Proposed AM Format

IEEE P802.3bs 400 Gb/s Ethernet Task Force

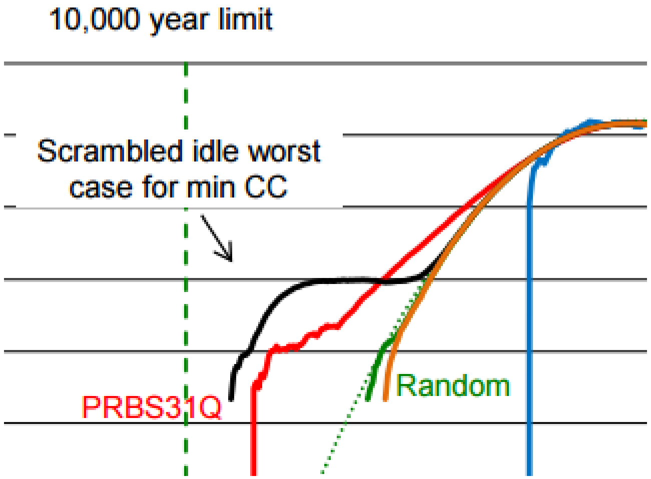
Pete Anslow – Ciena
Eric Baden - Broadcom
Adrian Butter – Global Foundries
Juan Carlos Calderon - Inphi
Mark Gustlin – Xilinx
Ben Jones – Xilinx
Dave Ofelt – Juniper
Rob Stone - Broadcom
Tongtong Wang - Huawei

Introduction

- In comment #146 Pete Anslow proposed shortening the AM CM to 48b, and the AM UM also to 48b (in draft 1.1 the CM was 64b and the CM 56b)
  - The comment was withdrawn due to some remaining issues to look at
- The 96b AMs fit nicely into 6x257b blocks, with 6b of pad
  - In draft 1.1 the AMs fit into 8x257b blocks, with 136b of pad
- There is a concern with the new format
  - The data does not nicely fit into nx10b symbols
- There has been a lot of email and some slide discussions since the last meeting, this slide deck looks at the consensus out of those discussions

Original Format

- This is the originally adopted AM format
- But as shown in anslow_01_1215_logic, there is an undesirable shoulder due to the muxing of the long common marker portion

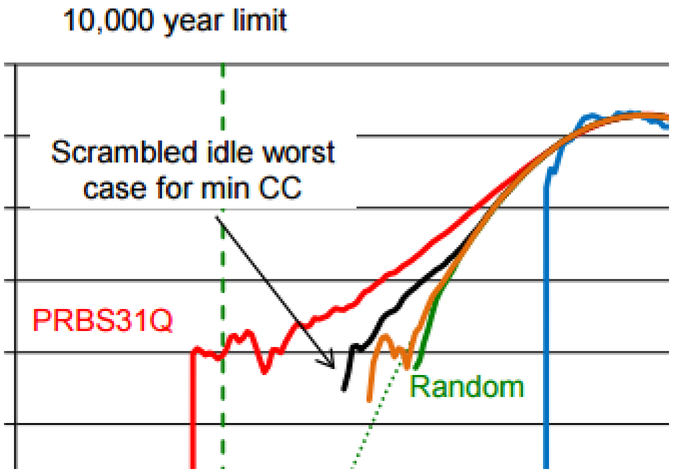


FEC Lane	Reed-Solomon symbol index (10 bit symbols	
	0 1 2 3 4 5 6	7 8 9 10 11 12
0	0 AM0 63	64 400G AM0 119
1	AM0	400G AM1
2	AM0	400G AM2
3	AM0	400G AM3
4	AM0	400G AM4
5	AM0	400G AM5
6	AM0	400G AM6
7	AM0	400G AM7
8	AM0	400G AM8
9	AM0	400G AM9
10	AM0	400G AM10
11	AM0	400G AM11
12	AM0	400G AM12
13	AM0	400G AM13
14	AM0	400G AM14
15	AM0	400G AM15

12 x 10b FEC symbols wide

Pete Anslow’s Revised Proposal

- From anslow_01_0116_logic, it shortens the common marker portion and proposes fitting within 6x257b blocks
- This ‘fixes’ the shoulder problem:



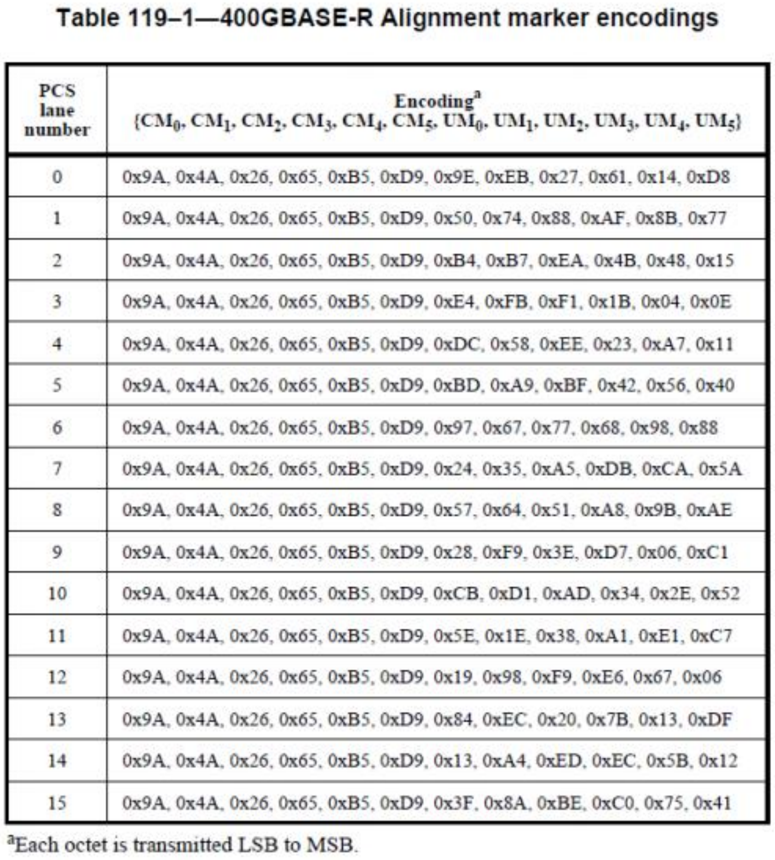


Table 119–1—400GBASE-R Alignment marker encodings

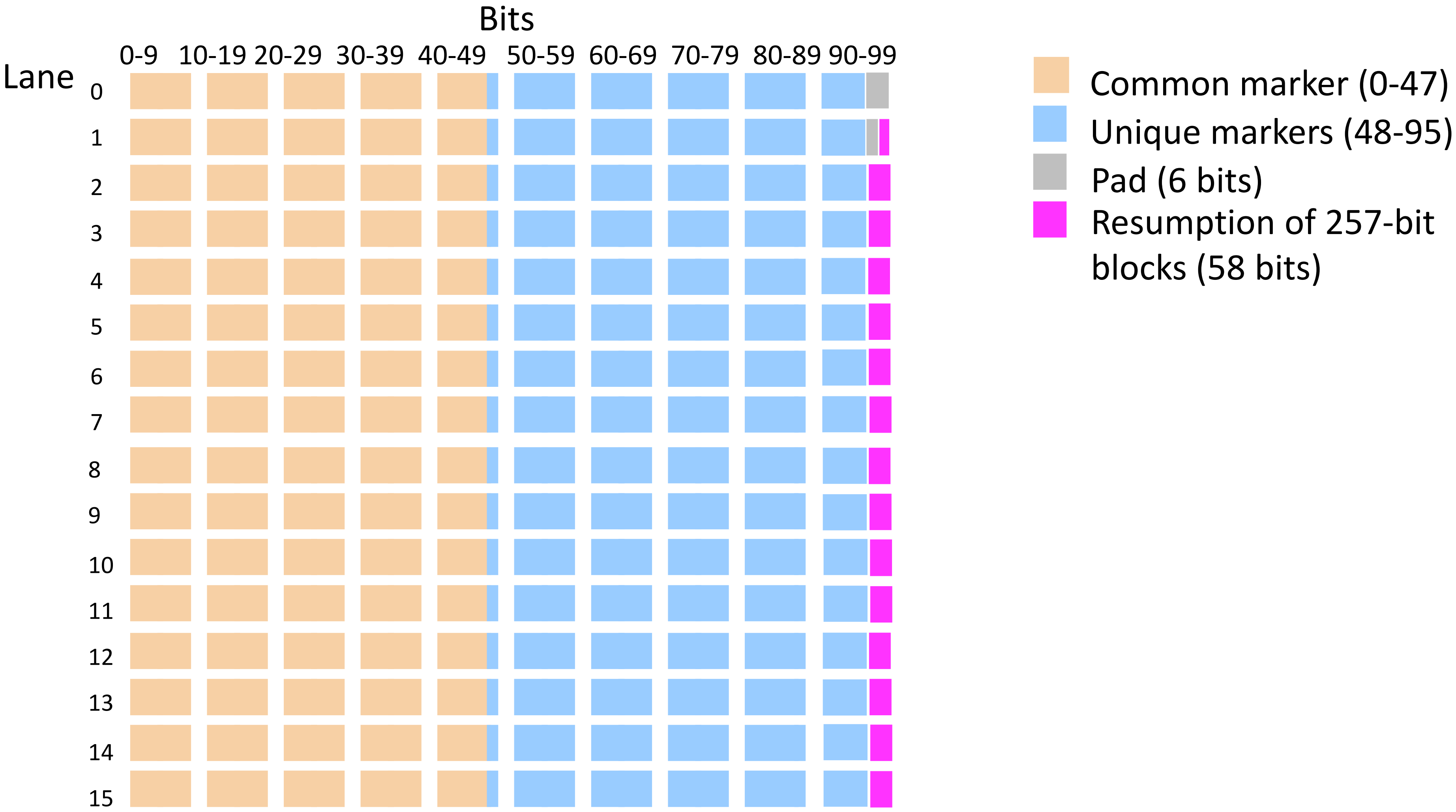
PCS lane number	Encoding[a] {CM_0, CM_1, CM_2, CM_3, CM_4, CM_5, UM_0, UM_1, UM_2, UM_3, UM_4, UM_5}
0	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x9E, 0xEB, 0x27, 0x61, 0x14, 0xD8
1	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x50, 0x74, 0x88, 0xAF, 0x8B, 0x77
2	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xB4, 0xB7, 0xEA, 0x4B, 0x48, 0x15
3	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xE4, 0xFB, 0xF1, 0x1B, 0x04, 0x0E
4	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xDC, 0x58, 0xEE, 0x23, 0xA7, 0x11
5	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xBD, 0xA9, 0xBF, 0x42, 0x56, 0x40
6	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x97, 0x67, 0x77, 0x68, 0x98, 0x88
7	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x24, 0x35, 0xA5, 0xDB, 0xCA, 0x5A
8	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x57, 0x64, 0x51, 0xA8, 0x9B, 0xAE
