P802.3cx Introduction

Jingfei Lv, Silvana Rodrigues
(Huawei Technologies Co., Ltd.)

IEC/IEEE 60802
IEEE 802.1 Interim meeting, September, 2022

Agenda

- Status of P802.3cx project
- Background of PTP timestamp generation
- Key points of P802.3cx
- Summary

Note that this presentation expresses the opinion and views of the authors, and not of the P802.3cx task force.

Status of P802.3cx project

- P802.3cx Task Force working on an amendment: “Media Access Control (MAC) service interface and management parameters to support improved Precision Time Protocol (PTP) timestamping accuracy”
  - Addresses shortcoming in current IEEE 802.3 standard that may affect timestamp accuracy
- Objective: “Define optional enhancements to IEEE Std 802.3 Clause 90 (support for time synchronization protocols) to support applications requiring sub-nanosecond performance requirements, e.g., “ITU-T Recommendation G.8273.2 'Class C‘ and 'Class D’ system time error performance requirements.”
- P802.3cx is in SA ballot phase, and it is planned to be published at beginning of 2023
- Website: https://www.ieee802.org/3/cx/index.html

Timestamp generation model of IEEE 1588

- Figure 18 of IEEE 1588-2019 shows a general model about timestamp generation.
- Firstly, it needs to specify a message timestamp point, which is used for device to record timestamps when it's detected at TX and RX sides.
- IEEE 1588-2019 defines the first symbol after SFD as the timestamp point, which is consistent with IEEE 802.1AS.

