Pre-FEC BER Monitoring Proposal

David Ofelt – Juniper Networks

Gary Nicholl – Cisco

Contributers

- Jeffery Maki – Juniper Networks

Supporters

- Tom Issenhuth – Microsoft
- Michael Lyngbøl - TDC
- Xinyuan Wang – Huawei
- Tong Tong Wang – Huawei

Background & History

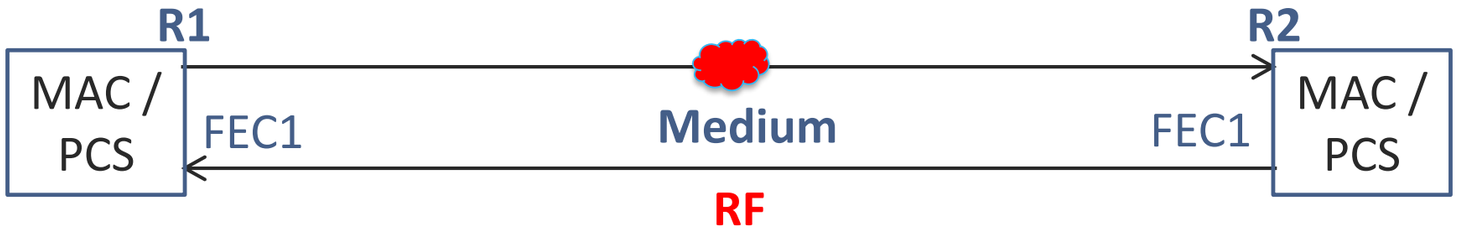
- For a longer introduction and motivation:
  - See maki_3bs_01a_1115.pdf from the 2015-11 Dallas meeting

- BIP8 added to 802.3ba to track bit errors on the wire
  - Added counters but no policy or signaling mechanism
  - Customers requested features to monitor BER
    - Receive notification when BER above a user-defined threshold
    - Signal in band to the transmit MAC when the receiver's BER passed a threshold
  - Customer can use these notifications to manage their network better
    - Ex: Router with the Transmit MAC could trigger a re-route before the link became too degraded
  - Added in non-interoperable proprietary ways by multiple vendors

FEC-Enabled Opportunity

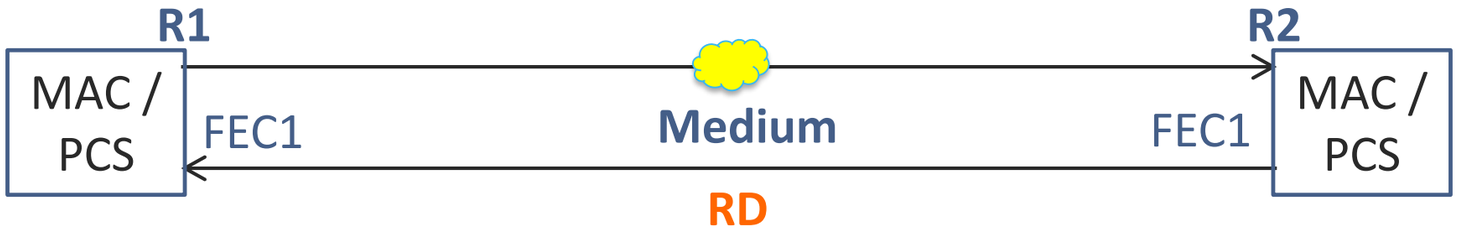
- FEC provides an opportunity to improve on this functionality
  - Pre-FEC BER can show link health before packet errors are even seen on the link
  - Customers are requesting the ability to:
    - Declare a link as failed at a user-settable threshold
    - Declare a link as degraded at a user-settable threshold 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.

Pre-FEC Fault Example



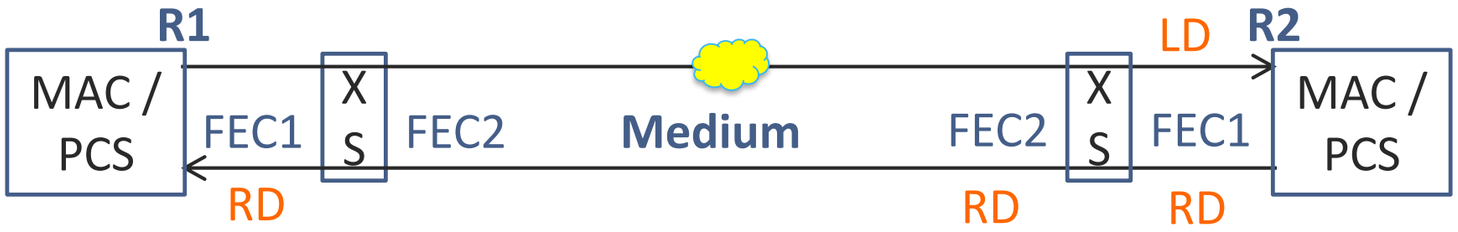
- Receive MAC in R2 detects pre-FEC count above the fault threshold
- RF sent to transmit MAC in R1
- Interface brought down

Pre-FEC Degrade Example



- Receive MAC in R2 detects pre-FEC count above the degrade threshold
- Remote Degrade (RD) signal sent to the transmit MAC in R1
- Traffic not affected

Pre-FEC Degrade Example with XS



- Receive XS at R2 detects pre-FEC error above degrade threshold
- Receive XS at R2 signals Receive MAC using LD (Local Degrade)
- MAC in R2 signals a RD (remote degrade) to the MAC in R1. The RD signal is regenerated by both the XS in R2 and in R1.
- Packet traffic not affected

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 == 12.8ns
- 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 42,483 errors in 781k 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

Pre-FEC Fault

- Essentially a generalization of the hi-ber functionality
  - If receiver detects $Fault_{threshold}$ errors in $Fault_{interval}$ 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 fault threshold – it can signal with LF/RF as appropriate.

