

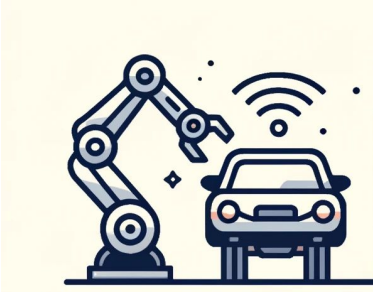
Perspectives on Network and Computer Timekeeping

WSTS 2024 / San Diego, CA / 2024-05

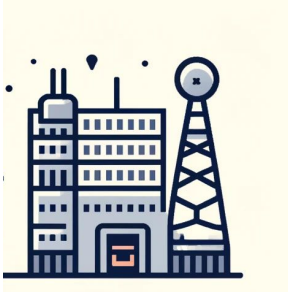
Kevin B. Stanton, Ph.D.
Stanton Consulting LLC
kevin.b.stanton@gmail.com
[Linkedin](#)

End-to-end Time Distribution

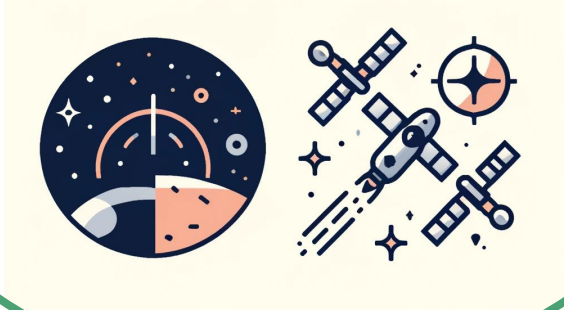
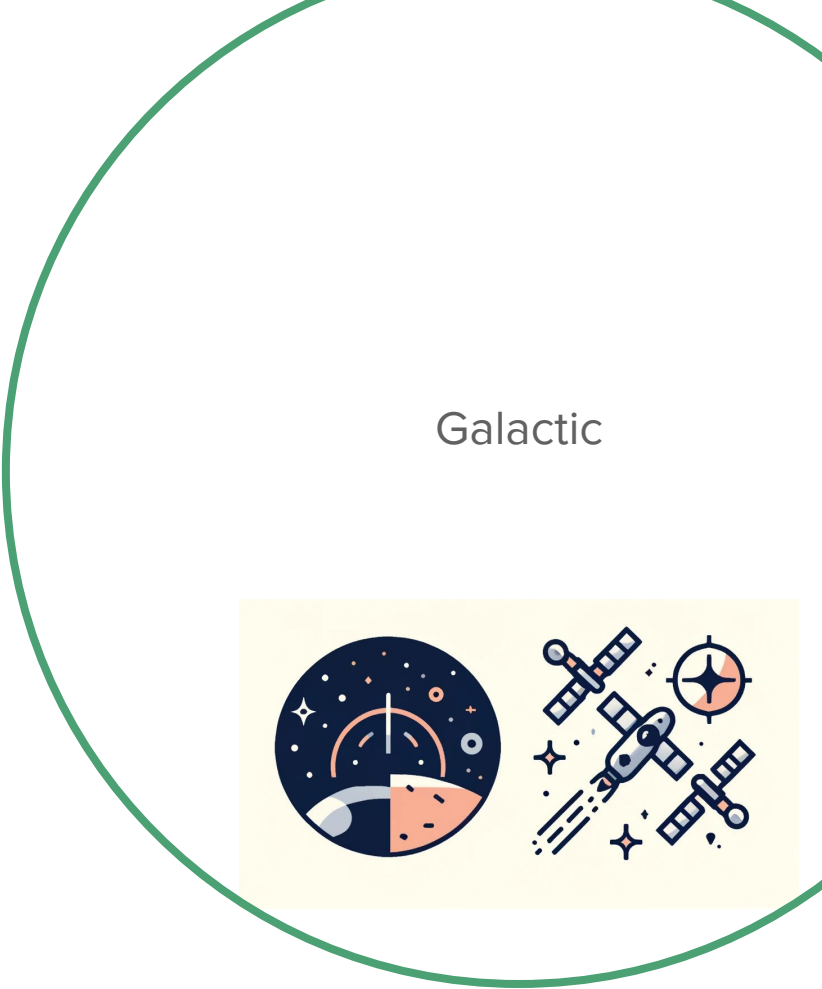
Local