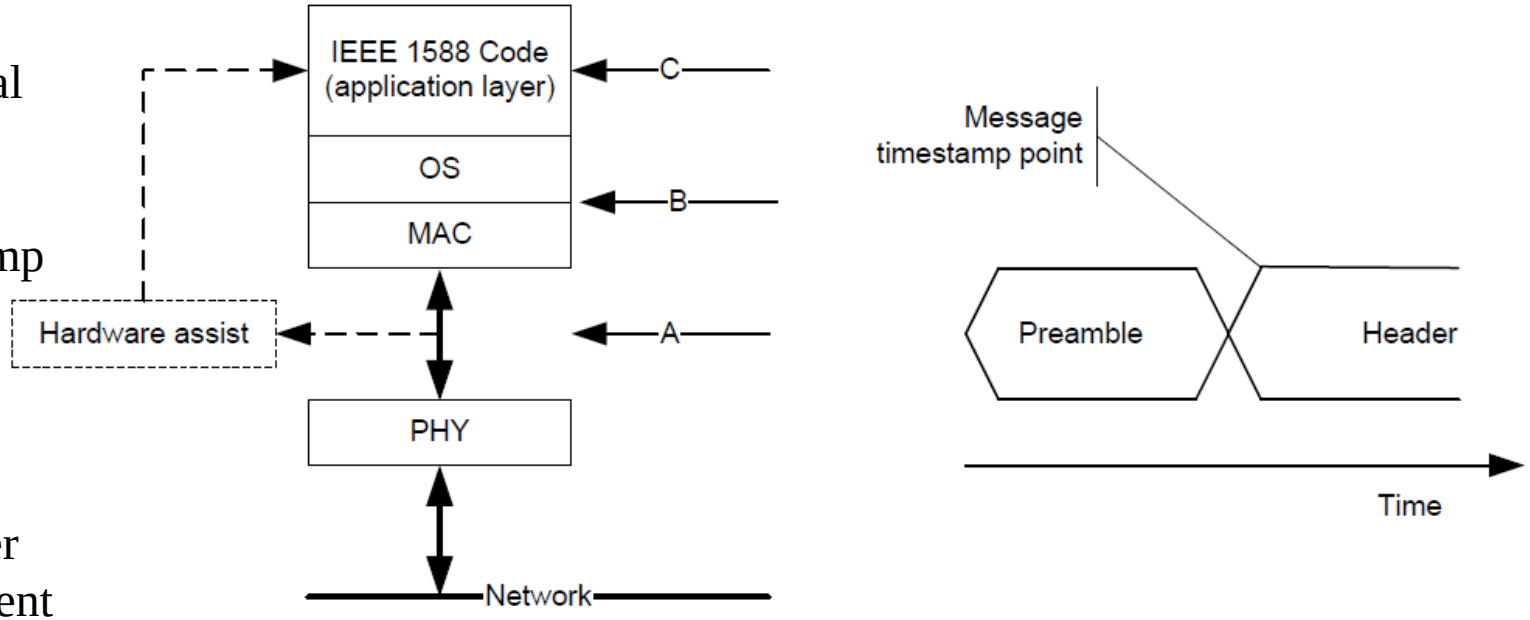


Figure 18—Timestamp generation model

- Secondly, a timestamp reference plane, which is a place that a timestamp is recorded when a message timestamp point passes it. As shown in the figure 18, the reference plane could be at "C", "B" or "A". This is generally implementation specific, and the lower point closer to PHY (even within the PHY) will get a higher accuracy.
- However, even if the reference plane is at "A", there could still be some error unless the delay between the point "A" and the media is symmetrical and constant.

Data delay measurement of current IEEE 802.3

- Since IEEE 802.3-2015, the figure 90 – 3 has defined a data delay model to improve the accuracy of timestamp.
- It assumes the timestamp plane is above PHY, e.g., the gRS layer, then the delay asymmetry and variation of PHY affect the timestamp accuracy.
- Therefore, the clause 90 of IEEE 802.3 required PHY to report its TX and RX delay to the upper layer (see table 90 – 1 of IEEE 802.3), the PTP engine in the upper layer could use these data delay to compensate PTP timestamps captured at the timestamp reference plane.
- However, Clause 90 of IEEE 802.3 only defines the maximum and minimum data delay, which means that the actual delay could be any value between the maximum and minimum.
- Then, a reasonable implementation at the PTP engine is to compensate the mean value of the maximum and minimum, and the error after compensation is +/- (maximum - minimum)/2. But, even with this rough estimation, it can still improve timestamp accuracy.

