Timestamping Models and Activities in P802.3cx

Richard Tse (Microchip Technology)
Richard.Tse@microchip.com

IEEE 802.1 TSN TG
IEC/IEEE 60802 meeting on 2021-04-09

Outline of the presentation

- Introduction to P802.3cx
- Timestamp Generation Models
  - IEEE 1588
  - IEEE 802.3
- High Accuracy Issues with IEEE 802.3
  - Mismatched Message Timestamp Point
  - Multi-PCS Lane Delay Variation
  - AM, CWM, Idle Insertion/Deletion
  - Tx Skew
- IEEE 802.3 Timestamp Error Performance
- Solutions Adopted by P802.3cx
- Summary

Introduction to P802.3cx

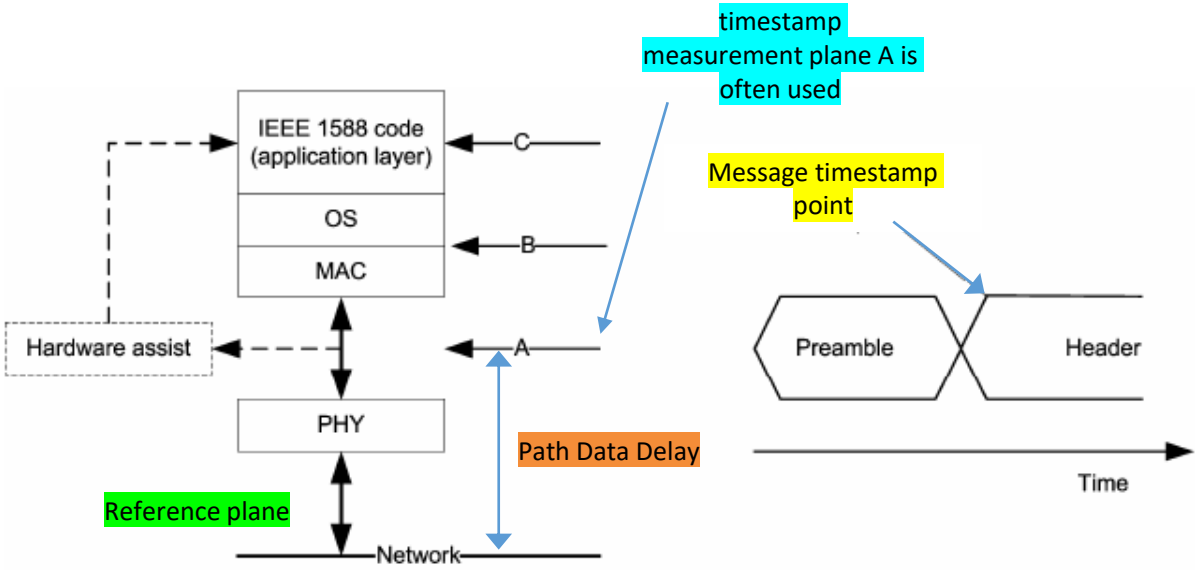
P802.3cx - Improving PTP Timestamping Accuracy Task Force

- Will address identified shortcomings in the IEEE 802.3 specification that could affect timestamping accuracy
- Targeted completion of technical aspects: Q3-2021
- Targeted publication date: mid-2022
- Website: https://www.ieee802.org/3/cx/index.html
- Contributions: https://www.ieee802.org/3/cx/public/index.html

The views expressed in this presentation are those of the author and not of the P802.3cx task force.

Timestamp Generation Model: IEEE 1588

- A timestamp is generated at the time the "message timestamp point" crosses "reference plane", which is the intersection between the network (i.e. the medium) and the PHY
- Timestamp capture is implemented at the "timestamp measurement plane", which, in practice, occurs at point A and must be moved back to the reference plane
- *Good estimate of the PHY delay* ("path data delay", the time between the reference plane and the timestamp measurement plane) *is needed → varying delays should be compensated for*
- *Every endpoint needs to have the same understanding of the above concepts and how compensation is done*



Timestamp Generation Model: IEEE 802.3

- Clause 90 of IEEE 802.3-2018 assumes PTP timestamp is generated at a layer above PHY, and provides a function for the PHY layer to report its TX and RX delay, but with static maximum and minimum values (not a latency value per packet)
- Therefore, after latency compensation by PTP timestamp function, some error might still be present (e.g., the difference of maximum and minimum latency)

