

Restricted muxing option (natural pairs)

Mark Gustlin – Xilinx, Kapil Shrikhande – Innovium, Dave Ofelt – Juniper, Gary Nichol - Cisco

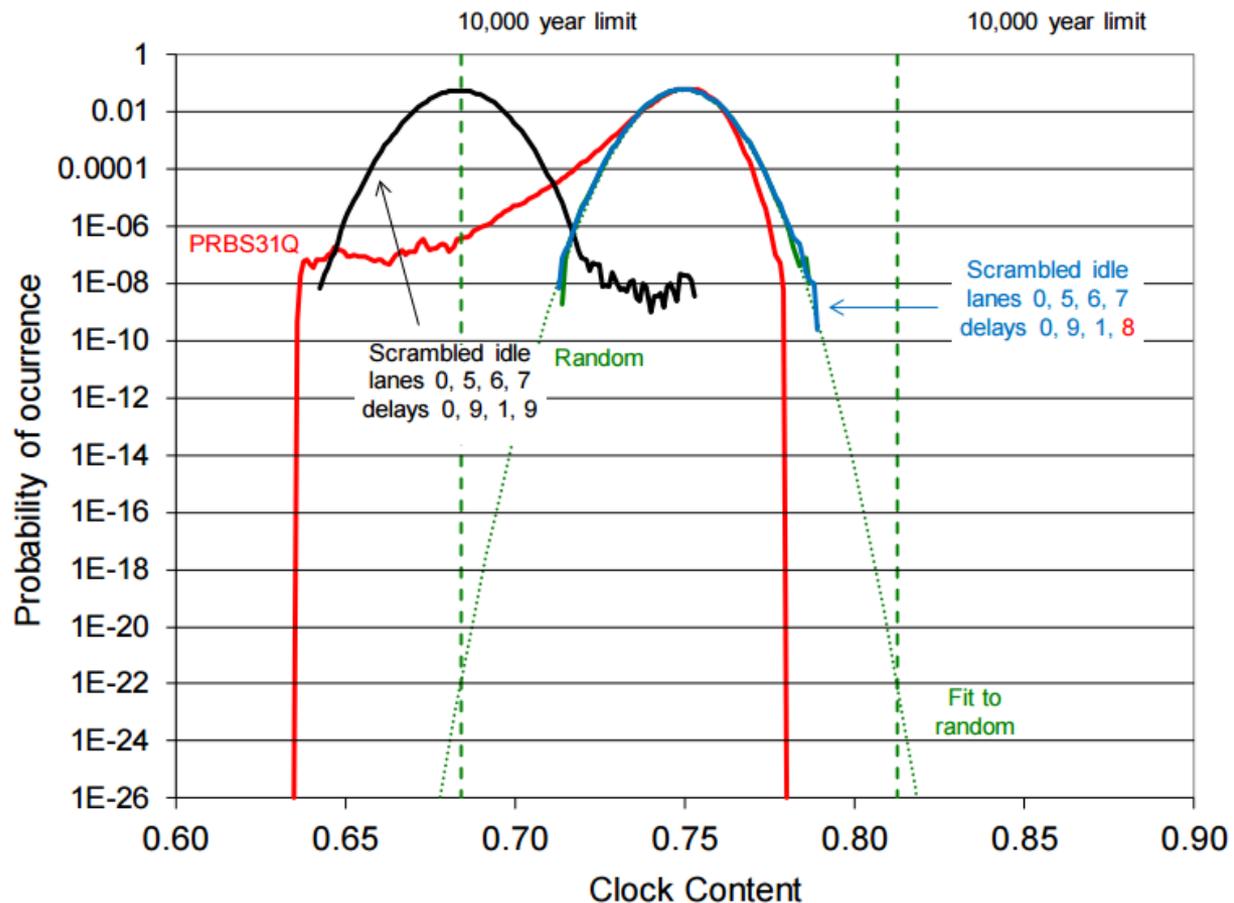
IEEE P802.3bs Task Force, Huntington Beach, January 2017

Overview

- Pete Anslow shared the clock content issue, originally found by Ryan Wong

http://www.ieee802.org/3/bs/public/adhoc/elect/19Dec_16/anslow_01_121916_elect.pdf

200GbE clock, all transitions, 0, 5, 6, 7



Possible Solution Direction

- Can we take advantage of the fact most ports will start with 50G lanes even at the MAC/PCS?
- But still support 16x25G lanes for early adopters?
- A 50G lane will have a natural muxing set, 0+1 and 2+3 and 4+5 etc.