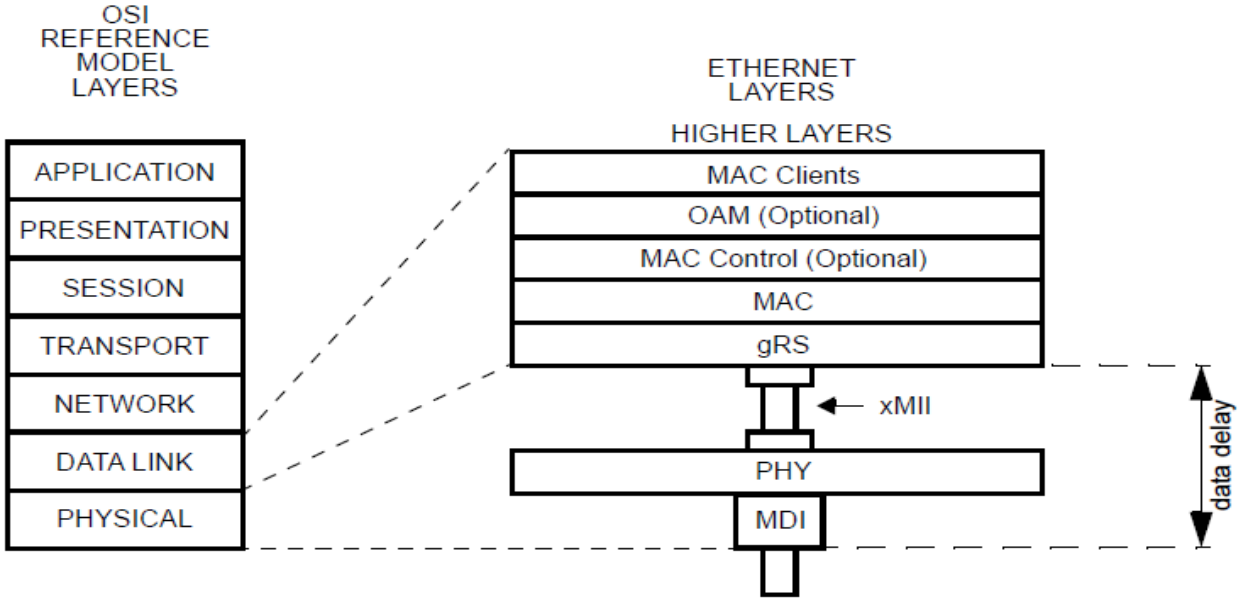


Figure 90–3—Data delay measurement

Timestamp reference plane

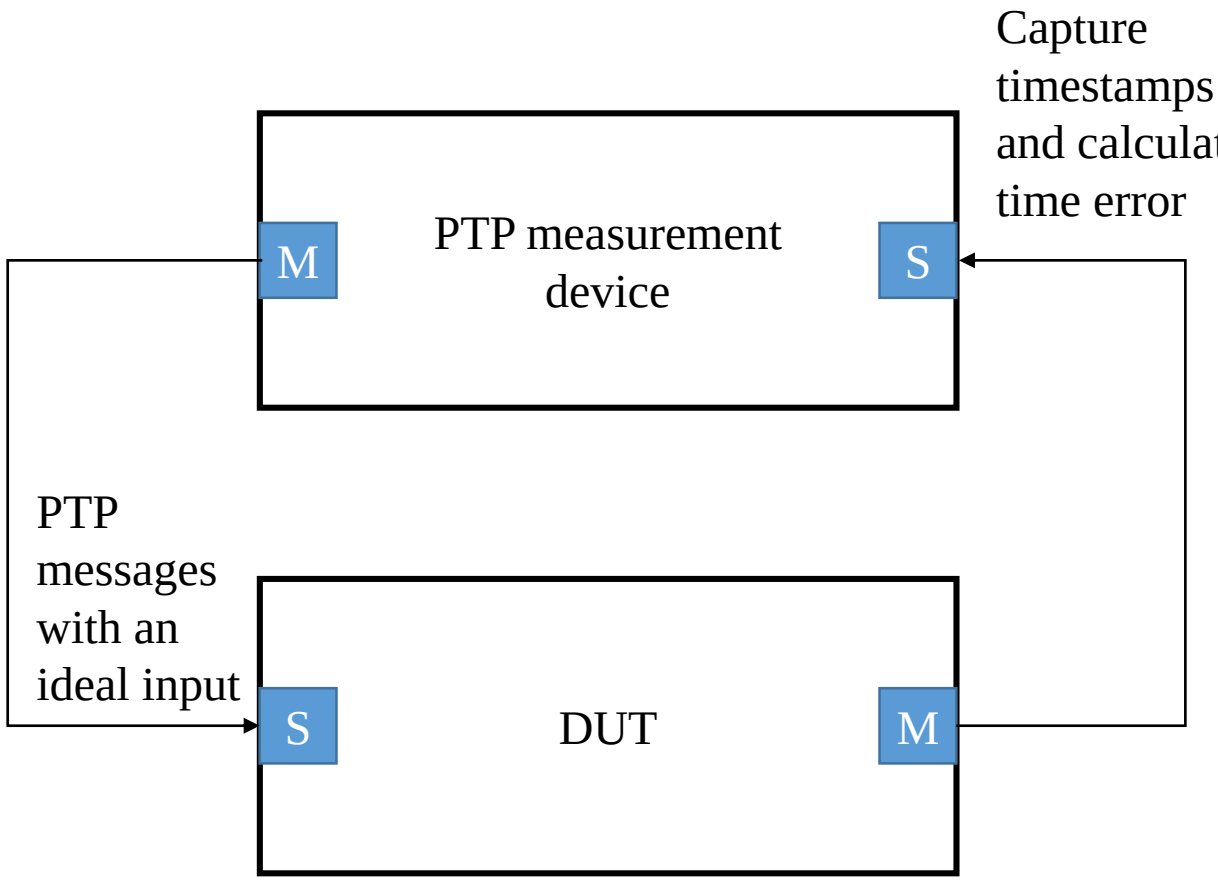
Table 90–1—Summary of TimeSync features in Clause 45