Global



Galactic



Not to scale

Timing in the Last Centimeters ⇒ *The Need*

Timing in the Last 2.54 Centimeters (“the last inch problem”)

Timing in cyber-physical systems: the last inch problem

John C. Eidson* Kevin B Stanton†

*University of California at Berkeley, Berkeley, CA, USA

†Intel Corporation, Hillsboro, OR, USA

Abstract—Distributed cyber-physical systems (CPS) are increasingly provided with an accurate and precise common sense. Section V then discusses how these primitives enable timing decisions that are correct by construction. Architectural choices



ISPCS 2015 International IEEE Symposium on
Precision Clock Synchronization for
Measurement, Control and
Communication

October 11 - 16, 2015

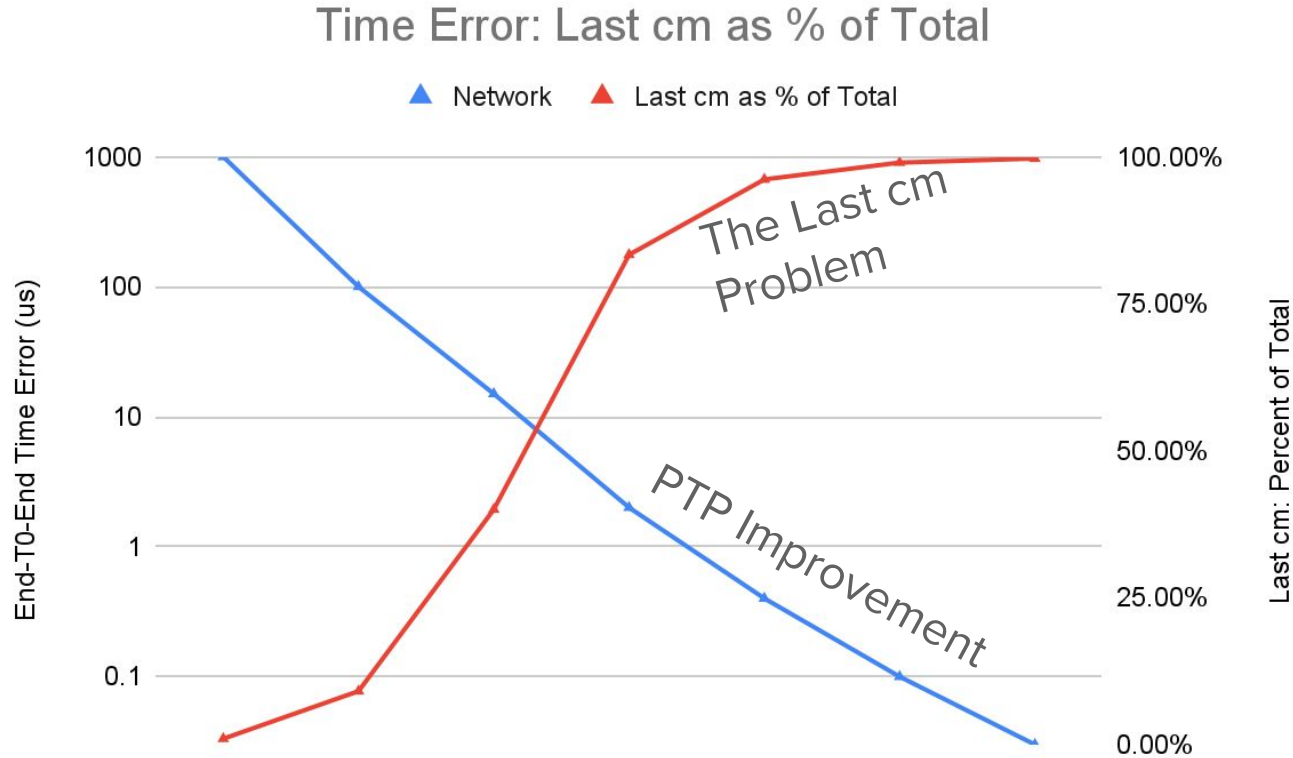
Mercure Wanshang Beijing, Beijing, China

“In either case, close attention must be paid to path latency between the PHY and the PTP clock in the microprocessor. In particular, **substantial degradation of accuracy** can occur when transferring PTP time from a **network interface card to the microprocessor** over a serial load/store interface such as PCI Express.”

“However, a recent **PCI Express enhancement** for in-band time transfer over PCI called **Precision Time Measurement (PTM)** promises to transfer time to the microprocessor with **accuracy in the nanoseconds**[35].”