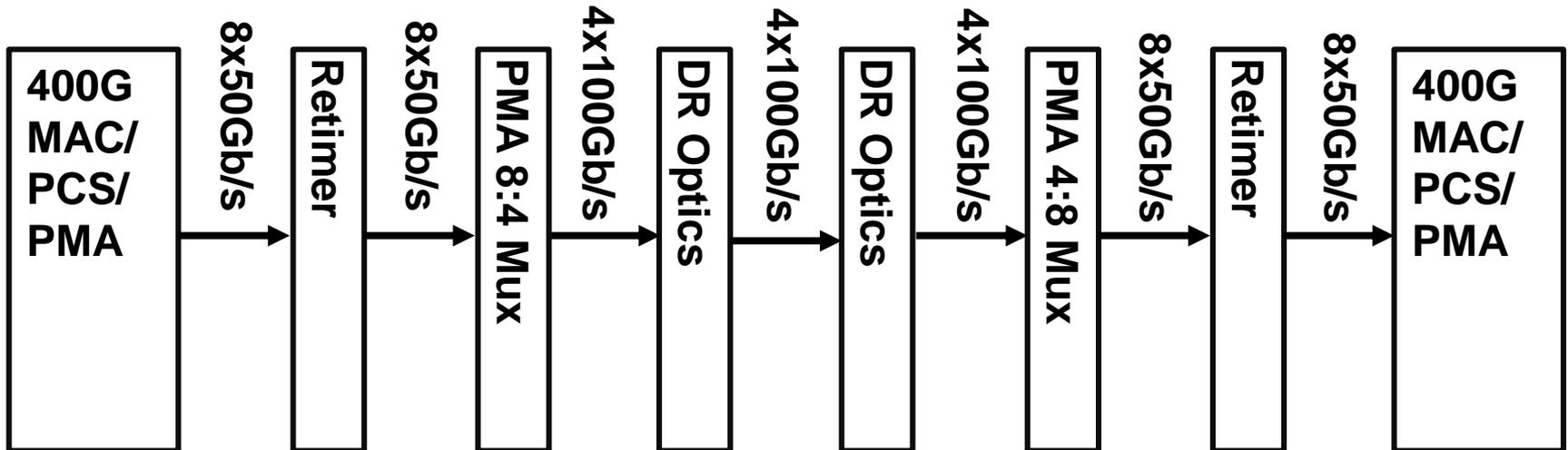
Analysis of Rogue Cases

- Pete Anslow created a spreadsheet with the rogue cases he has found
 - http://www.ieee802.org/3/bs/public/adhoc/elect/19Dec_16/anslow_02_121916_elect.xlsx
- In this spreadsheet there are many examples of naturally muxed lane pairs (0+1 etc), but there are no cases with two naturally muxed pairs (0+1 and 2+3).
- Highlight below....

		Natural Pair		Non-natural Pair										
2558	lanes	12	13	0	6	delays	0	5	-4	-3	average clock	0.217149	0.499628	0.716676
2559	lanes	12	13	0	6	delays	0	5	-3	-2	average clock	0.216747	0.499228	0.716115
2560	lanes	12	13	1	7	delays	0	-5	-9	-8	average clock	0.215482	0.500016	0.715619
2561	lanes	12	13	1	7	delays	0	-5	-8	-7	average clock	0.216532	0.500801	0.715629
2562	lanes	12	13	6	0	delays	0	5	-3	-4	average clock	0.216743	0.498685	0.715953
2563	lanes	12	13	7	1	delays	0	-5	-8	-9	average clock	0.215472	0.501171	0.715488