Register	Name	Reference
1.1800	TimeSync PMA/PMD capability register	45.2.1.128
1.1801 through 1.1804	TimeSync PMA/PMD transmit path data delay	45.2.1.129
1.1805 through 1.1808	TimeSync PMA/PMD receive path data delay	45.2.1.130
2.1800	TimeSync WIS capability register	45.2.2.20
2.1801 through 2.1804	TimeSync WIS transmit path data delay	45.2.2.21
2.1805 through 2.1808	TimeSync WIS receive path data delay	45.2.2.22
3.1800	TimeSync PCS capability register	45.2.3.48
3.1801 through 3.1804	TimeSync PCS transmit path data delay	45.2.3.49
3.1805 through 3.1808	TimeSync PCS receive path data delay	45.2.3.50
4.1800	TimeSync PHY XS capability register	45.2.4.12
4.1801 through 4.1804	TimeSync PHY XS transmit path data delay	45.2.4.13
4.1805 through 4.1808	TimeSync PHY XS receive path data delay	45.2.4.14
5.1800	TimeSync DTE XS capability register	45.2.5.12
5.1801 through 5.1804	TimeSync DTE XS transmit path data delay	45.2.5.13
5.1805 through 5.1808	TimeSync DTE XS receive path data delay	45.2.5.14
6.1800	TimeSync TC capability register	45.2.6.14
6.1801 through 6.1804	TimeSync TC transmit path data delay	45.2.6.15
6.1805 through 6.1808	TimeSync TC receive path data delay	45.2.6.16

Background of P802.3cx

- As the PAR of P802.3cx states, the objective of P802.3cx is to support the ITU-T G.8273.2 Class C and Class D system time error, which the max|TE| of Class C is 30ns, Class D is 5ns.
  - Note: ITU-T G.8273.2 also defines Class A and Class B, and max|TE| of Class A is 100ns, Class B is 70ns. The existing function of Clause 90 of IEEE 802.3, which reports the maximum and minimum delay of PHY, could be able to meet Class A and B, but not C and D.

- P802.3cx provides several new functions to improve the report of data delay of the PHY layer.



ITU-T G.8273.2 PTP test platform

Key points of P802.3cx

- Table 90A-1 of P802.3cx shows the potential accuracy impairments without support of P802.3cx, and it includes four items that affect the timestamp accuracy.

❶ ❷ ❸ ❹

Table 90A–1—Magnitude of potential timestamp accuracy impairments

Ethernet rate	Magnitude of potential timestamp accuracy impairments per transmit or receive port (ns)			
	Mismatched data delay measurement point [a]	Idle insertion / removal [b,c]	Alignment marker/ codeword marker insertion / removal [c]	PCS lane distribution / merging
10M	800	400	N/A	N/A
100M	80	40		N/A
1G	8	16[d], 8[e]		0 [e], N/A[d]
2.5G	3.2	12.8		N/A [g]
5G	1.6	6.4		N/A [g]
10G	0.8	3.2		N/A [d, f, g] , 0 [e]
25G	0.32	1.28	10.24	N/A
40G	0.2	1.6	6.4	4.8
50G	0.16	1.28	5.12	3.84
100G	0.08	0.64	12.8	12.16
200G	0.04	0.32	2.56	N/A [g]
400G	0.02	0.16	2.56	N/A [g]

