oFEC-based 800GBASE-LR1/ER1 proposal with Synchronous Mapping

Gary Nicholl, Mike Sluyski, Jim Duda – Cisco

Ted Sprague – Infinera

P802.3dj joint optics/logic ad hoc meeting, 13 April 2023

Background

- williams_3dj_01a_2303.pdf provided a baseline proposal to address the 800GbE 10km and 40km objectives, based on a single lambda coherent solution using oFEC

- This proposal leveraged industry investments in adjacent applications (such as OIF 800ZR) and used asynchronous GMP mapping to map the 800GBASE-R signal into the oFEC frame.

- During discussions at the March meeting, it was pointed out that the async GMP mapping process results in additional delay variation that impacts PTP time sync accuracy and that cannot be easily compensated for by the techniques defined in Clause 90 and 802.3cx project.

- This contribution proposes using a synchronous mapping (bypassing GMP) to eliminate the above concern, while still maintaining commonality with the 800ZR solution

PTP Overview

- Common Terms:
  - PTP – Precision Timing Protocol. e.g., IEEE-1588v1 or v2.
  - PDV – Packet Delay Variation. **PDV is the difference in end-to-end one-way delay between selected packets in a flow (essentially "packet jitter")**
- Types of latency that impact PTP accuracy
  - **Static latency - design or measure fixed latency (known and easily compensated for, e.g. 802.3cx)**
  - **Quasi-static latency - static latency that varies on power cycle (typically unknown and difficult to compensate for)**
  - **Dynamic latency - variable latency (can be bounded, filtered or in some cases compensated for)**
- Quasi-static latency variations from power cycle to power cycle
  - This type of delay variation can be caused by FIFO initialization anywhere in the datapath (implementation specific)
  - This has caused issues with both coherent and IMDD products
  - Can be addressed/minimized by careful FIFO management/initialization
- Various functions will contribute to dynamic latency (PDV)
  - GMP mapping
  - Idle Insertion/Removal, AM Insertion/Removal, etc
  - Can (in some cases) be compensated for (e.g. TX/RX_NUM_BIT_CHANGE parameters added in 802.3cx)
  - Egress filtering can (in some cases) mitigate/minimize PDV
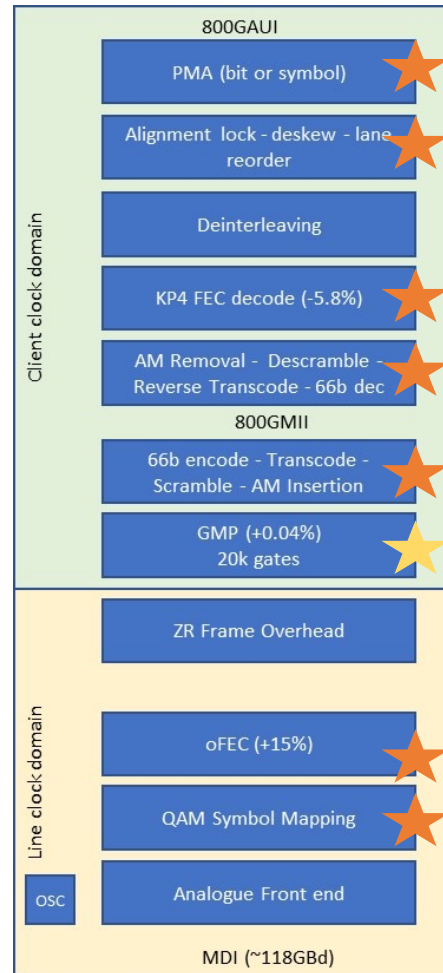
PTP and FIFOs (in previous proposals)

