Consideration on 100Gbps/lane PCS/FEC for 800GbE/1.6TbE

Leon Bruckman (Huawei), Yan Zhuangyan (Huawei)

Purpose

- Present baseline thoughts on 100Gbps/lane based 800GbE and 1.6TbE PCS, FEC and PMA.
- Highlight issues that need attention when adapting the existent architecture to 802.3df

The 100G/lane based PMD Objectives

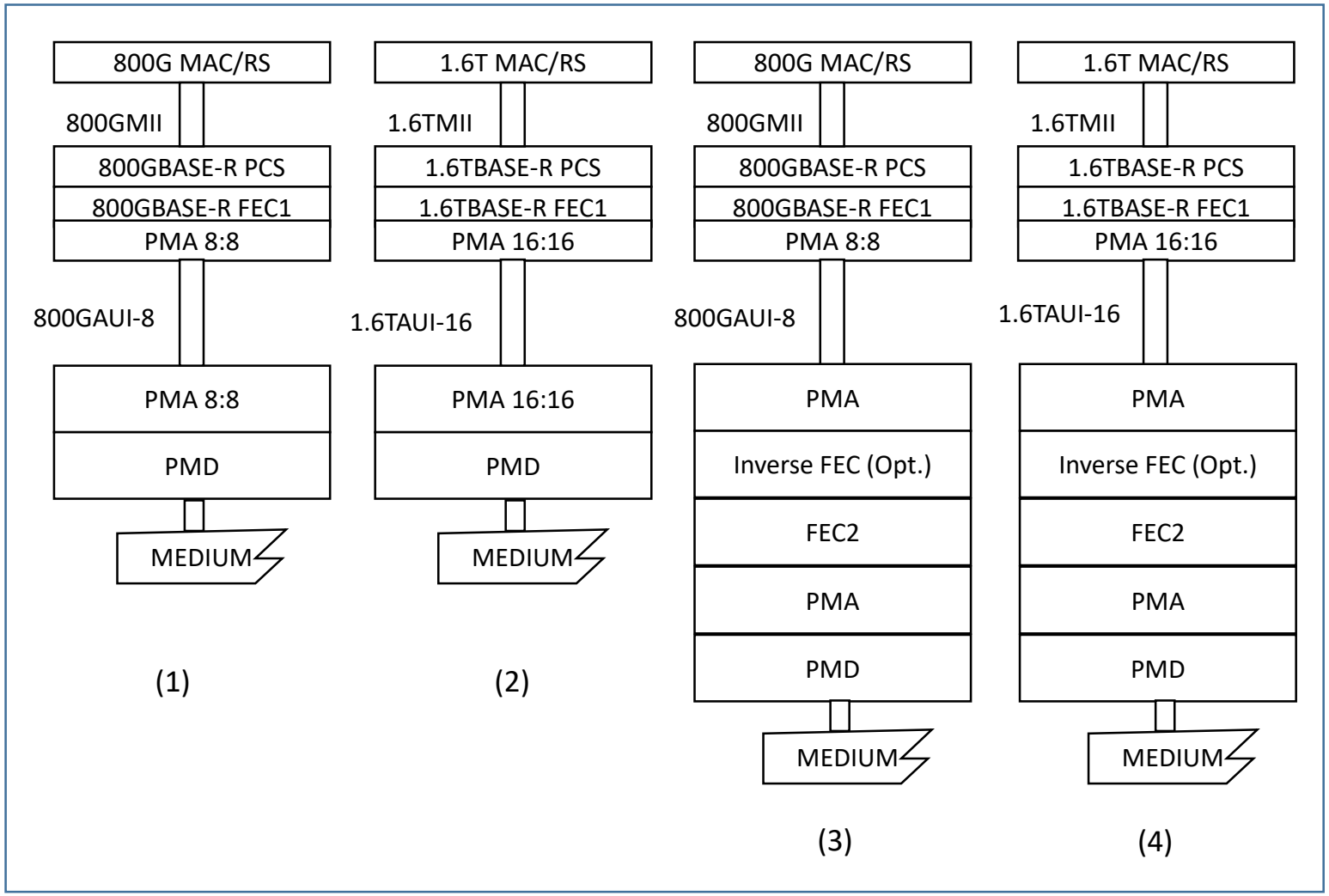
- **Rate: 800Gb/s Ethernet**
  - AUI: 800GAUI-8
  - Backplane: 800GBASE-KR8
  - Copper cable: 800GBASE-CR8
  - MMF 50m: 800GBASE-VR8
  - MMF 100m: 800GBASE-SR8
  - SMF 500m over 8 pairs: No agreed name yet
  - SMF 2Km over 8 pairs: No agreed name yet
- **Rate: 1.6Tb/s Ethernet**
  - AUI: 1.6TAUI-16

