RS Symbol Muxing Option for 802.3ck

IEEE P802.3ck

San Diego July 2018

Mark Gustlin – Xilinx
Jeff Slavick – Broadcom

Introduction

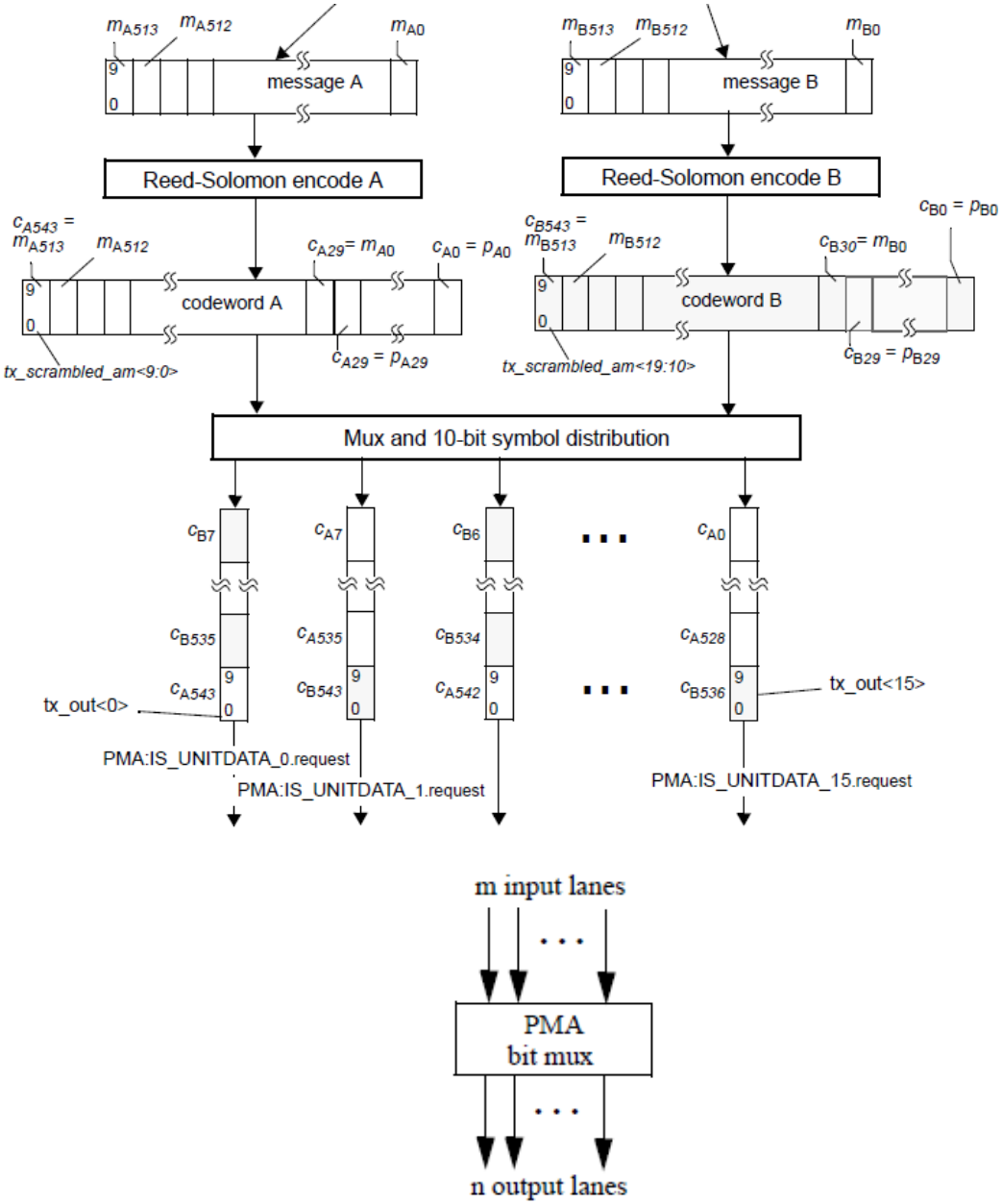
- In gustlin_3ck_01_0518.pdf, it is suggested that we should reuse the current 802.3bs/cd PCSs for this project, especially for the C2M interface to maintain compatibility with existing PMDs
- This includes 4:1 bit muxing in the PMA
- What happens if we can't close the harder channels, ie. backplane and copper cable links?
- This presentation explores the possibility of symbol muxing instead of bit muxing for these more difficult channels
  - For this presentation Symbol refers to a 10b RS FEC symbol

