Synchronization over Optical Technologies

Steve Gorshe, Microchip,
Rapporteur ITU-T Q11/15, Chair IEEE P802.3cx

13 July 2024 Workshop, Montréal

Outline

- PTP over GMP
- Review of items addressed by IEEE P802.3cx
- Subsequent challenges emerging from the IEEE P802.3dj project

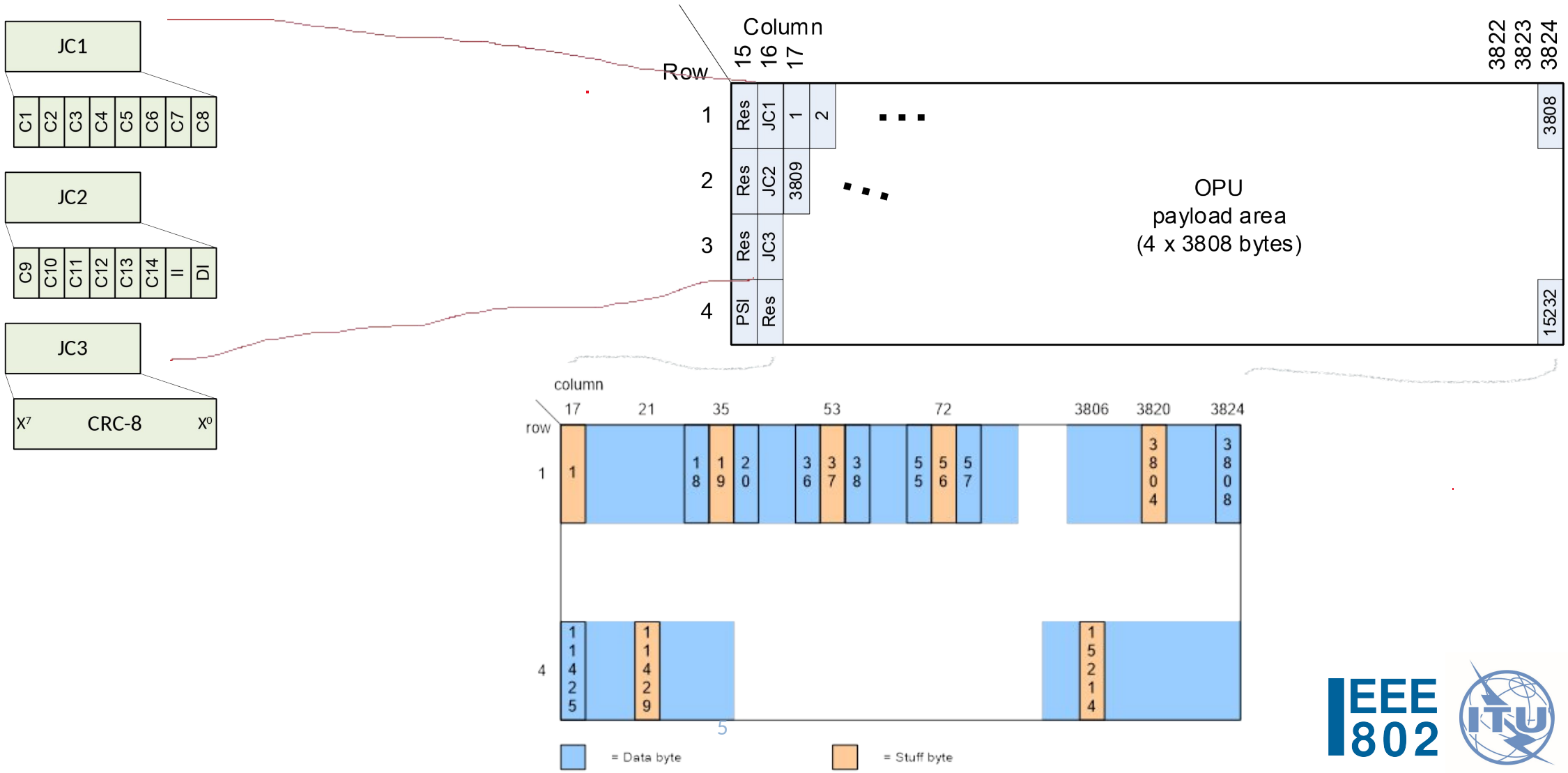
PTP over GMP

- Q11/15 developed the Generic Mapping Procedure (GMP) for mapping constant bit rate (CBR) clients into OTN channels
  - GMP accommodates the difference between the client signal and the OTN server payload rates
  - GMP was adopted for its flexibility and the excellent jitter performance it inherently enabled
- GMP was subsequently adopted into IEEE P802.3ct (100GBASE-ZR), and OIF 400ZR and 800ZR