Leverage existing or work-in-progress 100 Gb/s per lane (e.g. 3bs, 3cu, 3ck, 3db) to higher lane counts

Baseline points for 100G/lane based 800GbE and 1.6TbE PCS, FEC and PMA

- Reuse as much as possible the 200GbE/400GbE architectures in 802.3bs.
- Separate the PCS and FEC sublayers.
  - gustlin_3df_01_220118.pdf
- End to End FEC, possibly reusing RS(544, 514) and interleaving.
- Define 800GMII and 1.6TMII.
- Increase FEC lane rate to 106.25Gb/s.
  - Take advantage of 802.3ck work
- 8 FEC lanes for 800Gb/s and 16 FEC lanes for 1.6Tb/s.
- Reuse Clause 120 PMA – Small adaptations

Architecture for 800GbE and 1.6TGbE



- (1) and (2) provides an end-to-end FEC architecture for 800GbE and 1.6TbE that can be applied to PMDs with 8 lanes of 100 Gb/s.
- If PMDs need a stronger FEC (e.g. for long reach or 200G/lane modules) , segmented FEC architecture (3) and (4) can be used to provided segmentedFEC or concatenated FEC for more difficult PMDs.
- Inverse FEC not needed for concatenated FEC

Separating the FEC sublayer

- **IEEE 802.3bs PCS services:**

  a) Encoding (decoding) of eight 200GMII/400GMII data octets to (from) 66-bit blocks (64B/66B).

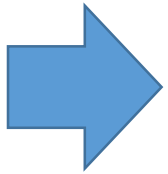
  b) Transcoding from 66-bit blocks to (from) 257-bit blocks.

  c) Reed-Solomon encoding (decoding) the 257-bit blocks.

  d) Transferring encoded data to (from) the PMA.

  e) Compensation for any rate differences caused by the insertion or deletion of alignment markers or due to any rate difference between the 200GMII/400GMII and PMA through the insertion or deletion of idle control characters.

  f) Determining when a functional link has been established and informing the management entity via the MDIO when the PHY is ready for use.



- **IEEE 802.3df PCS services:**

  a) Encoding (decoding) of eight 200GMII/400GMII data octets to (from) 66-bit blocks (64B/66B).

  b) Compensation for any rate differences caused by the insertion or deletion of alignment markers or due to any rate difference between the 200GMII/400GMII and PMA through the insertion or deletion of idle control characters.

  c) Determining when a functional link has been established and informing the management entity via the MDIO when the PHY is ready for use.

- **IEEE 802.3df FEC sublayer services:**

  a) Transcoding from 66-bit blocks to (from) 257-bit blocks.

  b) Reed-Solomon encoding (decoding) the 257-bit blocks.

  c) Transferring encoded data to (from) the PMA.

