

# Ethernet Metadata Services PCS and FEC baseline proposal

Eugene Opsasnick, Broadcom

Adee Ran, Cisco

IEEE P802.3dt

IEEE 802.3 2026 May Interim – Munich, GER

# Supporters

- Kapil Shrikhande (Upscaleai)
- Viet Tran (Keysight)
- Gary Nicholl (Cisco)
- Xiang He (Huawei)
- Ulf Parkholm (Ericsson)

# Outline

- General Approach and Definition Considerations
- Changes needed in 64B/66B coding
- Additional changes needed in the PCS and FEC sublayers to protect preamble and CtIOS using RS-FEC
- New EMS subclause contexts in each xBASE-R PCS
- Backup: Complimentary changes needed in Reconciliation sublayer

# General Approach and Definition Considerations

# To allow metadata to pass through the PCS

- Per-Packet Metadata: preamble
  - No PCS changes needed
  - Preamble is modified in Reconciliation Sublayer (RS) – it is data to PCS
- Packet-independent Metadata: 8-byte Control Ordered Sets (CtIOS)
  - Need to define CtIOS encoding on xMII (in RS)
  - Must modify the 64B/66B PCS coding to support CtIOS
  - Must allow CtIOS to interrupt packet data
    - Requires Stateless 64B/66B Encoder/Decoder as defined in latest PCSs
- Must prevent RX “False CtIOS arrivals” and “bad preamble metadata”
  - Requires discard of all data in received uncorrectable RS-FEC codewords (set to EBLOCK\_R)
  - Requires not enabling FEC bypass correction
  - Requires not enabling FEC bypass indication

# General Approach to Support CtIOS

- Any changes to current PCS and FEC clauses must be optional
  - So current implementations do not become non-compliant
- In each relevant PCS sublayer
  - Add a subclause that specifies the required options to support EMS
    - Similar approach as new timesync subclauses in 802.3dj
  - For example:
    - Must use stateless 64B/66B encoder/decoder
    - Must use RS-FEC – to protect preamble and CtIOS
    - Must use option to mark all 66-bit blocks of uncorrectable FEC codewords as ERROR block
  - Add any missing functionality as optional to PCS and FEC sublayers
- Any needed changes in RS will be added to a new EMS clause
  - PCS changes need some complimentary changes in RS
  - Assume a "generic RS" (gRS) will be defined in the new EMS Clause

# Current 802.3 Clauses of Interest

- 100GBASE-R and 100GBASE-P
  - Clause 81 – RS
  - Clause 82 – PCS
  - Clause 91 – RS-FEC (non-interleaved)
  - Clause 161 – RS-FEC (interleaved)
- 50GBASE-R
  - Clause 132 – RS
  - Clause 133 – PCS
  - Clause 134 – RS-FEC
- 200GBASE-R and 400GBASE-R
  - Clause 117 – RS
  - Clause 119 – PCS (includes RS-FEC)
- 800GBASE-R and 1.6TBASE-R
  - Clause 170 – RS for 800GbE and 1.6TbE
  - Clause 172 – PCS for 800GbE (includes RS-FEC)
  - Clause 175 – PCS for 1.6TbE (includes RS-FEC)

# PCS 64B/66B Definition Considerations

- Clause 82 defines the 64B/66B block formats in Figure 82-5
  - All other PCS clauses of interest refer to it.
  - Need to add optional block code changes in Clause 82 to support CtIOS
- Stateless encoder/decoder is defined in Clause 119 (in 802.3dj)
  - Stateless encoder is optional in CL 119 PCS for 200/400GbE
  - Clauses 172 and 175 refer to the definition in Clause 119.
  - Clause 82, 100G PCS, needs to add option to use stateless encoder.
  - Clause 133, 50G PCS, refers to Clause 82.
    - Stateless encoder/decoder option would be included by the current reference after it is added to Clause 82.

