

OTN Support Proposal

P802.3bs 400 Gb/s Ethernet Task Force

Steve Trowbridge
Alcatel-Lucent

Supporters

- Pete Anslow (Ciena)
- Gary Nicholl (Cisco)
- Tongtong Wang (Huawei)
- Xinyan Wang (Huawei)
- David Ofelt (Juniper)
- Andre Szczepanek (Inphi)

Key Elements of OTN Support

- See “[OTN Support: What is it and why is it important?](#)”, July 2013
 - A new rate of Ethernet (e.g., 400 Gb/s) fits into the corresponding rate OTN transport signal
 - All Ethernet PHYs of a given rate are mapped the same way and can be interconnected over the OTN (e.g., same PCS for all 100 Gb/s PHYs gives a single canonical format (“characteristic information” in ITU-T terminology) that can be mapped
 - Optical modules for Ethernet can be reused for OTN IrDI/client interfaces at the corresponding rate

A new rate of Ethernet (e.g., 400 Gb/s) fits into the corresponding rate OTN transport signal

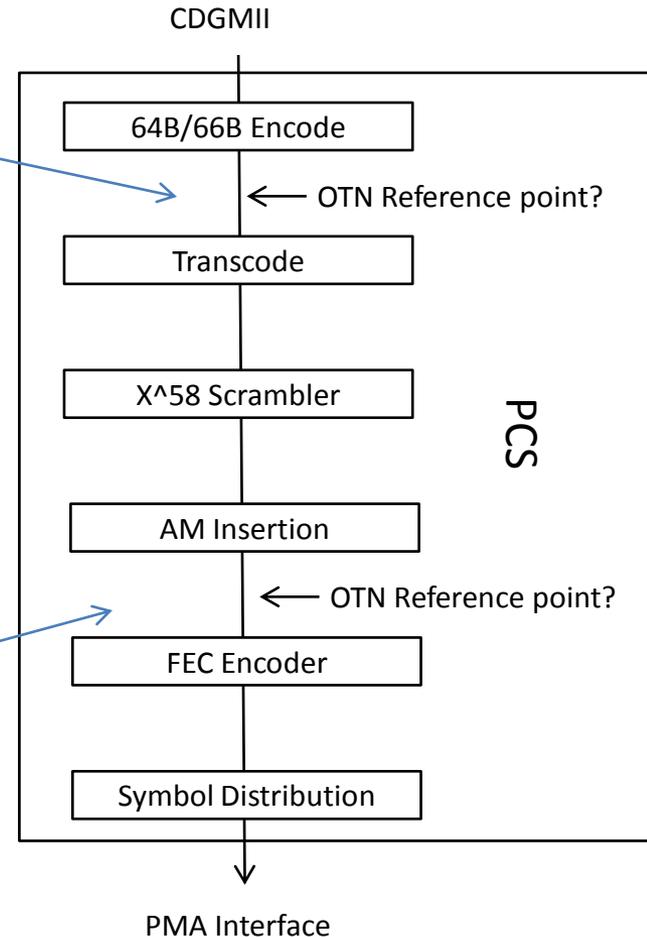
- Assumption – the OTN mapper/demapper will terminate and regenerate any Ethernet FEC code, correcting errors at the OTN ingress since the FEC is chosen to correct single-link errors but not double-link errors
- Assumption – the OTN mapper/demapper may trans-decode/trans-encode back to 64B/66B to avoid MTTFPA reduction for OTN transported signal
- Based on these assumptions, the encoded data rate of the OTN-mapped 400 Gb/s Ethernet would be no more than $400 \text{ Gb/s} \times 66 / 64 = 412.5 \text{ Gb/s} \pm 100\text{ppm}$. Since the 400 Gb/s OTN container would presumably be designed to also transport four “lower order” ODU4s, there should be no concern that it is large enough to carry 400 Gb/s Ethernet based on the assumption that the canonical form is near this rate.
- Any Ethernet bits in excess of this rate are likely to be part of a FEC that is not carried over OTN

All PMDs of a given rate are mapped in the same way into OTN:

Candidate Locations for the OTN Reference point

Option B

Option A



OTN Reference Point Option A

- Pro
 - Most similar to what ITU-T chose for mapping of 40GbE into OPU3 and 100GbE into OPU4 based on 802.3ba assumption that all PMDs of a given rate used exactly the same logical lane striping
 - Capable of carrying end-to-end information in the alignment markers if all PMDs are striped the same
- Con
 - Even if all P802.3bs PMDs use the same logical lane striping, if all future 400GbE PMDs are not striped in the same way, requires re-striping into the canonical format before mapping into OTN
 - Relies on OTN FEC for satisfactory MTTFPA

OTN Reference Point Option B

- Pro
 - High certainty to be a common format regardless of whether all 400GbE PMDs (P802.3bs and future) use the same logical lane striping
 - No need to convert between logical lane formats in OTN mapper/demapper
 - Robust MTTFPA regardless of what ITU-T does with FEC
- Con
 - Can't carry end-to-end information in alignment markers. But note that there doesn't seem to be a compelling reason to support BIP inside of FEC-enabled Ethernet PMDs
 - Higher bit-rate to map as 64B/66B rather than mapping the transcoded signal.