Pre-FEC Degrade

- If receiver counts $Degrade_{threshold}$ errors in $Degrade_{interval}$ 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 MAC turns around LD and sends RD
  - Proposal is to put them in the alignment marker fields
    - Limits min detection time to ~100us
- 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 hi_ser calculation in 802.3bj (91.5.33)
  - Policy and signaling are different than in 802.bj

Summary of Proposal

- Optional Feature
- Add new configuration registers:
  - $Fault_{threshold}$
  - $Fault_{interval}$
  - $Degrade_{threshold}$
  - $Degrade_{interval}$
- Add new state:
  - $Fault_{count}$
  - $Degrade_{count}$
- State machines:
  - Fault tracking
  - Degrade tracking
- Alignment marker bits
  - Local Degrade (LD)
  - Remote Degrade (RD)
- Ability bits
  - Fault Threshold Ability
  - Degrade Threshold Ability
- Receiver turns around LD and transmits RD
- Add new status bits
  - Local $Fault_{count}$ above $Fault_{threshold}$
  - Local $Degrade_{count}$ above $Degrade_{threshold}$
  - LD detected in Alignment Marker
  - RD detected in Alignment Marker

Next Steps

- Decide Consensus
- Put together detailed comments against D1.2

Straw Poll

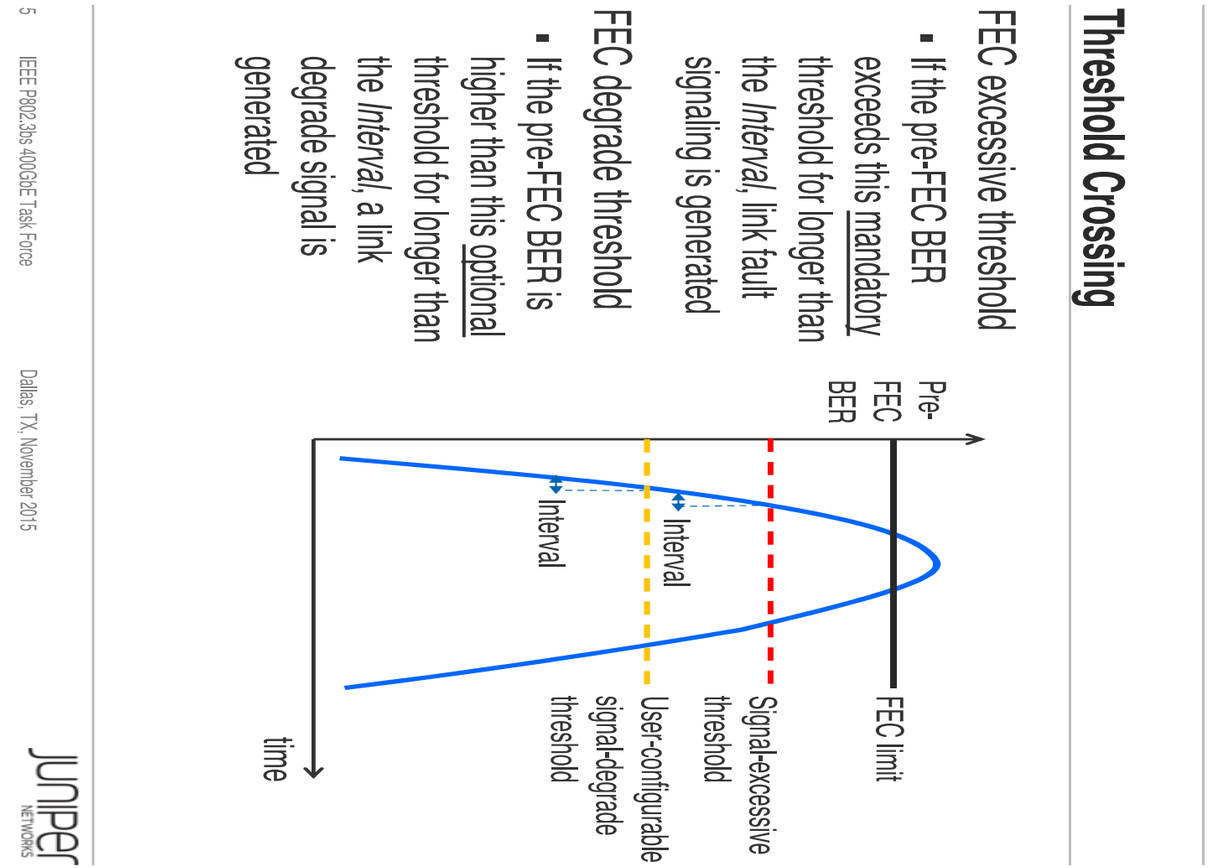
- I support the addition of the optional pre-FEC BER degrade feature proposed in ofelt_3bs_01a_0116.pdf.
- (vote for one)
  - Yes - 24
  - No – 0
  - Abstain - 14

Thanks!

Backup

Thresholds

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



From maki_3bs_01a_1115.pdf