★ Possible synchronous FIFO location (potential to cause significant power cycle to power cycle variable static latency, if not managed/initialized correctly)
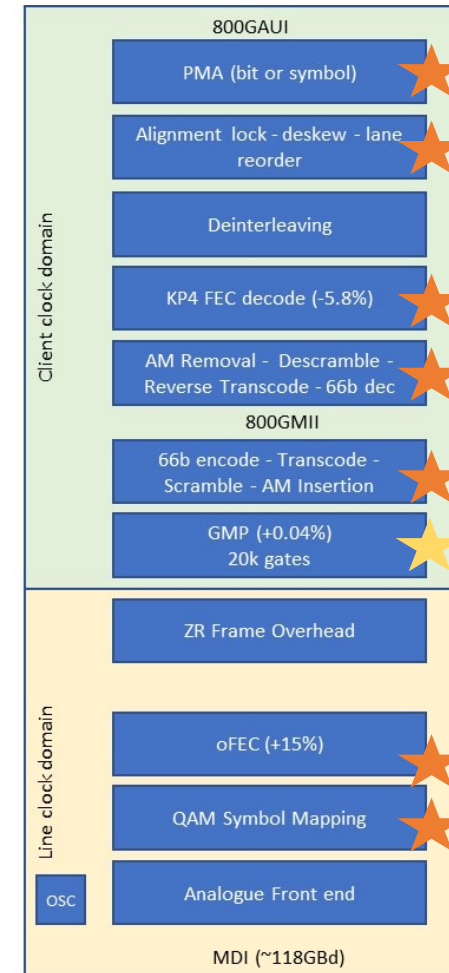
★ Asynchronous GMP mapping FIFO (generates bounded periodic varying latency, i.e. saw tooth)

- All solutions (whether synchronous or asynchronous mappings) are likely to contain multiple "synchronous" FIFOs
- If not managed/initialized correctly synchronous FIFOs can cause significant indeterminate power cycle to power cycle latency variation (that cannot be compensated for)
- Power cycle latency variation reported in some initial 400ZR designs was caused by synchronous FIFOs (not GMP)
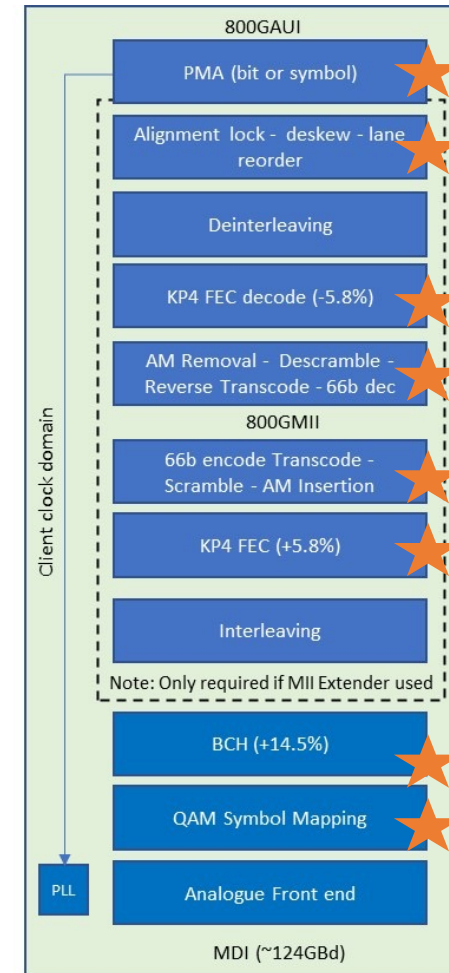

OIF 800ZR

800GAUI
- PMA (bit or symbol) ★
- Alignment lock - deskew - lane reorder ★
- Deinterleaving
- KP4 FEC decode (-5.8%) ★
- AM Removal - Descramble - Reverse Transcode - 66b dec ★

800GMII
- 66b encode - Transcode - Scramble - AM Insertion ★
- GMP (+0.04%) 20k gates ☆

- ZR Frame Overhead
- oFEC (+15%) ★
- QAM Symbol Mapping ★
- Analogue Front end

OSC

MDI (~118GBd)