9	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x28, 0xF9, 0x3E, 0xD7, 0x06, 0xC1
10	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0xCB, 0xD1, 0xAD, 0x34, 0x2E, 0x52
11	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x5E, 0x1E, 0x38, 0xA1, 0xE1, 0xC7
12	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x19, 0x98, 0xF9, 0xE6, 0x67, 0x06
13	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x84, 0xEC, 0x20, 0x7B, 0x13, 0xDF
14	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x13, 0xA4, 0xED, 0xEC, 0x5B, 0x12
15	0x9A, 0x4A, 0x26, 0x65, 0xB5, 0xD9, 0x3F, 0x8A, 0xBE, 0xC0, 0x75, 0x41

[a] Each octet is transmitted LSB to MSB.

10b Issue

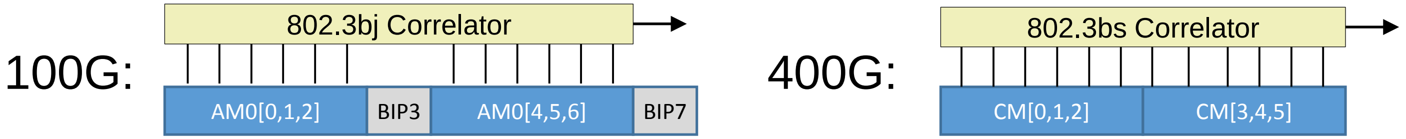
The analysis done in anslow_01_0116_logic requires even distribution of the AM across all PCS lanes...
However, an AM which is not 10-bit aligned requires unusual striping among the AM, Pad (for 257-bit alignment) and 257-bit block which complicates integration of AM processing with the 257-bit data flow.

