200 Gb/s per lane AUI

Performance summary and implications on the architecture

Adee Ran, Cisco

Introduction

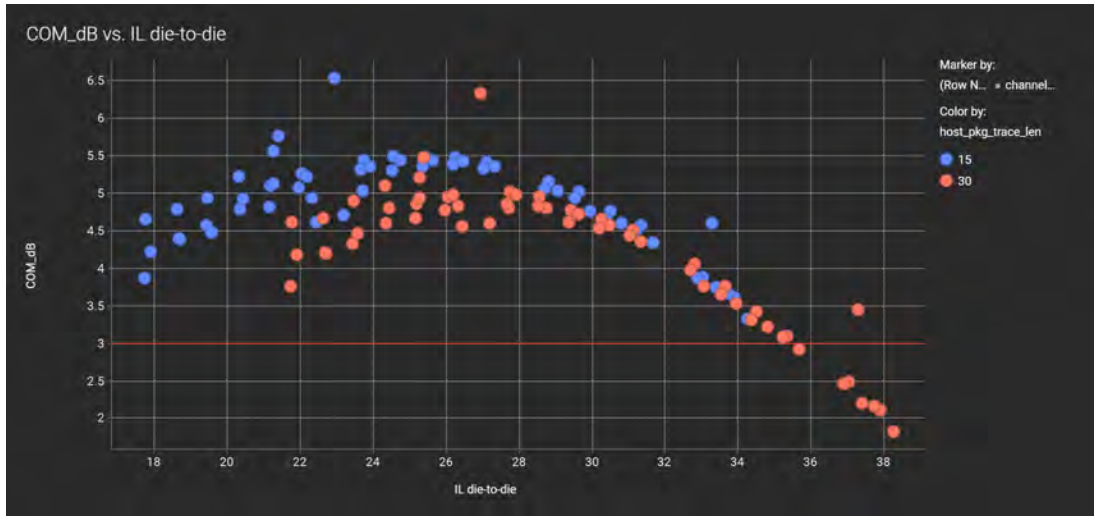
- This presentation is a report out of a more detailed presentation provided in the electrical ad hoc – see ran_3df_elec_01b_220921
- In this presentation I address the implications on architecture
- Also, a brief report of FEC performance with correlated errors is provided

Summary of electrical ad hoc presentation

- A set of COM parameters was proposed for 200 Gb/s AUI application for operating with:
  - IL range up to 36-38 dB die-to-die as suggested in kareti_3df_01a_2207 for large-scale switch applications
  - Channels contributed to the task force so far (Akinwale, Mellitz, Rabinovich)
  - Package model proposed in benartsi_3df_01b_2207 for large-scale switch, with host package IL up to 9.7 dB, module package with IL of 3 dB
  - Die model proposed in mli_3df_02a_220316
- The proposed COM parameters represent a feasible (if challenging) endpoint implementation
- Sensitivity analysis of key parameters was presented

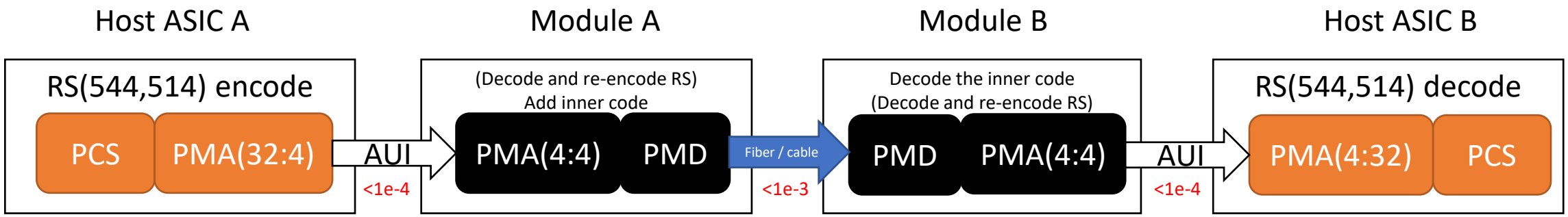
Summary of electrical ad hoc presentation

- Feasibility of operation over contributed C2M channels with die-to-die IL of 18-37 dB was demonstrated by COM analysis
- Assumptions include:
  - DER (detector random error rate) of ~1e-4
  - Strong DFE equalization: $b_{max}(1)=1$ (or close)
- The required Tx/Rx specs are similar to those of “long reach” (CR/KR) previous generations
- Implications beyond electrical specifications must be considered:
  - FEC scheme
  - Symbol muxing
  - Link training on AUI