Background

- These longer channels will require strong equalization, including DFE
- DFE can induce burst errors which cause a degradation of the FEC gain due to the 4:1 bit muxing
  - A 4-bit burst error can consume 4 symbols worth of correction capability
  - A RS544 codeword can fix 15 symbols per codeword
- Precoding can help with some receiver architectures, but for sparse error bursts it might not help much
- An incremental improvement is to RS Symbol multiplex in the PMA instead of bit muxing, how does this help the concern?
  - With RS Symbol multiplexing, a single burst error is contained into fewer FEC symbols compared to bit muxing
  - Resistance to doing RS Symbol muxing previously is due to the complexity it adds to the PMA inside a module (making them PCS aware)
  - But if we did this only for the backplane and copper cable PHYs, it would eliminate that concern

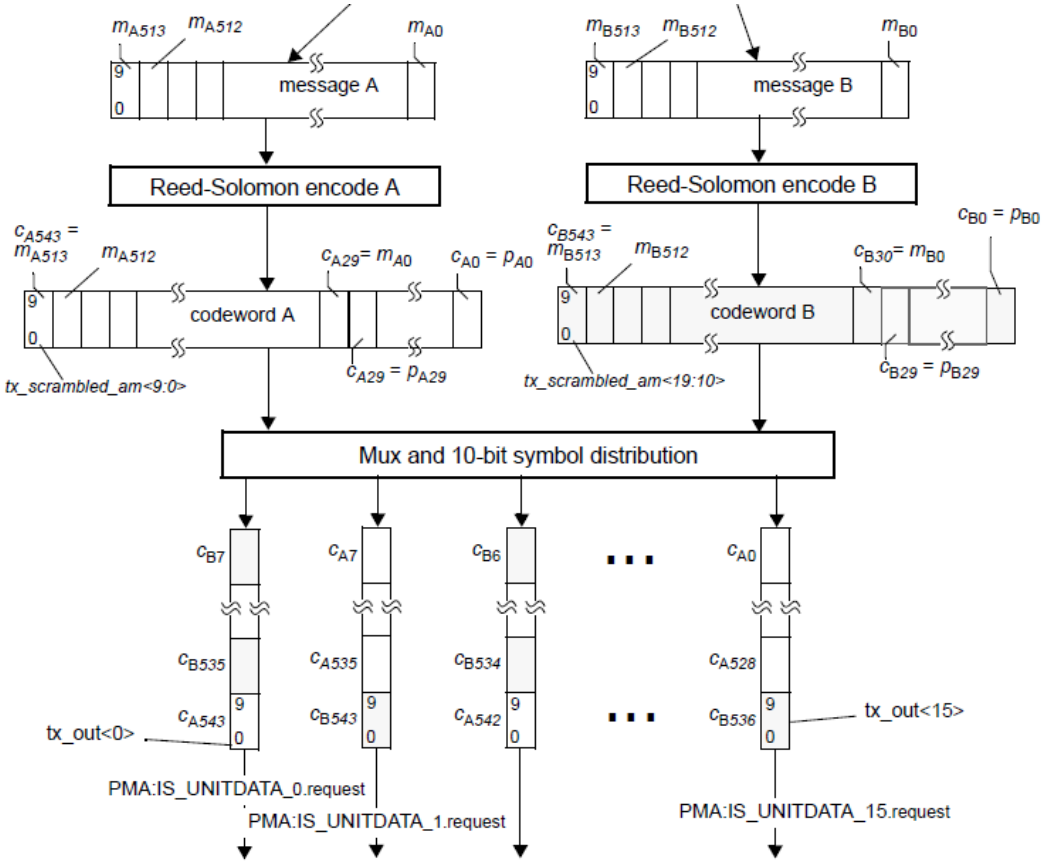
Today's PCS/PMA

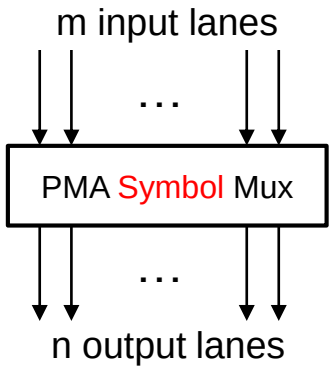
- The 802.3bs PCS plays out two FEC codewords into 16 PCS lanes (400GbE)
- The 10b symbols are distributed in a checkboard fashion to the PCS lanes
- The existing PMA then does bit muxing to get to fewer lanes
  - 4:1 for 100Gb/s per lane
  - 2:1 for 50Gb/s per lane



PMA for Harder Channels?

