Updated logic baseline proposal for an 800GbE coherent PHY based on oFEC/C-band

Gary Nicholl, Mike Sluyski, Tom Williams - Cisco

IEEE 802.3dj Task Force Meeting, July 2023

Supporters

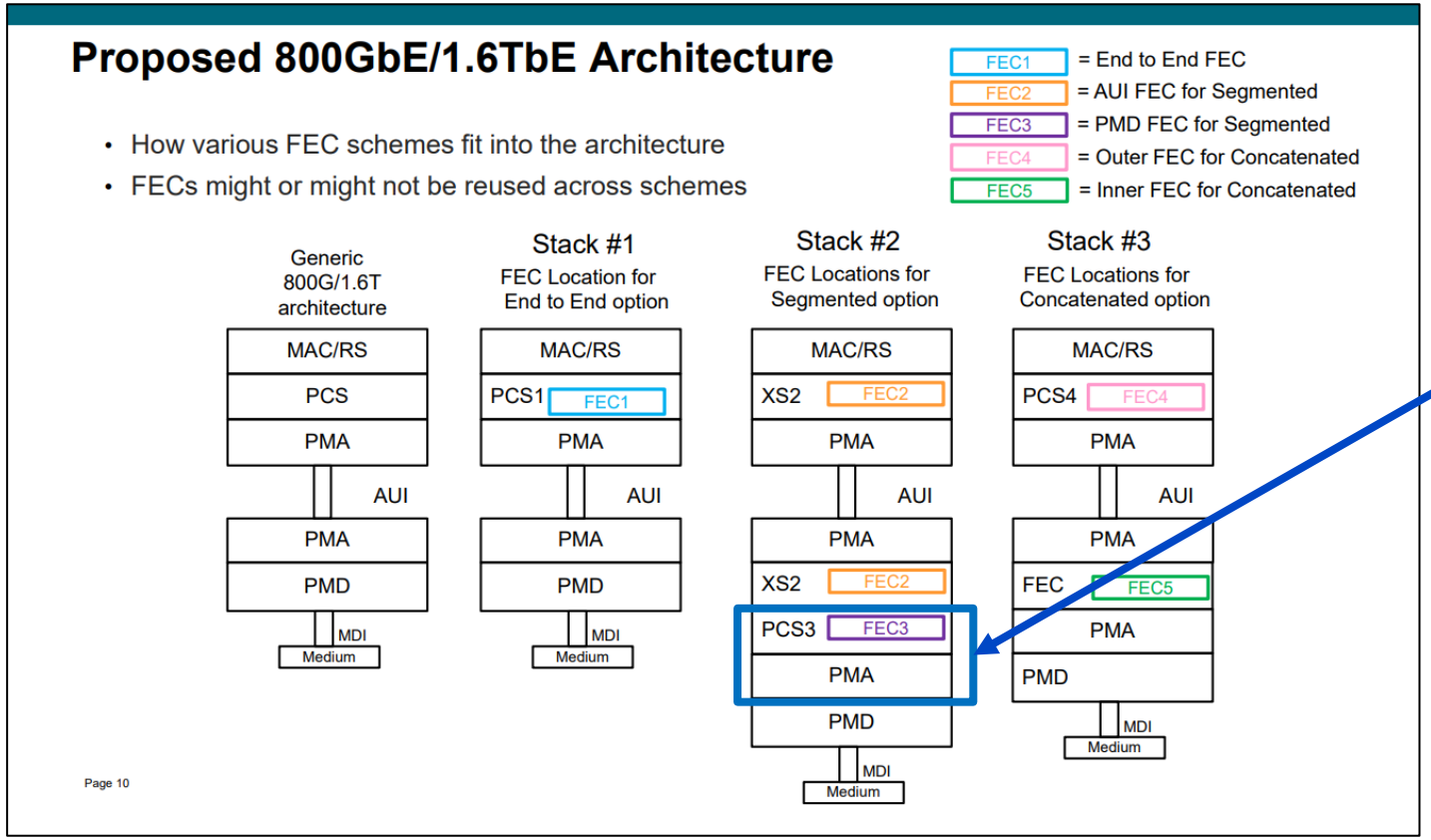
- Jörg-Peter Elbers, Adtran/Adva
- Ross Saunders, Adtran/Adva
- Vasudevan Parthasarathy, Broadcom
- Mark Gustlin, Cisco
- Ray Nering, Cisco
- Vipul Bhatt, Coherent
- Roberto Rhodes, Coherent
- Tomoo Takahara, Fujitsu Labs
- Ted Sprague, Infinera
- Jeffery Maki, Juniper Networks
- Jerry Pepper, Keysight
- Kumi Omori, NEC
- Yann Loussouarn, Orange
- Erwan Princeman, Orange
- Frank Chang, Source Photonics
- Dave Estes, Spirent
- Paul Brooks, Viavi
- Huijun Sha, Viavi
- Haojie Wang, China Mobile
- Xue Wang, H3C
- Yu Zhu, Hengtong Group
- Zhan Su, Ruijie Networks
- Rangchen Yu, SiFotonics
- ChiYuan Chen, Spirent
- Aihua Liu, ZTE
- Chengbin Wu, ZTE

Introduction

- nicholl_3dj_01a_2305 proposed a complete logic baseline for an 800GbE coherent PHY based on oFEC.
- This contribution provides a minor update to the proposal to reinstate the asynchronous GMP mapping (after extensive discussions during the May 23 meeting in San Antonio indicated that the original concerns over the impact of GMP mapping on PTP performance had been overstated).
- The baseline is primarily aimed at a single lambda 800GbE 40km application (800GBASE-ER1) but could also be used for other 800GbE coherent PHYs.

Adopted logic architecture for reference

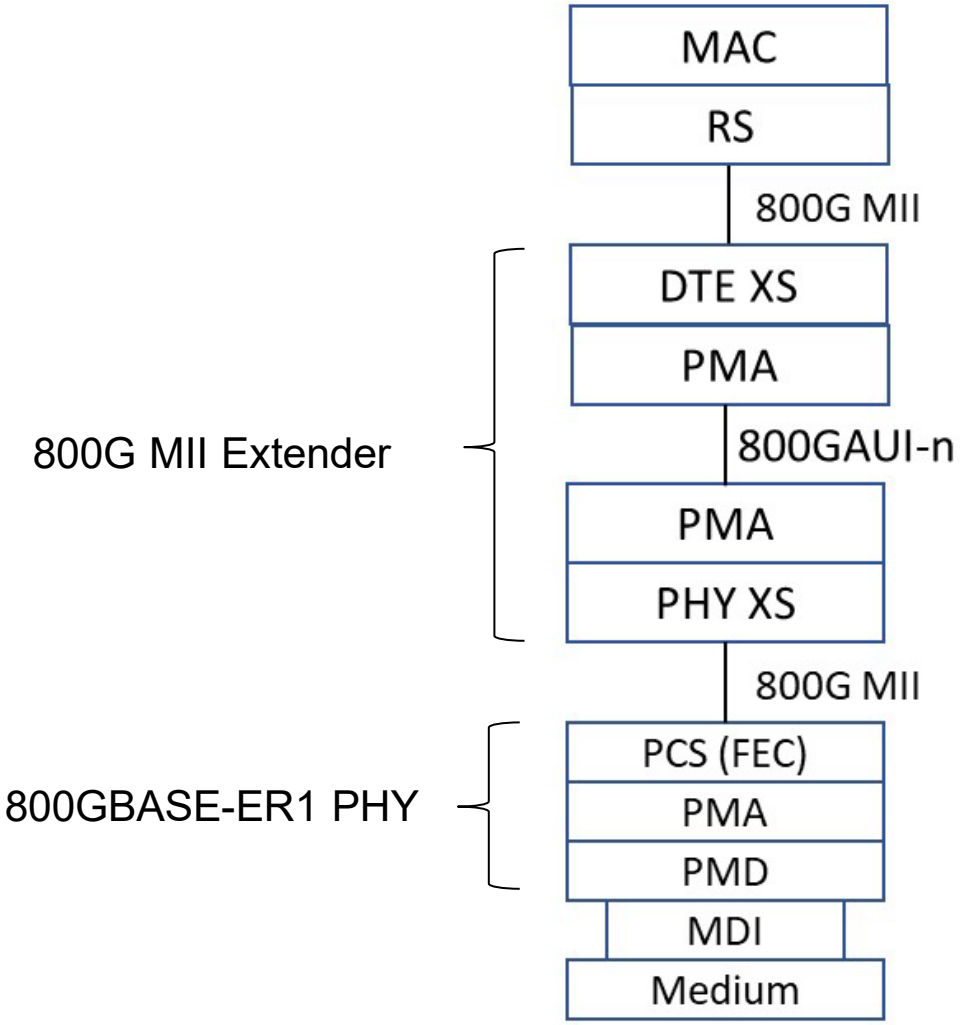
This proposal fits within the adopted 802.3dj logic architecture