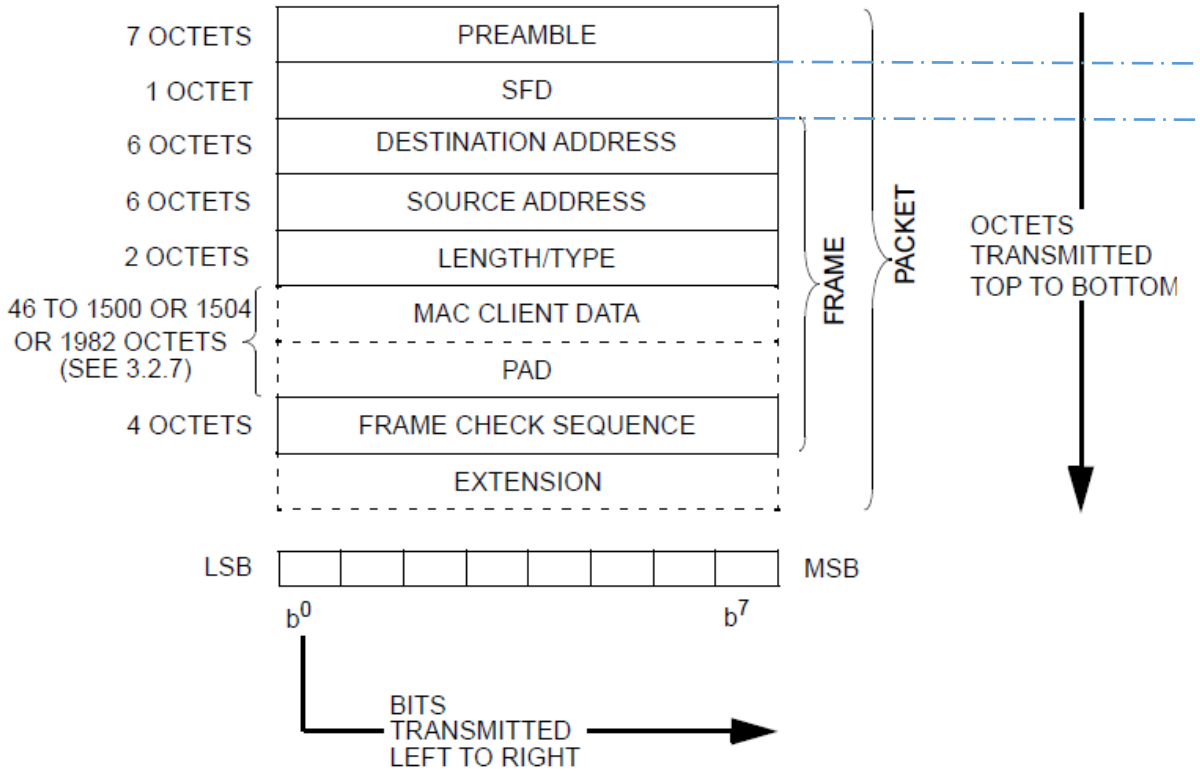
[a] The value shown only accounts for the time between the two data delay measurement point options when they are adjacent. See Annex 90A.3 for other factors that can affect some of these values.
[b] The value shown corresponds to the effect of a single Idle insertion/removal.
[c] The path data delay of a packet can be affected when its data delay measurement point occurs after an alignment marker, codeword marker, or Idle insertion/removal event.
[d] For 1000BASE-X or 10GBASE-R
[e] For 1000BASE-T or 10GBASE-X
[f] For 10GBASE-T
[g] For PHYs including FEC, the lane distribution/merging operation belongs only to the forward error correction (FEC) function. The FEC lane distribution/merging operation is not subject to potential timestamp accuracy impairments because its path data delay determination was already clearly defined, and not subject to implementation flexibilities.

- Note: the value in the table is impairment per timestamp. According to the time error equation [(t2 – t1) – (t4 – t3)]/2, as a worst case, the time error impairment could be twice of the value in the table.

Key points of P802.3cx

① **Mismatched data delay measurement point**

- As introduced by one previous slide, IEEE 1588 and IEEE 802.1AS defines the first symbol after SFD as the message timestamp point
- And, the existing IEEE 802.3 specifies the beginning of the SFD as its message timestamp point.