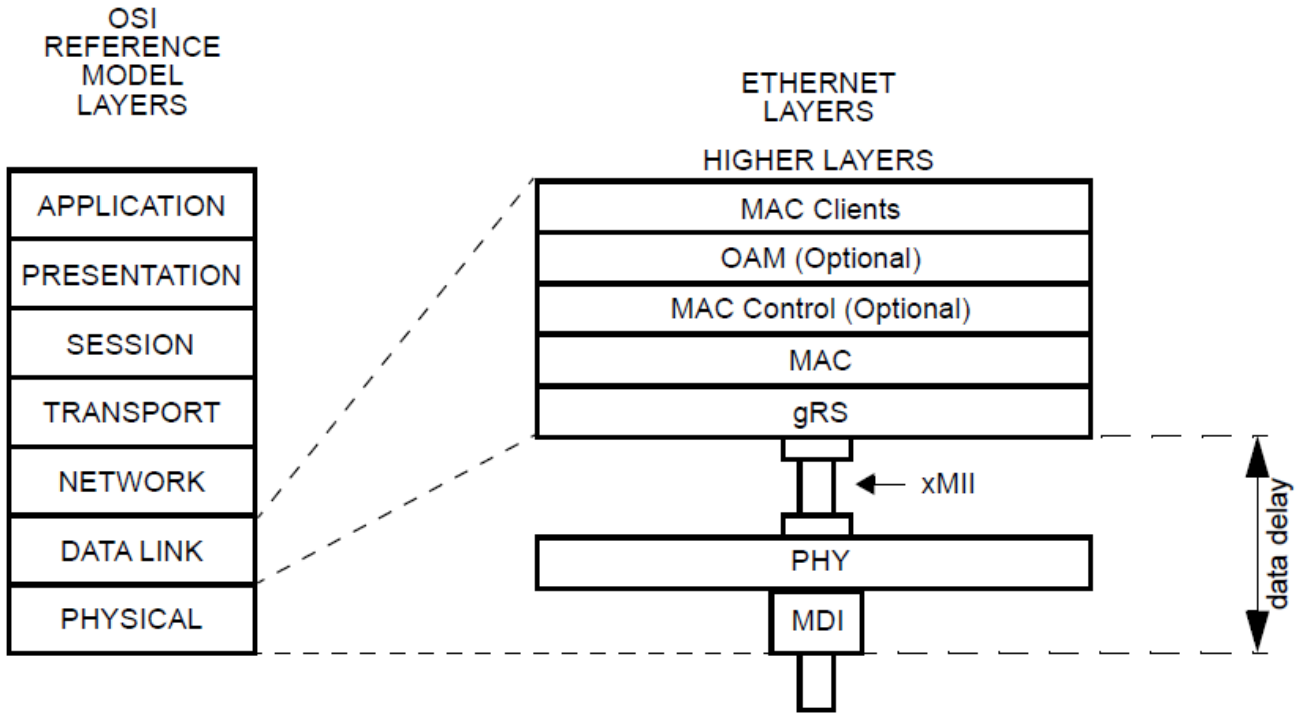


Figure 90–3—Data delay measurement

Timestamp Generation Model: IEEE 802.3

- IEEE 802.3 Clause 90 provides support for a TimeSync Client
  - The optional Time Synchronization Service Interface (TSSI) supports protocols that require knowledge of packet egress and ingress time
  - Timestamping is done in the gRS, where the timestamp is captured when the message timestamp point crosses the xMII

