

# Pre-FEC BER Monitoring Proposal

---

David Ofelt – Juniper Networks

# Contributors

- Jeffery Maki – Juniper Networks

# Supporters

# Background & History

- For a longer introduction and motivation:
  - See maki\_3bs\_01a\_1115.pdf from the 2015-11 Dallas meeting
- BIP8 added to 802.3ba to track bit errors on the wire
  - Added counters but no policy or signaling mechanism
  - Customers requested features to track BER and signal the transmitter when the receiver's BER passed a threshold
  - Transmitter then could trigger a re-route before the link became too degraded
  - Added in non-interoperable proprietary ways by multiple vendors

# FEC-Enabled Opportunity

- FEC provides an opportunity to improve on this functionality
  - Pre-FEC BER can show link health before packet errors are seen
  - Customers are requesting the ability to:
    - Declare a link as failed at a user-settable threshold
    - Declare a link as degraded and then signal the transmit side
  - These signals need to be carried through a FEC change in the extender sublayer (XS)
- Adding these features to the standard allows for interoperability and a consistent feature set.
- Some things proposed here are similar to existing features & proposals
  - Please let me know if I've missed things (since I'm sure I have...)

# How to count BER

- There are existing per-lane counters for RS symbol errors
  - In 802.3bs these are in: 45.2.3.47a,b
  - Proposal is to use symbol errors as proxy for bit-errors
- Can use FEC codeword/block as the unit of time
  - 5440 bits for KP4 == 13.6ns
- Specific BER set by specifying a threshold for the number of errors in a given number of FEC blocks.
  - Ex – a BER of  $1e-5$  using a 10ms window is 4000 errors in 735k FEC blocks
- This is how the hi\_ser functionality in 802.3bj (91.5.33) works.
  - .3bj has a fixed interval of 8192 codewords and fixed thresholds of 417/6380 for KR4/KP4

# Thresholds

- Two thresholds:
  - Degradate
  - Error
- Configurable Values for:
  - Threshold
  - Interval
- Sensitivity can be controlled by averaging over a longer or shorter period

## Threshold Crossing

FEC excessive threshold

- If the pre-FEC BER exceeds this mandatory threshold for longer than the *Interval*, link fault signaling is generated

FEC degrade threshold

- If the pre-FEC BER is higher than this optional threshold for longer than the *Interval*, a link degrade signal is generated

5 IEEE P802.3bs 400GbE Task Force Dallas, TX, November 2015 JUNIPER NETWORKS

From maki\_3bs\_01a\_1115.pdf

# Configuration range

- Open question is what the lowest pre-FEC BER you would want to trigger either error or degrade at?
  - A 32b interval counter counting FEC codewords is 58s
  - A pre-FEC BER of  $1e-13$  is 2.3 errors in this window
- Reaction timescales are likely all  $>1ms$ , so a larger interval granularity would be fine.
  - A 1us interval granularity gets to  $1e-15$
  - A 1ms interval granularity gets to  $1e-18$



# Pre-FEC Error

- Essentially a generalization of the hi\_ber functionality
  - If receiver detects  $\text{Error}_{\text{threshold}}$  errors in  $\text{Error}_{\text{interval}}$  codewords it declares the link down
- This is very similar to the hi\_ser functionality in 802.3bj (91.5.33)
- No new signaling necessary to inform transmitter
  - Receive side decides it is over threshold and generates LF
  - Normal mechanisms turn around LF and send RF to other end
- If FEC in XS goes over the error threshold – it can signal with LF/RF as appropriate.

# Pre-FEC degrade

- If receiver counts  $\text{Degrade}_{\text{threshold}}$  errors in  $\text{Degrade}_{\text{interval}}$  codewords it declares the link degraded
- Receiver signals the transmitter using new signaling mechanism
  - XS FEC transitions participate
- New signaling bits - Local/Remote Degrade
  - Spiritually similar to LF/RF
  - Receiver turns around LD and sends RD
  - Proposal is to put in the alignment marker fields somewhere
- No policy attached to degrade
  - Status reported through register and system is responsible for what actions are taken
- Degrade detect is very similar to the  $\text{hi\_ser}$  calculation in 802.3bj (91.5.33)
  - Policy and signaling are different than in 802.bj

# Summary of Proposal

- Add new configuration registers:
  - $\text{Error}_{\text{threshold}}$
  - $\text{Error}_{\text{interval}}$
  - $\text{Degrade}_{\text{threshold}}$
  - $\text{Degrade}_{\text{interval}}$
- Add new state:
  - $\text{Error}_{\text{count}}$
  - $\text{Degrade}_{\text{count}}$
- State machines:
  - Error tracking
  - Degrade tracking
- Alignment marker bits
  - Local Degrade
  - Remote Degrade
- Receiver turns around LD and transmits RD
- Add new status bits
  - FEC above error threshold
  - FEC above degrade thresholds
  - LD Set
  - RD Set

# Next Steps

- Decide on optional/mandatory for these features
- Work out BER dynamic range issues
  - Size & granularity of interval setting
  - Size of threshold setting
- Put together detailed comments against D1.2

# Thanks!

---