IEEE 802.3 Figure 3–1—Packet format

IEEE 802.3 message timestamp point

IEEE 1588 and 802.1AS message timestamp point

- So, the timestamp difference with these different points will be the delay of one octet regarding to the SFD, but this should be a static error, which is easily to compensate.
- The worst case is 800ns for 10M interface

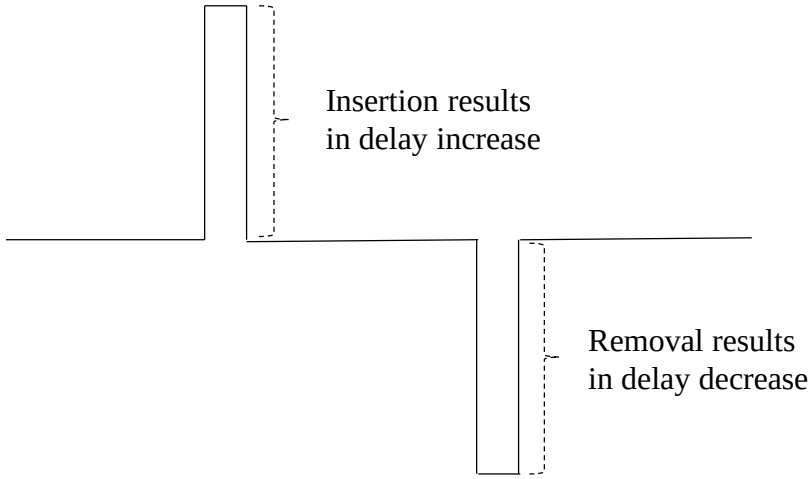
- **P802.3cx allows both of them, and specifies that the point of IEEE 1588 and 802.1AS are recommended.**