Focus of this proposal

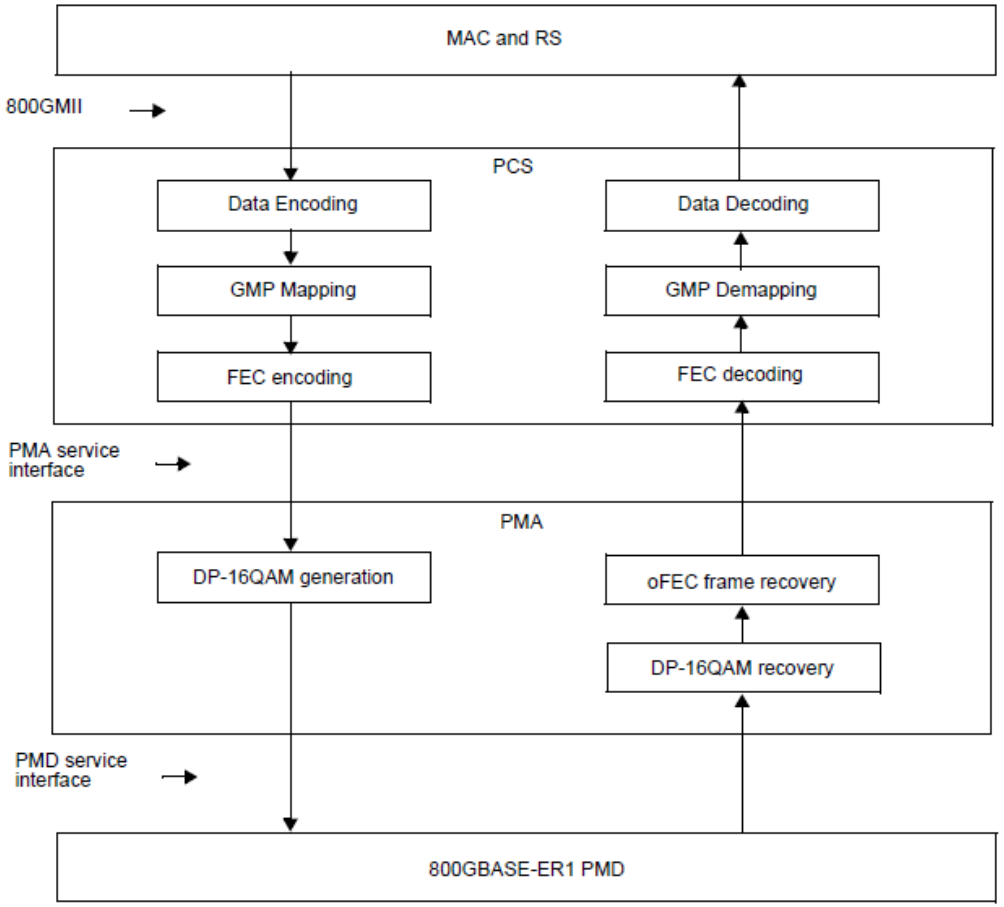
From: gustlin_3df_01a_220517

Alignment with 802.3dj architecture



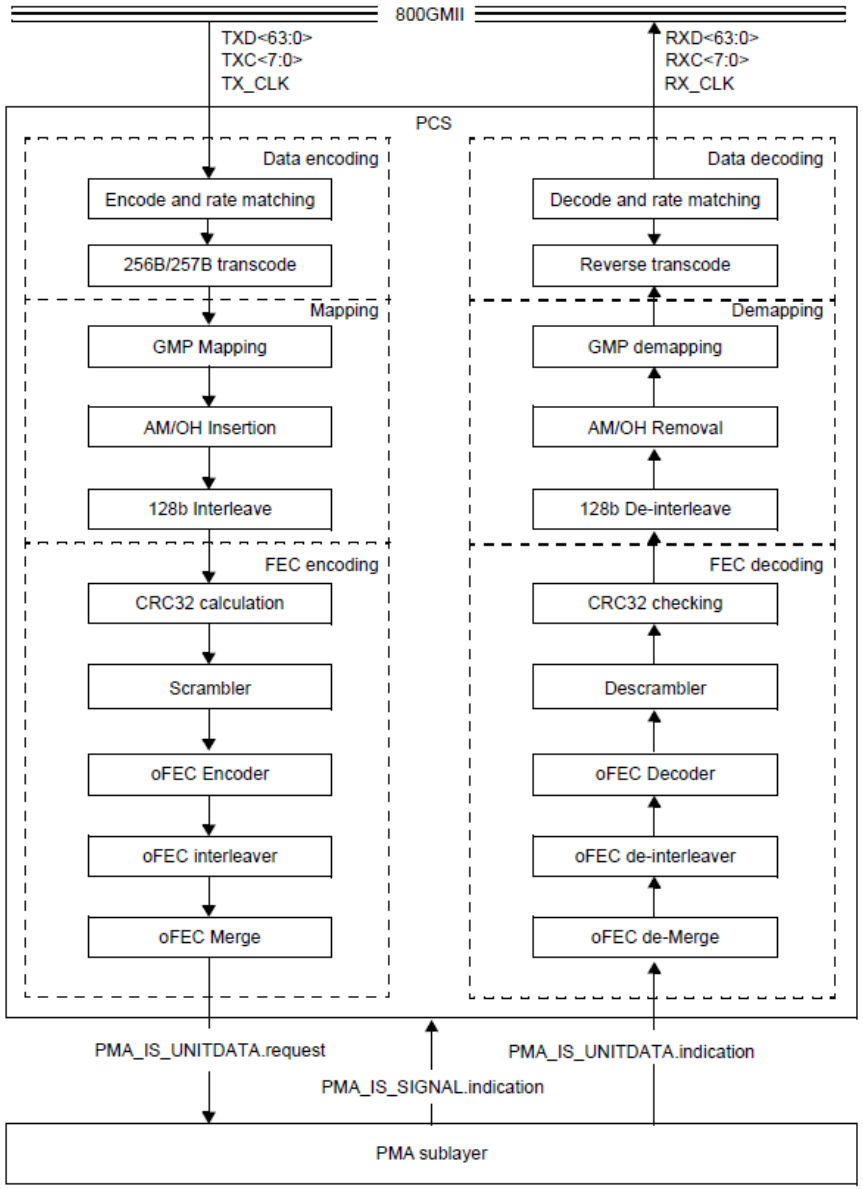
- Based on a Type 3 FEC/PHY scheme (“segmented FEC”) as described in [brown_3dj_optx_adhoc_01a_230222]
  - Same architecture as 400GBASE-ZR (802.3cw)
- The 800GBASE-ER1 PHYs only cover the optical links
  - No optional AUIs are supported within the PHY
- AUIs (if required) are supported using the 800GMII Extender
  - 800GAUI-8 defined in 802.3df (already done)
  - 800GAUI-4 will be defined in 802.3dj

