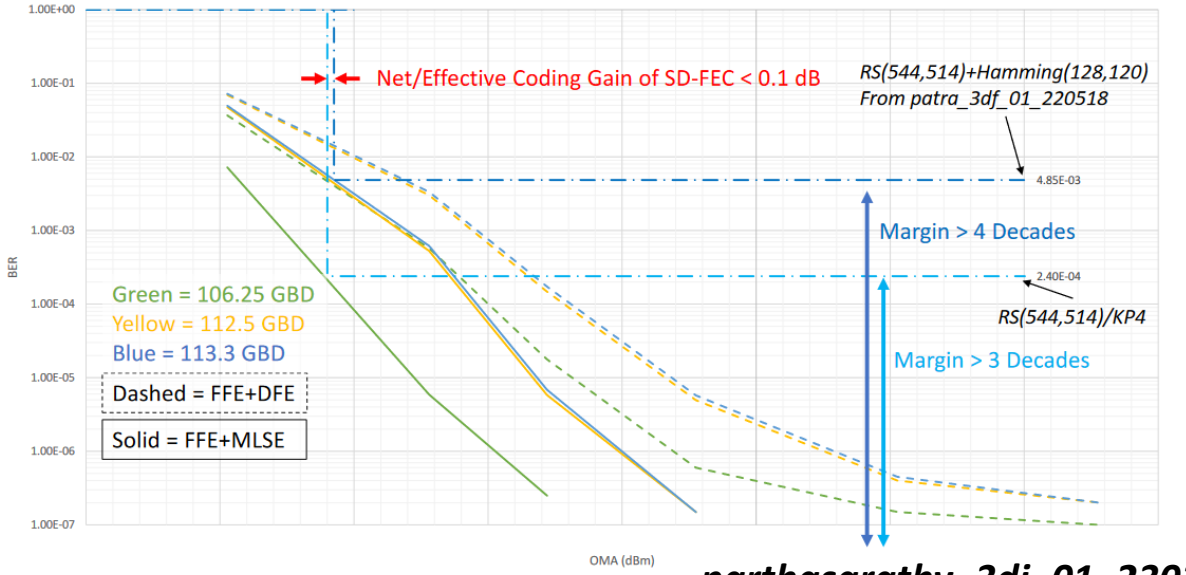
- Note: Inner code bypass option also likely to have meaningful power benefits.