# RS-FEC Definition Considerations

- RS-FEC is a required to prevent bad EMS data reception
  - Clause 91 has “original” RS[544] definition
    - Includes FEC bypass correction option
  - Clause 161 (100G interleaved FEC) refers to Clause 91 with some changes
  - Clause 134 (50G FEC) refers to Clause 91 with some small modifications
    - Number of PCS lanes and alignment marker changes
    - FEC bypass correction feature is not supported
  - Clause 119 PCS contains complete RS[544] definition (no x-refs to CL 91)
    - Some changes to the RS-FEC decoder error handling
  - Clause 172 and 175 PCS/FEC refer to Clause 119 with small exceptions
- Significant differences between the CL 91 and CL 119 RS-FEC decoders
  - FEC bypass correction feature is removed
  - Clause 119 FEC decoder invalidates all blocks of uncorrectable codewords
  - Clauses 91, 134, and 161 only invalidate every 8<sup>th</sup> 66-bit block of uncorrectable codewords
    - An option needs to be added to invalidate all blocks, either in the FEC or PCS sublayer.
    - These FEC clauses include the 256b/257b transcoder.
    - The descrambler and 64b/66b decoder are in the CL 82 PCS.

# Changes needed in 64B/66B coding

# Changes to 64B/66B Encode/Decode (Clause 82) - 1

Input Data	S y n c	Block Payload									
Bit Position:	0 1 2	65									
<b>Data Block Format:</b>											
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> D <sub>5</sub> D <sub>6</sub> D <sub>7</sub>	01	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>	D <sub>7</sub>		
<b>Control Block Formats:</b>											
<b>Block Type Field</b>											
C <sub>0</sub> C <sub>1</sub> C <sub>2</sub> C <sub>3</sub> C <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0x1E	C <sub>0</sub>	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	
S <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> D <sub>5</sub> D <sub>6</sub> D <sub>7</sub>	10	0x79	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>	D <sub>7</sub>		
O <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> Z <sub>4</sub> Z <sub>5</sub> Z <sub>6</sub> Z <sub>7</sub>	10	0x4B	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	O <sub>0</sub>	0x000_0000				
T <sub>0</sub> C <sub>1</sub> C <sub>2</sub> C <sub>3</sub> C <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0x87			C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> T <sub>1</sub> C <sub>2</sub> C <sub>3</sub> C <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0x99	D <sub>0</sub>			C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> T <sub>2</sub> C <sub>3</sub> C <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0xAA	D <sub>0</sub>	D <sub>1</sub>			C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> T <sub>3</sub> C <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0xB4	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>			C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> T <sub>4</sub> C <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0xCC	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>			C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> T <sub>5</sub> C <sub>6</sub> C <sub>7</sub>	10	0xD2	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>			C <sub>6</sub>	C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> D <sub>5</sub> T <sub>6</sub> C <sub>7</sub>	10	0xE1	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>			C <sub>7</sub>
D <sub>0</sub> D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> D <sub>5</sub> D <sub>6</sub> T <sub>7</sub>	10	0xFF	D <sub>0</sub>	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>		

- Block Type 0x4B:
  - Current definition works for Sequence Ordered Sets but not for CtIOS

- Add optional support for new definition of block type 0x4B which works for both CtIOS and Sequence ordered sets.

Figure 82-5—64B/66B block formats

O <sub>0</sub> * D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub> D <sub>5</sub> D <sub>6</sub> D <sub>7</sub>	10	0x4B	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>	D <sub>7</sub>
---	----	------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

# Changes to 64B/66B Encode/Decode (Clause 82) - 2

Table 82-1—Control codes

Control character	Notation	XLGMII/ CGMII control code	40/100GBASE-R O code <sup>a</sup>	40GBASE-R and 100GBASE-R control code
idle	/I/	0x07		0x00
LPI	/LI/	0x06		0x06
start	/S/	0xFB		Encoded by block type field
terminate	/T/	0xFD		Encoded by block type field
error	/E/	0xFE		0x1E
Sequence ordered set	/Q/	0x9C	0x0	Encoded by block type 0x4B plus O code, control codes are set to 0x00

<sup>a</sup>INCITS T11 Fibre Channel uses O code 0xF for the signal ordered set. OIF uses O code 0x5 for the FlexE [B59] ordered set. ITU-T uses O code 0xC for the G.8312 [B47] MTN path ordered set.