800GBASE-ER1 PHY Overview



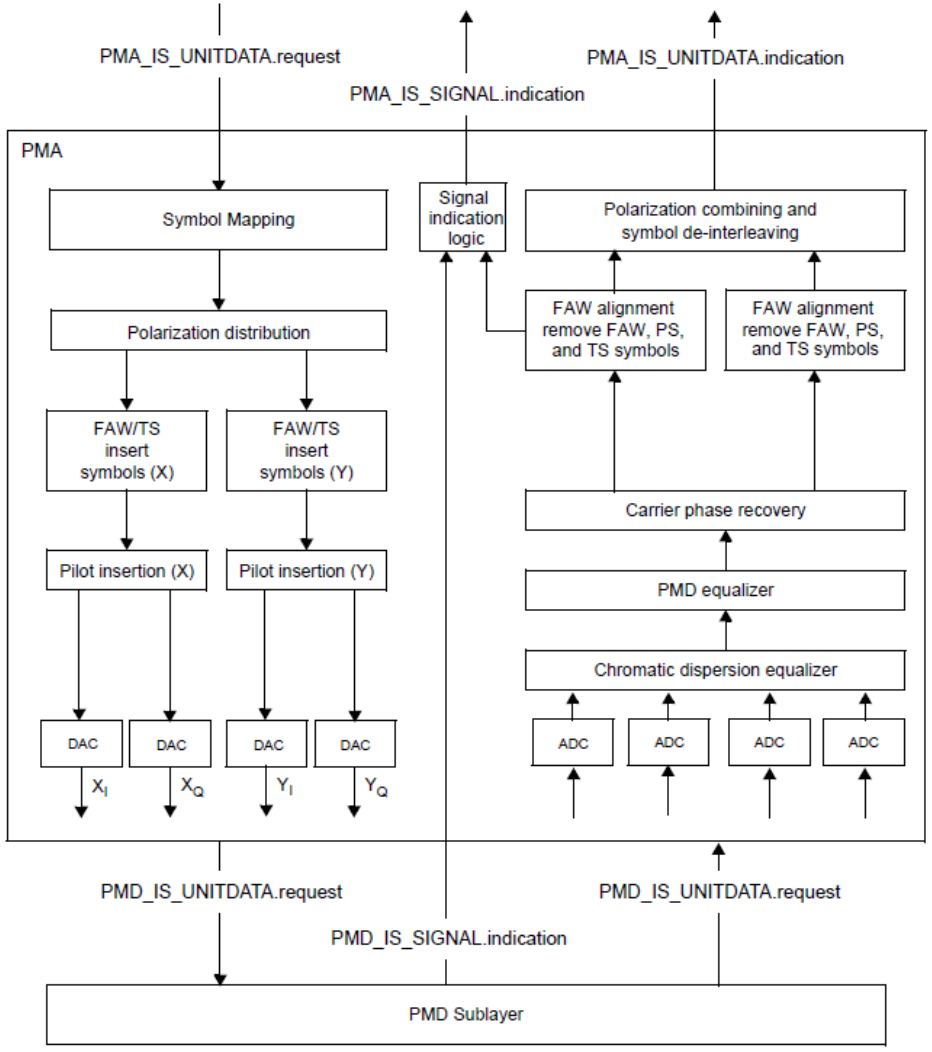
- Builds upon the efforts in 802.3cw to define an 802.3 PHY documentation structure to support a coherent optical interface
  - Split of functionality between PCS, PMA and PMD
  - Definition of PMA and PMD services interfaces
- PCS
  - 256/257b data encoding/decoding
  - GMP mapping/demapping
  - FEC encoding/decoding
    - based on oFEC defined for 800ZR/ZR+
- PMA
  - DP-16QAM generation/recovery

PCS - Functional Block Diagram



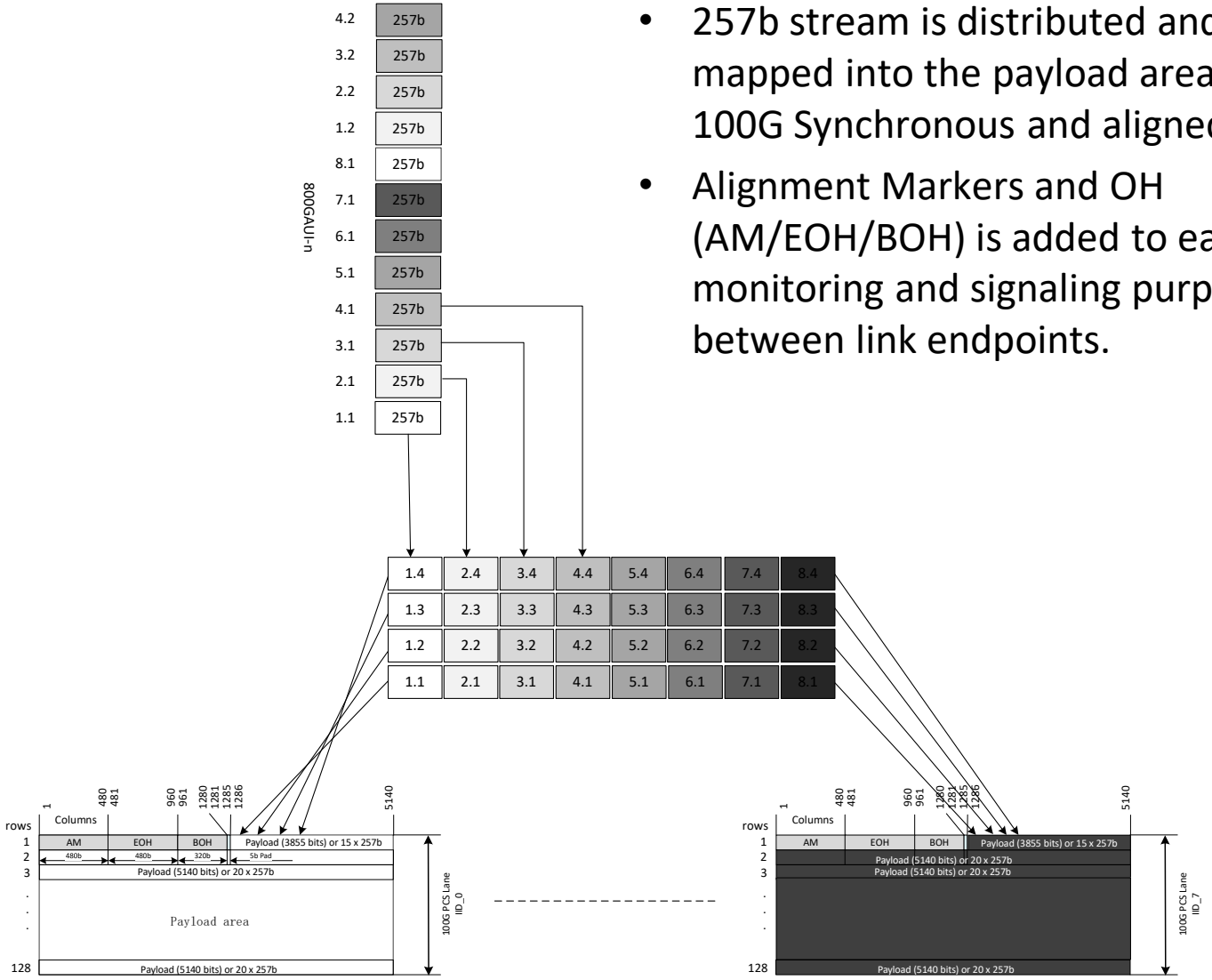
- Leverages efforts from IEEE 802.3cw, OIF 400ZR, OIF 800ZR, Open ROADM, and ITU-T SG15.
- Based on oFEC
- PCS functions:
  - 64B/66B encode, 256B/257B transcode
  - 257b de-interleave/interleave + GMP to/from 8 x 100G synchronous and aligned lanes
  - AM/OH (EOH/BOH) Insert/Remove
  - 128b interleave/de-interleave 100G lanes to/from 800G structure
  - CRC32 generate/check
  - oFEC Encode/decode

