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

Piers Dawe

IEEE P802.3bm, Jan 2014, Indian Wells

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
Tom Palkert	Molex
Brian Misek	Avago Technologies

Introduction

- In 802.3, an interface is specified "logically" (what bits and coding) as well as, often, timing and electrical specifications
- There are two different retimed chip-to-module interfaces needed:
  - Host to/from 100GBASE-LR4 module. CFP4, perhaps QSFP in future. No FEC
    - Bits and coding from Clause 82 and Clause 83, timing and electrical from Annex 83E
  - Host to/from 100GBASE-SR4 module. Probably QSFP, possibly CDFP. Always with FEC
    - Bits and coding from Clause 91 and Clause 83
    - But what are the timing and electrical requirements?
    - Annex 83E can be used for dual-purpose (100GBASE-SR4, 100GBASE-LR4) ports
    - Need something suitable for single-purpose 100GBASE-SR4 and dual-purpose (100GBASE-CR4, 100GBASE-SR4) ports
- This presentation addresses the electrical requirements
  - As to timing: this specification addresses Eye Width. Any CDR issues such as jitter accumulation, low frequency jitter tolerance, are assumed to be the same for the two retimed chip-to-module interfaces

Are such ports of commercial interest?

- Yes
- A switch for a high density data centre use will support 100GBASE-CR4 and 100GBASE-SR4
  - Both with FEC
  - A future 16-lane 400G chip-to-module interface is expected to use FEC also
- A **data centre switch** might or might not support 100GBASE-LR4
  - I.e., it might not support non-FEC modules at all, except for 40GBASE-SR4
- A **100GBASE-SR4 module** never has to work without FEC
- 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 setting realistic specs
  - Host (switch) makers will take advantage of this as it suits them, standard or not
    - Possible fragmented market with unnecessary confusion
- 100GBASE-SR4 modules will have to support this / these customer requirements, standard or not
  - **For all the usual reasons, it could be better to have a common spec for this FEC-protected interface**

Lower cost CAUI-4

- This presentation proposes 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
- Advantages vs. 1e-15 "CAUI-4 heavy"
  - Relaxed host channel and IC package requirements
  - Reduced test costs (host and module)
  - Reduced design costs (mainly host)
  - Minimise unnecessary power consumption
- Compatibility
  - 100GBASE-SR4 coexisting with 100GBASE-CR4 in adjacent ports
  - Compatibility with nPPI and full-strength C2M CAUI-4
- Method
  - Choose an appropriate BER spec
  - Use a similar BER for Eye Height and Eye Width
  - Don't require too large an SR4 Rx electrical signal

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 Assume no better receiver silicon
  - 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: benefits both host and module
    - **More tolerant to e.g. channel ILD and loss: benefits mainly host output**
  - For the future: if lower limits can be identified that do not require better receivers, use them
  - Reduced test time for electrical outputs
    - Extrapolation is not required with the 3 x 4 million samples (as for C2M CAUI-4 heavy) but this takes at least 2 minutes per lane on a sampling scope. With a criterion of 1e-6, implementers will be able to extrapolate from much quicker measurements. Benefits both host and module.
- Host and module input testing at BER <= 2.5e-6
  - **Much reduced test time and cost (host and module inputs)**
  - **More tolerant to e.g. channel ILD and loss: benefits mainly host input**

- Because the module has a much shorter, simpler electrical channel than the host, most of the benefit of the FEC protection should go to the host

