FEC Bypass: Procedural Considerations

IEEE P802.3dj Task Force
Joint Optics / Logic Ad Hoc
15 Aug 2023

John D'Ambrosia,
Futurewei, U.S. Subsidiary of Huawei

David Law
HPE

My Role as Chair

- Per the IEEE 802.3 Ethernet WG Operations Manual (http://www.ieee802.org/3/rules/P802_3_rules.pdf)
  - The operation of the TF has to be balanced between democratic procedures that reflect the desires of the TF members and the TF Chair's responsibility to produce a draft standard, recommended practice, or guideline in a reasonable amount of time for review and approval by the WG. Robert's Rules of Order shall be used in combination with these operating rules to achieve this balance.
  - The full responsibilities of the chair are specified in 3.4.3 Task Force Chair’s Responsibilities.

Introduction

- IEEE P802.3dj has DRx, DRx-2, FR4, and LR4 “physical layer specification” objectives @ 200 GbE, 400 GbE, 800 GbE, and 1.6 TbE.
- Mar 2023 – Motion #5
  - A concatenated FEC approach (see patra_3dj_01b_2303) has been adopted for DRx, DRx-2, FR4, and LR4 for relevant objectives @ 200 GbE, 400 GbE, 800 GbE, and 1.6 TbE.
- July 2023 - Motion #9
  - Adopted direction to “adding an option to support only RS544 FEC (aka Bypass Inner FEC) for the single wavelength 500m and 2km optical PMDs“
- Subsequent conversations have raised the issue if the proposal is one or two PHYs

My Understanding - “FEC Bypass” Proposal

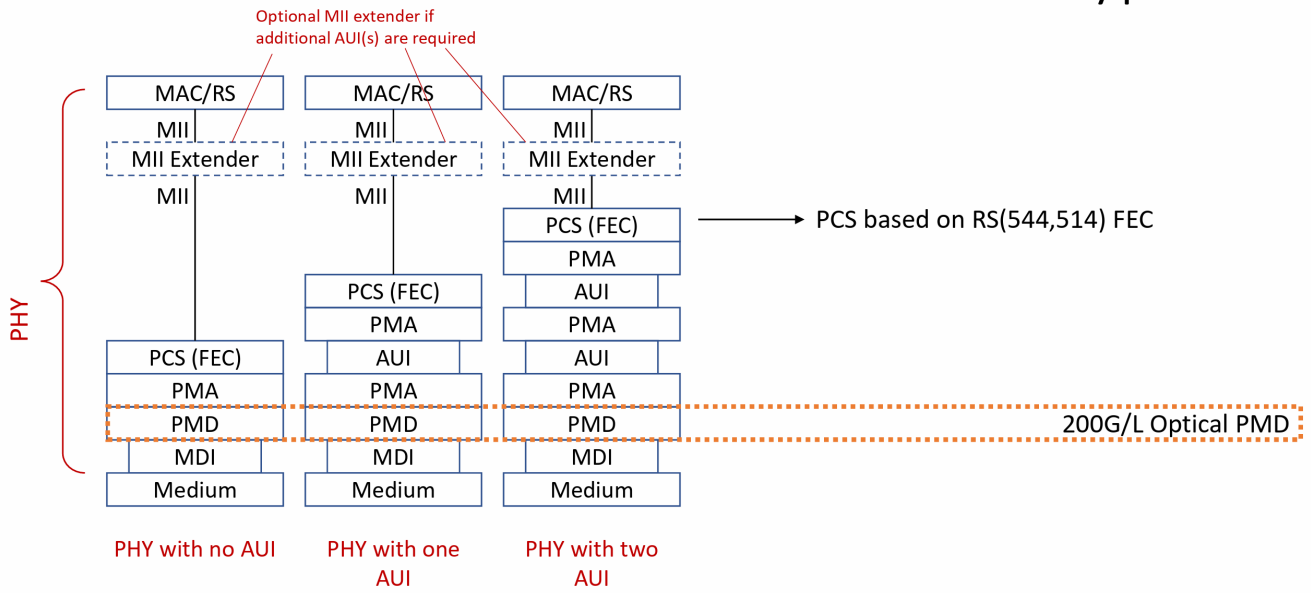
- Inner-FEC bypass is an option for latency sensitive applications
  - Per welch_3dj_03c_2305: a