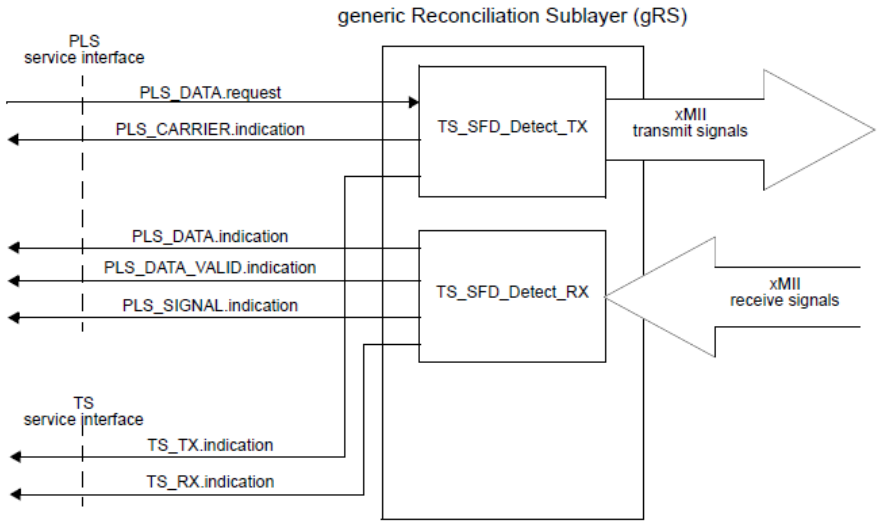
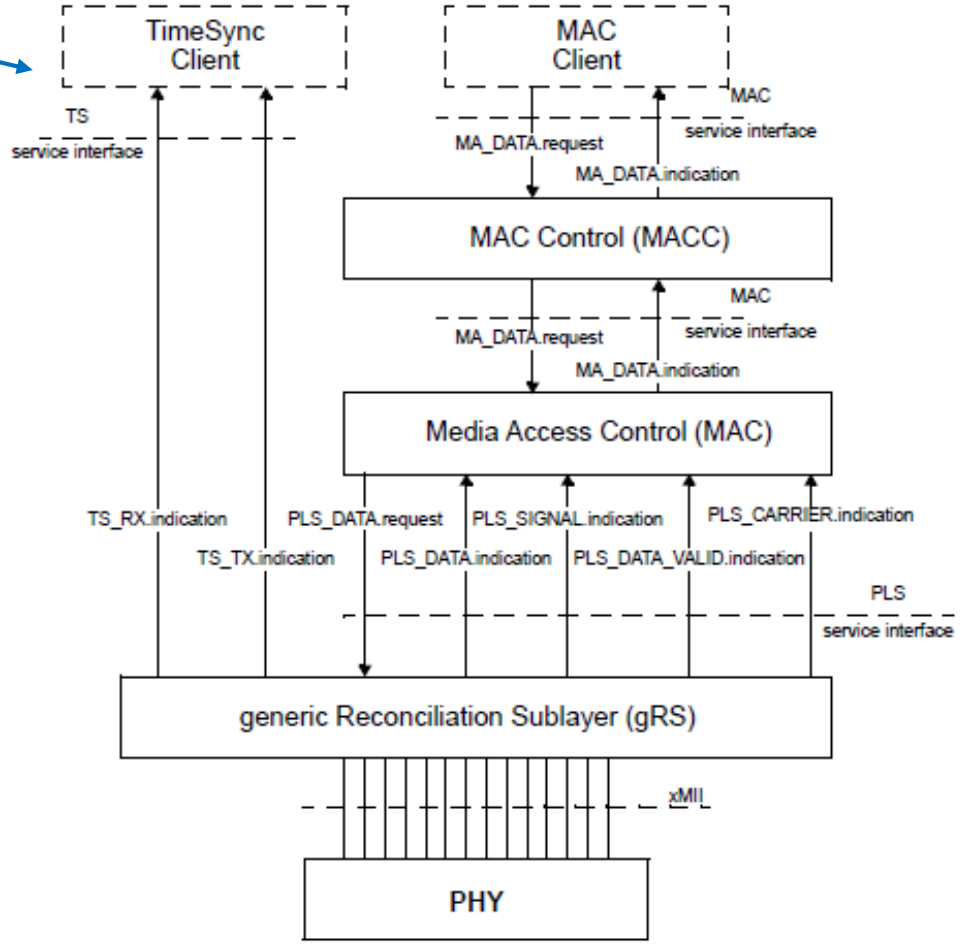


Figure 90–2—TS_SFD_Detect_TX and TS_SFD_Detect_RX functions within the generic Reconciliation Sublayer (gRS)



High Accuracy Issues with IEEE 802.3-2018

- PTP timestamping is done at the MDI
- IEEE 802.3's timestamp is captured at the xMII
- PHY data delay must be known to move the timestamp from xMII (measurement plane) to MDI (reference plane)
- Many newer 802.3 PHYs have fundamental dynamic variations in their data delay

- But
  - Data delay variations in the PHY are not inherently visible at the xMII

- Thus
  - IEEE 802.3's current timestamping mechanism does not inherently support high accuracy on PHYs with data delay variations
  - Specifications are needed on how to deal with each data delay variation → **done by P802.3cx**