Control ordered set <sup>b</sup>	/C/	0x5C	0x1 to 0xF	Encoded by block type 0x4B and a nonzero O code
----------------------------------	-----	------	------------	--

- Add text to footnote that “UEC uses O-code 0x6 for UE-defined Control ordered sets”.
- Add new definition and optional support for “Control ordered set”

# Changes to 64B/66B Encode/Decode (Clause 82) - 3

- Add option to use stateless PCS encoder/decoder in CL 82
  - Use same definition as used in CL 119 (not the CL 172 definition)
- Note: RS-FEC (RS[544]) must be active to protect CtIOS and preamble
  - Stateless encoder/decoder is not allowed without RS-FEC.

Additional changes needed in the PCS and FEC sublayers  
to protect EMS data with RS-FEC

# PCS and RS-FEC Decoder Functional Changes – 200+ GbE

- 200/400GbE (CL 119), 800GbE (CL 172), 1.6TbE (CL 175)
  - No functional changes needed.

# PCS and RS-FEC Decoder Functional Changes – 50/100GbE

- Two possible approaches to invalidate all blocks of uncorrectable FEC codewords
  1. Using Clause 119 as model with changes mainly in FEC
  2. No changes to FEC sublayer – Add error extension in PCS
- 1. Changes mainly in FEC Clauses (CL 91, CL 161, CL 134)
  - Add option to invalidate all blocks of bad codeword(s) with extension to first 4 blocks of next codeword (wording similar to 119.2.5.3).
  - Add option to PCS (CL 82) to disable BER monitor
- 2. Changes in PCS only (CL 82)
  - Add option to extend any invalid sync header by 7 more blocks in 64B/66B decoder.
  - No changes to FEC clauses or BER monitor

# PCS and RS-FEC Decoder Changes – 100GbE – Option 1

- 100 GbE RS-FEC (CL 91 and CL 161)
  - Add option to mark all blocks of uncorrectable codewords as invalid
  - Modifications needed to subclause 91.5.3.3 and 161.5.3.3
  - Note: Error extension due to descrambler must also be considered and the following 257-bit block after the uncorrectable codeword(s) must also be marked as invalid.
- 100 GbE PCS (CL 82)
  - Add option to disable the BER monitor process (Figure 82-15 state diagram)
    - Clause 82 sets hi\_ber to true after reception of 97 bad blocks.
    - Option needed so marking of all blocks of uncorrectable codewords does not set hi\_ber.
    - PCS clauses 119, 172, 175 do not have a BER monitor process.
    - The BER monitor is not needed in CL 82 when RS-FEC is used.
  - Note: No change needed to Block lock process (Figure 82-12 state diagram)
    - Block lock is lost after 65 bad sync headers are received on any one PCS lane
    - 50GbE and 100 GbE FEC decoders will not hit this threshold if RS-FEC invalidates all uncorrectable blocks (per-codeword: 4 blocks per 100G PCSL, and 20 blocks per 50G PCSL)

# PCS Changes – 100GbE – Option 2

- 100 GbE RS-FEC (CL 91 and CL 161)
  - No changes
- 100 GbE PCS (CL 82)
  - Add option to use stateless encoder from Clause 119
  - Add option to use stateless decoder from Clause 172, but with changes to error the extension
    - Clause 172 stateless decoder (incorrectly) extends ERRORS by 1 block.
    - Change error extension from 1 block to 7 blocks (as defined in UEC).

# PCS and RS-FEC Decoder Changes – 50GbE – Option 1

- 50GbE RS-FEC (CL 134)
  - Add option to mark all blocks of uncorrectable codewords as invalid
  - Modifications needed to subclause 134.5.3.3
  - Note: Error extension due to descrambler must also be considered and the following 257-bit block after the uncorrectable codeword(s) must also be marked as invalid.
- 50GbE PCS (Clause 133)
  - All PCS functions are identical to those of 40GBASE-R in CL 82 by cross-reference with some exceptions for the difference in data rate and AMs (see 133.2.1).
  - The option to disable hi\_ber monitor process would be included by the current cross-reference after the option is added to CL 82.