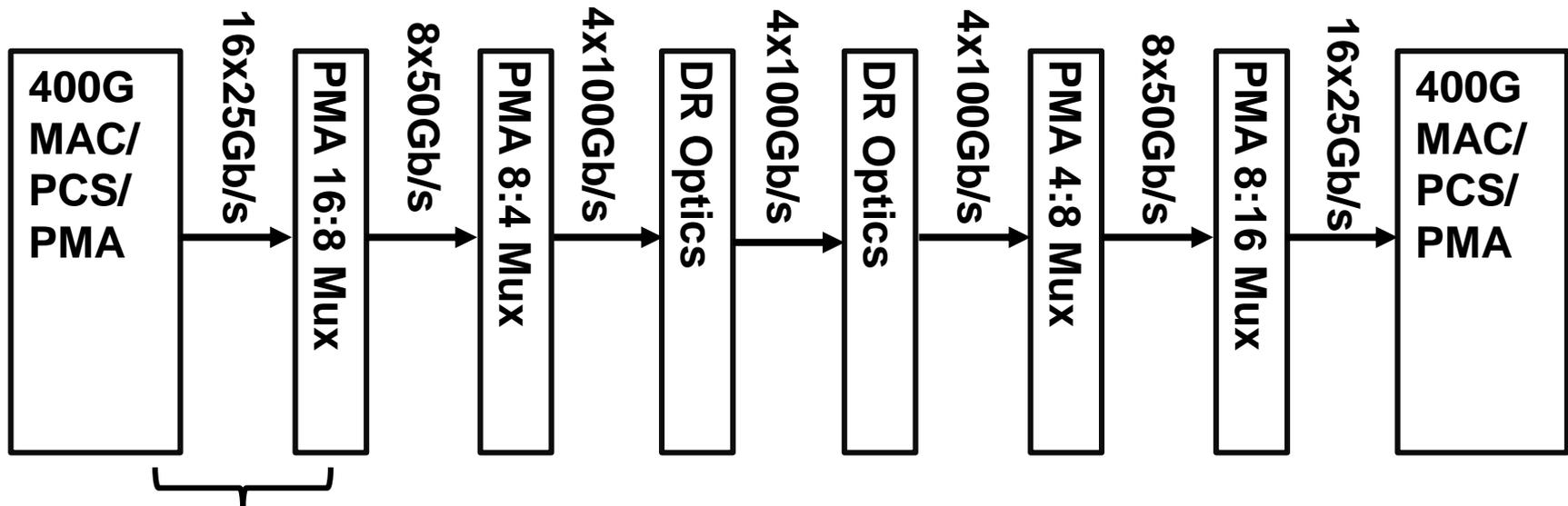
50Gb/s or faster Lanes only

- State that the TX PMA (16:8) must bit mux PCS Lane 0+1, 2+3, 4+5 etc (natural pairings)
- If there is a retimer in the path, it must keep the same paired PCS lanes together
 - This is natural anyhow
- TX PMA (8:4) mux (50G to 100G) must keep natural pairs of PCS lanes together
- RX PMA (4:8) from 100G to 50G will be blind and won't necessarily keep the desired PCS lane pairings, but at that point it won't matter, we don't have 2:1 muxing concerns



Systems with 25Gb/s Lanes

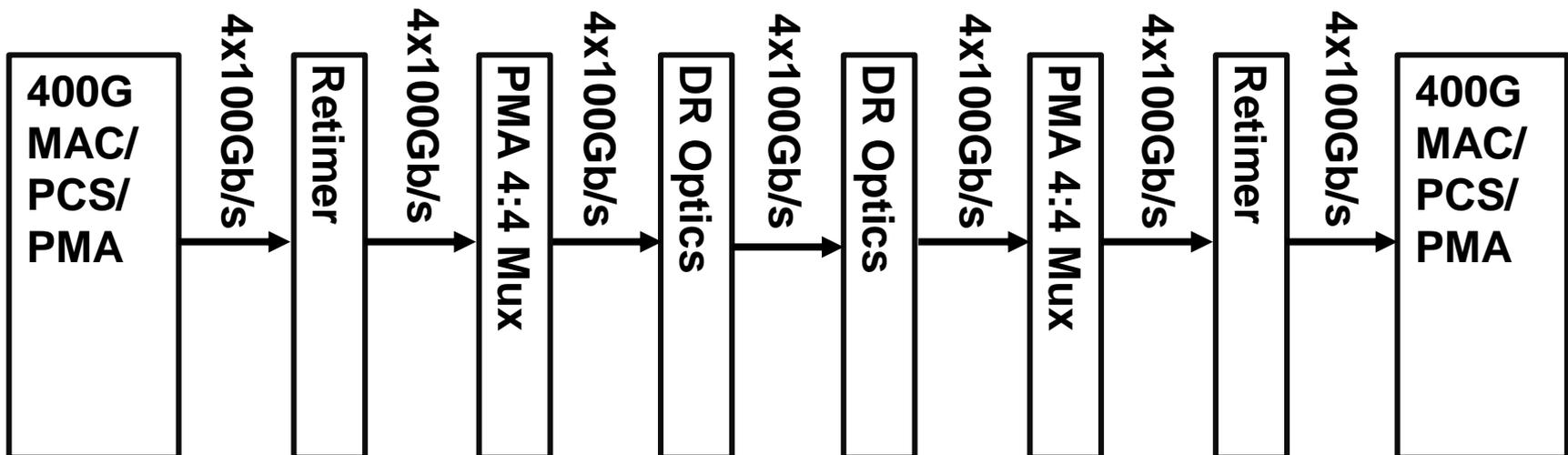
- The TX MAC/PCS/PMA (16:16) must have properly constrained lane mapping/routing to the TX PMA (16:8) device
- The TX PMA (16:8) must then bit mux PCS Lane 0+1, 2+3, 4+5 etc (natural pairings)
- See previous slide for the other constraints
- Supporting these restrictions seems easy in real implementations