Key points of P802.3cx

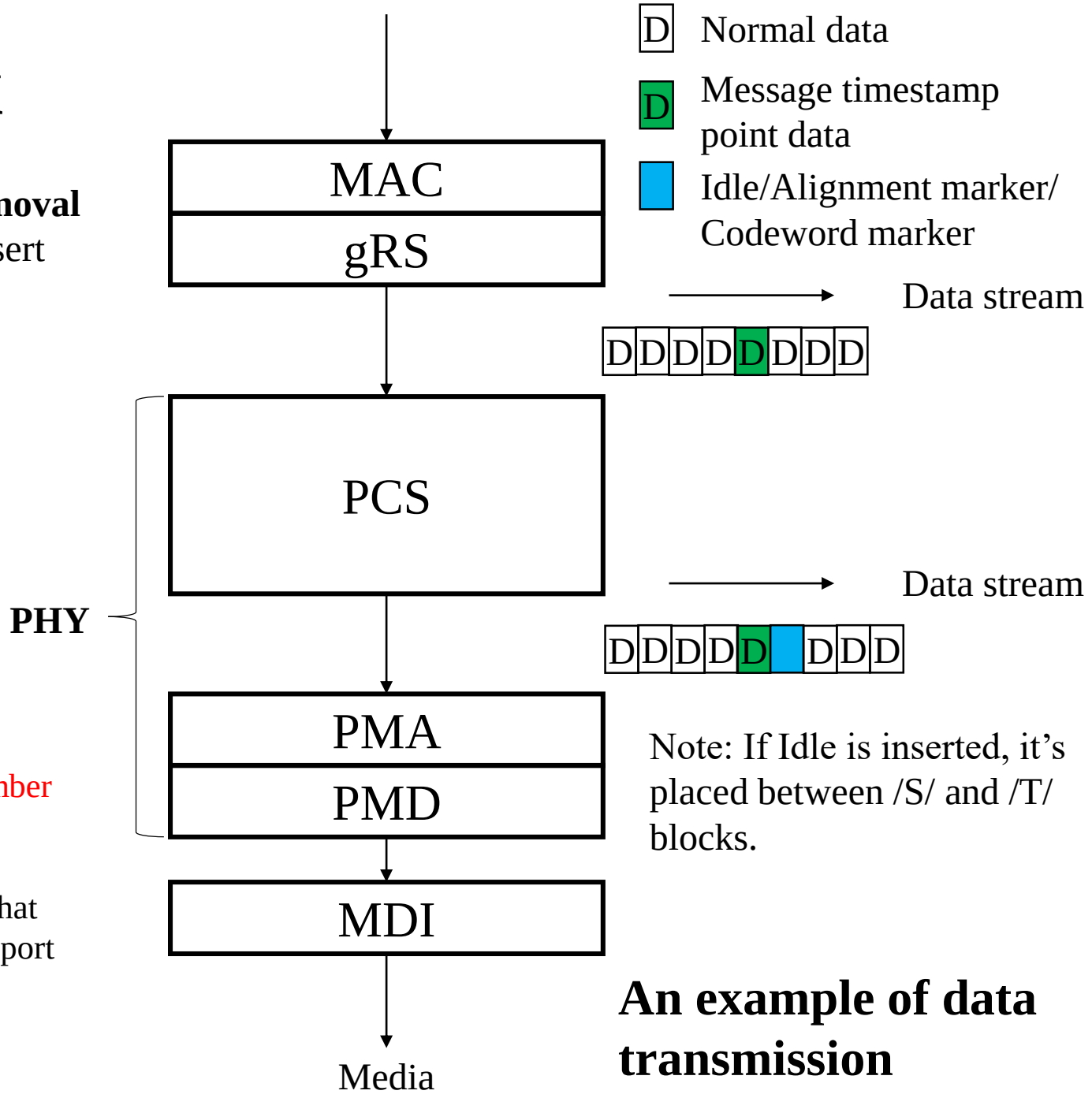
② **Idle insertion / removal**

③ **Alignment marker/codeword marker insertion/removal**

- According to the specification of PCS layer, it will insert or remove Idle/Alignment marker/codeword marker;



- The insertion/removal is random, and P802.3cx defines TX/RX_NUM_BIT_CHANGE that is used to report the number of bits inserted or removed at PCS. The PTP engine can compensate into timestamps.
- This function may be only available for an implementation that integrate MAC and PCS into a single chip. Otherwise, the report function could not work properly.



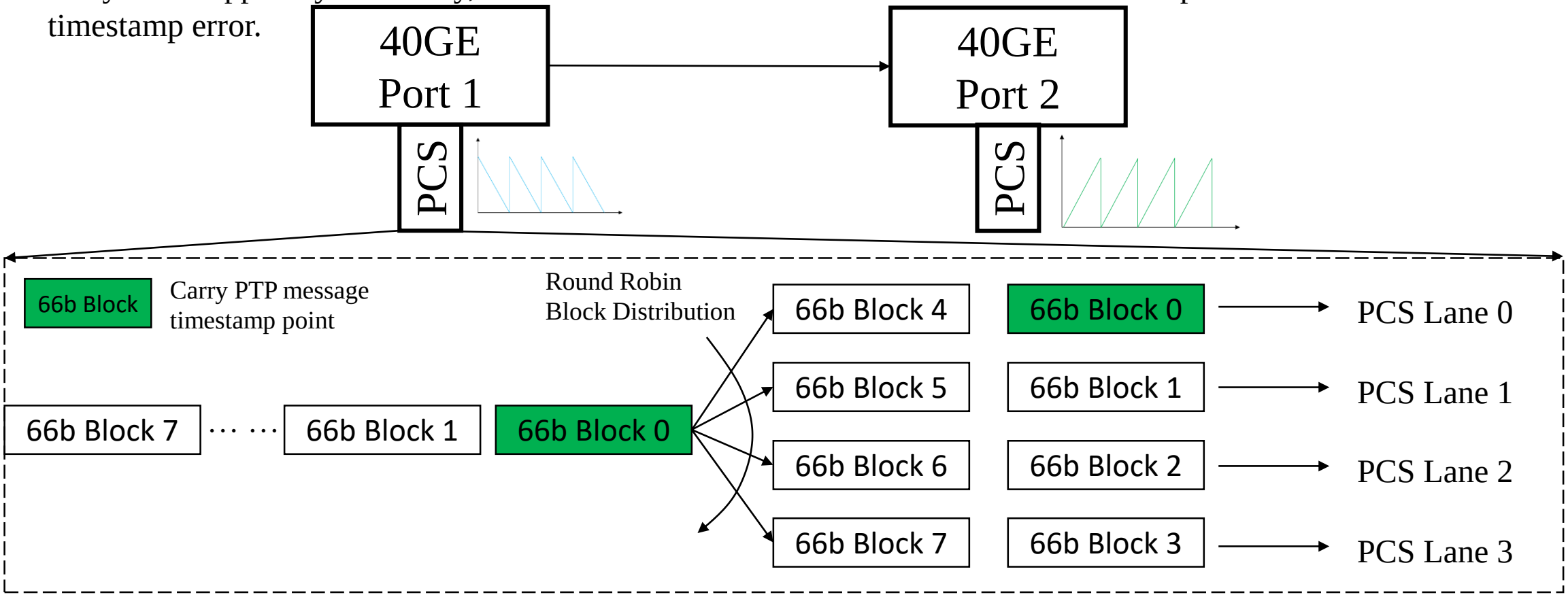
Note: If Idle is inserted, it's placed between /S/ and /T/ blocks.