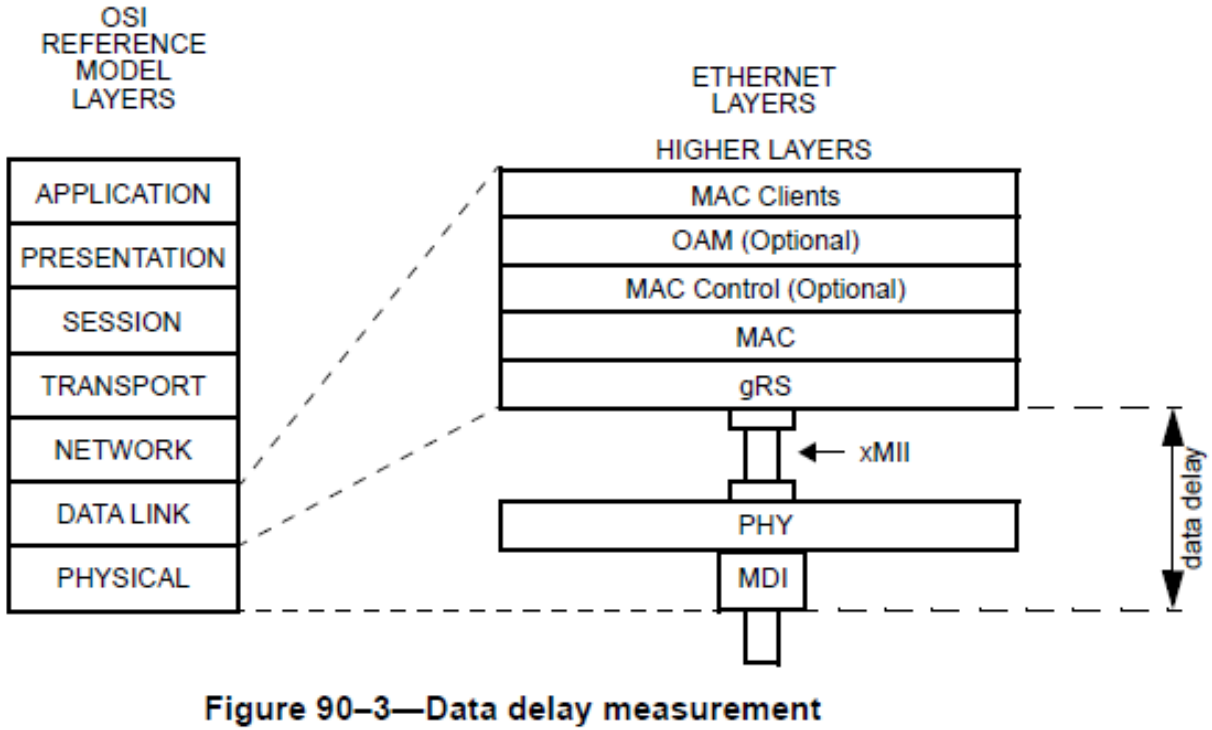


Figure 90–3—Data delay measurement

Issue #1: Mismatched Message Timestamp Point

Subclause 90.7 of IEEE 802.3 states

- “The transmit path data delay is measured from the input of the beginning of the SFD at the xMII to its presentation by the PHY to the MDI. The receive path data delay is measured from the input of the beginning of the SFD at the MDI to its presentation by the PHY to the xMII.”

however…

Subclause 7.3.4.1 of IEEE 1588v2 and subclause 11.3.9 of IEEE 802.1AS define the message timestamp point as follow:

- “the message timestamp point for an event message shall be the beginning of the first symbol after the Start of Frame (SOF) delimiter”
- “the message timestamp point for a PTP event message shall be the beginning of the first symbol following the start of frame delimiter”

PTP link delay measurements will be incorrect if different message timestamp points are used by the PTP endpoints

Issue #2: Multi-PCS Lane Delay Variation

- Many Ethernet interfaces have multiple PCS lanes
- The delay experienced on each lane might be different
  - E.g., the block going to PCS lane 0 must wait for the block going to PCS lane n-1 to catch up so they can all be transmitted at the same time

