OTN Support for 25GbE
Support of Comments 136, 137, 138, 139, 190

Steve Trowbridge

Alcatel-Lucent

Supporters

- Pete Anslow (Ciena)
- Jonathan King (Finisar)
- Ralf-Peter Braun (Deutsche Telekom)
- William Szeto (Xtera)
- Ghani Abbas (Ericsson)
- Noburu Yoshikane (KDDI)
- John McDonough (NEC America)
- Steve Gorshe (PMC Sierra)
- Martin Carroll (Verizon)

OTN Support for 25GbE – 1/2

Reference trowbridge_3by_01_0115.pdf

- Slide 4 - If there are FEC and non-FEC 25G PMDs, the OTN-mapped format should be the same. FEC correction, Trans(-de)-coding, and FEC generation in the OTN mapper/demapper is OK and expected. Idle insertion/deletion shouldn't be required in the OTN mapper/demapper to convert between the FEC and non-FEC formats
- Slide 9 - Fits into 20 TS of OPU4, 21 TS of OPU3, or five 5G TS of an OTN B100G signal as currently envisioned – OK
- Slide 10 - Reuse modules for equivalent rate OTN signals - NA

OTN Support for 25GbE – 2/2

Reference trowbridge_3by_01_0115.pdf

- Slide 12 - Ideal case is if all 25GbE PMDs have a marker or none of them do. A case of including a marker in all PMDs, and FEC PMDs use 256B/257B transcoding with RS(528,514) FEC would allow easy interconnection of a FEC and a non-FEC PMD across OTN with the same physical layer clock at the OTN ingress and egress, suitable for Sync-E
- Slide 12 - If some PMDs have a marker and others don't, then providing the same payload bit-rate for all PMDs would require either under-clocking non-marker PMDs by a factor of 1-1/16K or over-clocking marker PMDs by a factor of 1+1/16K to allow interconnection of FEC and non-FEC PMDs across an OTN (note that the 1/16K factor has evolved to 1/20K since the January presentation)

ITU-T Q11/15 Discussion

- Current working assumption is that the OTN mapping will use the PCS format (64B/66B without CWM) as the mapping format for all 25GbE PMDs. RS-FEC may be present on the ingress and/or egress link, and idles compensate the CWM removed from the data stream across the OTN link
- Consequence is that the mapping is Frame, Preamble, and Timing transparent, but NOT PCS codeword transparent (the market expectation developed during P802.3ba after evolving from the P802.3ae expectation of “bit transparency for 10G)

Is PCS codeword transparency possible?

- In some cases, but unattractive. Either split the 25GbE universe into two non-interconnectable subsets (RS FEC mode and non-RS FEC mode operating PMDs), or provide 3 different mappings:
  - Provide a PCS codeword transparent mapping for RS-FEC to RS-FEC interfaces that replaces the CWM with a four 66B block version of the CWM in place rather than inserting idles
  - Provide a PCS codeword transparent mapping for interconnecting two non RS-FEC interfaces that simply maps as 64B/66B
  - Provide a non-PCS codeword transparent mapping for interconnecting an RS FEC interface to a non-RS FEC interface

What are the consequences of not having PCS codeword transparency?

- Certain applications like FlexE can't be supported over 25GbE since idle insertion/deletion is done below the PCS. Even an implementation that integrates the PCS/FEC into the same device couldn't work with a FlexE unaware transport network
- Most agree that PTP is tolerant to the idle redistribution that would occur with the single mapping, but some envision some future precision time application that might not be able to tolerate the Packet Delay Variation (PDV)

Market Perception of non PCS-codeword transparent mapping

- On the one hand, the properties not preserved by the single mapping of 25GbE over OTN are properties that a 25GbE signal is not guaranteed to have;
- On the other hand, this is viewed by some as limiting future applications
- Why do people perceive this as a case of the OTN mapping being broken rather than the fundamental nature of 25GbE?

Options for providing PCS codeword transparency

- Option 1: Remove CWMs from all 25GbE PHYs
  - Proposed and rejected from previous presentations slavick_3by_01a_0315.pdf and slavick_3by_01a_0515.pdf
- Option 2: Put CWMs in every 25GbE PHY
  - More detail on slides 11-13
- Option 3: Don't do idle deletion in the RS FEC block, but overclock to insert CWMs at the same PCS block rate
  - Awkward for SyncE, but not impossible as there is a fixed ratio
  - Not the preferred option as it could require either a PLL or a large divider if the RS FEC is implemented in a different device from the PCS

History of Option 1:

- Berlin Motion #3: Move to adopt slide 8 of slavick_3by_01a_0315.pdf to remove CWMs from 25GE RS-FEC data streams.
  - M: Jeff Slavick
  - S: Andre Szczepanek
  - Technical (>= 75%),
  - Y: 31 , N: 11 , A: 35
  - Result: Fails
- Pittsburgh Motion #3: Move to adopt slides 7-12 of slavick_3by_01a_0515.pdf to remove CWMs from 25GE RS-FEC data streams.
  - M: Jeff Slavick
  - S: Adrian Butter
  - Technical (>= 75%)
  - Y: 8 , N: 36 , A: 27
  - Result: fails

Option 2 – Changes required to put CWMs in every PHY – 1/2

- Move clauses 108.5.2.2 and 108.5.3.6 (rate compensation) to clause 107 (the PCS). Remove the “Rate Compensation” blocks from Figure 108-2
- Move most of the contents of 108.5.2.4 “Codeword marker insertion” to the PCS. The codeword markers are inserted as a group of four 66B blocks, which are the full alignment markers from Table 82-2 and Table 82-3 for PCS lanes 0, 1, 2, and 3 with the BIP_3 field replaced by 0x33 and the BIP_7 field replaced by 0xCC. These four blocks are inserted once per 81920 66B blocks in normal operation, or once per 80 66B blocks when coming up out of deep sleep operation and sending rapid CWMs. Move 108.5.3.4 “Codeword marker removal” to the PCS, which is removing four 66B blocks.

Option 2 – Changes required to put CWMs in every PHY – 2/2

- Add a sub-clause to 108.5.2.3 about how to transcode a CWM, which is done by removing the sync headers from the four 66B blocks and adding a trailing 0 in the 257th bit position.
- Add a sub-clause to 108.5.3.5 about how to transcode a CWM to 66B, which involves inserting the sync headers and deleting the trailing 0 in the 257th bit position.
- Add a sentence to 108.5.2.5 (similar to clause 91) to indicate that the transcoded CWM is mapped into the first 257 bits of each 1024th codeword in normal operation, or in the first 257 bits of every codeword when sending rapid codeword markers to emerge from deep sleep.

Proposal

- Implement Option 2 (CWMs in every 25GbE PHY) as the most straightforward way to allow for PCS codeword transparency of 25GbE transport over OTN with a single mapping for every 25GbE PHY

THANKS!