<https://ptolemy.berkeley.edu/projects/chess/pubs/1156.html>

We Knew that PTP Would both Solve and Create Problems



Time Transfer To Software (Within the Compute System)

The Way We've Always Done It (Software-Based)

1. Read **CPU** Counter
2. Read **PTP** Counter in PCIe NIC
Wait (for the long & variable delay)
3. Read **CPU** Counter

⇒ What could possibly go wrong?

Changing the Rules with PTM Hardware-Based

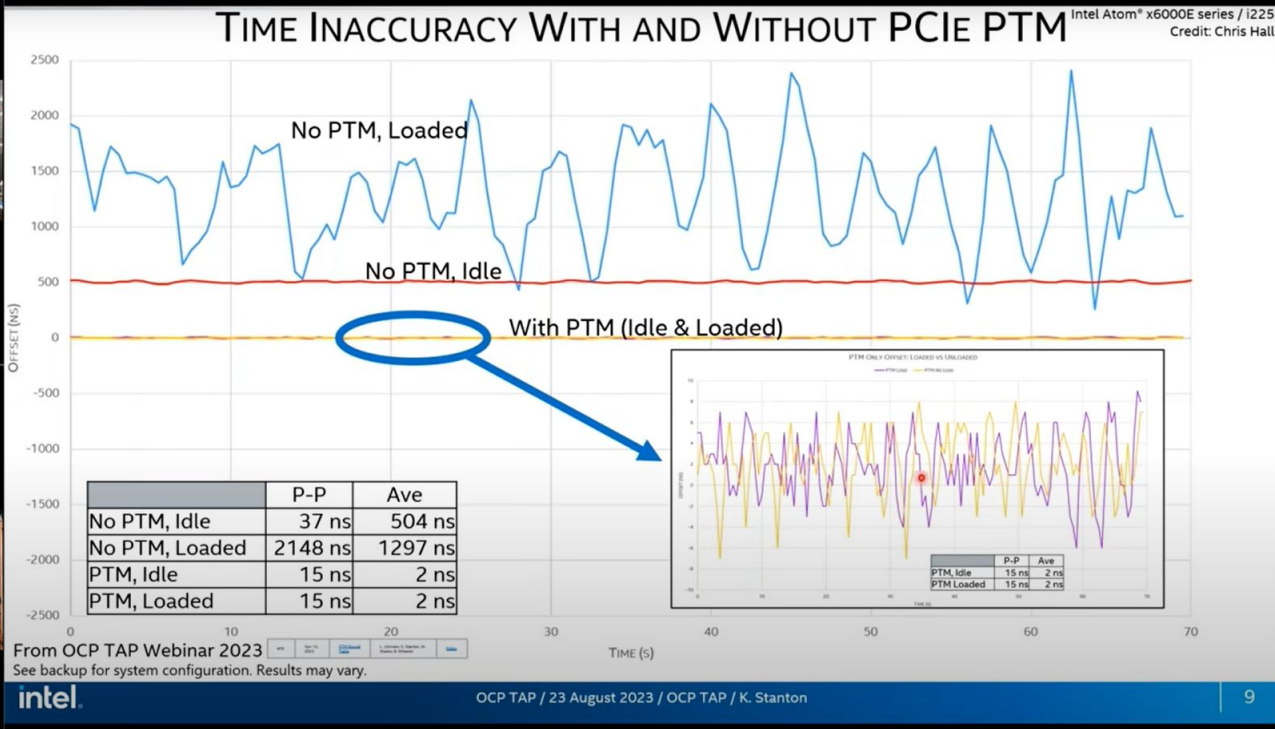
1. Extend the **CPU** Counter to PCIe NIC
2. NIC “simultaneously” captures
 - a. **CPU** Counter
 - b. **PTP** Counter in PCIe NIC

⇒ *Synchronous Logic:*
“Simultaneous”

Timing in the Last Centimeters ⇒ *The Cure*

Software Time versus PTP Time via PPS Outputs

OCP TAP CALL: "Precision Time in the Last Centimeters with PCIe PTM: A Deeper Dive" August 23, 2023



← This Chris H. has a Poster at WSTS 2024. Check it out!