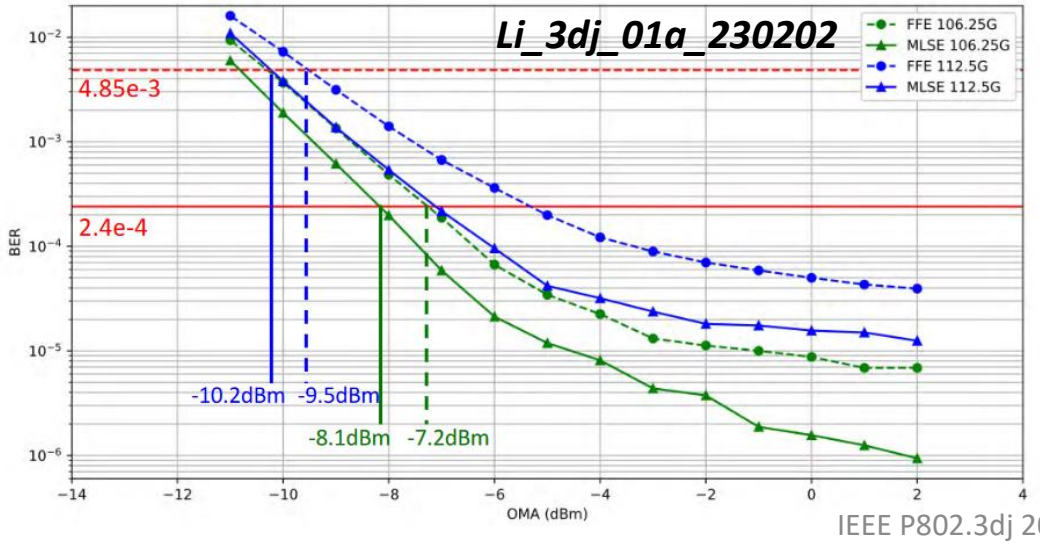
Thank You

Technical Feasibility

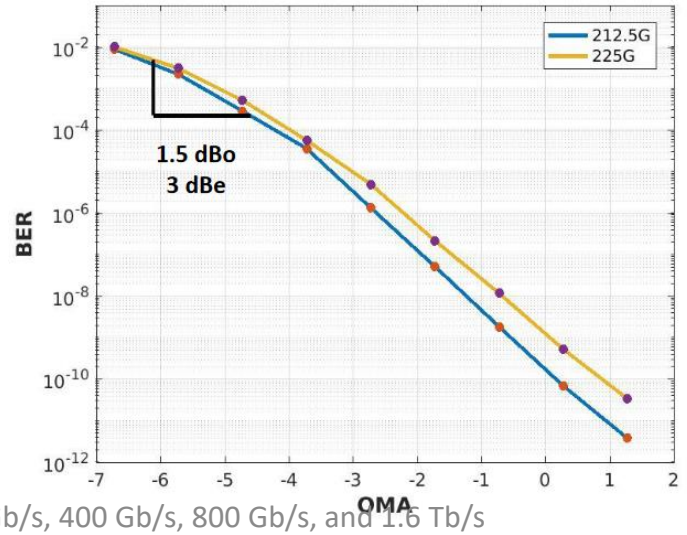


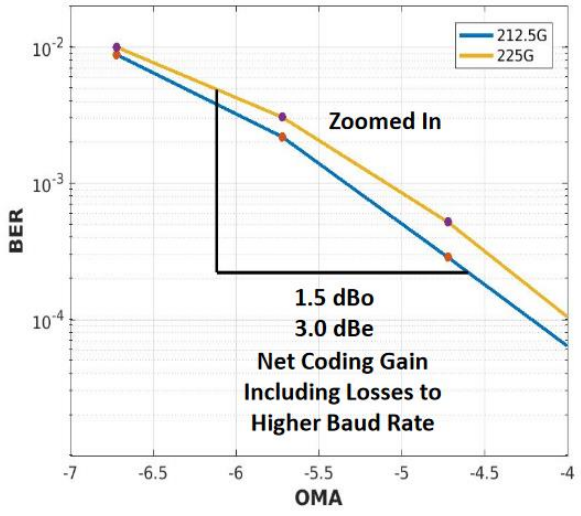
Welch_3df_01a_221011



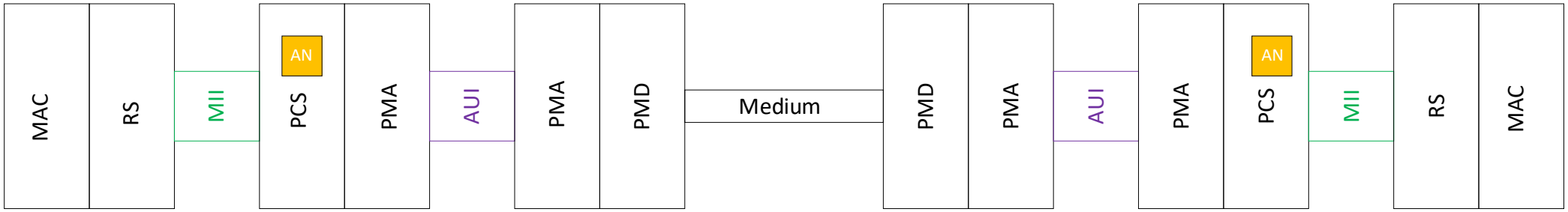


parthasarathy_3dj_01_2303





AN between PCSs



- Similar concept as in Clause 37
  - Define 64/66 ordered set which carries AN information.
  - Clause 73 Paged structure with the same handshake variables as in clause 73 (Ack, Next-Page etc.)
- Leverage existing PCSs (Clause 119, 172 and 1.6T PCS) and Clause 73 state machines.
  - Potential for short definition process.
- **No implications to traditional Module's sub-layering (PMA/PMD)**
- AN pages protected by KP4 FEC and leverages 64/66 encoding - very reliable!
- Require MDIO register/interference primitive to bypass (or enable) inner-FEC