# PCS and FEC sublayer considerations

Tom Huber, Nokia

802.3df Task Force

18 May 2022

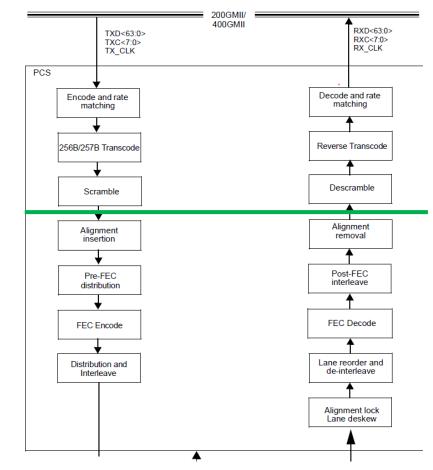# Supporters

Steve Gorshe, Microchip

# Introduction

- There is broad agreement that there is a need to use different FEC codes and FEC architectures with different PMDs
  - Which FEC code(s) to use with 200G lanes is not known
  - Every 800G and 1.6T PHY will be required to implement at least one FEC code, but is not known if any FEC code will be used by every PHY, so in effect every FEC is optional in the context of the overall architecture
    - Each PHY will have to indicate which FEC(s) are mandatory
    - Analogous to what was done with 100G PHYs, where FEC was optional in the context of the overall architecture, but mandatory for many PHYs
- While it is desirable to reuse as much of the clause 119 PCS as is practical, more discussion is needed regarding some of the processes
  - Rate adaptation
  - Alignment markers
  - FEC
  - OTN mapping reference point

# RS(544,514) vs other FECs in existing PHYs

- RS(544,514) as used with 200GBASE-R and 400GBASE-R
  - The ratio of 257b blocks to FEC codewords is fixed at 20
    - In other words, the FEC frame and payload are locked to each other
  - 1/20k blocks is enough space for alignment markers, which enables the space for them to be created via the Idle insert/delete process that is used for rate adaptation
- Other FECs (e.g., those used by 100GBASE-ZR, 400GBASE-ZR)
  - The codeword size is not an easy integer multiple of 257b, and may be quite large
  - There is a separate FEC frame into which the PCS blocks are mapped as payload, using an asynchronous mapping process for rate compensation
  - Alignment markers may not have the 1/20k ratio; the ratio will depend on the details of FEC frame
- Lane striping will almost certainly need to be different for different FECs

# PCS and FEC as separate sublayers

- Since there are multiple FEC codes and architectures, and none are required for all PHYs, it makes sense to describe the PCS and FEC separately
  - One PCS for 800G and 1.6T with aspects that are common
  - One sublayer per FEC with the aspects that vary per-FEC
- Since alignment markers and lane striping are potentially different for each FEC, a reasonable starting point for splitting the layers is between the scrambler and AM insertion processes
- If the scrambling also varies per-FEC, then that process should also be part of the FEC sublayer
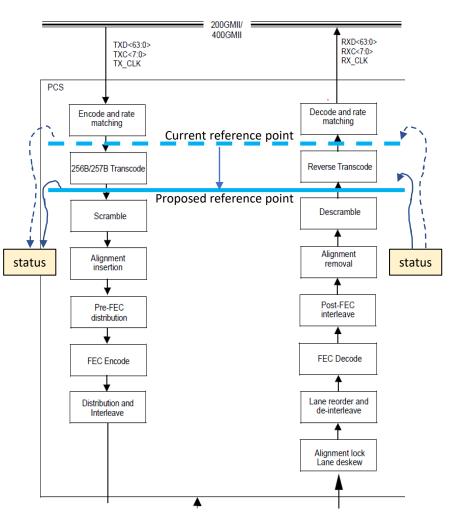
# Preserving flexibility in 800G and 1.6T