    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.
- Transmitter
  - Two classes of transmitters (different specifications, including bit rate)
    - Tx_A _ inner FEC OFF
    - Tx_B _ inner FEC ON
  - No stated requirement to support both classes of transmitters
  - No stated requirement whether support of inner FEC is mandatory
- Common receiver that accommodates both transmitters
  - Receivers must support RS544 and RS544+inner code mode
- PMD BER
  - For Tx_A: 2.4x10-4
  - For Tx_B: 3.0x10-3
- Different proposals to switch between modes –
  - Management Interface
  - Plug-n-play Interop: Auto-Detect, Auto-Negotiation

Clarity of "FEC Bypass" Proposal Needed

- Concept of "FEC Bypass" needs clarification
  - Is the inner FEC present and being bypassed?
  - Is the inner FEC optional to implement?
  - Is the concept a by-product of the chronological order of motions?
- Per welch_3dj_03c_2305.pdf – "The proposal is not to remove the inner-code, but to supplement with a bypass mode."
- Per welch_3dj_03a_2307.pdf – see diagram to right for Inner FEC Bypassed – Inner FEC not shown (perhaps just graphical decision?)

Location in Ethernet Stack: Inner FEC Bypassed



The remainder of this presentation will assume inner FEC is present for all implementations, and the proposal is to actually bypass it.

Referenced Precedents

- Per welch_3dj_03c_2305.pdf
  - 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

Referenced Precedent – Optical Tx OMA vs. TDECQ

Table 151–7—400GBASE-FR4 and 400GBASE-LR4-6 transmit characteristics

Description	400GBASE-FR4	400GBASE-LR4-6	Unit
Signaling rate, each lane (range)	53.125 ± 100 ppm		GBd
Modulation format	PAM4		—
Lane wavelengths (range)	1264.5 to 1277.5 1284.5 to 1297.5 1304.5 to 1317.5 1324.5 to 1337.5		nm
Side-mode suppression ratio (SMSR), (min)	30		dB
Total average launch power (max)	10.4	11.1	dBm
Average launch power, each lane (max)	4.4	5.1	dBm
Average launch power, each lane[a] (min)	–3.2	–2.7	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (max)	3.7	4.4	dBm
Outer Optical Modulation Amplitude (OMA_{outer}), each lane (min) for TDECQ <1.4 dB for 1.4 dB ≤ TDECQ ≤ 3.4 dB	 –0.2 –1.6 + TDECQ	 0.3 –1.1 + TDECQ	 dBm dBm
Difference in launch power between any two lanes (OMA_{outer}) (max)	3.9	4	dB
Transmitter and dispersion eye closure for PAM4 (TDECQ), each lane (max)	3.4	3.4	dB
Transmitter eye closure for PAM4 (TECQ), each lane (max)	3.4	3.4	dB
| TDECQ – TECQ | (max)	2.5	2.5	dB
Over/under-shoot (max)	22	22	%
Transmitter power excursion (max)	1.8	2.5	dBm
Extinction ratio, each lane (min)	3.5	3.5	dB
Transmitter transition time (max)	17		ps
Average launch power of OFF transmitter, each lane (max)	–16	–16	dBm
$RIN_{17.1}OMA$ (max)	–136	—	dB/Hz
$RIN_{15.6}OMA$ (max)	—	–136	dB/Hz
Optical return loss tolerance (max)	17.1	15.6	dB
Transmitter reflectance[b] (max)	–26		dB

[a] Average launch power, each lane (min) is not the principal indicator of signal strength. A transmitter with launch power below this value cannot be compliant; however, a value above this does not ensure compliance.
[b] Transmitter reflectance is defined looking into the transmitter.

- **Single specification per PHY type**
- "The values for OMAouter each lane (min) in Table 151–7 vary with TDECQ. The relationships are illustrated in Figure 151–3 along with the values for OMAouter each lane (max)."