Functional block diagram

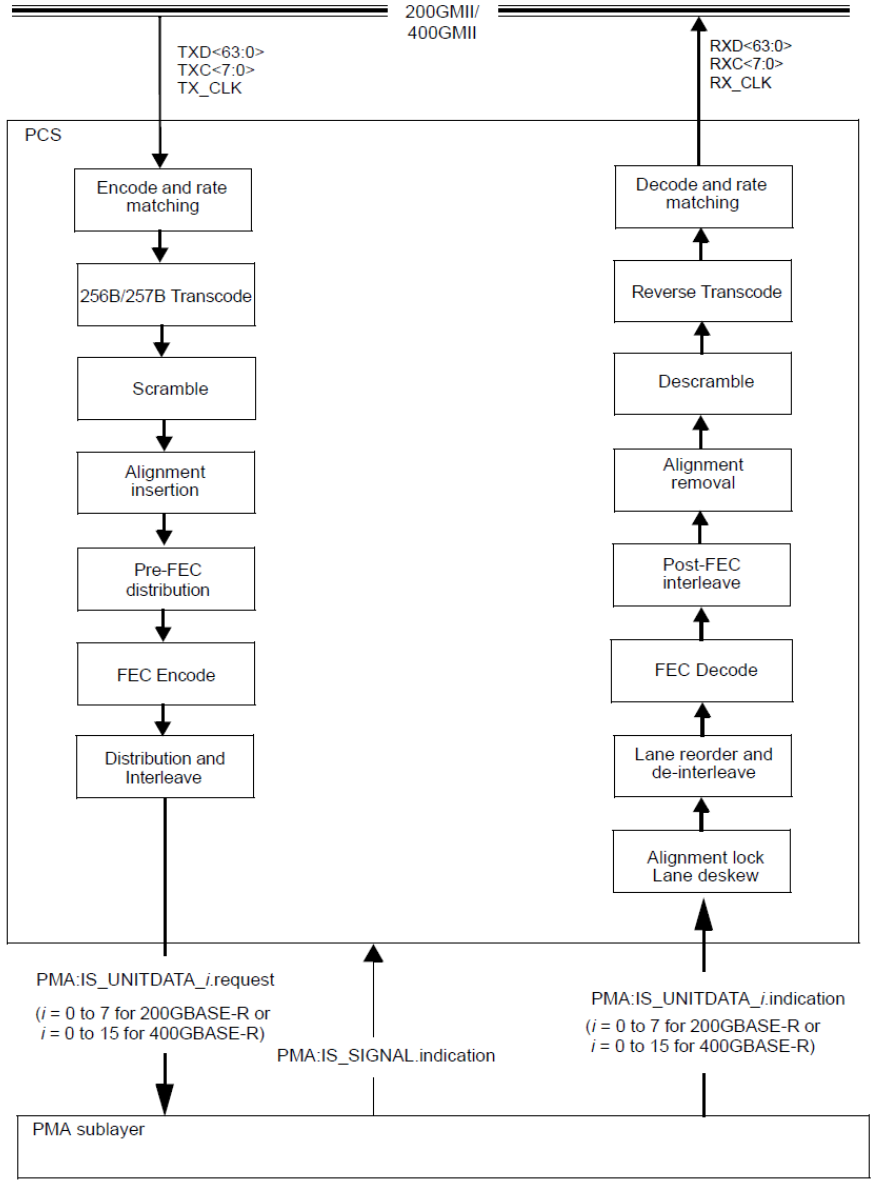