Straw Polls relevant to OTN support

- Straw Poll #1 - I support FEC for optical PMDs
 - **FEC mandatory – 69**
 - FEC optional – 7
 - Some PMDs may not need FEC – 0
 - Mandatory for some, optional for others – 10
 - Need more information – 10
- Straw Poll #9 - If all PMDs developed in P802.3bs include mandatory FEC and FEC error statistics are available, do we also require BIP?
 - Y: 4; **N: 24; A: 69**
- Straw Poll #10 – If BIP is required, should it be:
 - Segment by segment (optimized for fault isolation) – 2
 - End to end (optimized for service assurance) – 6
 - **Need more information – 35**
 - **Not required/don't care - 33**

Observations

- Strong majority believe that all optical PMDs developed by P802.3bs will have mandatory FEC
- Most think that if all optical PMDs have mandatory FEC, that BIP is not necessary, i.e., you get a better view of link quality from FEC corrected errors
 - Note that any BIP inside of a FEC exhibits a “cliff” behavior, going from zero errors to quite a lot the instant the error ratio exceeds the correction capability of the FEC. In addition, this information is available from the FEC uncorrected codewords counter
- For those who still think there would be BIP, the number who express an opinion on how it is used (fault isolation or service assurance) is statistically insignificant
- For those who still think they need more information, please study slides 5-12 of [trowbridge 3bs 01 0714.pdf](#) and ask questions!

OTN Reference Point Proposal

- Propose that the OTN reference point for 400GbE is a logically serial stream of 64B/66B blocks, unstriped, without lane alignment markers (Option B). An implication of this choice is that there is no BIP carried end-to-end (note that a segment BIP for a particular lane striping and FEC would still be possible).
- Any idle insertion/deletion to provide room for striping overhead must occur between the CDGMII and the OTN reference point. No idle insertion/deletion should occur between the OTN reference point and the PMD. Assuming 16K frequency of lane markers (regardless of the lane count or lane rate), the effective rate of this logical interface is the *nominal MAC rate* $\times 66/64 \times (1 - 1/16384)$ so that any physical instantiation has room to insert lane markers as needed without idle insert/delete elsewhere in the stack

Module Reuse

- Module reuse was facilitated by the fact that nothing below a CAUI chip-to-module interface cared about the or manipulated the bit values on the lanes – as long as OTN was striped into the same number of logical lanes as Ethernet, everything would work
- The following likely can be preserved: no idle insertion/deletion occurs below a CDAUI chip-to-module interface
- The following are possibly not be precluded by the 400GbE architecture:
 - Logical to physical lane multiplexing in a module may be on a block or FEC symbol basis rather than a bit basis
 - One (possibly Ethernet Frame Format dependent) FEC code may be replaced with another)

Options for Module Reuse

- Option I: Preserve the 802.3ba rule that no sublayers below a CDAUI care about bit values or manipulate the bit values on logical lanes (bit multiplexing only). Any FEC is done on the host board above a CDAUI. OTN may use a different FEC than Ethernet if it needs a stronger FEC to compensate for the higher bit-rate
- Option II (most general, described in Norfolk) encode the OTN frame as 66B blocks (all data) and use whatever striping and FEC encoding mechanisms are used for Ethernet. OTN and Ethernet use the same FEC
- Option III (potentially very complex) Allow OTN to use a different (stronger) FEC than Ethernet but do not require bit multiplexing of logical lanes. This would constrain that OTN and Ethernet choose FEC (or pairs of FECs, if not all 400GbE PMDs use the same FEC) with the same symbol size and that the marking to discover the FEC symbol alignment is common between OTN and Ethernet

Option II Amplification

- Use the fact that the OTN reference point, as proposed, is a logically serial stream of 64B/66B blocks.
- Note that before this reference stream can be physically instantiated, it must be striped over multiple physical or logical lanes
- Maintain the principle, as in 802.3ba, that idle insertion/deletion is not done below this reference point.