Constrained routing/pairing: Opt 1, constrained PCB routing that keeps natural pairs together. Opt 2, Arbitrary routing on PCB possible if using a programmable lane Swizzle in MAC/PCS/PMA

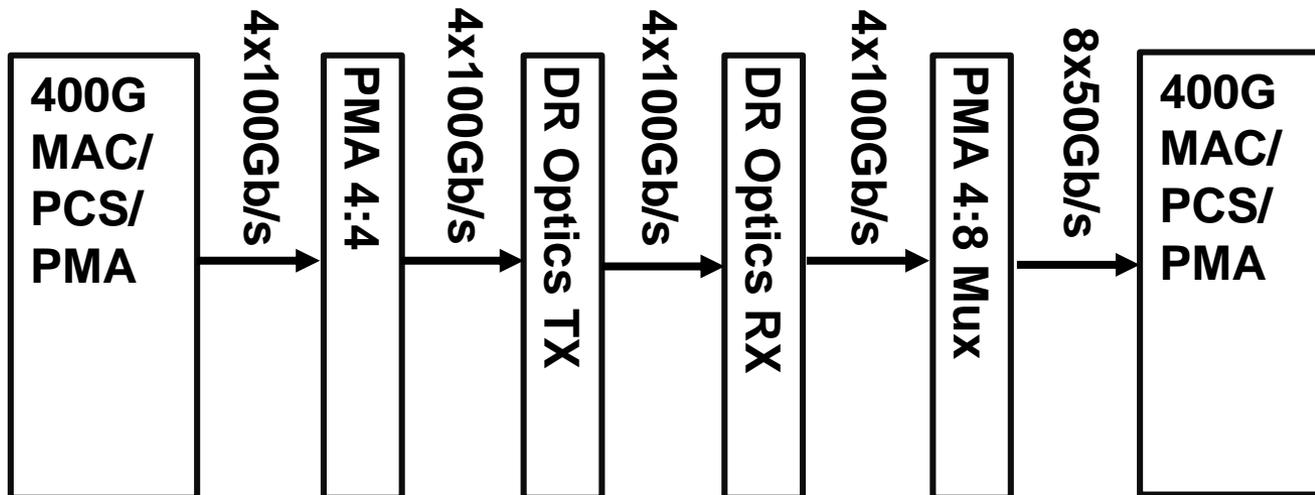
Only 100Gb/s Lanes (future)

- What happens in the future, if 100G electrical lanes can use the same FEC/PCS?
- This scenario is ok, as long as natural pairs are not disturbed
 - Most paths are retimers essentially, so why would they remux stuff?