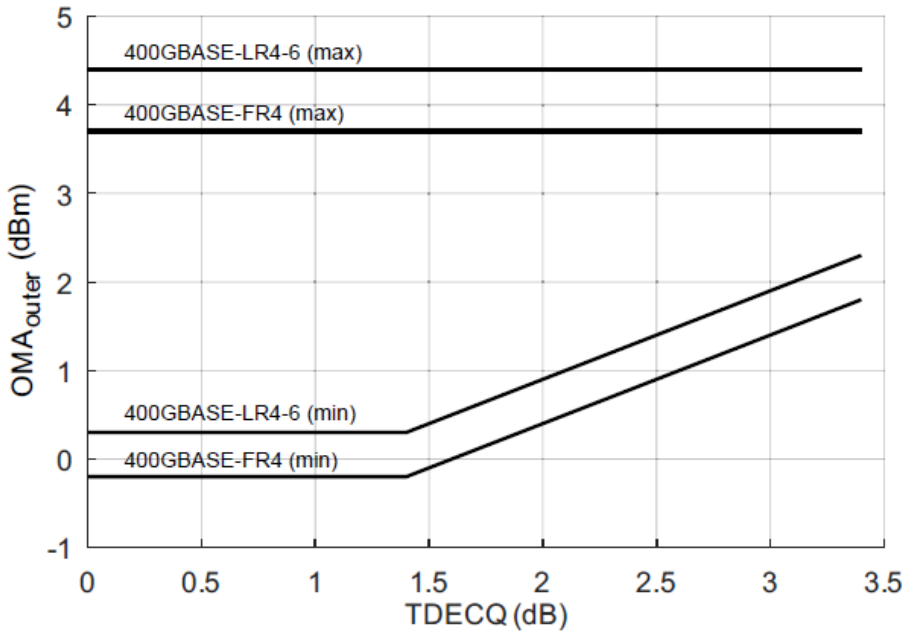


Figure 151–3—OMA_{outer} each lane (max) and OMA_{outer} each lane (min) versus TDECQ for 400GBASE-FR4 and 400GBASE-LR4-6

Historical Perspective - 10GBASE-KR

- One Tx / one Rx specification
- BASE-R FEC is optional to implement – Per 74.1 –
  - "… optionally use the FEC sublayer to increase the performance on a broader set of backplane channels than are defined in Clause 69. The FEC sublayer provides additional margin to account for variations in manufacturing and environmental conditions."
- One PHY with optional FEC

Historical Perspective – 25GBASE-KR/CR, KR-S/CR-S

- 25GBASE-KR / 25GBASE-KR-S – same Tx specs, different Rx specs (These are different PHYs w different FEC modes)
- 25GBASE-CR / 25GBASE-CR-S – same Tx specs, different Rx specs (These are different PHYs w different FEC modes)

- FEC Requirements
  - BASE-R FEC mandatory: 25GBASE-KR-S, 25GBASE-KR, 25GBASE-CR-S, 25GBASE-CR
  - 25GBASE-R RS-FEC mandatory: 25GBASE-KR, 25GBASE-CR

- Three FEC modes are supported:
  a) When the 25GBASE-R RS-FEC sublayer is enabled, the PHY is defined to operate in the RS-FEC mode.
  b) When the BASE-R FEC sublayer is enabled, the PHY is defined to operate in the BASE-R FEC mode.
  c) When no FEC sublayer is enabled, the PHY is defined to operate in the no-FEC mode.

- A 25GBASE-KR / CR PHY can operate in RS-FEC, BASE-R FEC, or no-FEC mode.

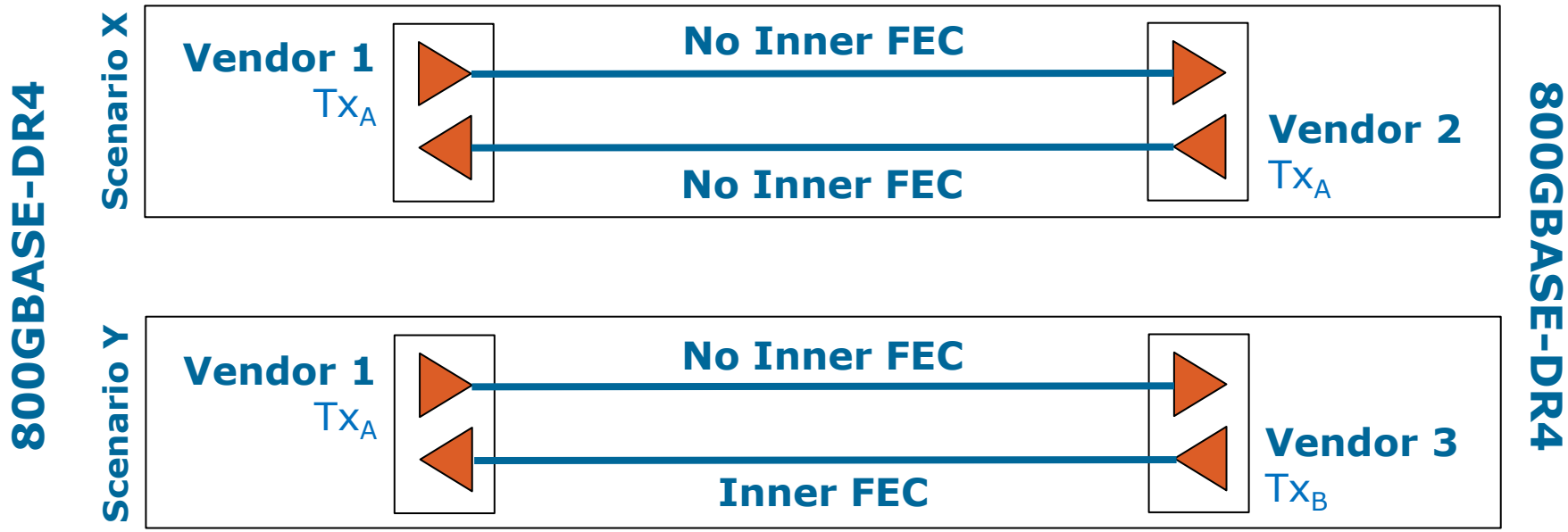
- A 25GBASE-KR-S / CR-S PHY can operate in either BASE-R FEC or no-FEC mode.