PTP over GMP – GMP Concept

- GMP overhead is inserted into the server channel on a strictly periodic basis (e.g., once per frame period) to provide client bit rate information
- The GMP overhead conceptually consists of:
  - A count (C_m) of the integer number of client data words that the source will be able to transmit in the next period, including robust C_m error protection
  - An indication of phase offset (i.e., the C_{nD} count value representing the amount of client data less than a word size that the source will not be able to transmit in the next period)
- Client data words are inserted into the payload area using a modulo(C_m) algorithm
  - This results in "stuff" / padding words being evenly spaced
- The sink uses the same modulo algorithm to extract the client data

PTP over GMP – GMP Illustration for OTN



PTP over GMP – Timing Consideration

- All delay variation at the source due to stuff locations are mirrored at the sink
  - Stuff locations are predictable with optimally smooth spacing
    - In sharp contrast to the unpredictable locations of idle blocks when using idle insertion / removal for rate compensation
  - Results in a fixed end-to-end delay per period with easily filtered jitter due to periodic C_m changes
    - E.g., the combination of the C_{nD} phase, tracking C_m history and/or PLL LPF
- Consequently, GMP provides the best PTP performance with respect to recovering the client signal rate with low jitter

PTP over GMP – Buffer considerations

- The mapping/demapping algorithm with periodic stuff locations minimizes buffer size requirements
- Regarding system reset impact on changes to the GMP buffer depth
  - Framer devices inherently have multiple FIFOs in the data path, with GMP being a relatively small contributor to the cumulative FIFO size
  - Non-GMP approaches have the same issues
  - Consequently, this is not an inherent drawback for GMP

IEEE Std 802.3cx overview

IEEE P802.3cx (published in 2023) was formed to address aspects of complex Ethernet PHYs that can impact PTP accuracy related to the latency variability between the xMII and MDI

- Timestamp reference point discrepancy between 1588/802.1AS and 802.3
  - "Timestamp reference point" represents both/either the 1588/802.1AS message timestamp point and/or the 802.3 data delay measurement point (DDMP)
- Impact of indeterminate Idle insertion/removal locations, esp. due to AM insertion/removal
- Delay variations through the PHY (such as from PCS lane distribution/merging)
- Transmit skew

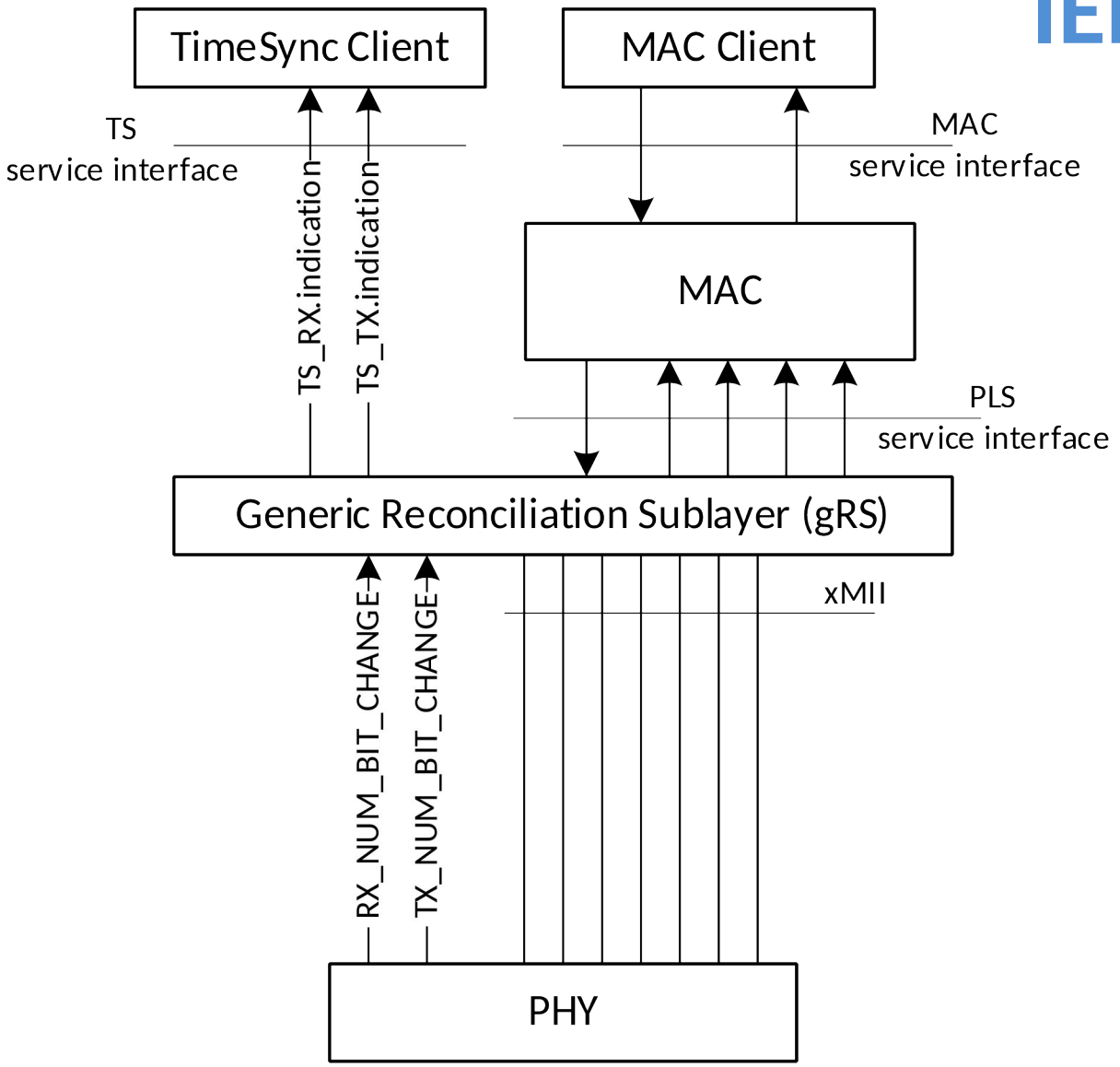
IEEE Std 802.3cx overview (continued)

