Need for FEC-protected chip-to-module CAUI-4 specification

Piers Dawe

Mellanox Technologies

Supporters

- Yonatan Malkiman Mellanox
- Adee Ran Intel
- Oded Wertheim Mellanox
- Arash Farhood Cortina Systems
- Richard Mellitz Intel
- Liav Ben-Artsi Marvell Technology Group
- Charles Moore Avago Technologies

Introduction

- There are two very different retimed chip-to-module interfaces:
  - Host to/from 100GBASE-LR4 module. CFP4, perhaps QSFP in future. No FEC
  - Host to/from 100GBASE-SR4 module. Probably QSFP. With FEC

- A switch for a high density data centre use will support 100GBASE-CR4 and 100GBASE-SR4
  - Both with FEC

- It might or might not support 100GBASE-LR4
  - Might not support non-FEC modules at all, except for 40GBASE-SR4

- The FEC is in the host (802.3bj silicon)
  - It protects the chip-to/from-module links as well as the optical link

- The non-FEC chip-to-module CAUI-4 specification is unnecessarily expensive for this switch
  - In particular, design and test costs driven by BER <= 1e-15 will be avoided by not using full strength chip-to-module CAUI-4
  - A lower cost option will be defined

- 100GBASE-SR4 modules will have to support this

Lower cost CAUI-4

- **Musts**
  - Reduced test costs
  - Reduced design costs
  - 100GBASE-SR4 coexisting with 100GBASE-CR4 in adjacent ports
  - Minimise unnecessary power consumption
- **Wants**
  - Compatibility with nPPI and full-strength C2M CAUI-4
- **Method**
  - Choose an appropriate BER spec
  - Consider reduced eye mask
    - Don't require too large an SR4 Rx electrical signal
- **This presentation investigates creating a C2M CAUI-4 lite with minimal differences to full-strength C2M CAUI-4**
  - Resulting in two options in Annex 83E
  - Possibly with two names
  - Keeping the same VSR methodology; nearly all the annex is common to both options

Evolving CAUI-4 to lower cost CAUI-4 lite for use with FEC

- **Define host and module output eye height and eye width at 1e-6**

  - EH6 and EW6. Same CTLE

  - For now: use EH6 and EW6 specs with the same limits as full strength CAUI-4's EH15 and EW15

  - Not counting irrelevant statistical tails allows somewhat lower voltage swings

    - Good for power and crosstalk

    - Also more tolerant to e.g. channel ILD

  - For the future: look to see what lower limits can be chosen that do not require better receivers

    - Noticing that extrapolation is not required, but 3 x 4 million samples takes at least 2 minutes per lane on a sampling scope, define eye mask that allows shorter test times

- **Host and module input testing at BER <= 2.5e-6**

  - Much reduced test time and cost

- **Other changes?**

  - Are the host reflection specs the same for CAUI-4 and 100GBASE-CR4?

  - Is it worth revisiting the module reflection specs?

  - Other?

- **Does this methodology deliver enough of an improvement?**

Choice of BER limit

- **Traditional non-FEC method: example**
  - XAUI spec 1e12
  - PMD spec 1e-12
  - XAUI spec 1e-12
  - BER varies very strongly with SNR. Although the BERs add, it is very unlikely that all three links have spec-worst SNR. Compound XAUI-optical-XAUI links turn out better than spec
  - 1e-12 or better delivered

- **With FEC, it's different**
  - Adding together pre-FEC BERs would give a super-linear increase in post-FEC BER, so be more cautious
  - Want a pre-FEC BER <= 5e-5 for 1e-12 after FEC (errors in optical link expected to be uncorrelated). Want to allow the optical link to make nearly all of the errors
  - Allow each CAUI-4 lite link to have a spec BER of 2.5e-6, or only 5% of the optical link's spec
    - The corrected BER 2.5e-6 is ~3e-23
    - Errors in CTLE-based CAUI-4 lite also expected to be uncorrelated (no DFE needed)
  - Pre-FEC BER varies strongly with SNR: the difference between 4.5e-5 and 5e-5 is 0.03 dB of optical power. It is very unlikely that all three links have spec-worst SNR. Compound CAUI-optical-CAUI links will turn out better than spec
  - 1e-12 or better delivered after FEC correction – determined by optical link

We already have two things in Annex 83E

- In 802.3, an interface is specified "logically" (what bits and coding) as well as, often, timing and electrical specifications

- Annex 83E contains two things, at present both using the same name
  - One with FEC,
  - One without

- We could name them CAUI-4p for the RS-FEC protected interface and CAUI-4u for the unprotected interface

Conclusions

- The non-FEC chip-to-module CAUI-4 specification is unnecessarily expensive for use with 100GBASE-SR4 modules

- A lower cost option is needed

- Create two options in Annex 83E:
  - EH6 and EW6
  - Stressed input test to maximum BER 2.5e-6

Thank You