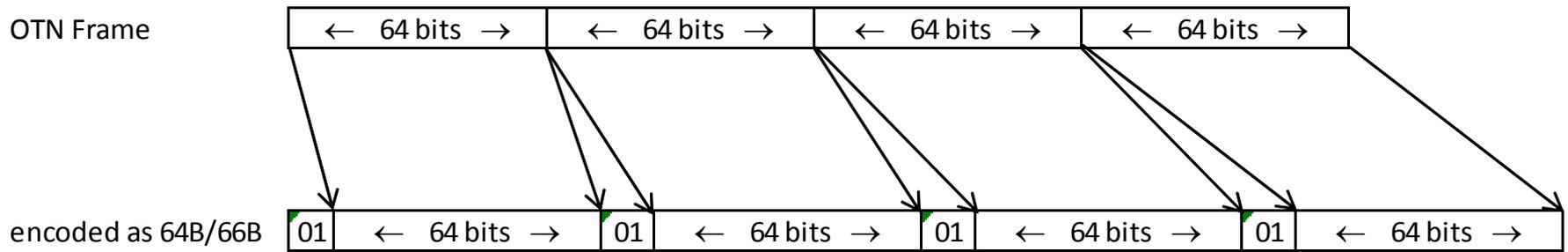
Option II Example

- Example physical instantiation could be something like [gustlin 400 02a 1113.pdf](#), produced by transcoding 64B/66B to 256B/257B, striping first into 100G groups, striping within each 100G group into 4 logical lanes on 10-bit symbol boundaries, inserting alignment markers on each lane, and applying an RS(528,514) code based on 10-bit symbols with alignment markers appearing in the first of each of 4096 Reed Solomon code blocks (essentially 4 instances of P802.3bj 100G FEC)

Option II Implications for OTN

- Likely only possible if the same FEC code can be used for OTN applications as for Ethernet applications at about 6% higher bit-rate
- Would need to make OTN look like 66B blocks. Easiest way to do this and not lose any information in transcoding is to insert a “01” sync header after every 64 bits (all data)
- Since this is just part of the logical frame format, this doesn't waste as many bits as it appears. 8 sync header bits are added to every 256 data bits in the “logical” frame format, but 7 of those bits are immediately recovered in 256B/257B transcoding and reused for the FEC code. So 0.39% net is added to the OTN frame to make it look like 66B blocks, then 2.724% overhead RS FEC added

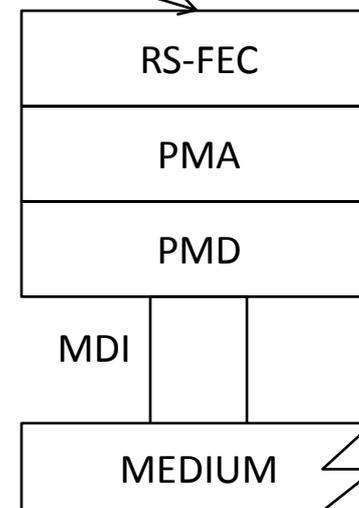
Option II - Illustration of turning OTN frame into 64B/66B blocks



Use the Ethernet Stack to stripe and FEC encode the OTN frame when carrying over an Ethernet Module for an OTN IrDI or client interface

Could be OTN frame aligned as an OTUC4 frame without FEC is exactly 7648×64 bits, but not essential with scrambling

Scramble



Option II: OTN Bit-rates using this scheme

	Working Assumption Bit-Rate
OTUC4 bit-rate without FEC	422.904 Gb/s
64B/66B encoded	436.120 Gb/s
256B/257B transcoded	424.556 Gb/s
Insert Lane Markers	424.582 Gb/s
Add RS(528,514) FEC	436.146 Gb/s
Logical Lane Rate (well within CEI-28G)	27.259 Gb/s
Ethernet Nominal Bit-rate	412.5 Gb/s
400G OTN Increase in bit-rate	5.73 %
100G OTN Increase in bit-rate	8.42 %

Smaller increase for 400G than for 100G, mainly due to RS(528,514) FEC rather than RS(255,239) FEC. Proportion remains the same even for Ethernet PMDs that use a higher overhead FEC

Option II – Recommended module reuse mechanism for OTN

- There is an Ethernet sublayer reference point such as the that is logically equivalent to a serial stream of 64B/66B blocks (the same as the recommended OTN reference point)
- No idle insertion/deletion occurs below the that reference point, and hence the rest of the stack can deal with a constant-bit-rate (CBR) bitstream that is effectively an infinite-length packet.
- Note that any logical to physical lane interleaving that works for Ethernet also works for OTN since they are encoded the same way
- The link parameters and FEC coding gain have sufficient margin to meet the error performance target when running at approximately 5.73% higher bit-rate than necessary for 400G Ethernet

THANKS!