- 802.3bs PCS stays the same, plays out two FEC codewords into 16 PCS lanes (400GbE)
- The 10b symbols are distributed in a checkboard fashion to the PCS lanes
- The new PMA then uses RS Symbol muxing to get to fewer lanes
  - 40b:10b for 100Gb/s per lane

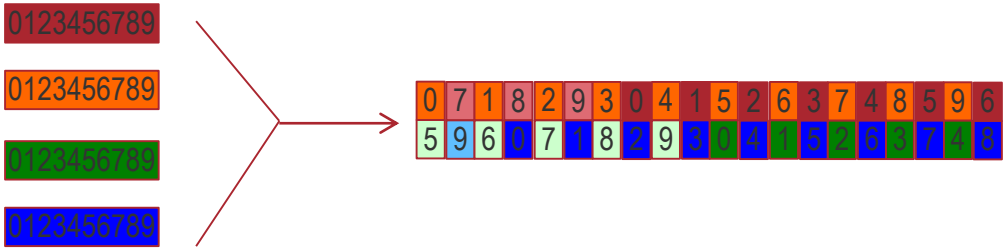




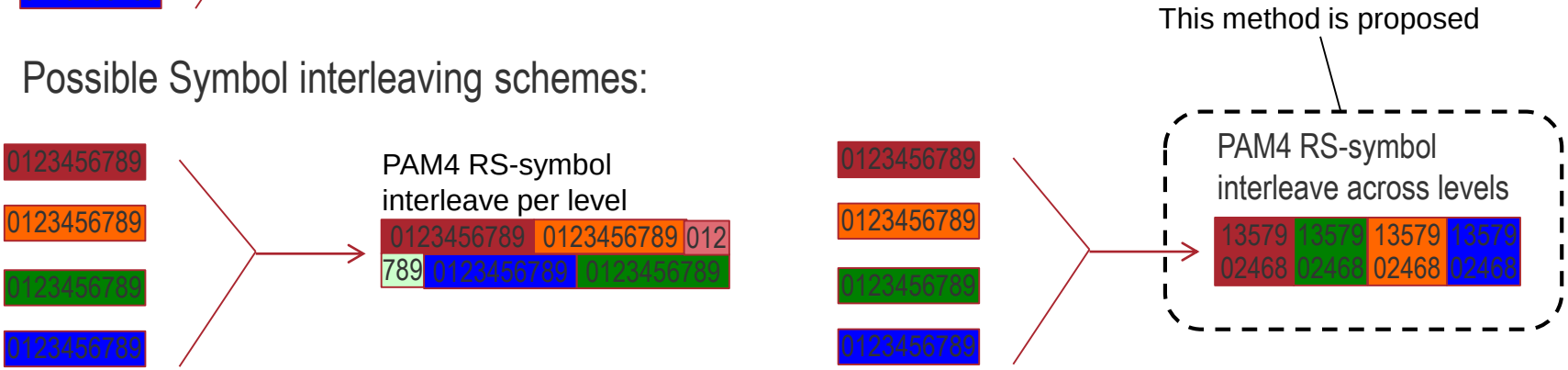
Review of the Options

- From http://www.ieee802.org/3/bs/public/15_01/slavick_3bs_01a_0115.pdf

802.3bs/cd PAM4 bit interleave structure:



Possible Symbol interleaving schemes:

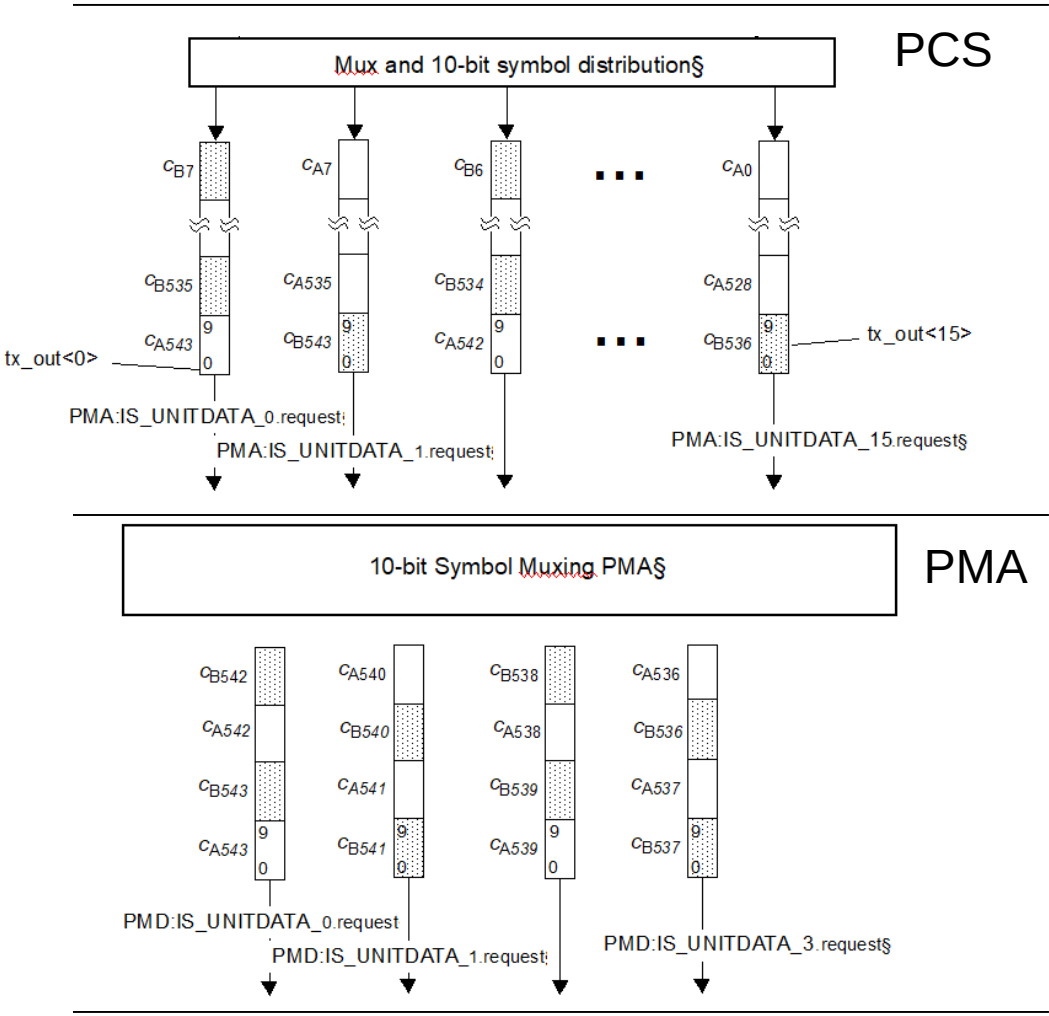


What do burst errors on a PAM4 link look like?

- Are they single level bursts
- Do they toggle between levels
- This impacts the multiplexing choice we would make

More TX Details

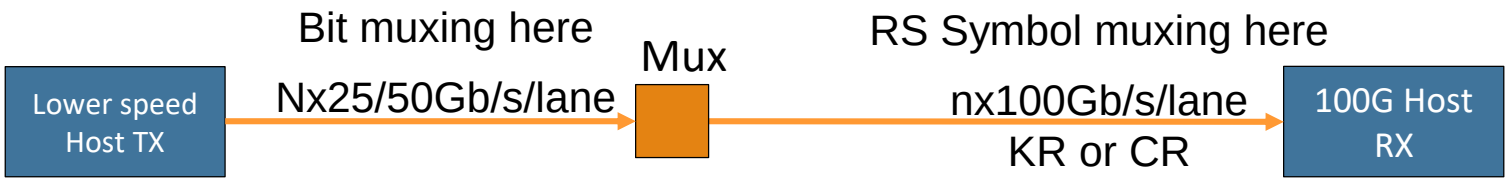
- An example of the RS symbol muxing
- 16 PCS lanes down to 4x100G lanes
- The rules are:
  - PMA achieves symbol lock (or it is provide by the PCS), common marker alignment is sufficient
  - In round robin fashion, play out RS symbols one at a time from each of the incoming PCS lane
  - Which PCS lane is played out where does not matter, as long as once muxing starts it continues in the same fashion