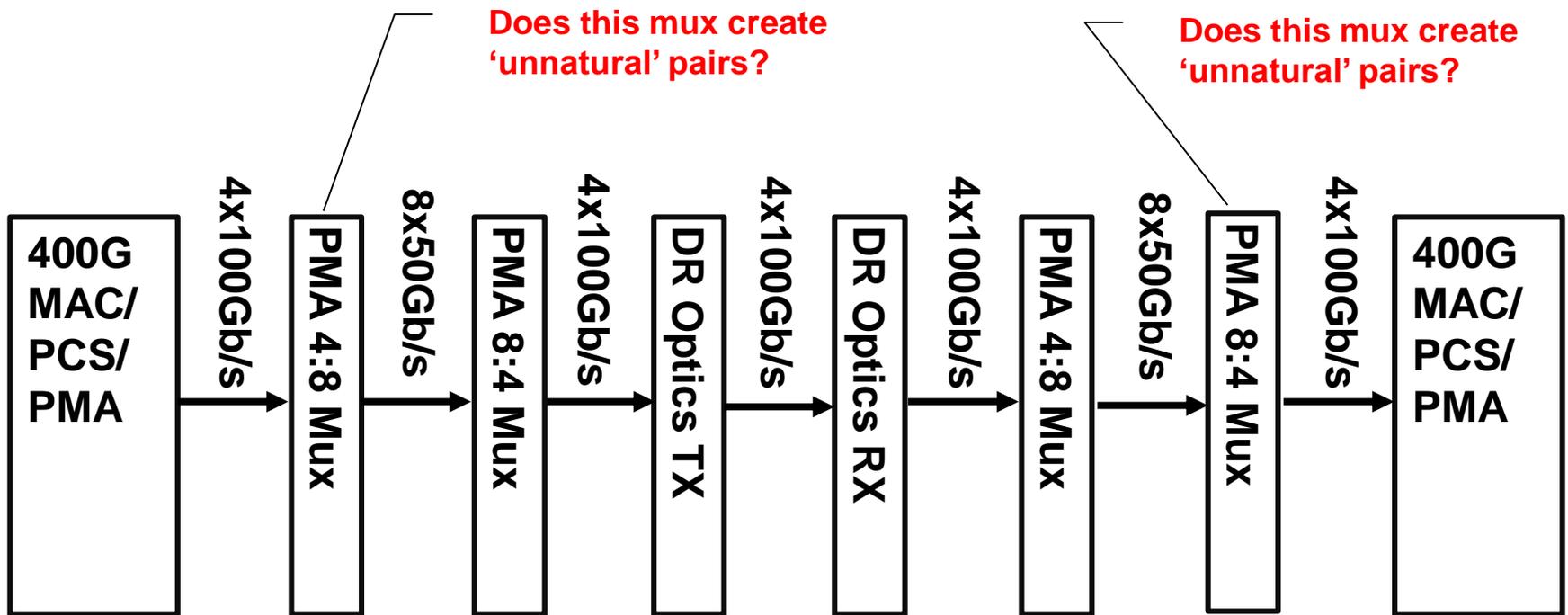
Mix of 50Gb/s and 100Gb/s Lanes (future)

- What happens in the future, if 100G electrical lanes can use the same FEC/PCS?
- This scenario is ok, as long as natural pairs are not disturbed
 - Most paths are retimers essentially, so why would they remux stuff?
- Last 4:8 PMA can do blind bit muxing and not impact that links performance



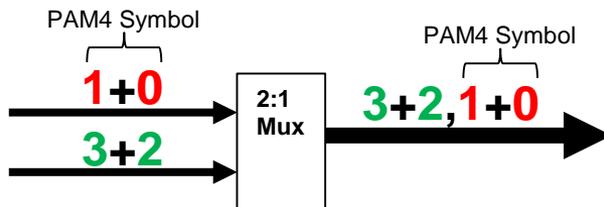
Mix of 50Gb/s and 100Gb/s Lanes (future)

- What happens in the future, if 100G electrical lanes can use the same FEC/PCS?
- Is the scenario below realistic?
- The 4:8 mux might create 'unnatural' pairs which can cause a problem

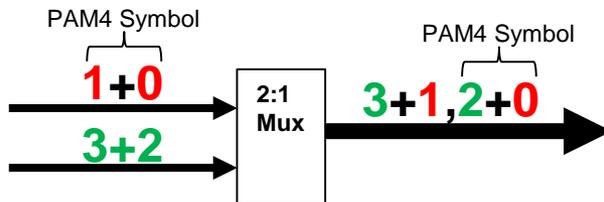


Dive Into 8:4 mux issues

- Propose that natural pairs are always kept together!
- When creating a 50G lane from 25G lanes (16:8), simply mux 0+1, 2+3, 4+5 etc. together (natural pairs)
- When creating a 100G lane from 50G lanes, and with PAM4 encoding, there are two possibilities:
 - Option A: Keep natural pairs together in each PAM4 symbol: 0+1 in one PAM4 symbol, 2+3 in 2nd PAM4 symbol, then back to 0+1