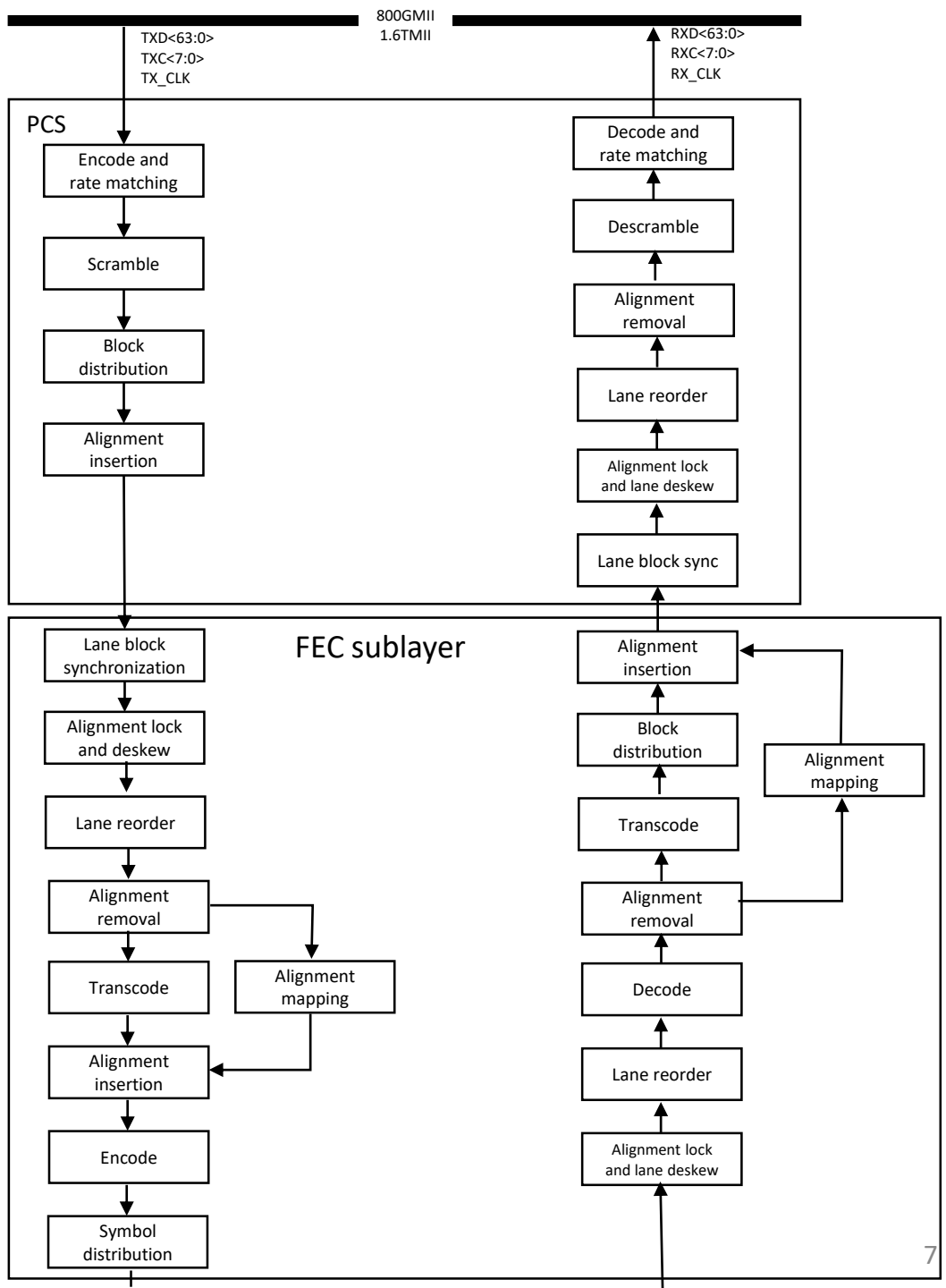