TX Complexity for Different Scenarios

- Simple for a 100G/lane host TX to RS symbol mux to 4 lanes vs. bit muxing

- Transmit side of mux must performs RS symbol muxing

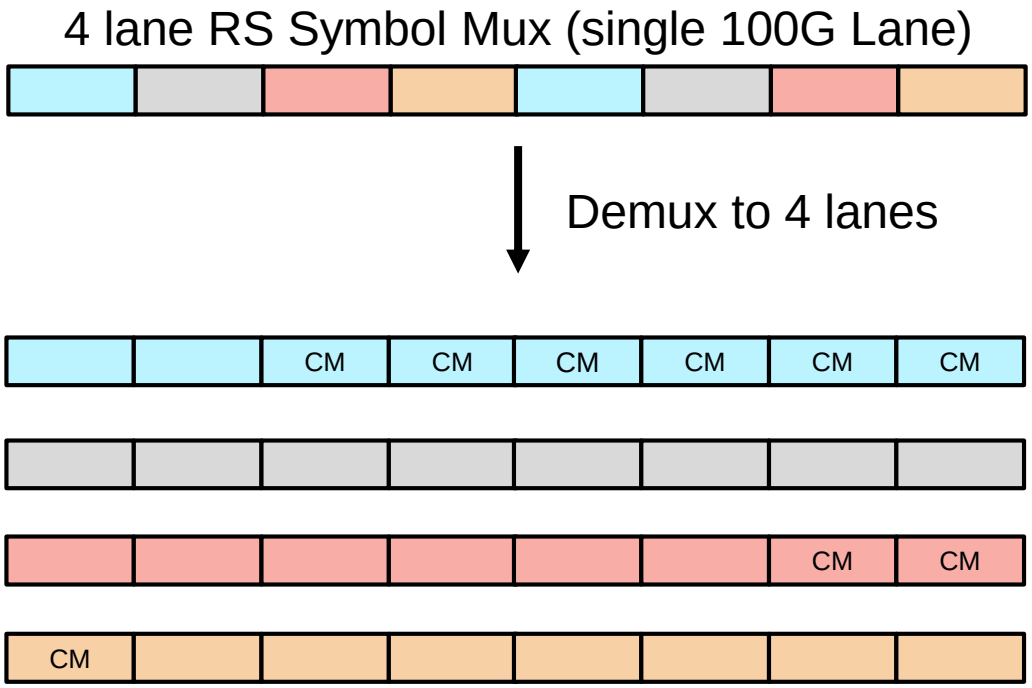


More RX PMA Details for 100G lanes

- The rules are:
  - PMA achieves symbol lock; common marker alignment to any one of the 4 PCS lanes, on a given PMA lane, is sufficient
    - You look for AMs across 10b chunks that are 40b apart
    - If you don't find lock, you shift 1b and try again
  - In round robin fashion, play out RS Symbols one at a time to each of the outgoing PCS lanes
    - If the PMA is adjacent to the PCS
  - Which PCS lane is played out where does not matter, as long as once demuxing starts it continues in the same fashion
    - No reordering or deskew between the PCS lanes is necessary