800GBASE-LR1/ER1 (GMP)
(williams_3dj_01a_2303)

800GAUI
- PMA (bit or symbol) ★
- Alignment lock - deskew - lane reorder ★
- Deinterleaving
- KP4 FEC decode (-5.8%) ★
- AM Removal - Descramble - Reverse Transcode - 66b dec ★

800GMII
- 66b encode - Transcode - Scramble - AM Insertion ★
- GMP (+0.04%) 20k gates ☆

- ZR Frame Overhead
- oFEC (+15%) ★
- QAM Symbol Mapping ★
- Analogue Front end

OSC

MDI (~118GBd)

800GBASE-LR1/ER1
(maniloff_3dj_01a_2303)

800GAUI
- PMA (bit or symbol) ★
- Alignment lock - deskew - lane reorder ★
- Deinterleaving
- KP4 FEC decode (-5.8%) ★
- AM Removal - Descramble - Reverse Transcode - 66b dec ★

800GMII
- 66b encode Transcode - Scramble - AM Insertion ★
- KP4 FEC (+5.8%) ★
- Interleaving

Note: Only required if MII Extender used

- BCH (+14.5%) ★
- QAM Symbol Mapping ★
- Analogue Front end

PLL

MDI (~124GBd)

Client clock domain / Line clock domain

New proposal (synchronous mapping)

Old Proposal (GMP)
williams_3dj_01a_2303
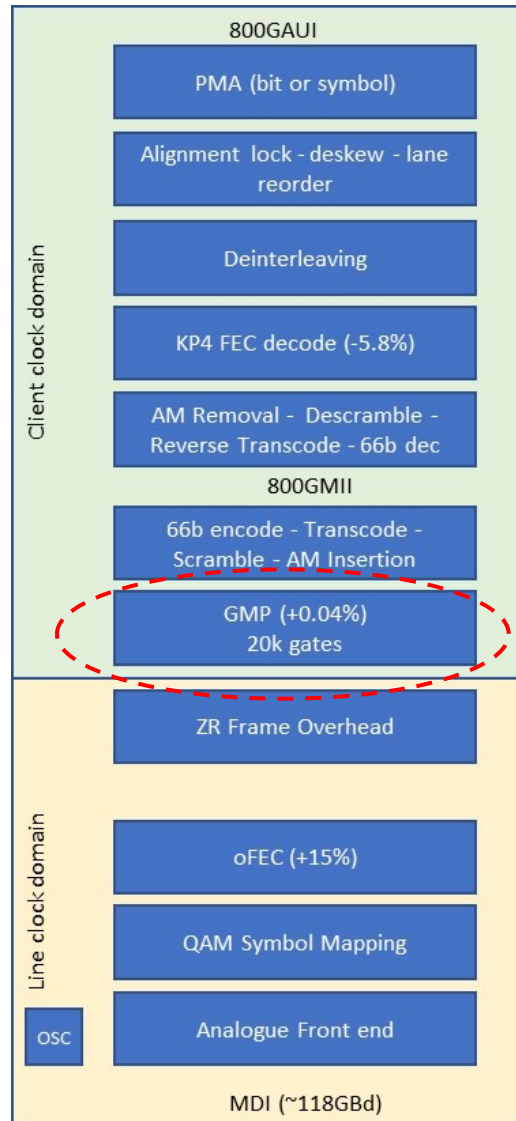
New Proposal (no GMP)