- When forming a complete 25GBASE-KR/CR or 25GBASE-KR-S/CR-S Physical Layer, the link BER requirements depend on the FEC mode (see 110.6) according to the following guidelines:
  a) If a PHY operates in the RS-FEC mode, and the RS-FEC decoder does not bypass error correction (see 108.5.3.2), the link is required to operate with a BER of 10–5 or better.
  b) If a PHY operates in the BASE-R FEC mode, the link is required to operate with a BER of 10–8 or better.
  c) If a PHY operates in the no-FEC mode, or in the RS-FEC mode with error correction bypassed, the link is required to operate with a BER of 10–12 or better.

Observations

- A single PHY approach:
  - No distinguishing between “transmitter classes”
  - Implementors will need to implement receivers supporting both modes forever going forward
  - “Low latency” implementations would require implementors to work with vendors / “data sheets” to identify interoperable low latency solutions

Example



Tx_A – inner FEC OFF
Tx_B – inner FEC ON

Questions to Consider

- Will the market accept this dual approach?
  - "Low-latency" approach will not be ensured by all solutions meeting the standard
- How will users identify "lower latency" PHYs?
  - No new objectives?
    - Will market accept a standard that doesn't differentiate PHYs and having to go to "data sheets" if a PHY is lower latency?
  - New Objectives?
    - See following pages (#12 - #15) for proposed approach
    - Could adding new objectives be justified?
      - Distinct identity?
      - Justification presentation needed
- Will all future optical transmitters meet Tx_A requirements i.e. no inner FEC necessary
- Will an optimized low latency only PHY (no inner FEC) be desired?

Proposed Objectives Update

- **200 Gb/s Related**
  - Define ~~a~~ physical layer specifications that supports 200 Gb/s operation over 1 pair of SMF with lengths up to at least 500 m
  - Define ~~a~~ physical layer specifications that supports 200 Gb/s operation over 1 pair of SMF with lengths up to at least 2 km
- **400 Gb/s Related**
  - Define ~~a~~ physical layer specifications that supports 400 Gb/s operation over 2 pairs of SMF with lengths up to at least 500 m
  - Define ~~a~~ physical layer specifications that supports 400 Gb/s operation over 2 pairs of SMF with lengths up to at least 2 km
- **800 Gb/s Related**
  - Define ~~a~~ physical layer specifications that supports 800 Gb/s operation over 4 pairs of SMF with lengths up to at least 500 m
  - Define ~~a~~ physical layer specifications that supports 800 Gb/s operation over 4 pairs of SMF with lengths up to at least 2 km
  - Define ~~a~~ physical layer specifications that supports 800 Gb/s operation over 4 wavelengths over a single SMF in each direction with lengths up to at least 2 km
- **1.6 Tb/s Related**
  - Define ~~a~~ physical layer specifications that supports 1.6 Tb/s operation over 8 pairs of SMF with lengths up to at least 500 m
  - Define ~~a~~ physical layer specifications that supports 1.6 Tb/s operation over 8 pairs of SMF with lengths up to at least 2 km

Proposed IEEE P802.3dj Objectives Update (1 of 3)