<https://youtu.be/9OILFLV-Sfc?si=U9EX2IB4PTiGiIV3&t=899>

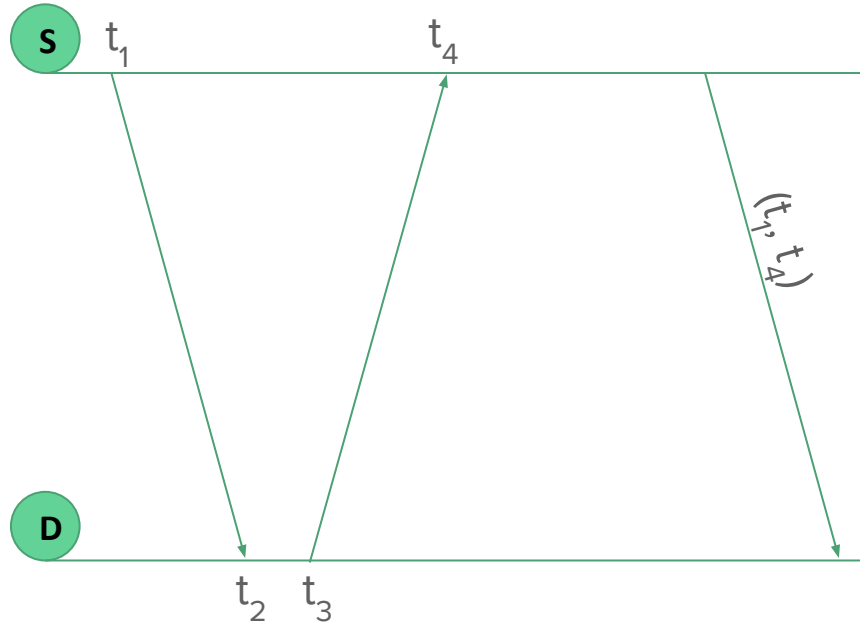
Last Centimeters Time-Sync...

...is “Solved”



Link to Youtube on Previous Page

Two-Way Transfer: Detail



Time **S**ource
Time **D**estination

$$\text{Link Delay} = [(t_4 - t_1) - (t_3 - t_2)] / 2$$
$$\text{Clock Offset} = [(t_2 - t_1) - (t_4 - t_3)] / 2$$

$$\text{Source time at } t_2 = t_1 + \text{Link Delay}$$
$$= t_1 + [(t_4 - t_1) - (t_3 - t_2)] / 2$$

Questions

1. Is the Delay Symmetrical
 - a. Over Ethernet?
 - b. Over PCIE tree
 - c. Over SOC Fabric?

Timing in the Last Centimeters ⇒ *Lessons Learned*

The Path to PTM

The Three Criteria



https://commons.wikimedia.org/wiki/File:Three_first_fingers.JPG

IEEE 1588 / PTP

1588

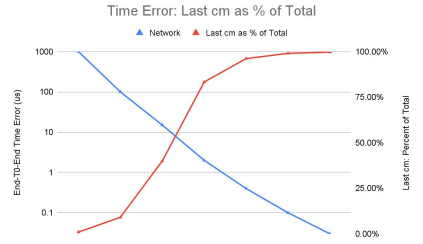
Ethernet NICs

List of Precision Time Measurement Results

Case	Manufacturer	Product	Quality	PTP
1	CPUC	Intel Core i7-9700	400	Support
2	CPUC	Intel Core i5-9600	400	Support
3	CPUC	Intel Core i3-9100	400	Support
4	CPUC	Intel Core i7-9700	400	Support
5	NET	Intel Ethernet	400	Support
6	NET	Intel Ethernet	400	Support
7	NET	Intel Ethernet	400	Support
8	NET	Intel Ethernet	400	Support
9	NET	Intel Ethernet	400	Support
10	NET	Intel Ethernet	400	Support
11	NET	Intel Ethernet	400	Support
12	NET	Intel Ethernet	400	Support
13	NET	Intel Ethernet	400	Support
14	NET	Intel Ethernet	400	Support
15	NET	Intel Ethernet	400	Support
16	NET	Intel Ethernet	400	Support
17	NET	Intel Ethernet	400	Support
18	NET	Intel Ethernet	400	Support
19	NET	Intel Ethernet	400	Support
20	NET	Intel Ethernet	400	Support
21	NET	Intel Ethernet	400	Support
22	NET	Intel Ethernet	400	Support
23	NET	Intel Ethernet	400	Support
24	NET	Intel Ethernet	400	Support
25	NET	Intel Ethernet	400	Support
26	NET	Intel Ethernet	400	Support
27	NET	Intel Ethernet	400	Support
28	NET	Intel Ethernet	400	Support
29	NET	Intel Ethernet	400	Support
30	NET	Intel Ethernet	400	Support
31	NET	Intel Ethernet	400	Support
32	NET	Intel Ethernet	400	Support
33	NET	Intel Ethernet	400	Support
34	NET	Intel Ethernet	400	Support
35	NET	Intel Ethernet	400	Support
36	NET	Intel Ethernet	400	Support
37	NET	Intel Ethernet	400	Support
38	NET	Intel Ethernet	400	Support
39	NET	Intel Ethernet	400	Support
40	NET	Intel Ethernet	400	Support
41	NET	Intel Ethernet	400	Support
42	NET	Intel Ethernet	400	Support
43	NET	Intel Ethernet	400	Support
44	NET	Intel Ethernet	400	Support
45	NET	Intel Ethernet	400	Support
46	NET	Intel Ethernet	400	Support
47	NET	Intel Ethernet	400	Support
48	NET	Intel Ethernet	400	Support
49	NET	Intel Ethernet	400	Support
50	NET	Intel Ethernet	400	Support