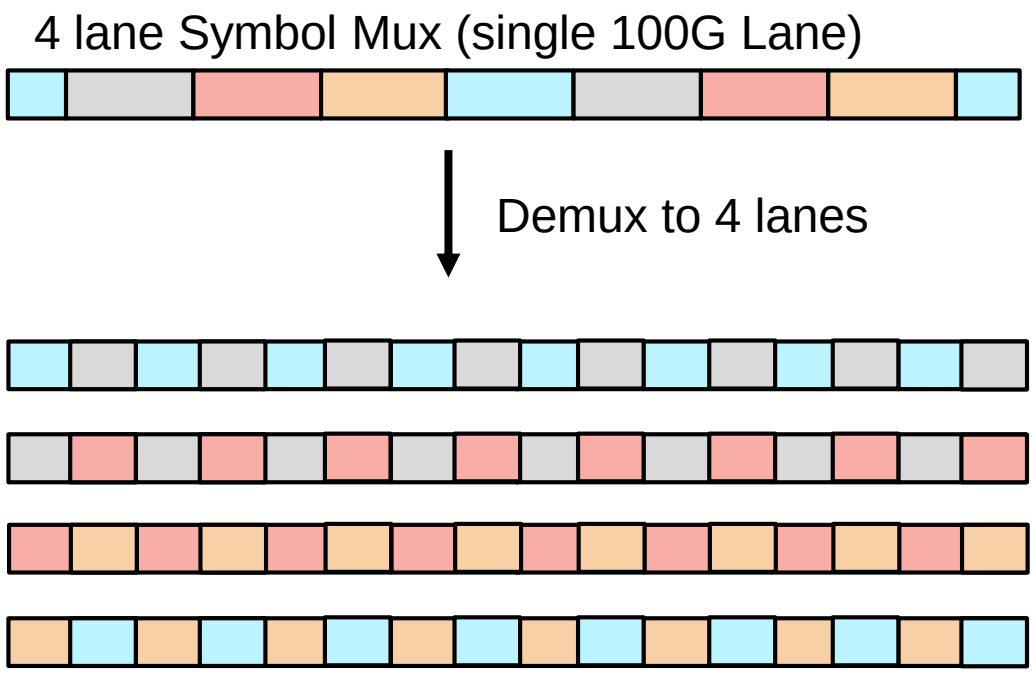
More RX Details

- An example of the symbol demux operation
- Demux each 100G lane to 4x25G lanes
  - RS Symbol demuxing
- Then hunt for AMs, common marker portion only
- If you can find AMs, then you shift your symbol demux by one bit
- You will find alignment after 10b shifts maximum



Example for starting on non 10b RS boundary

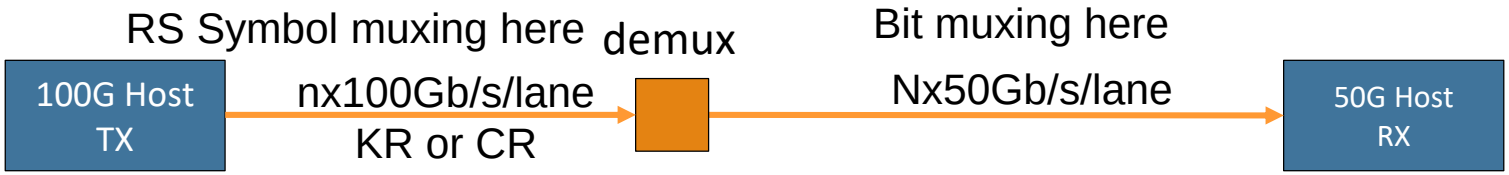
- An example of the symbol demux operation
- Demux each 100G lane to 4x25G lanes
  - RS Symbol demuxing
- Then hunt for AMs
- In this case you started off of a non 10b RS Symbol boundary
  - You won't find alignment
- You need to shift 1b, try again etc.



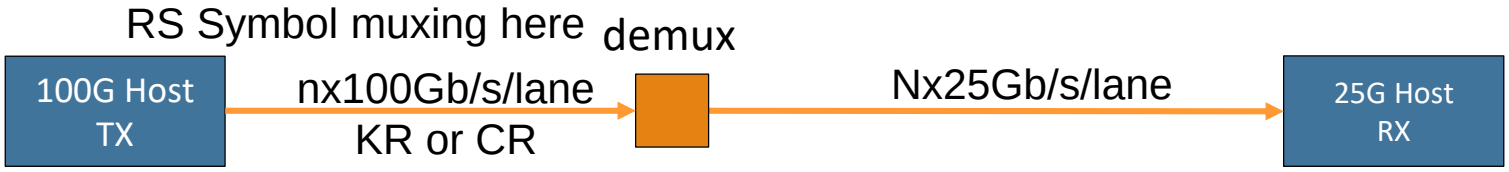
RX Complexity for Different Scenarios

- Straight forward for a 100G/lane host RX to RS symbol demux to 4 lanes vs. bit muxing
  - Though more complicated than bit demuxing, PMA must be PCS aware, lock to AMs etc.

- Receive side of demux must sync up to AMs, and convert data from RS Symbol to bit muxing



- Receive side of demux must sync up to AMs, and convert data from RS Symbol muxing



More Work…

- How much can RS Symbol muxing help?
  - Can it help close our longer channels?
- What error models assumptions should we make?
- We need simulations and analysis to see where we stand
- Need to look at false lock probability etc.

Conclusion

- In gustlin_3ck_01_0518.pdf, it is suggested that we should reuse the current 802.3bs/cd PCSs for this project, especially for the C2M interface to maintain compatibility with existing PMDs
  - This includes 4:1 bit muxing in the PMA
- This presentation explored the possibility of RS Symbol muxing instead of bit muxing for our more difficult channels
- We need to quantify how much this would help preserve FEC gain and if it helps close the adopted objectives (assuming they can't be closed with bit muxing)
- Options to choose from for a given PHY type (in order preference):
  1. Current bit muxing - Simplest
  2. RS Symbol muxing - Adds a little complexity, not backward compatible for optical
  3. New FEC scheme - Unknown higher complexity

Thanks!