- Bypass GMP
- Synchronously map into ZR Frame
- A very minor change to the 800ZR design
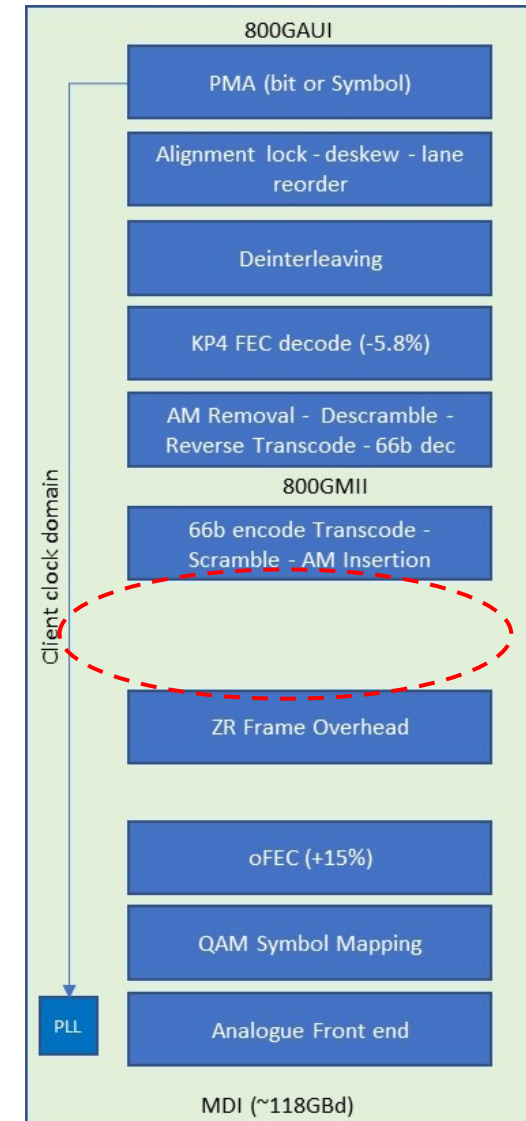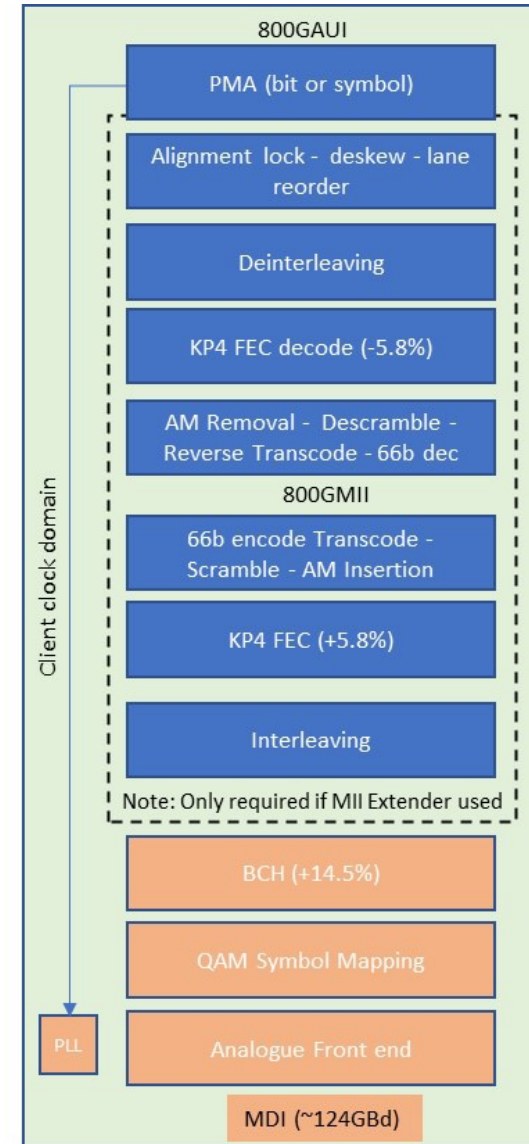
Commonality with 800ZR

Different from 800ZR

OIF 800ZR

800GAUI
- PMA (bit or symbol)
- Alignment lock - deskew - lane reorder
- Deinterleaving
- KP4 FEC decode (-5.8%)
- AM Removal - Descramble - Reverse Transcode - 66b dec