# PCS Changes – 50GbE – Option 2

- 50GbE RS-FEC (CL 134)
  - No changes
- 50GbE PCS (CL 133)
  - All PCS functions are identical to those of 40GBASE-R in CL 82 by cross-reference with some exceptions for the difference in data rate and AMs (see 133.2.1).
  - An optional stateless encoder/decoder is also required (same as 100GbE on slide 18) and would be included by the current cross-reference after the option is added to CL 82.

# New EMS subclause contexts in each xBASE-R PCS

# EMS Subclause for 200G/400G/800G/1.6T

- Add a subclause to each PCS to describe EMS support
  - Clauses 119, 172 and 175
- Introduction:
  - Ethernet Metadata Services provides a mechanism to transmit and receive per-packet and packet-independent metadata for PHYs that meet the following requirements and configuration restrictions
- Requirements and Restrictions:
  - Requires use of extended gRS functionally defined in new EMS clause
  - Shall use the stateless 64b/66b encoder/decoder
  - Shall use the alternate definition of 66b Block Type 0x4B as defined in CL 82
  - Shall not enable FEC Bypass Indication mode

# EMS Subclause for 100GbE

- Add a subclause to CL 82 PCS to describe EMS support
- Introduction:
  - Ethernet Metadata Services provides a mechanism to transmit and receive per-packet and packet-independent metadata for 100GBASE-R and 100GBASE-P PHYs that meet the following requirements and configuration restrictions
- Requirements and Restrictions:
  - Requires use of extended gRS functionally defined in new EMS clause
  - Shall use the stateless 64b/66b encoder/decoder
    - Option 1 uses the CL 119 stateless decoder, and option 2 uses a new stateless decoder.
  - Shall use the alternate definition of 66b Block Type 0x4B as defined in CL 82
  - The optional invalid AM checker (amp\_bad\_count) in Figure 91-8 is required.
  - Clause 91 or Clause 161 RS-FEC is required
  - RS-FEC decoder restrictions:
    - Shall not enable FEC Bypass Correction mode
    - Shall not enable FEC Bypass Indication mode
    - If option 1: Shall enable the option to invalidate all blocks of uncorrectable FEC codewords
  - If option 1: Shall disable the hi\_ber monitor process

# EMS Subclause for 50GbE

- Add a subclause to CL 133 PCS to describe EMS support
- Introduction:
  - Ethernet Metadata Services provides a mechanism to transmit and receive per-packet and packet-independent metadata for 50GBASE-R PHYs that meet the following requirements and configuration restrictions
- Requirements and Restrictions:
  - Requires use of extended gRS functionally defined in new EMS clause
  - Shall use the stateless 64b/66b encoder/decoder
    - Option 1 uses the CL 119 stateless decoder, and option 2 uses a new stateless decoder.
  - Shall use the alternate definition of 66b Block Type 0x4B as defined in CL 82
  - The optional invalid AM checker (amp\_bad\_count) in Figure 91-8 is required.
  - Clause 134 RS-FEC is required
    - Shall not enable FEC Bypass Indication mode defined in CL 134
    - If option 1: Shall enable the option to invalidate all blocks of uncorrectable FEC codewords
  - If option 1: Shall disable the hi\_ber monitor process

Thank You

# Backup

# Complimentary changes to the Reconciliation Sublayer

- RS Service Interface definition – modified preamble and CtIOS
  - In a separate baseline for a new gRS for EMS
- Changes/exceptions required to support CtIOS in PCS
  - Optional for current implementations
  - Clause 81:
    - Table 81-3 – Permissible encodings of TXC and TXD
      - Add support for TXC=1, TXD=0x5C : Control (only valid on lane 0)
    - Table 81-4 – Permissible lane encodings of RXD and RXC
      - Add support for RXC=1, RXD=0x5C : Control (only valid on lane 0)
    - Table 81-5 – Sequence ordered sets
      - Add similar table to define format of Control ordered sets – define O-code position