PMA - Functional Block Diagram

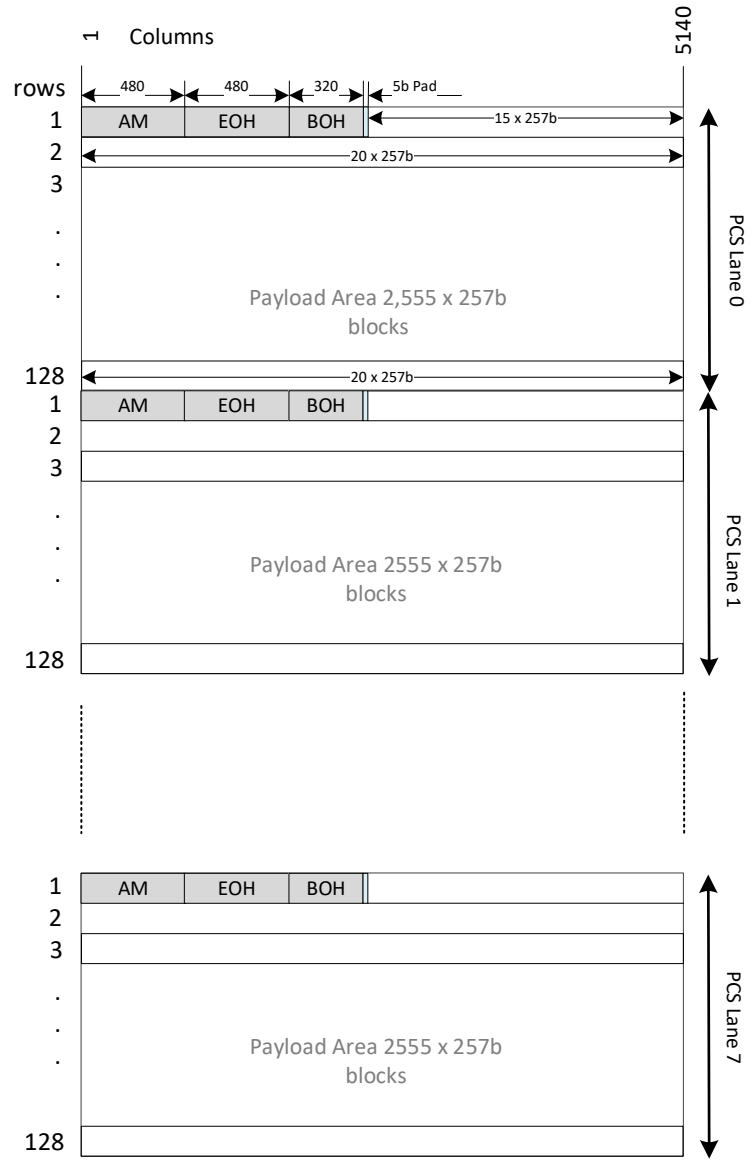


- DP-16QAM PMA
- Leverages heavily from the DP-16QAM PMA defined for 400GBASE-ZR (802.3cw, Clause 155)
- PMA functions:
  - 16QAM symbol generation and polarization distribution (X and Y)
  - DSP frame – 175,104 DP-16QAM symbols per X/Y polarization. Organized as 24 DSP sub-frames @ 7296:
    - Pilot Symbol (PS) spacing every 64 symbols - 2736 symbols
    - Training Symbols (TS) – 24 × 11 Symbols (first symbol shared as PS).
    - Frame Alignment Word (FAW) - 22 symbols
    - Reserved/User Defined (RES) - 74 symbols

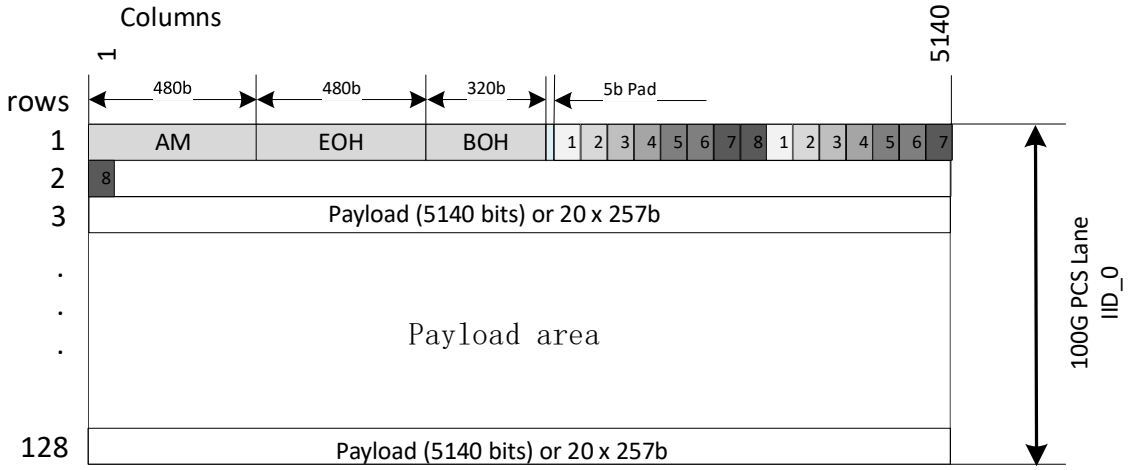
PCS - 257b distribution & GMP mapping



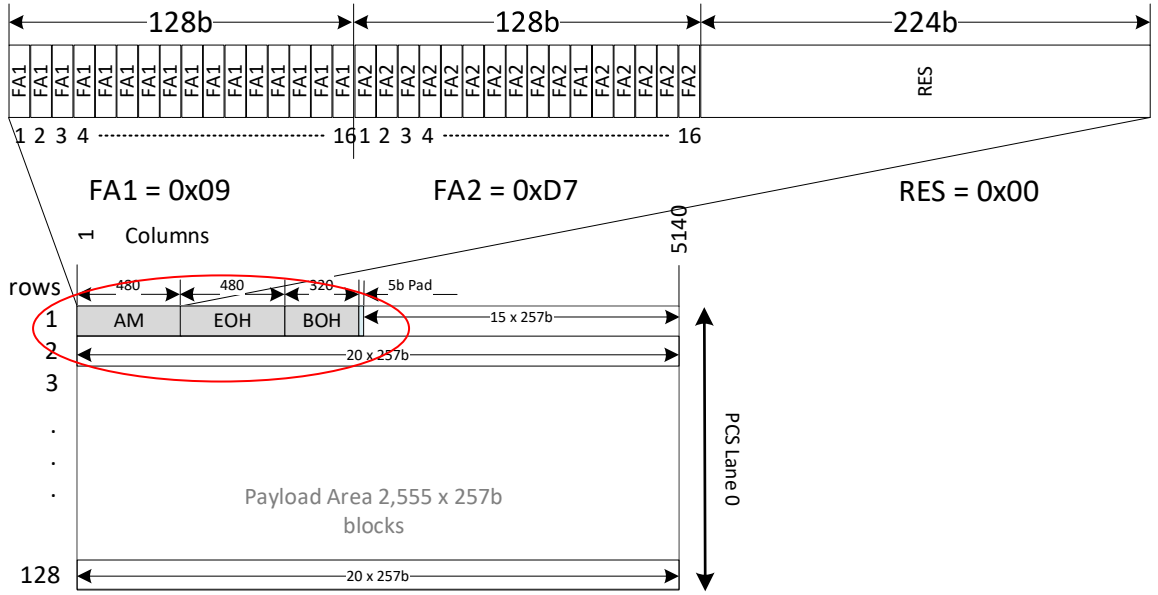
- 257b stream is distributed and GMP mapped into the payload area of 8 x 100G Synchronous and aligned lanes
- Alignment Markers and OH (AM/EOH/BOH) is added to each lane for monitoring and signaling purposes between link endpoints.