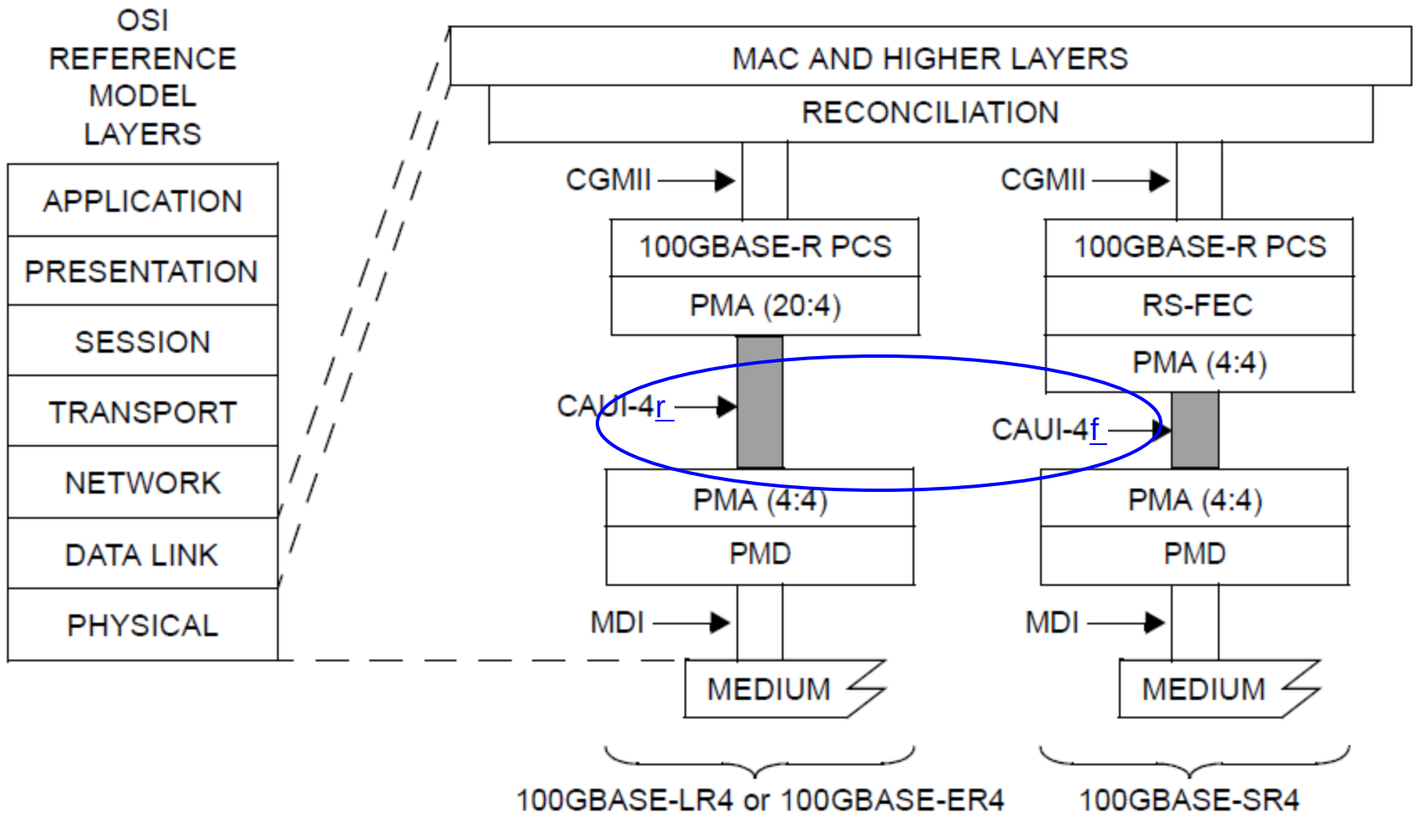
Choice of BER limit

- Compare traditional non-FEC method: example from 802.3ae
  - XAUI spec 1e-12 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 total 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 from pre-FEC 2.5e-6 is ~3e-23; for 5e-6 it's ~9e-21
    - 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, so we don't need to change the optical specs. 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
  - Outside the standard, if 1e-15 or 1e-18 is desired, the CAUI-4 lite is still a negligible contributor to BER

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-4**f** for the RS-**F**EC protected interface and CAUI-4**r** for the unprotected (**r**aw, **R**-encoded) interface
  - Or CAUI-4p and CAUI-4u
  - The FEC status across the host-to-module connection (FEC coded or not?) needs to be determined anyway: must be same for host and module

Detailed changes 1



CAUI-4 = FOUR-LANE 100 Gb/s ATTACHMENT UNIT INTERFACE
CGMII = 100 Gb/s MEDIA INDEPENDENT INTERFACE
MAC = MEDIA ACCESS CONTROL
MDI = MEDIUM DEPENDENT INTERFACE
PCS = PHYSICAL CODING SUBLAYER
PMA = PHYSICAL MEDIUM ATTACHMENT
PMD = PHYSICAL MEDIUM DEPENDENT
RS-FEC = REED-SOLOMON FORWARD ERROR CORRECTION

Figure 83E–1—CAUI-4 relationship to the ISO/IEC Open System Interconnection reference model and the IEEE 802.3 CSMA/CD LAN model

Detailed changes 2

83E.1 Overview

... The nominal signaling rate for each lane is 25.78125 GBd. There are two variants of chip-to-module CAUI-4; in CAUI-4r the signal is not RS-FEC encoded, and in CAUI-4f the signal is RS-FEC encoded. An electrical interface that is compliant to CAUI-4r would also be compliant to CAUI-4f if the signal were RS-FEC encoded.