- While it is desirable to reuse elements of the clause 119 PCS, we need to be careful to not retain elements that would overly constrain the solution for FECs other than RS(544,514)
- Allocating the rate compensation process to the PCS sublayer and the alignment marker processing to the FEC sublayer has some undesirable properties:
  - Rate compensation based on Idle insert/delete in the PCS requires looking at every block, which becomes more challenging as speeds increase, and may create increased challenges for PTP
  - 1/20k happens to be a good ratio for alignment markers when RS(544,514) FEC is used, but is likely suboptimal for other FEC codes
  - The rate of the logical unstriped PCS signal that is mapped into any other FEC frame besides RS(544,514), or into an OTN frame, is an awkward 'underclocked' rate of (20k-1)/20k times the MAC rate that is convenient for RS(544,514) FEC but not for other FECs
- Insertion of alignment markers before inserting the FEC and striping to lanes leads to awkward description to ensure that right AMs end up in the right places after lane striping
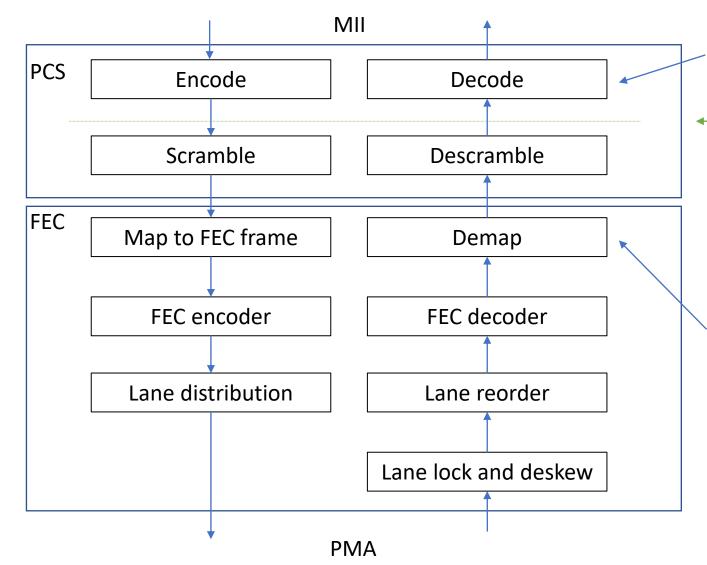
# OTN mapping reference point

- The OTN reference point defines what information must be carried over OTN
  - The most important aspect of the OTN mapping reference point (from SG15's perspective) is that it has a canonical format for all PHYs at a given MAC rate, so a single mapping can be defined

- 66b codewords have been used as the 'OTN mapping reference point' for 100G, 200G, and 400G Ethernet clients
  - This was done at 100G because 66b blocks are the canonical format for 100G PMDs
  - The reference point was retained for 200G and 400G, except that the 3-bit status carried in the AMs (but not the AMs themselves) is also expected to be carried over OTN
    - Enabled reuse of the mapping for 100G, which was convenient, but not required
    - Increased the bit rate difference between OTN and native Ethernet, which is not desirable

- For 800G and 1.6T, it would be better to select a reference point that uses whatever line coding native Ethernet uses to reduce the difference in bit rates
  - Clearly the FEC and AMs do not need to be carried over the OTN
  - OTN has its own scrambler, so better to descramble the Ethernet signal before mapping into OTN
  - A sensible reference point would be at the input to the scrambler/output of the descrambler
  - Further discussion with Q11/15 would be appropriate

# Proposal

- Specify the PCS and FEC as separate sublayers to maximize flexibility in which FEC(s) are used by which PHYs
  - PCS contains processes that are common, FEC has those that are specific to the FEC code
- Remove rate compensation from the PCS and put in the FEC sublayers
  - PCS rate is locked to the MAC rate based on the PCS coding that is presented to the FEC sublayer (e.g., 257/256 times the MAC rate, if 257b coding is retained)
  - Treat the 'output' of the PCS as a logical signal at the encoded bit rate
  - In the special case of the RS(544,514) FEC, no rate compensation is needed; the PHY rate is (1+20k)/20k times the MAC rate after the alignment markers are added
  - In all other cases, the PHY rate is based on the FEC frame, and an asynchronous mapping of the PCS into that frame is used
- Arrange the FEC encoding, rate compensation, alignment marker, and lane distribution processes within the FEC sublayer to facilitate simple description, including the need to forward FEC degrade information to the OTN mapping reference
- Define the OTN mapping reference point to be at the input of the scrambler and output of the descrambler, including the FEC degrade information
- The figure on the next slide provides a high-level illustration of this architecture

# High-level process architecture

MII

**PCS**

| Encode | Decode |

*OTN mapping reference point*

| Scramble | Descramble |

**FEC**

| Map to FEC frame | Demap |

| FEC encoder | FEC decoder |

| Lane distribution | Lane reorder |

| Lane lock and deskew |

PMA

Encode/decode can point to existing clauses on 66b coding plus 257b transcoding, assuming 257b code is retained

Map and demap are specific to each FEC, but includes mapping to the FEC frame, inserting alignment markers and rate adaptation
- For RS(544,514), the FEC frame is created by inserting AMs every 20k blocks, and no rate adaptation is needed.
- For other FECs, something like GMP is used to map into the FEC frame and provide rate compensation

# Thanks!