PCS – 100G lane format and overhead

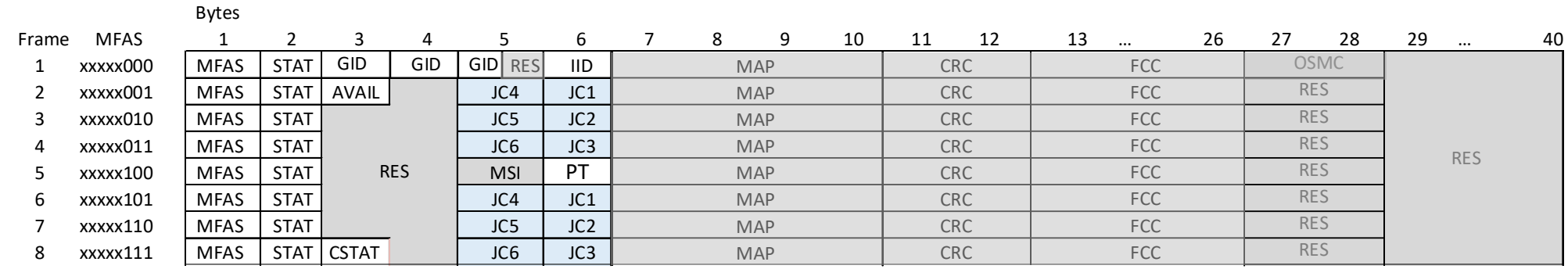


- 100G lane is 128 rows x 5140b
- 1st row contains 480b AM; 480b EOH and 480b of BOH
- The 1st 5-bits of payload area is an all-zero pad to align to 257b.
- 1st row carries 15 x 257b payload
- Rows 2-128 carry 20 x 257b payload



- 480b Alignment Marker (AM) Field
  - 16 Octet FA1= 0x09
  - 16 Octet FA2=0xD7
  - 28 Octet RES = 0x00
- 480b Extended Overhead (EOH) – Not used by 800GBASE-ER1
- 320b Basic Overhead – Mostly not used for 800GBASE-ER1 (except for link degrade signaling)
- 5b of pad for 257b alignment

PCS - BOH fields



1	2	3	4	5	6	7	8
RPF	MNT			RES			

Link status overhead

1	2	3	4	5	6	7	8
CSF	MNT			RES		RD	LD

Client status overhead

GMP Overhead

	1	2	3	4	5	6	7	8
JC1	C1	C2	C3	C4	C5	C6	C7	C8
JC2	C9	C10	C11	C12	C13	C14	II	DI
JC3	CRC8							
JC4	RES					D1	D2	D3
JC5	RES					D4	D5	RES
JC6	RES					CRC3		

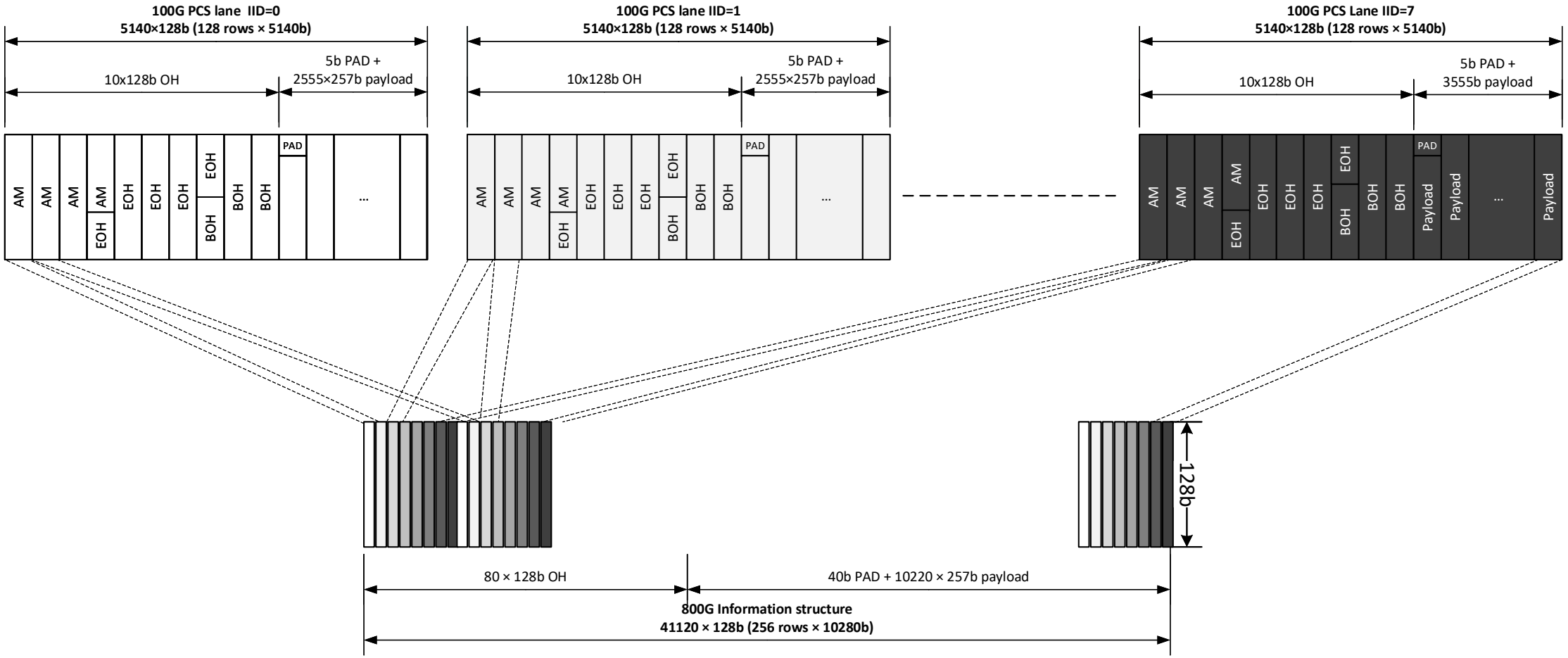
Undefined, Unused, or Reserved fields, not used for 802.3dj

Defined by 802.3dj using fields that either reserved or defined as mapping specific in ITU-T G709.1