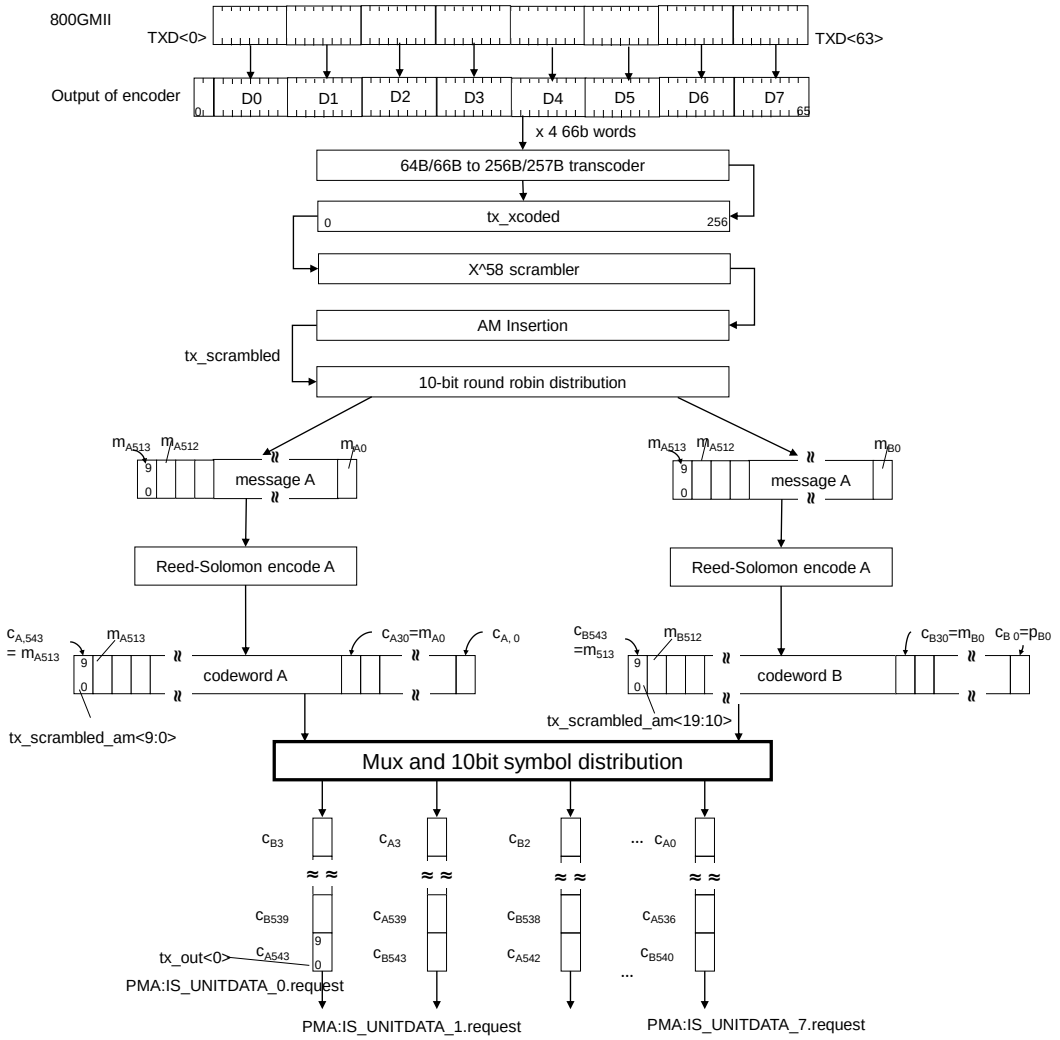
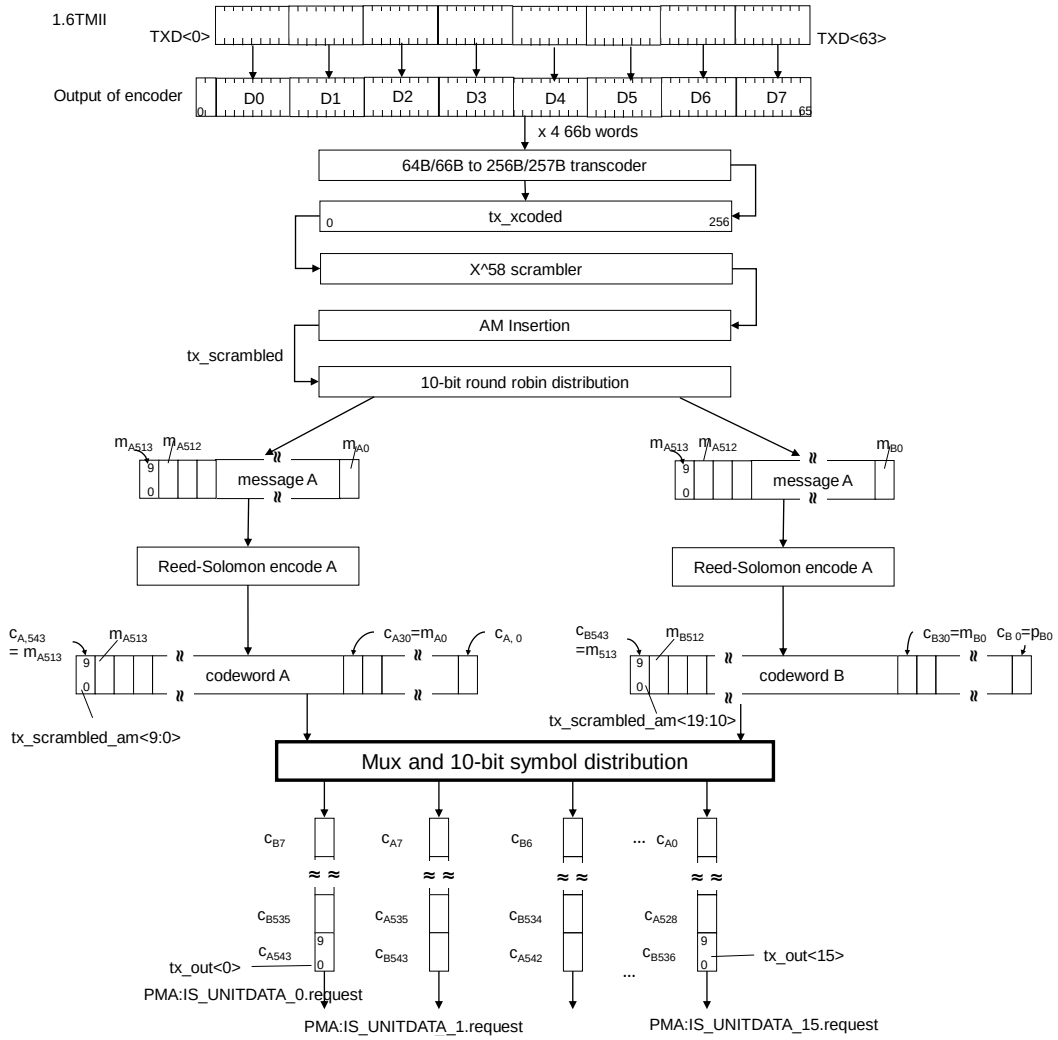