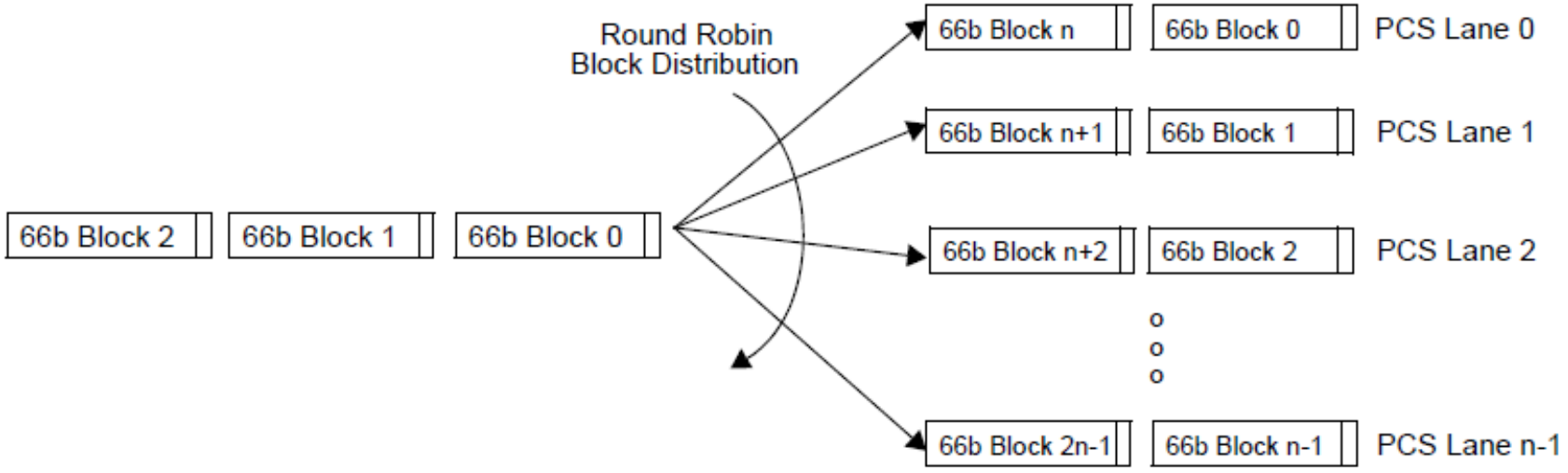


Figure 82–6—PCS Block distribution

Issue #3: AM, CWM, Idle Insertion/Deletion

- Alignment Markers (AM) or Codeword Markers (CWM) are used to align multi-lanes and to help Forward Error Correction (FEC) coding
  - AM/CWM blocks are inserted periodically into a Tx Ethernet stream and Idles are deleted to compensate for the bandwidth used by the AM/CWM
- AMs, CWMs, and Idles may be inserted or deleted by a PHY and affects the instantaneous path data delay
  - Insert/delete of AM/CWM or Idle momentarily increases/decreases the path data delay by T_{AM} or T_{Idle}, respectively
  - Idle insert/delete operate independently at Rx and Tx so delay changes do not have deterministic relationship

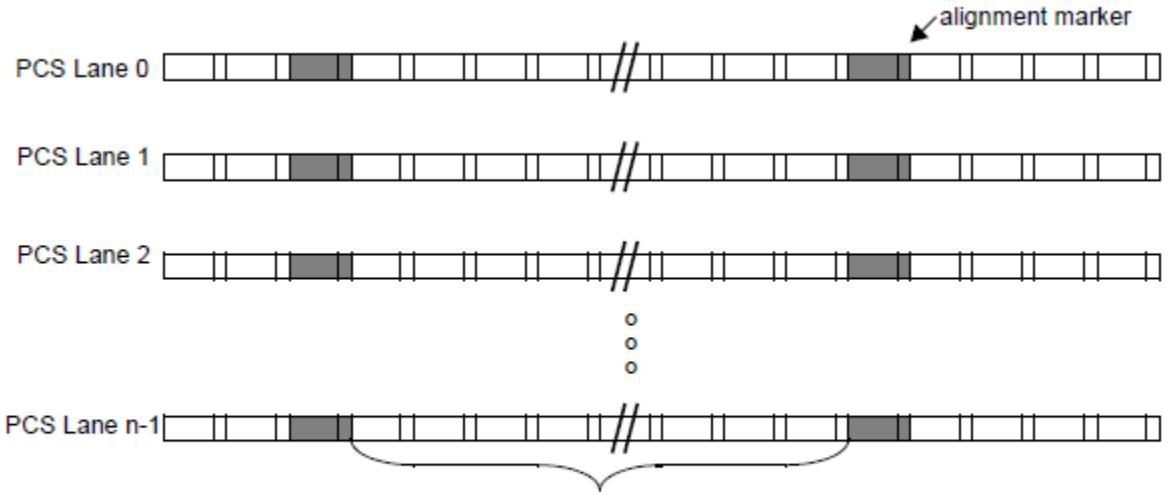
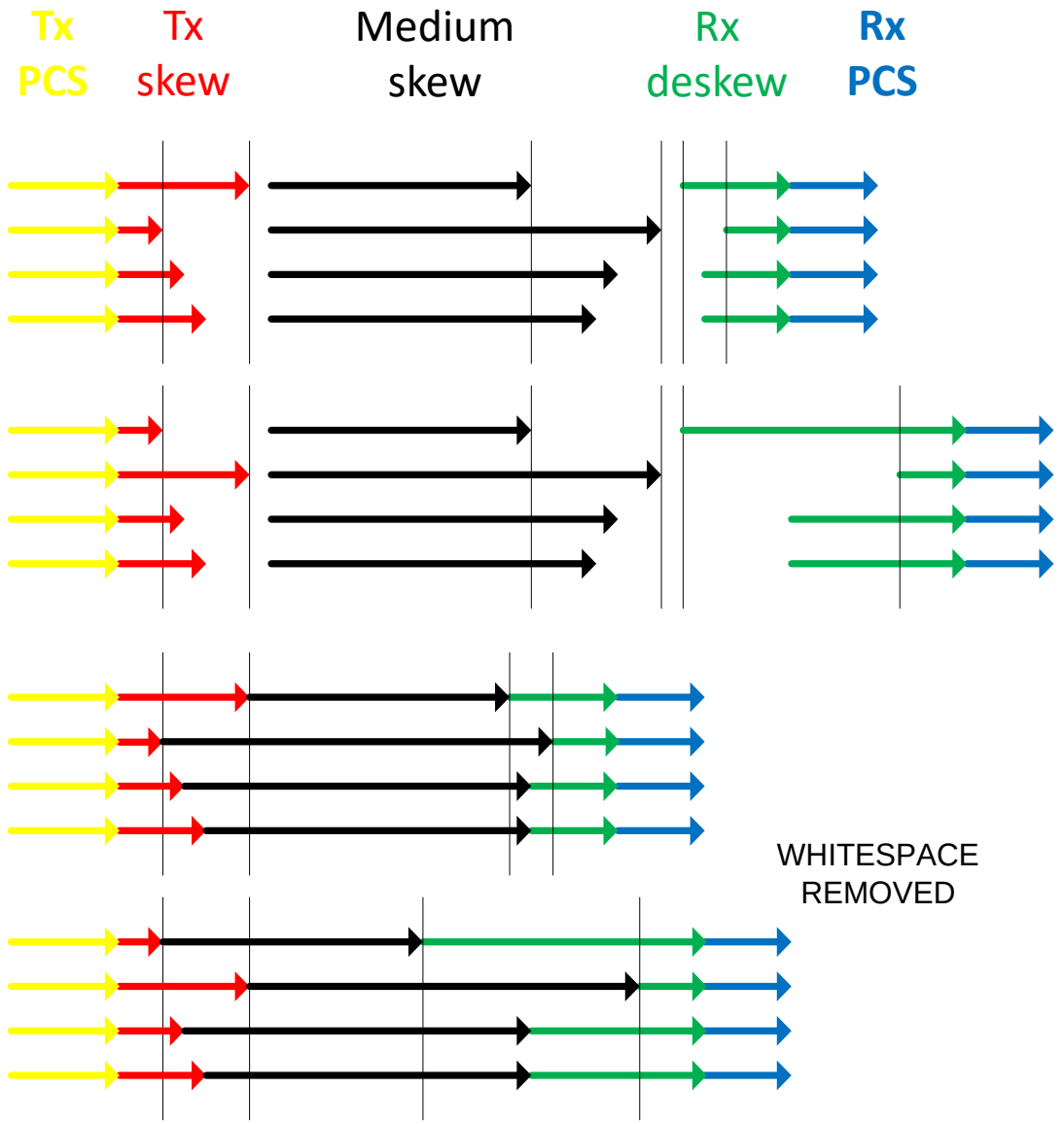