- Option B: Bit mux the natural pairs with each other: 0+2 in one PAM4 symbol, or 1+3 in the next PAM4 symbol (0+3 and 1+2)



- Whichever way we decide (and we would have to decide one consistent way if we go down this path), when you demux back to 50G lanes you must do the reverse to keep the natural pairs together

Option A: Mix of 50Gb/s and 100Gb/s Lanes

- Example assuming we chose the keep natural pairs within a PAM4 symbol:

50G Lanes		100G Lanes			50G Lanes	
PCS Lanes		PCS Lanes			PCS Lanes	
PAM4 Symbols		PAM4 Symbols (double baud rate)			PAM4 Symbols	
Lane 0	0+1				Lane 0	2+3
Lane 1	2+3				Lane 1	0+1
Lane 2	4+5	Lane 0	2+3	0+1	Lane 2	4+5
Lane 3	6+7	Lane 1	6+7	4+5	Lane 3	6+7
Lane 4	8+9	Lane 2	10+11	8+9	Lane 4	10+11
Lane 5	10+11	Lane 3	14+15	12+13	Lane 5	8+9
Lane 6	12+13				Lane 6	12+13
Lane 7	14+15				Lane 7	14+15
					Order of pairs does not matter	

Option B: Mix of 50Gb/s and 100Gb/s Lanes

- Example assuming bit mux between PAM4 symbols

50G Lanes		100G Lanes			50G Lanes	
PCS Lanes		PCS Lanes			PCS Lanes	
PAM4 Symbols		PAM4 Symbols (double baud rate)			PAM4 Symbols	
Lane 0	0+1				Lane 0	2+3
Lane 1	2+3				Lane 1	0+1
Lane 2	4+5	Lane 0	1+3	0+2	Lane 2	4+5
Lane 3	6+7	Lane 1	5+7	4+6	Lane 3	6+7
Lane 4	8+9	Lane 2	9+11	8+10	Lane 4	10+11
Lane 5	10+11	Lane 3	13+15	12+14	Lane 5	8+9
Lane 6	12+13				Lane 6	12+13
Lane 7	14+15				Lane 7	14+15
					Order of pairs does not matter	

Proposed Rules (for option A)

1. When muxing from 25G lanes to 50G lanes (PMA16:8), keep natural pairs together
 - 0+1, 2+3 etc. (0+1 form a single pam4 symbol on a given lane etc.)
 - Which physical lane a given natural pair appears does not matter (no constraints on the physical routing of 50G lanes)
2. When muxing from 50G lanes to 100G lanes (PMA8:4), keep natural pairs intact within a PAM4 symbol
 - PAM4 symbols from a given 50G lane are kept intact when sent to a 100G lane
 - Which physical lane a given pair of natural pairs appears does not matter (no constraints on the physical routing of 100G lanes)
3. When demuxing from 100G lanes to 50G lanes (PMA4:8), keep PAM4 symbols together
 - Alternate PAM4 symbols from a given 100G lane are played out to the 2x50G lanes
- When demuxing from 50G lanes to 25G lanes (PMA8:16), no proposed constraints

Proposed Rules (for option B)

1. When muxing from 25G lanes to 50G lanes (PMA16:8), keep natural pairs together
 - 0+1, 2+3 etc. (0+1 form a single pam4 symbol on a given lane etc.)
 - Which physical lane a given natural pair appears does not matter (no constraints on the physical routing of 50G lanes)
 2. When muxing from 50G lanes to 100G lanes (PMA8:4), bit mux between natural pairs
 - PAM4 symbols from a given 50G split up by bit muxing to get two PAM4 symbols for the 100G lane
 - No constraints on the physical routing of 100G lanes)
 3. When demuxing from 100G lanes to 50G lanes (PMA4:8), recreate natural pairs by reversing the muxing that occurred
- When demuxing from 50G lanes to 25G lanes (PMA8:16), no proposed constraints

Next Steps

- Are the proposed rules to prevent the problem acceptable?
- Choose option A vs. option B?

Thanks!