Defined by [ITU-T G.709.1] and used for 802.3dj

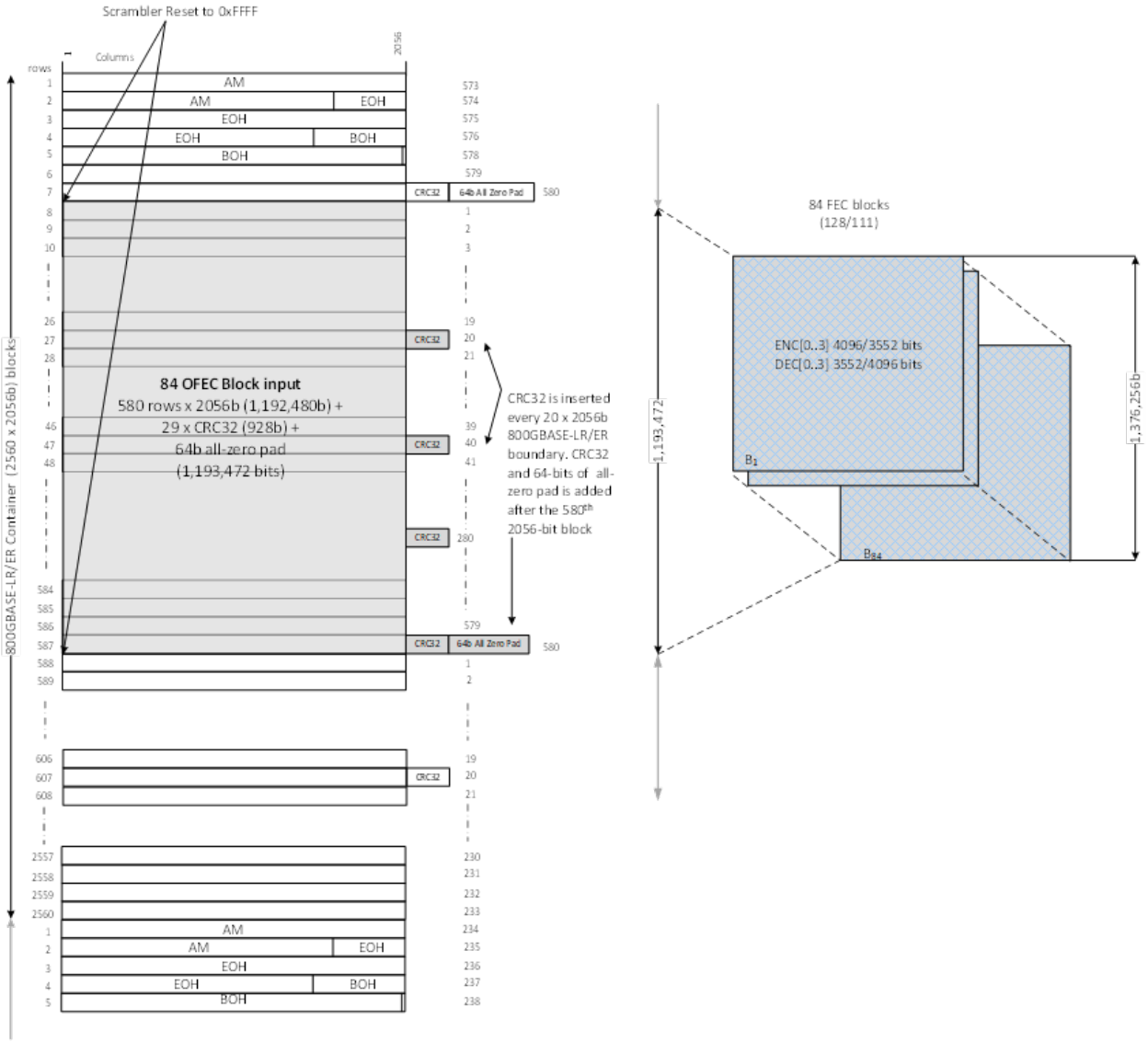
Reference ITU-T.G.709.1 for GMP parameter values

PCS – 128b Interleaving



8 x 100G lanes are 128b interleaved to an 800G structure

PCS - 800G structure adapted to oFEC

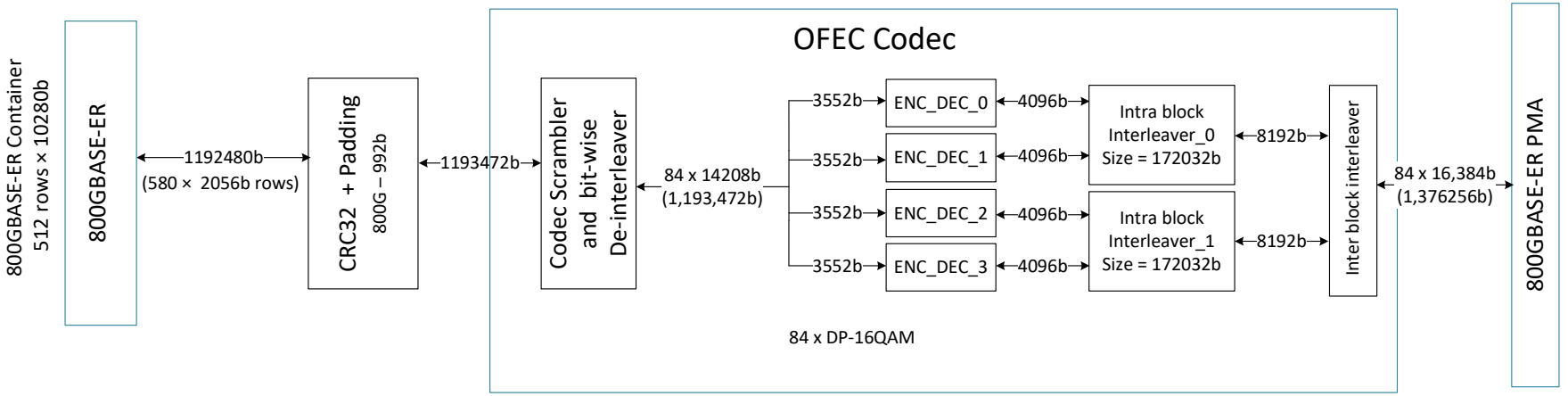


- The 800G structure at the output of the 128b interleaver can be redrawn as 2560 rows x 2056b (5,263,360b) total.
- 580 x 2056b rows + 29 x CR32 + 64-bits of pad align to 84 oFEC blocks (1,193,472b)
- A CRC32 is generated covering every 20 x 2056b rows to ensure MTTFPA.
- 84 oFEC blocks correspond to a symbol based optical frame (DSP Superframe).
- The oFEC block includes a scrambler which is reset at the beginning every 580 x 2056b rows of information bits.
- The oFEC blocks produce 1,376,256b of data + parity to the PMA.

PCS - oFEC overview

- The proposed oFEC is detailed in ITU-T G.709.3 Appendix III, IV, and V, Open ROADM MSA 6.0 W-Port Digital Specifications, and OpenZRplus v2.0. Suggested to reference to one of these documents in the specification
- The oFEC engine is a block-based encoder and iterative Soft-Decision (SD) decoder.
- With 3 SD iterations the Net Coding Gain is 11.6 dB @ 10-15 (DP-16QAM), with pre-FEC BER threshold of 2.0 × 10-2.
- The latency of the oFEC encoder/decoder pair, including interleaving and deinterleaving, is 800,000 bits. 800,000 bits. For this 800GBASE-ER1 implementation this equates to ~1.5µs.

PCS - oFEC encoder/interleaver/merge



- The oFEC block includes a scrambler which is reset at the beginning every 580 x 2056b rows of information bits.
- The oFEC block is organized as 4 oFEC encode/decoders (ENC_DEC_[0..3]) followed by two Intra block interleavers (Intra_block [0..1]) and an inter-block interleaver. These functions operate in parallel to produce an oFEC codeword.
- A codeword is a semi-infinite set of bits organized in a matrix with a semi-infinite number of rows and N columns (N=128).
- 84 oFEC codec blocks are ratio locked and aligned to the media-side DSP super frame.
- The interleaver structure is organized as an (84,8) array of 16b x 16b square blocks and contains two mechanisms:
- An intra-block interleaver (170,032b) reorders the bits in each 16b x 16b square block to ensure that the bits in each row and column of a square block at the encoder output are remapped uniformly in the square block for transmission.
- An inter-block interleaver ensures that nearby symbols on the link contain bits that are widely separated from the encoder output.
- The oFEC codec block consumes 1,193,472b and produces 1,376,256b of data and parity, (128/111) expansion ratio

PMA – 16QAM Symbol Mapping

The input bits of the symbol mapper are denoted by c_k (k=0…1376255). The symbol mapper shall receive 8-bit blocks from each interleaver in time order in a round-robin fashion.