Figure 82–8—Alignment marker insertion period

Issue #4: Tx Skew

- For multi-lane interfaces, per-lane skew at Tx MDI adds to timing error
- Intermingles with per-lane skew of the medium
  - Tx skew and medium skew can be additive or subtractive relative to max total skew
  - Tx skew is implementation-specific and, thus, not symmetric between endpoints



IEEE 802.3 Timestamp Error Performance

Ethernet Rate	Path Data Delay Variation per Tx/Rx Interface (ns)				Max|TE| per Tx or Rx Interface (ns)	Max|TE| contribution per PTP Boundary Clock (ns)
	mismatched SFD timestamp point[1]	Idle insert/delete[2,3]	AM/CWM insert/delete[3]	PCS Lane distribute/ merge		
10M	800	400	N/A	N/A	1200	2400
100M	80	40	N/A	N/A	120	240
GE	8	16[4], 8[5]	N/A	N/A[4], 0[5]	24[4], 16[5]	48[4], 32[5]
2.5GE	3.2	12.8	N/A	5[6]	21	42
5GE	1.6	6.4	N/A	2.5[6]	10.5	21
10GE	0.8	3.2	N/A	N/A	4	8
25GE	0.32	1.28	2.56	N/A	4.16	8.32
40GE	0.2	1.6	6.4	4.8	13	26
100GE	0.08	0.64	12.8	12.16	25.68	51.36
200GE	0.04	0.32	2.56	N/A[7]	2.92	5.84
400GE	0.02	0.16	2.56	N/A[7]	2.74	5.48

1. Not applicable for IEEE Std 802.1AS as it specifies use of just one message timestamp point.
2. Value shown corresponds to the minimum effect of an Idle insert/delete.
3. Only PTP packets that coincide with an AM, CWM, or Idle insert/delete event have their path data delay affected.
4. For 1000BASE-X
5. For 1000BASE-T
6. For illustrative purposes on these rates, lane distribute/merge operation is shown as belonging to the PCS rather than to the FEC.
7. For these rates, lane distribute/merge operation belongs only to the FEC and not to the PCS.