Figure 119–2—Functional block diagram



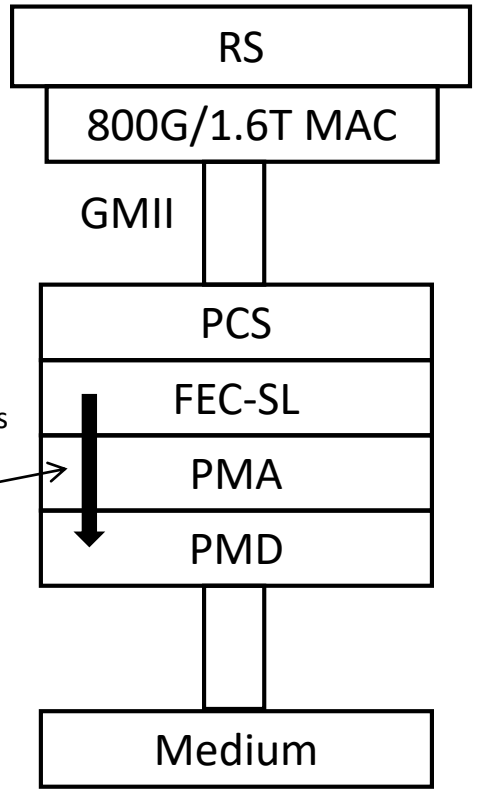
800GBASE-R and 1.6TBASE-R transmit bit distribution





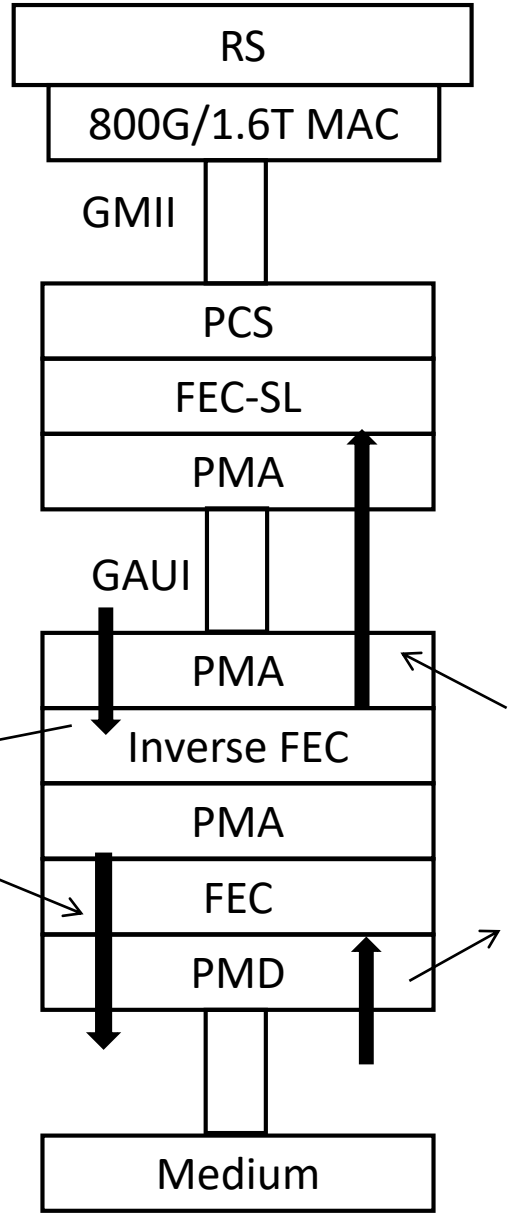
- Reuse 200GBase-R and 400GBase-R bit distribution defined in Clause 119.
- 8 FEC lanes for 800G and 16 FEC lanes for 1.6T.