Source: [ran_3df_elec_01b_220921](), page 25

FEC scheme



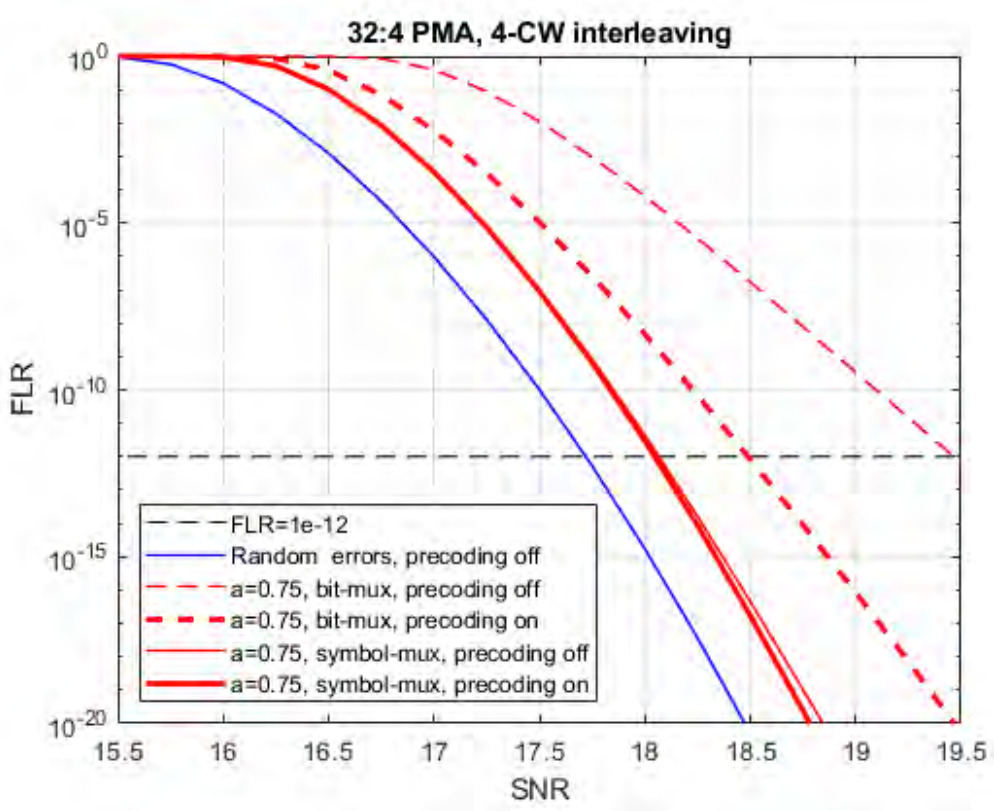
- High loss hosts such as high-radix switches may not be able to achieve DER (PAM4 random error ratio) of ~1e-5 as assumed in previous generations
- We can assume they achieve DER≈1e-4, which requires RS error correction and re-encoding in the module
- The optical link likely needs stronger error correction, provided by concatenated code
- Some hosts may have low enough BER to allow bypassing error correction in the module
  - This would enable significant reduction in power and latency, when it's possible
- A concatenated FEC with **termination in the module, that can optionally be bypassed**, would be a good solution for covering both cases.

Bit/symbol muxing

- Using 8:1 bit muxing (e.g. from 32 PCS lanes to 4 AUI/PMD lanes) would severely degrade the RS-FEC performance with correlated errors
  - Long error bursts are expected due to equalization of high loss channel; close to the maximum a=0.75
- With strongly correlated errors, the required BER would be much lower than previously assumed
  - For an end-to-end link: 2e-4 →1e-5
  - Allocating only a part of the BER to the AUI is not feasible
- Symbol muxing eliminates the error propagation penalty
  - Best utilization of the 4-way codeword interleaving in the PCS

FEC performance: bit vs. symbol muxing

- With bit muxing, the coding gain degradation from correlated errors can be 1.75 dB
  - Requires 1/20 factor in pre-FEC BER to get the same FLR
  - Or, a factor of a billion in FLR
- Precoding reduces the degradation to 0.75 dB
  - A factor of ~4 in pre-FER BER
  - Precoding with bit muxing has a penalty with any error profile
- With symbol muxing, The maximum penalty is 0.3 dB
  - A factor of 2 in pre-FEC BER
  - Precoding has a small benefit and may not be required
  - Almost no penalty with lower values of a
- **The effect is too large to be dismissed**
- **We should use symbol muxing.**



Summary

- Chip-to-module is comparable to “LR” in previous generations.
- A combined segmented/concatenated FEC structure is preferred.
- Symbol muxing instead of bit muxing of PCS lanes is necessary.

Thank you

Questions / Comments