Solutions Adopted by IEEE 802.3cx

- Solutions for the previously identified issues have been adopted by P802.3cx (currently at draft D0.5)
- Technical contributions and meeting minutes can be found at:
  - https://www.ieee802.org/3/cx/public/index.html

Solution Adopted by P802.3cx for Issue #1

- P802.3cx allows both “SFD” and “symbol after SFD” to be used as its message timestamp point
  - “SFD” is preserved for legacy IEEE 802.3 implementations
  - “Symbol after SFD” is recommended for new implementations
- Note: This issue shouldn’t affect IEEE 802.1AS-based solutions because the “symbol after SFD” message timestamp point is specified

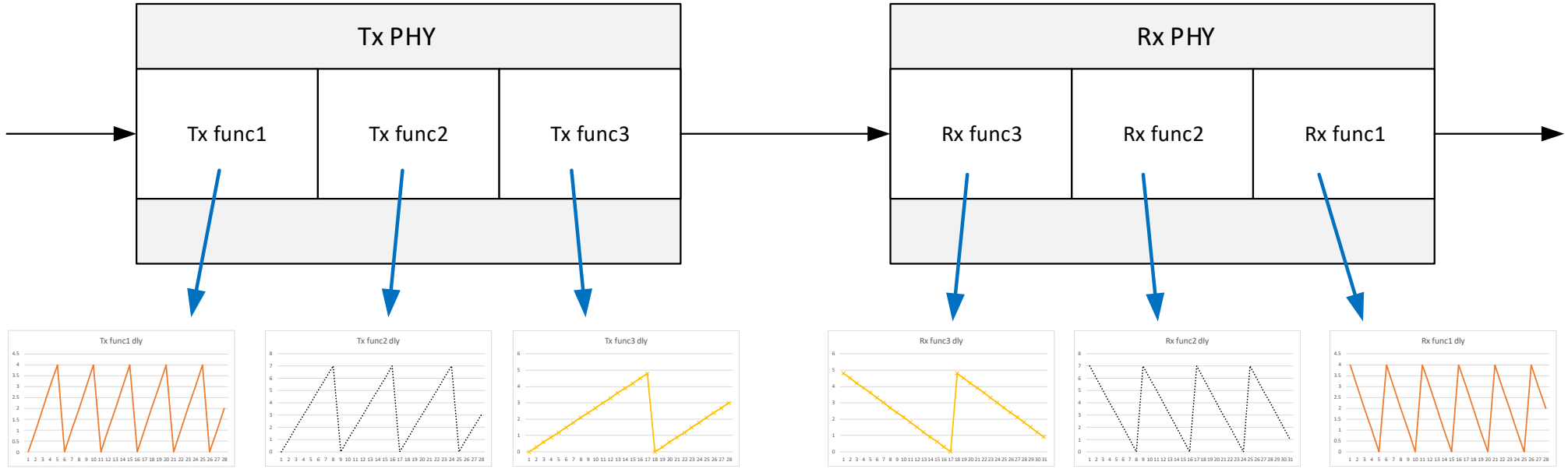
Solution Adopted by P802.3cx for Issue #2 - 1

- Intrinsic delay variations of FEC and PCS-lane distribution functions are periodic and deterministic and result in a constant end-to-end delay.
- For FEC, clause 90.7 of IEEE 802.3-2018 deals with this as follows:
  - *For a PHY that includes an FEC function, the transmit and receive path data delays may show significant variation depending upon the position of the SFD within the FEC block. However, since the variation due to this effect in the transmit path is expected to be compensated by the inverse variation in the receive path, it is recommended that the transmit and receive path data delays be reported as if the SFD is at the start of the FEC block.*
- P802.3cx addresses the multi-PCS lane delay variation using the same method.

P802.3cx will address this "SFD" message timestamp point issue

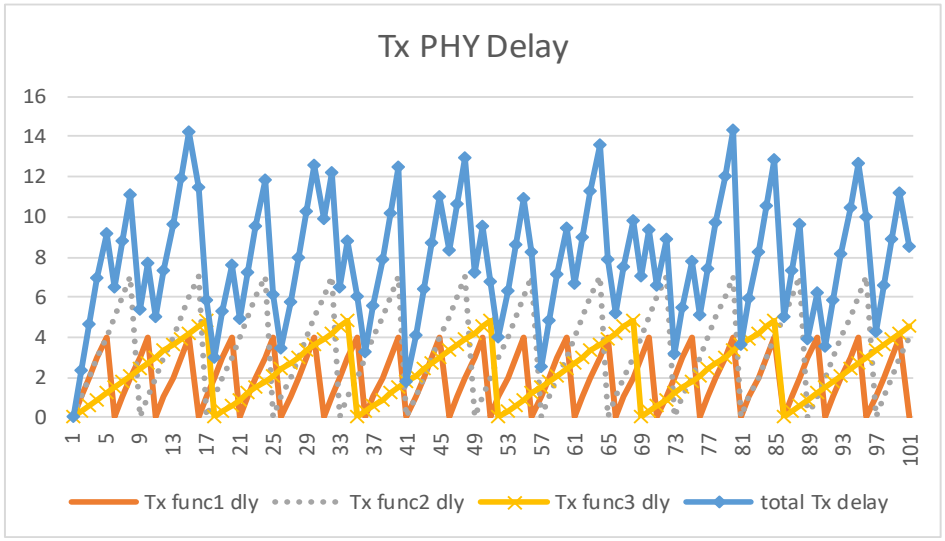
Solution Adopted by P802.3cx for Issue #2 - 2