- **Non-Rate Specific**
  - Support full-duplex operation only
  - Preserve the Ethernet frame format utilizing the Ethernet MAC
  - Preserve minimum and maximum FrameSize of current IEEE 802.3 standard
  - Support a BER of better than or equal to 10 -13 at the MAC/PLS service interface (or the frame loss ratio equivalent)
  - Provide support to enable mapping over OTN
- **200 Gb/s Related**
  - Support a MAC data rate of 200 Gb/s
  - Support optional single-lane 200 Gb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
  - Define a physical layer specification that supports 200 Gb/s operation:
    - over 1 lane over electrical backplanes supporting a die-to-die insertion loss <= 40 dB at 53.125 GHz **
    - over 1 pair of copper twin-axial cables in each direction with a reach of up to at least 1.0 meter
  - Define physical layer specifications that supports 200 Gb/s operation:
    - over 1 pair of SMF with lengths up to at least 500 m
    - over 1 pair of SMF with lengths up to at least 2 km
- **400 Gb/s Related**
  - Support a MAC data rate of 400 Gb/s
  - Support optional two-lane 400 Gb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
  - Define a physical layer specification that supports 400 Gb/s operation:
    - over 2 lanes over electrical backplanes supporting a die-to-die insertion loss <= 40 dB at 53.125 GHz **
    - over 2 pairs of copper twin-axial cables in each direction with a reach of up to at least 1.0 meter
  - Define physical layer specifications that supports 400 Gb/s operation:
    - over 2 pairs of SMF with lengths up to at least 500 m
    - over 2 pairs of SMF with lengths up to at least 2 km

Proposed IEEE P802.3dj Objectives Update (2 of 3)

- **800 Gb/s Related**
  - Support a MAC data rate of 800 Gb/s
  - Support optional four-lane 800 Gb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
  - Define a physical layer specification that supports 800 Gb/s operation:
    - over 4 lanes over electrical backplanes supporting a die-to-die insertion loss <= 40 dB at 53.125 GHz **
    - over 4 pairs of copper twin-axial cables in each direction with a reach of up to at least 1.0 meter
    - over 1 wavelength over a single SMF in each direction with lengths up to at least 10 km *
    - over 4 wavelengths over a single SMF in each direction with lengths up to at least 10 km *
    - over a single SMF in each direction with lengths up to at least 40 km
  - Define physical layer specifications that supports 800 Gb/s operation:
    - over 4 pairs of SMF with lengths up to at least 500 m
    - over 4 pairs of SMF with lengths up to at least 2 km
    - over 4 wavelengths over a single SMF in each direction with lengths up to at least 2 km

*** - Approved by IEEE 802.3 WG 16 Mar 2023**
**** - Approved by IEEE 802.3 WG 18 May 2023**

Proposed IEEE P802.3dj Objectives Update (3 of 3)

- **1.6 Tb/s Related**
  - Support a MAC data rate of 1.6 Tb/s
  - Support optional sixteen-lane 1.6 Tb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
  - Support optional eight-lane 1.6 Tb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
  - Define a physical layer specification that supports 1.6 Tb/s operation:
    - over 8 lanes over electrical backplanes supporting a die-to-die insertion loss <= 40 dB at 53.125 GHz **
    - over 8 pairs of copper twin-axial cables in each direction with a reach of up to at least 1.0 meter
  - Define physical layer specifications that supports 1.6 Tb/s operation:
    - over 8 pairs of SMF with lengths up to at least 500 m
    - over 8 pairs of SMF with lengths up to at least 2 km

*** - Approved by IEEE 802.3 WG 16 Mar 2023**
**** - Approved by IEEE 802.3 WG 18 May 2023**

Options Moving Forward

1. Continue with current objectives
   - The standard would not provide any guidance regarding a given "class" of the transmitter
     - The current proposal allows two classes of transmitters
   - No path towards a unique name to identify:
     - Class of transmitter
     - Whether a PHY provides bypass mode
2. Add new objectives to allow development of new PHY types
   - Determine "distinctness"
     - Latency? (Supporting presentation would be needed)
       - Focus on reduced latency of missing layer?