The symbol mapper shall map the input bits c_k (k=0…1376255) to DP-16QAM symbols denoted s_i (i=0…172031), where,

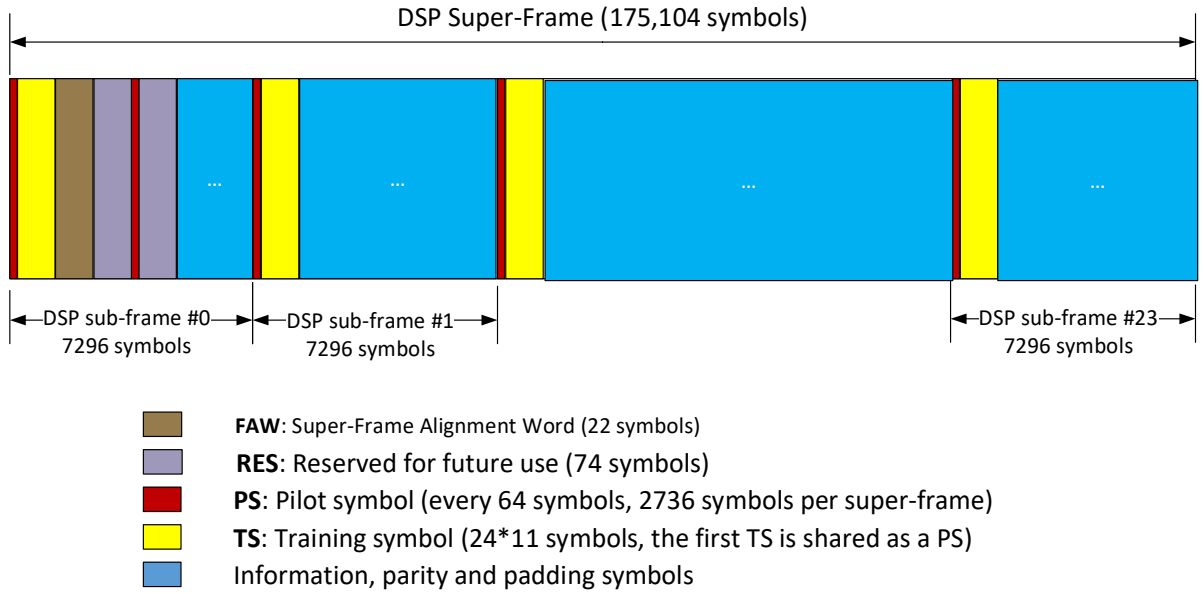
- (c_{8i},c_{8i+2}) shall map to the in-phase (I) component of the X-polarization of s_i
- (c_{8i+4},c_{8i+6}) shall map to the quadrature-phase (Q) component of the X-polarization of s_i
- (c_{8i+1},c_{8i+3}) shall map to the in-phase (I) component of the Y-polarization of s_i
- (c_{8i+5},c_{8i+7}) shall map to the quadrature-phase (Q) component of the Y-polarization of s_i

In each signaling dimension (XI/XQ/YI/YQ), the mapping from binary label to relative symbol amplitude shall be:

$$(0,0) \rightarrow -3,\ (0,1) \rightarrow -1,\ (1,1) \rightarrow +1,\ (1,0) \rightarrow +3$$

(c_{8i},c_{8i+2}, c_{8i+4}, c_{8i+6}) or (c_{8i+1},c_{8i+3}, c_{8i+5}, c_{8i+7})	I	Q
(0,0,0,0)	-3	-3
(0,0,0,1)	-3	-1
(0,0,1,0)	-3	+3
(0,0,1,1)	-3	+1
(0,1,0,0)	-1	-3
(0,1,0,1)	-1	-1
(0,1,1,0)	-1	+3
(0,1,1,1)	-1	+1
(1,0,0,0)	+3	-3
(1,0,0,1)	+3	-1
(1,0,1,0)	+3	+3
(1,0,1,1)	+3	+1
(1,1,0,0)	+1	-3
(1,1,0,1)	+1	-1
(1,1,1,0)	+1	+3
(1,1,1,1)	+1	+1

PMA – DSP frame format



- DSP frame format:
- 175,104 DP-16QAM symbols
- 24 x DSP sub-frames
- Each sub-frame is 7296 symbols

Pilot symbols shall be inserted every 64 symbols, starting with the first symbol of each DSP super-frame. The first 11 symbols of each DSP sub-frame can also be used for training (e.g, frame acquisition). The first symbol of the Training Sequence (TS) is a Pilot Symbol (PS).

- Every DSP subframe has the same structure based on a fixed TS with the first symbol processed as a pilot.
- The TS includes 11 QPSK symbols in each polarization. The TS is different between X and Y polarizations.
- The PS sequence includes (1+113) QPSK symbols based on PRBS. The first TS symbol is also the first symbol of the PS sequence

The first DSP sub-frame includes a 22 symbol Frame Alignment Word (FAW) used to align the 84 OFEC Code block frame.

- 22 symbols used as the Super Frame Alignment Word (FAW). The FAW is different between X and Y polarizations
- 74 symbols are reserved for future use. The symbols should be randomized to avoid strong tones

PMA – FAW/Training Sequence (TS)

- The FAW sequence at the beginning of each DSP super-frame shall match the sequence in the table below.
- The TS sequence shall match the sequence in the table below:

Index	FAW X	FAW Y	Index	FAW X	FAW Y
1	3-3j	3+3j	**12**	3-3j	-3+3j
2	3+3j	-3+3j	**13**	-3-3j	-3+3j
3	3+3j	-3-3j	**14**	-3-3j	3+3j
4	3+3j	-3+3j	**15**	-3+3j	-3-3j
5	3-3j	3-3j	**16**	3+3j	3+3j
6	3-3j	3+3j	**17**	-3-3j	-3-3j
7	-3-3j	3-3j	**18**	3-3j	-3+3j
8	3+3j	3-3j	**19**	-3+3j	3-3j
9	-3-3j	-3-3j	**20**	3+3j	-3-3j
10	-3+3j	3-3j	**21**	-3-3j	3-3j
11	-3+3j	3+3j	**22**	-3+3j	-3+3j

FAW sequence

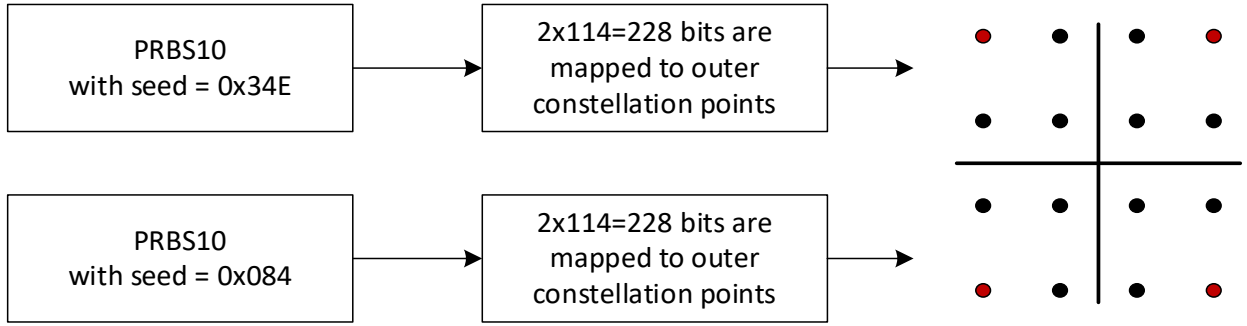
Index	Training X	Training Y
1*	-3+3j	-3-3j
2	3+3j	-3-3j
3	-3+3j	3-3j
4	3+3j	-3+3j
5	-3-3j	-3+3j
6	3+3j	3+3j
7	-3-3j	-3-3j
8	-3-3j	-3+3j
9	3+3j	3-3j
10	3-3j	3+3j
11	3-3j	3-3j

TS sequence

PMA - Pilot Sequence (PS)

A pilot sequence, formed from the outer symbols of the DP-16QAM constellation, is inserted every 64 symbols in each DSP frame. The pilot sequence is a fixed 114 symbol sequence formed from a PRBS10 sequence mapped to DP-QPSK with different seeds for the X/Y polarizations

- Seeds are selected so that the pilot and training sequence combined are DC balanced
- Seeds are selected so that the first symbol in the training sequence is also the first symbol in the pilot sequence
- The seed is reset at the start of every DSP sub-frame