- Discrepancy between the 1588/802.1AS (beginning of the first symbol after the SFD) and legacy 802.3 (beginning of the SFD) timestamp reference points
  - SFD = Start of Frame Delimiter
  - Creates a timing error of one octet time if the two endpoints use different timestamp reference points
    - AM (lane Alignment Marker) insertion between the two timestamp reference points and multi-lane functions can increase the timing error beyond one octet time
  - Resolution was to add optional support for an 802.3 DDMP that matches the 1588/802.1AS message timestamp point
    - Use of this new mode is recommended by 802.3

IEEE Std 802.3cx overview (continued)

- AM and Idle insertion/removal
  - Idle blocks are removed at the TX in order to create the bandwidth (space) for inserting AMs and idles are inserted at the RX after AM removal.
  - Since there are no rules to specify the relative locations of the removed and inserted Idles, significant timing error can be introduced, especially when the AM location is near the timestamp reference point.
  - The solution was to provide information across the TimeSync service interface (TSSI) about the dynamic path data delay (DPDD) variation between the xMII and the MDI
    - TX_NUM_UNIT_CHANGE and RX_NUM_UNIT_CHANGE values added to TS_TX.indication and TS_RX.indication primitives

IEEE Std 802.3cx overview (continued)



- Note that TX_NUM_UNIT_CHANGE and RX_NUM_UNIT_CHANGE are not physical signals on the xMII. They are information communication paths typically within the same device.

IEEE Std 802.3cx overview (continued)

- PCS lane distribution/merging
  - This is an example of a scenario where the different lane distribution/merging alignments at the TX and RX will impact the delay at each, but mirror each other such that the end-to-end delay is constant.
  - Addressed by allocating constant delay values to TX and RX (similar to how 802.3bf addressed FEC encode/decode delay variations)
    - Maximum delay of the lane distribution allocated to TX
    - Minimum delay of the lane merging allocated to RX
- Generally:
  - The above principle (max variable delay allocated to Tx, min to Rx) works for any PHY function with mirrored variable delays.