Detailed changes 3

83E.3.3 CAUI-4 host input characteristics

Table 83E–4—CAUI-4 host input characteristics (at TP4a)

Parameter	Subclause Reference	Value	Units
Bit error ratio (max)[a]	83E.3.3.1	10^{-15}	
CAUI-4r		10^{-15}	
CAUI-4f		2.5×10^{-6}	

83E.3.3.1 Input bit error ratio

The CAUI-4 chip-to-module host input is defined to operate at a bit error ratio (BER) as specified in Table 83E–4 better than 10^{-15} for an input signal defined by 83E.3.3.3.

83E.3.4 CAUI-4 module input characteristics

Table 83E–7—CAUI-4 module input characteristics (at TP1)

Parameter	Subclause Reference	Value	Units
Bit error ratio (max)[a]	83E.3.4.1	10^{-15}	
CAUI-4r		10^{-15}	
CAUI-4f		2.5×10^{-6}	

83E.3.4.1 Input bit error ratio

The CAUI-4 module input is defined to operate at a bit error ratio (BER) as specified in Table 83E–7 better than 10^{-15} for an input signal defined by 83E.3.4.2.

Detailed changes 4

83E.4.2 Host / Module eye contour measurement method

2) Apply respective reference receiver CTLE to captured signal. Any single CTLE setting which meets both eye width and eye height requirements is acceptable.

3) Use the differential equalized signal from step 2 to construct the CDF of the jitter zero crossing for both the left edge (CDFL) and right edge (CDFR), as a distance from the center of the eye. Calculate the eye width (EW6) as the difference in time between CDFR and CDFL with a value of 10^{-6}. CDFL and CDFR are calculated as the cumulative sum of histograms of the zero crossing samples at the left and right edges of the eye normalized by the total number of sampled bits. For a pattern with 50% transition density the maximum value for the CDFL and CDFR will be 0.5. The CDFL and CDFR are equivalent to bath tub curves where the BER is plotted versus sampling time.

4) For CAUI-4f, the eye width EW is EW6.

For CAUI-4r, lLeveraging the Dual-Dirac jitter model described in 48B.1.1, estimate the random jitter. Calculate the best linear fit in Q-scale over the range of probabilities of 10^{-4} to 10^{-6} of the CDFL and CDFR to yield the random jitter on the left edge (RJL) and the random jitter on the right edge (RJR) respectively. The eye width EW is then given by Equation (83E–7)

$$EW = EW15 = EW6 - 3.19 * (RJR + RJL) \quad (83E–7)$$

where

EW15 is the eye width extrapolated to 10^{-15} probability

EW6 is the eye width at 10^{-6} probability

RJL is the RMS value of the jitter estimated from CDFL

RJR is the RMS value of the jitter estimated from the CDFR

5) Use the differential equalized signal from step 2 to construct the CDF of the signal amplitude in the middle 5% of the eye, for both logic 1 (CDF1) and logic 0 (CDF0), as a distance from the center of the eye. Calculate the eye height (EH6) as the difference in amplitude between CDF1 and CDF0 with a value of 10^{-6}. CDF0 and CDF1 are calculated as the cumulative sum of histograms of the amplitude at the top and bottom of the eye normalized by the total number of sampled bits. For a well balanced number of ones and zeros the maximum value for CDF0 and CDF1 will be 0.5.

6) Apply the Dual-Dirac and tail fitting techniques to CDF1 and CDF0 to estimate the noise at the middle of the eye. Calculate the best linear fit in Q-scale over the range of probabilities of 10^{-4} and 10^{-6} of CDF1 and CDF0 to yield relative noise one (RN1) and relative noise zero (RN0).

For CAUI-4f, the eye height EH is EH6.

For CAUI-4r, tThe eye height EH is then given by Equation (83E–8)

$$EH = EH15 = EH6 – 3.19 * (RN0 + RN1) \quad (83E–8)$$

where

EH15 is the eye height extrapolated to 10^{-15} probability

EH6 is the eye height at 10^{-6} probability

RN1 is the RMS value of the noise estimated from CDF1

RN0 is the RMS value of the noise estimated from CDF0

83E.4.2.1 Vertical eye closure

Vertical eye closure is calculated using Equation (83E–9)

$$VEC = 10 \log(AV/EH15) \quad (83E–9)$$

where

VEC is vertical eye closure in dB

AV is the eye amplitude of the equalized waveform. Eye amplitude is defined as the mean value of logic one minus the mean value of logic zero in the central 5% of the eye

EH15 is given in item 6 of 83E.4.2equation Equation (83E–8)

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
  - Host makers will derate the spec, standard or not
  - Can avoid a fragmented market with unnecessary confusion by standardizing the lower cost option

- Create two options in Annex 83E:
  - As well as existing EH15, EW15, 1e-15:
  - EH6 and EW6
  - Stressed input test to maximum BER 2.5e-6
  - For use when signal is FEC encoded

- Simple editorial changes

Thank You