800GMII
- 66b encode - Transcode - Scramble - AM Insertion
- GMP (+0.04%) 20k gates

(Client clock domain)

- ZR Frame Overhead
- oFEC (+15%)
- QAM Symbol Mapping
- Analogue Front end

(Line clock domain)
OSC

MDI (~118GBd)
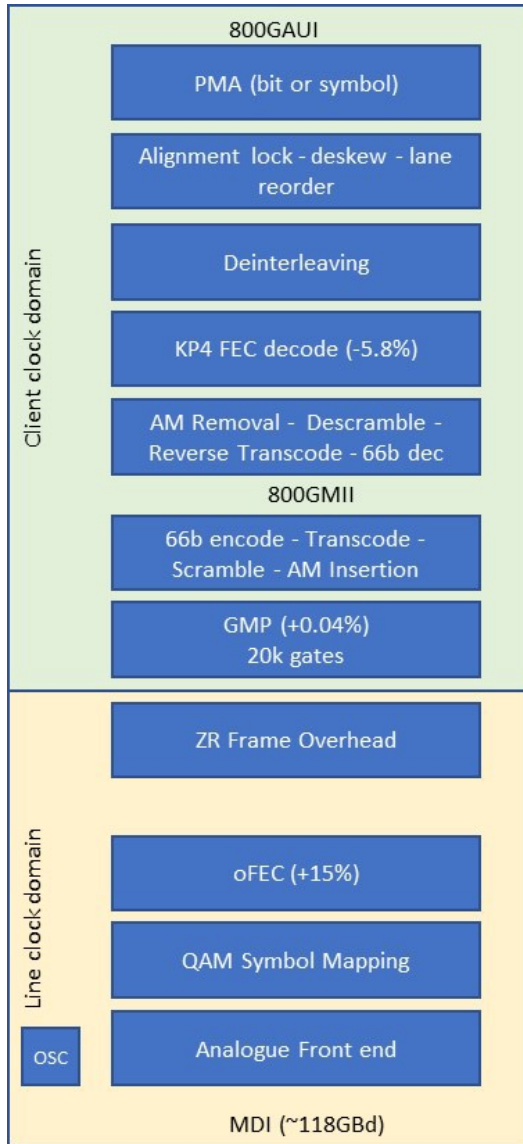
800GBASE-LR1/ER1 (no GMP)
Updated proposal

800GAUI
- PMA (bit or Symbol)
- Alignment lock - deskew - lane reorder
- Deinterleaving
- KP4 FEC decode (-5.8%)
- AM Removal - Descramble - Reverse Transcode - 66b dec

800GMII
- 66b encode Transcode - Scramble - AM Insertion

(Client clock domain)

- ZR Frame Overhead
- oFEC (+15%)
- QAM Symbol Mapping
- Analogue Front end

PLL

MDI (~118GBd)

800GBASE-LR1/ER1
(maniloff_3dj_01a_2303)

800GAUI
- PMA (bit or symbol)
- Alignment lock - deskew - lane reorder
- Deinterleaving
- KP4 FEC decode (-5.8%)
- AM Removal - Descramble - Reverse Transcode - 66b dec

800GMII
- 66b encode Transcode - Scramble - AM Insertion
- KP4 FEC (+5.8%)
- Interleaving

Note: Only required if MII Extender used

(Client clock domain)

- BCH (+14.5%)
- QAM Symbol Mapping
- Analogue Front end

PLL

MDI (~124GBd)

6

Summary

- This contribution proposes a very small tweak to williams_3dj_01a_2303.pdf, to replace the async GMP mapping with a synchronous mapping

- The proposal still leverages ongoing industry investment with a common oFEC solution and a common data rate of ~ 118Gbd, supporting LR, ER and ZR applications

- A full logic baseline proposal will be provided at the May interim meeting in San Antonio

Thanks