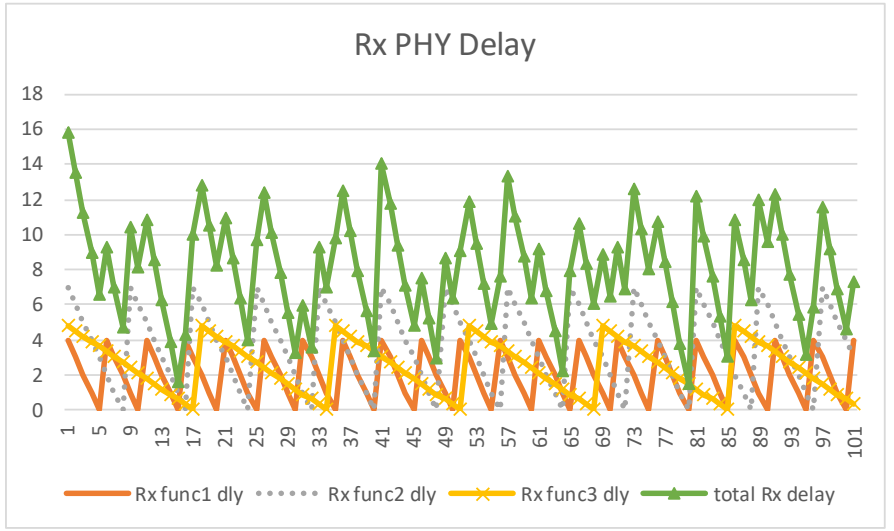
- Example PHY with Tx and inverted Rx intrinsic delay variations
  - Multi-PCS lane distribution/merging
  - FEC encoding/decoding
  - other

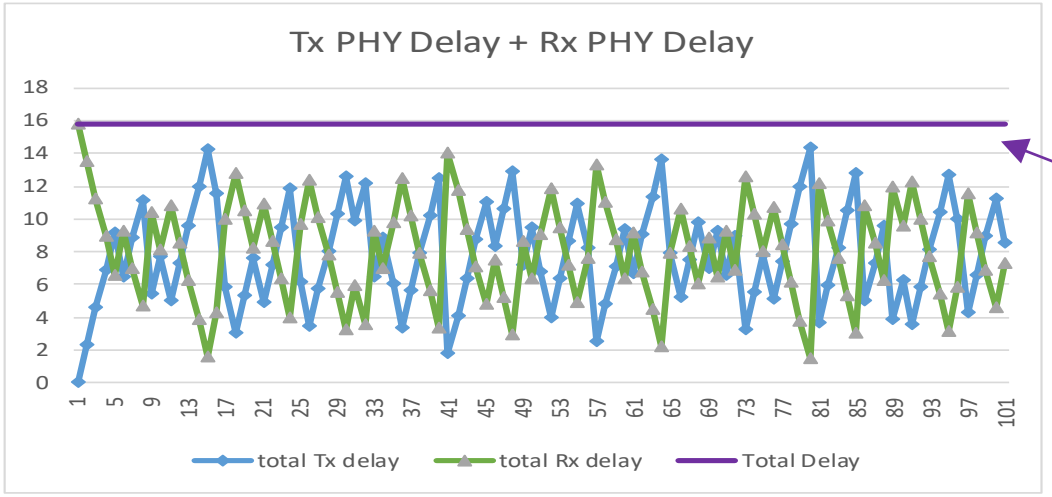


Solution Adopted by P802.3cx for Issue #2 - 3

- Resulting end-to-end delay is constant







Need to specify what portion of this constant delay is allocated to Tx PHY and to Rx PHY

Solution Adopted by P802.3cx for Issue #3

- Delay variations that are independent between Tx and Rx PHYs are to be compensated for individually
- The timestamping model enables the compensation for individual events

Solution Adopted by P802.3cx for Issue #4

- Minimization of Tx skew is recommended

Summary

- Incompatible timestamping implementations create timing errors when they interact with each other
- The IEEE 802.3 and upcoming P802.3cx timestamping models and methods should be used to raise the likelihood of compatibility between implementations

Thank you.