Proposed 96-bit AM per anslow_01_0116_logic:

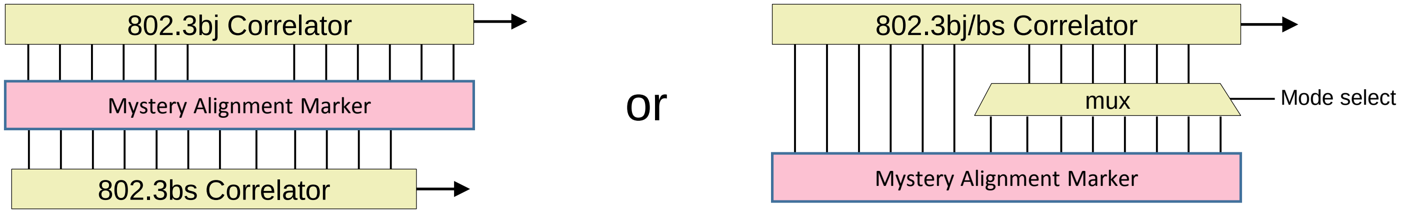


New Direction

- Keep the 96b AM proposed by Pete Anslow, 48b Common Marker (CM) and 48b Unique Marker (UM)
- Fit within 8x257b blocks to solve the 10b striping issue
- There is also a desire to organize the groups of bits so they mimic the organization of a 100GbE lane as defined in 802.3bj
  - Allows sharing of logic, this was part of the original format
  - If format/layout is different between the standards, e.g:

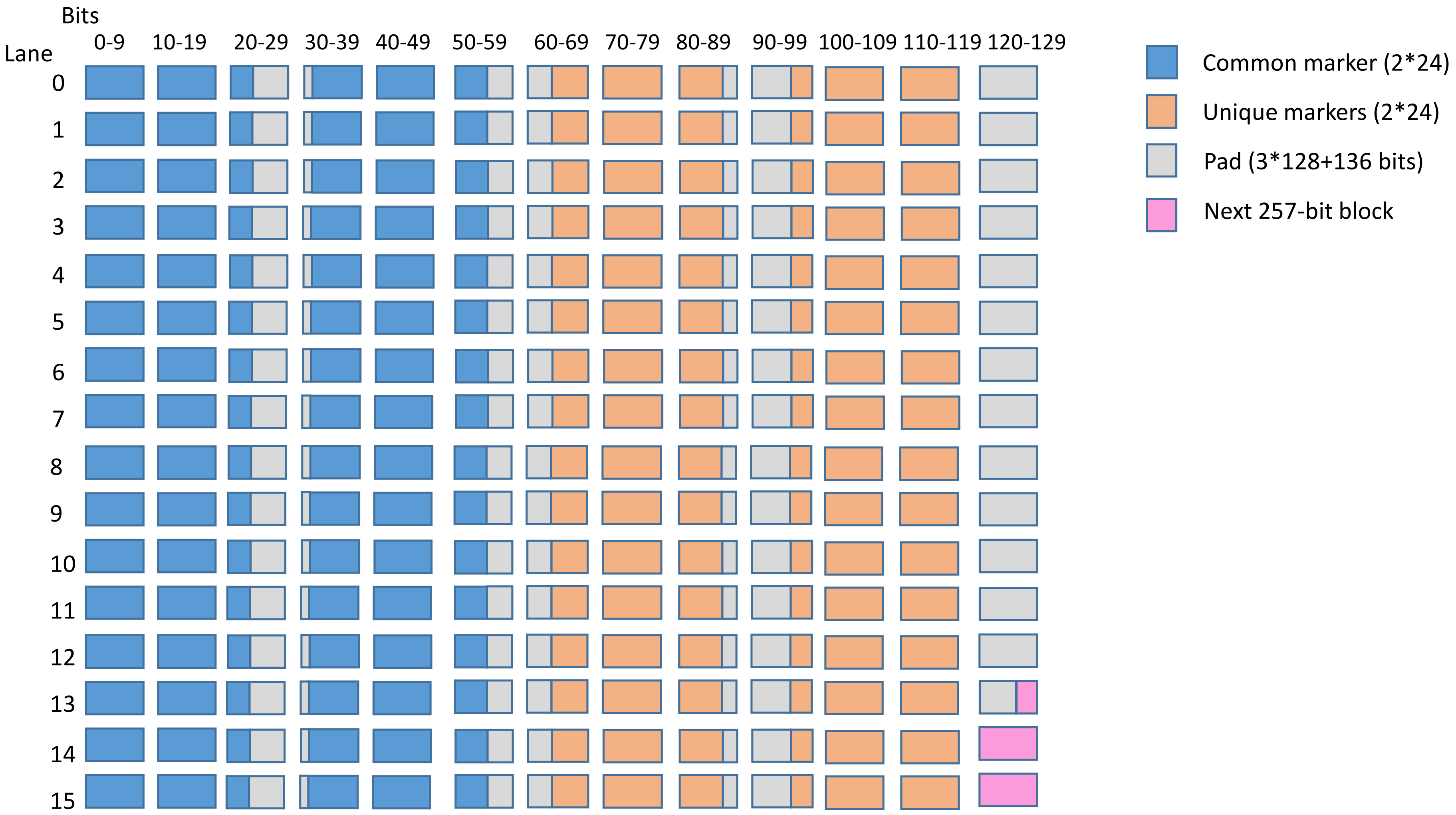


  - …then to switch dynamically between 100G and 400G operation we must have extra correlation circuitry, or at least pre-select the mode:



  - A large (e.g. x40) array of correlators is typically used to ensure fast lock times, so correlator implementation complexity is significant

400G Proposed Format



Lane \ Bits	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129
0	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
1	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
2	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
3	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
4	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
5	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
6	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
7	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
8	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
9	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
10	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
11	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
12	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜
13	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	⬜🟪
14	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	🟪
15	🟦	🟦	🟦⬜	⬜🟦	🟦	🟦⬜	⬜🟧	🟧	🟧⬜	⬜🟧	🟧	🟧	🟪

🟦 Common marker (2*24)

🟧 Unique markers (2*24)

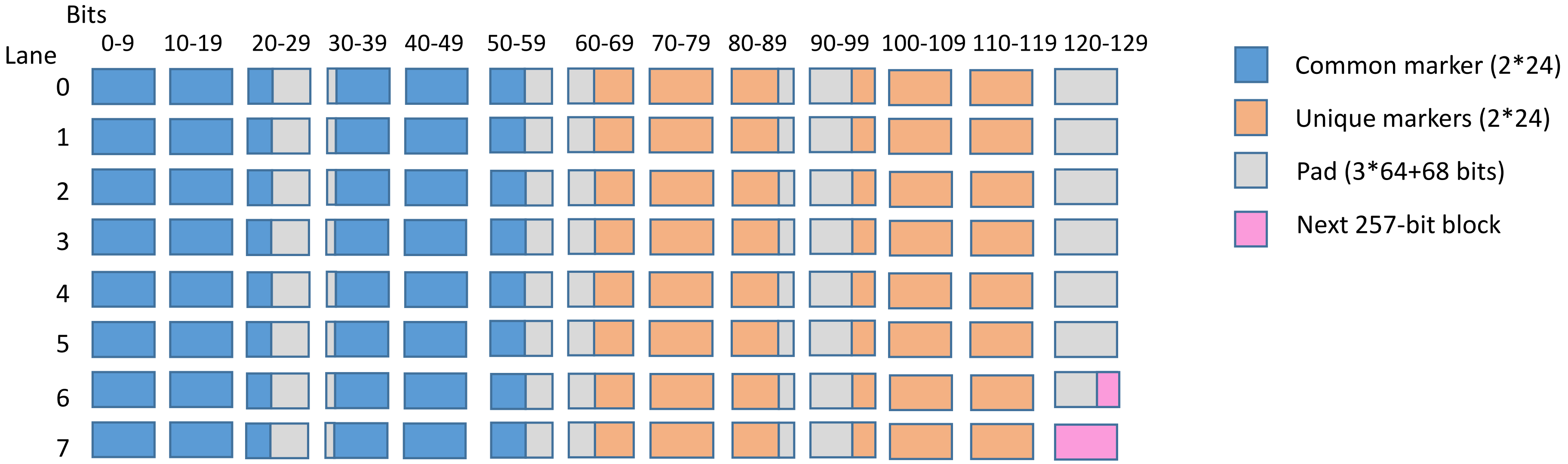
⬜ Pad (3*128+136 bits)

🟪 Next 257-bit block

8x257b blocks, Pad is filled with a PRBS9, free running

Pad columns coincide with position of BIP in 802.3bj

200G Proposed Format



4x257b blocks, Pad is filled with a PRBS9, free running

Summary

- The new format works well for both 200G and 400G
- And provides some logic compatibility with 100GbE lanes

- Still need to decide how to add in degrade signaling if that is adopted

Thanks!