Generator polynomial	Seed X	Seed Y
$x^{10}+x^{7}+x^{3}+x+1$	0x34E	0x084

Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y
1	-3+3j	-3-3j	30	3-3j	-3+3j	59	3-3j	-3+3j	88	-3+3j	-3-3j
2	3+3j	3-3j	31	3+3j	3+3j	60	3-3j	-3+3j	89	3+3j	3+3j
3	-3-3j	-3-3j	32	3-3j	3+3j	61	-3-3j	-3-3j	90	3-3j	3+3j
4	3-3j	-3+3j	33	-3+3j	-3-3j	62	3+3j	-3-3j	91	3-3j	3-3j
5	3+3j	-3-3j	34	-3-3j	-3+3j	63	3-3j	3+3j	92	-3-3j	3+3j
6	-3+3j	3+3j	35	3-3j	-3+3j	64	-3-3j	-3-3j	93	-3+3j	3+3j
7	3-3j	3+3j	36	-3+3j	-3-3j	65	3+3j	-3+3j	94	-3+3j	-3-3j
8	3+3j	-3+3j	37	-3-3j	3+3j	66	3+3j	-3+3j	95	-3+3j	3+3j
9	-3-3j	-3-3j	38	-3-3j	3+3j	67	3+3j	-3+3j	96	-3-3j	-3+3j
10	-3-3j	-3-3j	39	-3+3j	3-3j	68	-3-3j	3+3j	97	-3-3j	-3-3j
11	3+3j	3-3j	40	-3-3j	-3+3j	69	-3-3j	-3+3j	98	-3+3j	3-3j
12	-3-3j	3-3j	41	-3+3j	3-3j	70	3+3j	3+3j	99	-3+3j	3-3j
13	3+3j	3-3j	42	-3-3j	-3-3j	71	3+3j	3+3j	100	-3+3j	3-3j
14	-3-3j	-3-3j	43	-3-3j	-3-3j	72	3-3j	3+3j	101	3-3j	3+3j
15	-3+3j	-3-3j	44	3+3j	-3+3j	73	3-3j	-3+3j	102	-3+3j	-3+3j
16	-3-3j	3-3j	45	3+3j	3-3j	74	3+3j	3+3j	103	-3+3j	-3-3j
17	-3+3j	3-3j	46	-3-3j	-3+3j	75	3-3j	-3-3j	104	3-3j	-3-3j
18	-3+3j	3+3j	47	3+3j	-3-3j	76	3-3j	-3-3j	105	-3+3j	-3+3j
19	-3-3j	-3-3j	48	3+3j	-3+3j	77	3+3j	3-3j	106	-3-3j	3+3j
20	-3+3j	3-3j	49	-3+3j	3-3j	78	-3+3j	-3+3j	107	3-3j	-3-3j
21	3-3j	3+3j	50	3-3j	3+3j	79	3+3j	-3+3j	108	3+3j	3-3j
22	3-3j	-3-3j	51	-3+3j	3-3j	80	-3+3j	3+3j	109	-3+3j	3-3j
23	3-3j	3-3j	52	3-3j	-3-3j	81	3+3j	3-3j	110	-3-3j	-3+3j
24	3-3j	-3+3j	53	-3+3j	3+3j	82	-3+3j	3+3j	111	3-3j	-3-3j
25	3-3j	-3+3j	54	3-3j	-3+3j	83	3-3j	-3+3j	112	-3-3j	3-3j
26	3+3j	3-3j	55	3-3j	3-3j	84	-3+3j	3+3j	113	3+3j	3-3j
27	3+3j	3-3j	56	3+3j	-3-3j	85	-3-3j	3-3j	114	3+3j	-3+3j
28	-3+3j	-3+3j	57	-3-3j	3+3j	86	-3-3j	3-3j			
29	-3-3j	-3-3j	58	-3+3j	3+3j	87	3-3j	3+3j			

PMA – 16QAM Symbol mapping to physical lanes

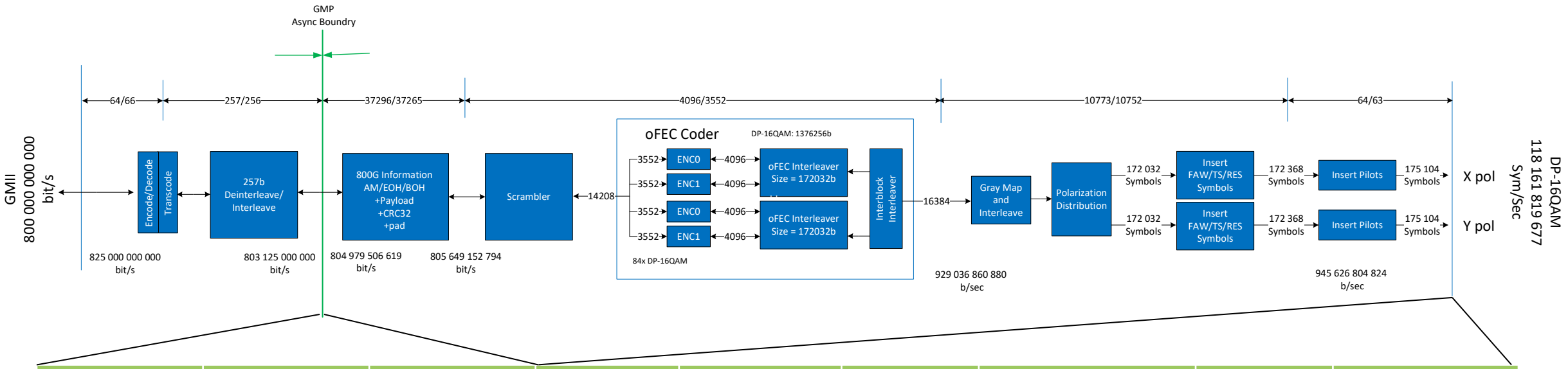
X and Y indicate a pair of mutually orthogonal polarizations of any orientation and I and Q are mutually orthogonal phase channels in each polarization. The four data path channels are therefore labeled XI, XQ, YI, and YQ.

All coherent channel mappings are allowed for the TX signal. The Rx should work in all cases because the Rx can unambiguously identify the signals polarization and phase, based on the FAW.

The Tx mapping is specified by two designations: [X:Y ; I,Q] where a ":" is used to separate X &Y, a "," is used to separate I & Q.

Mapping	X:Y	I,Q	Notes
[0,x]	X:Y		Polarization cannot be interleaved
[1,x]	Y:X		
[x,0]		I,Q:I,Q	Same across Polarizations
[x,1]		Q,I:Q,I	
[x,2]		I,Q:Q,I	Flip across Polarizations
[x,3]		Q,I:I,Q	

Data rates overview



GMII	64/66	Transcode	800ER Container	CRC32+ PAD	OFEC	FAW + TS + RES	PS	16QAM symbol rate
800 000 000 000	825 000 000 000	803 125 000 000	804 979 506 619	805 649 152 794	929 036 860 880	930 851 385 999	945 626 804 824	118 203 350 603

800GBASE-ER1 = DP-16QAM @ 118.203350603 GSym/Sec

OIF 800ZR and ITU-T FlexO-8e = DP-16QAM @ 118.203350603 GSym/Sec (Interoperable).

Summary

- The proposal provides a complete PCS and PMA baseline for a single lane 800 Gb/s coherent PHY capable up to at least 40km over SMF using C-band and based on oFEC
- Fully interoperable with OIF 800ZR and ITU-T G.709 FlexO-8e
- This proposal leverages well-understood technology (oFEC), broad industry investment (800ZR/ZR+) and meets all the requirements for 800GBASE-ER1
- The proposal also builds upon the ongoing efforts in 802.3cw to define an 802.3 PHY documentation structure to support a coherent optical interface
- Recommendation:
  - Use C-Band and oFEC for 800GBASE-ER1 PHY

Thanks