An example of data transmission

Key points of P802.3cx

④ PCS lane distribution/merging

- For some multi-lane interfaces, e.g., 40GE, there is a block distribution at PCS layer of TX side (see Clause 82.2.6 of IEEE 802.3), and a block merging at PCS layer of RX side.
- Although the delay across the PCS layer is variable due to block distribution and merging, the sum of delay caused by block distribution and merging are constant and symmetrical. Therefore, P802.3cx specifies PCS should not report this delay to the upper lay. Currently, devices from different vendors could use different implementation that results in timestamp error.



This is a block distribution at TX side, and there is another reverse block merging at RX side.

Summary

- In conclusion, without support of P802.3cx, it results in a larger timestamp error, especially for the lower Ethernet rate, which exceeds the current simulation assumption of timestamp error +/-8ns.

- It seems that a PHY with P802.3cx is needed in order to meet the 60802 target (+/-1us over 64 or 100 hops).

Table 90A–1—Magnitude of potential timestamp accuracy impairments

Ethernet rate	Magnitude of potential timestamp accuracy impairments per transmit or receive port (ns)			
	Mismatched data delay measurement point [a]	Idle insertion / removal [b,c]	Alignment marker/ codeword marker insertion / removal [c]	PCS lane distribution / merging
10M	800	400	N/A	N/A
100M	80	40		N/A
1G	8	16[d], 8[e]		0 [e], N/A[d]
2.5G	3.2	12.8		N/A [g]
5G	1.6	6.4		N/A [g]
10G	0.8	3.2		N/A [d, f, g] , 0 [e]
25G	0.32	1.28	10.24	N/A
40G	0.2	1.6	6.4	4.8
50G	0.16	1.28	5.12	3.84
100G	0.08	0.64	12.8	12.16
200G	0.04	0.32	2.56	N/A [g]
400G	0.02	0.16	2.56	N/A [g]

[a] The value shown only accounts for the time between the two data delay measurement point options when they are adjacent. See Annex 90A.3 for other factors that can affect some of these values.
[b] The value shown corresponds to the effect of a single Idle insertion/removal.
[c] The path data delay of a packet can be affected when its data delay measurement point occurs after an alignment marker, codeword marker, or Idle insertion/removal event.
[d] For 1000BASE-X or 10GBASE-R
[e] For 1000BASE-T or 10GBASE-X
[f] For 10GBASE-T
[g] For PHYs including FEC, the lane distribution/merging operation belongs only to the forward error correction (FEC) function. The FEC lane distribution/merging operation is not subject to potential timestamp accuracy impairments because its path data delay determination was already clearly defined, and not subject to implementation flexibilities.

References

[1] IEEE 802.3cx™/D2.2, Draft Standard for Ethernet Amendment 5: Media Access Control (MAC) Service Interface and Management Parameters to Support Improved Precision Time Protocol (PTP) Timestamping Accuracy

(https://www.ieee802.org/1/files/private/liaisons/IEEE_P802_3cx_D2_2_for_IEEE_802d1.pdf)

[2] IEEE 802.3 – 2022, *IEEE Standard for Ethernet*

[3] 60802-Tse-Timestamping-models-0421-v01.pdf, Timestamping Models and Activities in P802.3cx, Richard Tse (https://www.ieee802.org/1/files/public/docs2021/60802-Tse-Timestamping-models-0421-v01.pdf)

Thank you!