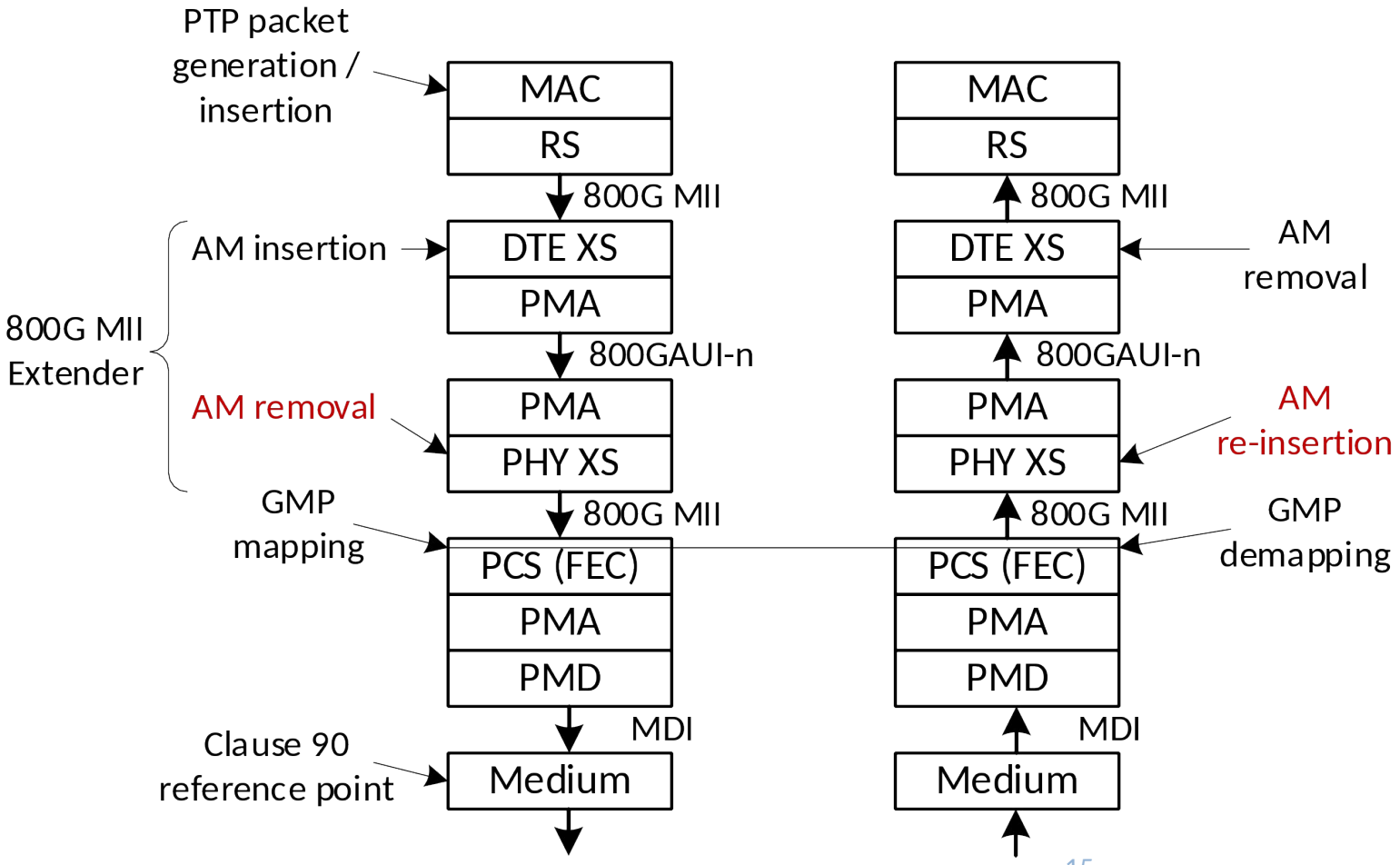
IEEE Std 802.3cx overview (continued)

- Transmit skew
  - For example, due to different serializer and deserializer implementations
  - Very difficult to isolate since is entwined with the skew of the medium
  - Resolution is the recommendation to design to minimize it

IEEE 802.3 timestamp accuracy issues subsequent to P802.3cx

- A challenge identified during P802.3dj is that some PHYs an additional removal and re-insertion idles and AMs in each direction. For example:
  - PHYs that include an extender sublayer (XS)
  - PHYs for coherent optical interface
- The problem is that the location from which the AMs are removed is independent of where they are re-inserted.
  - Consequently, the 802.3cx approach of dynamic delay indication will not be adequate, especially if the removal and insertion are performed on different devices
  - The note in 90.7.2 of 802.3cx hints that the solution for this scenario is to ensure that the AM position at the input to the Tx (PHY_XS + PCS) is the same as the AM position at the output of the Rx (PCS + PHY_XS).
    - The next slide shows the method to accomplish this that was adopted by P802.3dj for 800GBASE-ER1
- These issues are under study in IEEE P802.3dj and may lead to a new Task Force and/or maintenance activity

IEEE 802.3 timestamp accuracy issues subsequent to P802.3cx – solution adopted for 800GBASE-ER1



- GMP frame is based on the AM cycle
- The original AM locations from the TX side are encoded into GMP overhead bytes so that the RX side can restore them to the same locations


IEEE
802


ITU