Architectural Options and Technical Feasibility of 100GbE

IEEE P802.3 Next Generation 100 Gb/s Ethernet & 200 Gb/s Ethernet Study Group

January 2016

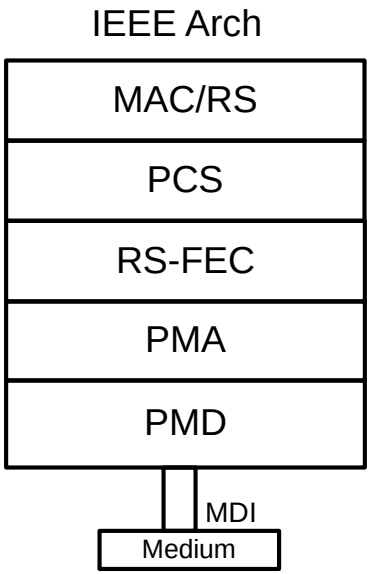
Mark Gustlin - Xilinx

Introduction

- This looks at an architectural option and technical feasibility of next generation 100GbE
  - Based on 50Gb/s per lane signaling
- The following assumes reuse from 802.3ba/bj architecture, and that FEC is always required
  - One person noted at the study group that they believed for the MMF PMD, KP4 (RS(544,514)) is needed
- Adopted objectives:
  - Define a 2 lane 100 Gb/s PHY for operation over copper Twinaxial cables
  - Define a 2 lane 100 Gb/s PHY for operation over printed circuit board backplane
  - Define a 2 lane 100 Gb/s PHY for operation over MMF with lengths up to at least 100m

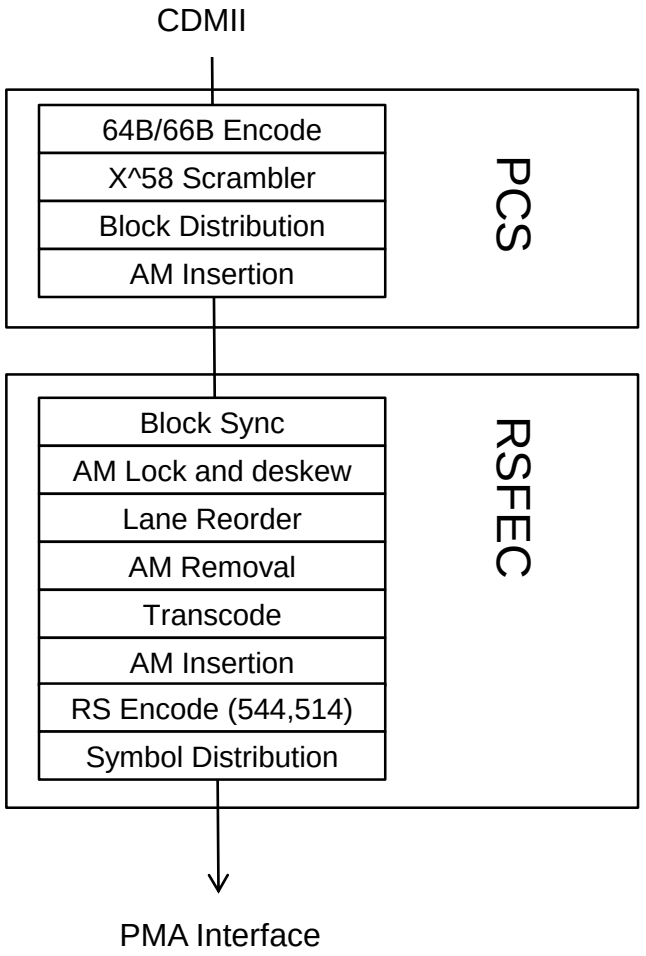
PCS/FEC Architecture

- Based on the draft 802.3bj system architecture



Possible TX PCS/FEC Data Flow

- Keep functional split from 802.3bj
- Same PCS with a separate RS-FEC sublayer
  - Backwards compatible with exiting 100GbE
  - 20 PCS lanes
  - Use RS(544,514) FEC
  - Same transcoding
- What is different?
  - Distribute FEC blocks to 2 lanes (vs. 4 for 802.3bj)
  - This preserves FEC gain compared to distribution to four lanes and then bit multiplexing, especially for correlated errors



Data Distribution

- This is the format for 802.3bj:

FEC Lane	Reed-Solomon symbol index (10 bit symbols					
0	AM0 (0 … 63)	AM4	AM8	AM12	AM16	5b pad
1	AM0	AM5	AM9	AM13	AM16	
2	AM0	AM6	AM10	AM14	AM16	
3	AM0	AM7	AM11	AM15	AM16	

- This is a possible format for 2x50G 100GbE:

FEC Lane	Reed-Solomon symbol index (10 bit symbols									
0	AM0 (0 … 63)	AM0	AM4	AM6	AM8	AM10	AM12	AM14	AM16	AM18
1	AM0	AM0	AM5	AM7	AM9	AM11	AM13	AM15	AM17	AM19

Backwards compatibility

- Does the PCS/FEC for the these projects need to run over 4 lanes (4x25G)?
- The RS(544,514) is not compatible with deployed interfaces
  - It has more overhead and therefore the lane rate is higher
  - And is not RS(528,514)
- The proposed format could be sent over 4 lanes, by splitting up the blocks to the same format at 802.3bj
- They could also be bit inverse muxed to 4 lanes
  - But to go back to 2 lanes you would need to be protocol aware

Options for Backwards Compatibility

- Re-use 802.3bj as is, with RS(528,514), bit mux to get to 2 lanes

FEC Lane	Reed-Solomon symbol index (10 bit symbols)					
0	AM0 [0:63]	AM4	AM8	AM12	AM16	
1	AM0	AM5	AM9	AM13	AM16	
2	AM0	AM6	AM10	AM14	AM16	
3	AM0	AM7	AM11	AM15	AM16	

5b pad

- T=7 for this code, worst case with a single 4-bit burst error you can use up T=4!

Options for Backwards Compatibility

- Re-use 802.3bj as is, with RS(544,514), bit mux to get to 2 lanes

FEC Lane	Reed-Solomon symbol index (10 bit symbols)					
0	AM0 [0:63]	AM4	AM8	AM12	AM16	
1	AM0	AM5	AM9	AM13	AM16	
2	AM0	AM6	AM10	AM14	AM16	
3	AM0	AM7	AM11	AM15	AM16	

5b pad

- T=15 for this code, worst case with a single 4-bit burst error you can use up T=4, not as bad

- But no implementations implement this today for 100GbE, so no real backwards compatibility?

Conclusion

- This presentation looks at one option for the 100GbE architecture
- This architecture is feasible, it follows 802.3ba/bj architectures which has been shown to be technically feasible
- This maximize reuse and allows one PCS/FEC design to support 4x25G and 2x50G, things that are different
  - RS(544,514) which is in 802.3bj but rarely used
  - Distribution to two lanes instead of four lanes
  - Per lane rate of 26.5625G
- Once technology is chosen for the PMD objectives, then we can validate that the gain is enough

Thanks!