IEEE 802.3df & Latency (Delay)

Table 169–4—Sublayer delay constraints (800GBASE)

Sublayer	Maximum (bit time)[a]	Maximum (pause_quanta)[b]	Maximum (ns)	Notes[c]
800G MAC, RS, and MAC Control	196 608	384	245.76	See 170.1.4.
800GBASE-R PCS or 800GXS[d]	640 000	1250	800	See 172.5.
800GBASE-R PMA	36 864	72	46.08	See 173.5.4.
800GBASE-KR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through backplane medium. See 163.5.
800GBASE-CR8 PMD	32 768	64	40.96	Includes allocation of 14 ns for one direction through cable medium. See 162.5.
800GBASE-VR8 PMD	32 768	64	40.96	Includes 2 m of fiber. See 167.3.1.
800GBASE-SR8 PMD	32 768	64	40.96	Includes 2 m of fiber. See 167.3.1.
800GBASE-DR8 PMD	32 768	64	40.96	Includes 2 m of fiber. See 124.3.1.
800GBASE-DR8-2 PMD	32 768	64	40.96	Includes 2 m of fiber. See 124.3.1.

- The standard specifies individual sub-layer delay
  - The aggregate of layers is the maximum latency.
  - This is specified for PAUSE buffer sizing to enable maximum implementation flexibility
  - The focus isn't low latency, i.e. "maximum" is specified
- Focus on layer not being present?
  - Applies if you are considering 200G / lane implementations
  - But….
    - Current 100G/lane based solutions does not have inner FEC, so how would 200G/lane solutions with FEC Bypass be lower latency?

Options Comparison

Option	Pro's	Con's	Questions?
1	• No new objectives necessary • Interoperability between two modes	• Full duplex "low-latency" link not ensured • Rx's (today / tomorrow) must support both modes	• Would market accept this approach? • Will some approach to identifying implementations supporting different Tx classes / low latency be desired?
2	• PHY names provide easy identification of implementations supporting different Tx classes, i.e. low latency • Optimized solutions	• New objectives needed	• Distinct Identity to support new objectives? • Is there a desire to achieve interoperability between objectives?

Summary

- Either Option #1 or Option #2 can be made to work
  - Each comes with different hurdles and questions to address
- If Option #2 is the preferred approach –
  - Presentation addressing distinct identity concerns will be necessary
  - Presentation proposing new objectives will be needed

Proposed Strawpoll

As chair, direction from ad hoc would be helpful!

I would support adding objectives to support physical layer specifications based on only RS544 FEC for:

- 200GBASE-DR1, 200GBASE-FR1,
- 400GBASE-DR2, 400GBASE-DR2-2,
- 800GBASE-DR4, 800GBASE-DR4-2, 800GBASE-FR4,
- 1.6TBASE-DR8, and 1.6TBASE-DR8-2

- Yes
- No
- Abstain

BACK-UP

Background – Related Motions

- Related Motions
  - Mar 2023 –
    - Motion #5 - Move to:
      - Adopt patra_3dj_01b_2303 slides 6 to 8, 13, 14, and 20 to 23 as part of the FEC approach for
      - 800GBASE-DR4, 800GBASE-DR4-2, 800GBASE-FR4
      - 400GBASE-DR2, 400GBASE-DR2-2* (Note: 400GBASE-DR2-2 pending WG objective approval)
      - 200GBASE-DR1, 200GBASE-FR1

      with FEC lane rate, convolutional interleaver details, and 1.6T support to be determined later
  - May 2023
    - Motion #9 - Move to:
    - Adopt patra_3dj_01b_2303 slides 6 to 8, 13, 14, and 20 to 23 as part of the FEC approach for 800GBASE-LR4 with FEC lane rate and convolutional interleaver details to be determined later
  - July 2023
    - Motion #4 - Move to adopt the direction of adding an option to support only RS544 FEC (aka Bypass Inner FEC) for the single wavelength 500m and 2km optical PMDs with the mechanism to enable it remaining TBD
    - Motion #9 - Move to adopt the same inner FEC architecture used for 200GbE/400GbE/800GbE for 1.6TbE SMF optical PMDs (500m/2km)
    - Motion #11 - Move to adopt the FEC_I sublayer architecture with 200G throughput convolutional interleaver as shown in slides 6-11 of he_3dj_01_2307 for 200G/400G/800G/1.6TbE

Relevant Presentations

- https://www.ieee802.org/3/dj/public/23_05/welch_3dj_03c_2305.pdf
- https://www.ieee802.org/3/dj/public/adhoc/optics/0623_OPTX/dudek_3dj_optx_01_230629.pdf
- https://www.ieee802.org/3/dj/public/23_07/welch_3dj_04a_2307.pdf
- https://www.ieee802.org/3/dj/public/23_07/welch_3dj_03a_2307.pdf