Partial FEC

Piers Dawe
Avago Technologies

Supporters

David Cunningham Avago Technologies

Introduction

- FEC usually used to improve BER
- Clause 74 FEC ("K-FEC") also corrects error bursts of up to 11 bits on the line
- Bursts of errors defeat the "guarantee" of the Ethernet CRC
  - The Ethernet CRC will detect ANY single line error in a frame, with 10GBASE-R coding.
  - It will detect all but around 1 in 2^32 or in 4.3 x 10^9 frames with multiple errors (32 bit CRC, might get a false match by chance)
  - If error rate were poor, mean time to false packet acceptance (MTTFPA) would be inadequate for our high standards
- Therefore, state machine features like "hi_ber" are included to interrupt a high-BER link
- If the errors are grouped in bursts, the "hi_ber" feature's safeguard is largely defeated (I believe)
- Hence K-FEC

Backplane vs. front side links

- Backplane Ethernet is seen as a "closed system"
  - Both cards and the backplane they plug into must be specified together e.g. by a single vendor
  - Not mix and match like front-panel Physical Layer types
  - That vendor can assure that the BER and/or error burst stats of a particular card/backplane combination is adequate for MTTFPA
  - Therefore, K-FEC is optional, not mandatory
- 40GBASE-CR4 and 100GBASE-CR10 is different
  - Anyone can put together a system using two box vendors and a third cable vendor
  - 802.3 has to provide the MTTFPA assurance
  - Does this mean that K-FEC must be mandatory for 40GBASE-CR4 and 100GBASE-CR10?
  - Only partly

FEC parts

- Transcode from non-FEC 64B/66B to FEC
  - Throw away half of each 2-bit sync header, make a CRC for 32, 65-bit blocks, append
  - Ideal latency 32 UI = 3.1 ns for 10GBASE-R or 40GBASE-R, double for 100GBASE-R
  - Low power. Simple logic (if fast)
- Transmit
- Error detection
  - Check the CRC by repeating the same procedure at the receiver
  - Up to 11-UI errors are detected
  - Perfect block goes into FIFO to wait its turn among correctable blocks
- Error correction
  - If implementation judges it feasible, correct an imperfect block
  - Latency on the order of a FEC block (2112 UI = 205 ns)
  - Thought to be the part that takes a significant part of FEC power
- Transcode back
  - Short latency, "low" power, "simple" as above
- Error marking
  - FEC must mark the known bad but uncorrected blocks
  - Optional in-band method in Clause 74 takes another one FEC block ~205 ns
  - Out-of-band method takes much less latency

(By comparison, 100 m of fibre, round trip, is 1000 ns)

Decide what you need to do

- If assuring MTTFPA for 40GBASE-CR4 and 100GBASE-CR10
- These Physical Layer types use long DFE with a comparatively high SNR
- When link is stretched to breaking, the errors will come in bursts
- ... **More than Backplane Ethernet** because there is less "unequalizable" impairment e.g. connector crosstalk and reflections, so link can be equalized more. More tuned up.
- P802.3ba may like to see more study on burstiness and MTTFPA
- Burst error detection is required
  - **But error correction is not**; if link has bad BER, the user can know the BER and do something about it
- **Make K-FEC encoding and detection mandatory for 40GBASE-CR4 and 100GBASE-CR10**
  - Full K-FEC (including correction) is already optional for 40GBASE-CR4 and 100GBASE-CR10
- There might be alternative fixes e.g. modifying the hi_ber state machine, imposing a test against bursty failure

100GBASE-ER4 is different

- 100GBASE-ER4 is the opposite of –CRn
- Random noise from the SOA is the challenge
- With the traditional SNR of an optical link, not much opportunity for excessive equalization
- Do not expect errors to be particularly bursty
- High power Physical Layer type (watts of heat) anyway
- FEC with correction seems attractive

Look no handshaking

- The receiver can choose autonomously whether to correct errors or just detect them
  - There are error counters so station management can see if a particular link's BER is consistently good and can support turning correction off
  - Could be something to be powered down at night for power reduction
  - If taking advantage of the improved latency, have to pay attention to latency change if FEC correction is bypassed
  - The two ends can have different FEC correction settings
- All that's needed is that the link partner transmit with K-FEC encoding on
- No auto-Negotiation needed

Really simple handshaking

- If non-FEC ability is optional
  - Need to get FEC-able and non-FEC PHYs talking to each other
  - This doesn't require Auto-negotiation (exchanging AN frames) either
- Can be done by "Parallel detection"
  - Receiver listens in both FEC and non-FEC modes, then transmits in the best mode that it received in
  - There is a pre-defined priority table to define "best"
  - See other presentation which compares this with Fibre Channel's very similar "Link Speed Negotiation"
- Once communication is established, the two ends can exchange OAM frames to agree a "non-best" mode of operation
- As K-FEC line rate is same as non-FEC line rate, CDRs can run continuously
  - No need to redesign CDR for 1/33 rate operation as AN needs

Summary

- Error bursts are a threat to MTTFPA for 40GBASE-CR4 and 100GBASE-CR10
- Protection against this threat is conveniently done by FEC error detection but does not require FEC error correction
- This reduces FEC power, latency and gate count noticeably
- Does not need Auto-negotiation nor CDR capable of 1/33 rate operation
- Prepare to make K-FEC error ***detection*** mandatory for 40GBASE-CR4 and 100GBASE-CR10