FEC Degrade handling with new FEC sublayer



When the FEC sublayer detects FEC degrade or receives local degrade, it signals remote degrade towards the PMD

Remote Degrade signaling without Extender Sublayer

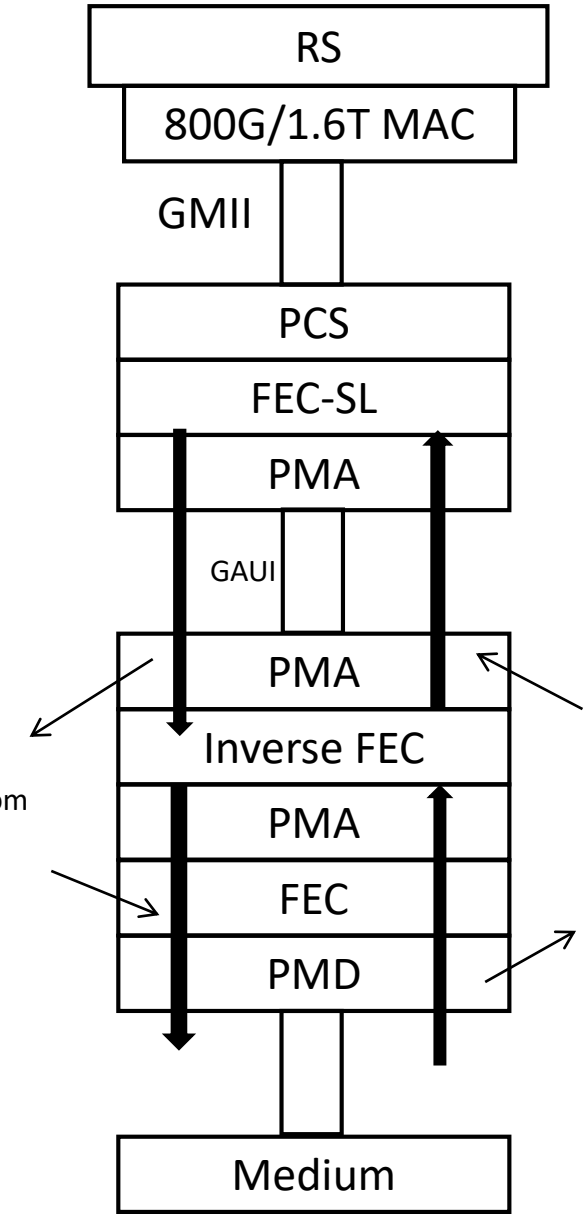


When the FEC sublayer detects FEC degrade, the signal is propagated to the adjacent FEC, which propagates that signal as local degrade towards the PMD

When the FEC layer detects FEC degrade or receives a local degrade indication from the far end of the link, the Inverse FEC propagates this as local degrade towards the FEC sublayer

Local Degrade Signaling with Extender Sublayer

FEC Degrade handling with new FEC sublayer - continued



When the FEC sublayer detects FEC degrade or receives local degrade from the FEC, remote degrade is signaled toward the FEC which propagates it toward the PMD.

When remote degrade or local degrade is detected by the FEC, the Inverse FEC propagates the signaling towards the FEC sublayer.

Remote Degrade Signaling with Extender Sublayer

Summary

- As fast path to 800G early adopters, 100G/lane based 800GbE and 1.6TbE should be simple:
  - In general, reuse the 200G/400G architecture defined in 802.3bs with increased lane rate (106.25Gb/s).
  - 8 FEC lanes for 800Gb/s and 16 FEC lanes for 1.6TGb/s.
  - Investigate burst error performance
- Start ASAP the work on 100G/lane based 800GbE and 1.6TbE baseline.
- Consider 100G/lane and 200G/lane communality of PCS

Thanks!