PCIe PTM Spec

CPU and SoC Timekeeping Upgrade



PTM Readiness

List of Precision Time Measurement Readiness [\[edit\]](#)

Class	Manufacturer	Product	Quanta	TGPIO	PTM Status
1 CPU	Intel	Core, Gen 11 (Tiger Lake)	4ns		Supported
2 CPU	Intel	Core, Gen 12 (Alder Lake)	4ns		Supported
3 CPU	Intel	Core, Gen 13 / Gen14 (Raptor Lake)	4ns		Supported
4 CPU	Intel	Xeon, Gen 4 (Sapphire Rapid)	2ns		Supported
5 NIC	Intel	I225 (Foxville)	4ns		Supported
6 NIC	Intel	I226 (Foxville)	4ns		Supported
7 FPGA	Intel	Intel Agilex	4ns		Supported
8 Microprocessor	Texas Instruments	AM64x / AM65x / AM68x / AM69x / DRA82x ARM SoCs	4ns		Supported
9 NIC	Nvidia	ConnectX-6	2ns		Beta Tested
10 NIC	Nvidia	ConnectX-7	2ns		Beta Tested
11 Time Card	OCP	Time Card 2	4ns		Supported
12 Time Card	Safran	ART2	4ns		In Progress
13 Time Card	Adtran	5400 SyncModule	4ns		In Progress
14 NIC	Liquid-Markets-Solutions	UberNIC	4ns		Supported
15 NIC	Broadcom	P2100G	4ns		In Progress
16 Corundum	UCSD	Corundum	4ns		In Progress
17 Motherboard	ASRock	SP2C741D16-2T	2ns	2 exposed	Supported
18 Motherboard	ASRock	W790D6UD-1L1N2T/BCM	2ns	2 exposed	Supported
19 Motherboard	Supermicro	X13SEDWF	2ns		Beta Testing
20 CPU	Intel	Core, Gen 14 (Raptor Lake Refresh)	4ns		Supported
21 Motherboard	Asus	W790SE	2ns	1 under BGA	Supported
22 Motherboard	Asus	Z790 Maximus Extreme	2ns	1 under BGA	Supported
23 Motherboard	Asus	Z690-I	2ns	1 under BGA	Supported
24 Motherboard	Asus	Z790-G WIFI	4ns	1 under BGA	Supported
25 Motherboard	Asus	Z690-Extreme	4ns	1 under BGA	Supported
26 Motherboard	Asus	W790-ACE	2ns	1 under BGA	Supported
27 Motherboard	Asus	Z790-I	4ns	1 under BGA	Supported

https://www.opencompute.org/wiki/PTM_Readiness

The Future of Time Synchronization ⇒ What's Next?

Google's Andrew Fikes on Time



“Everybody tells you [that] you **can't trust Time**. The clock on your server is **absolutely** something that you **can't trust**, and you develop this **innate fear** of that thing. [*Distrust of Time*] **is Legendary...**

I can tell you that I **completely trust our time system**, and it has made my life so much easier. Having a **source of global ordering** that I don't have to reach out to talk to, that I can **reference locally**, is just a **huge huge huge thinking shift**.

... It took two or three versions [of the Google database, Spanner] to get over that [fear of relying on time]”

— Andrew Fikes (VP of Infrastructure @ Google, CoudNext 2019) (Emphasis Mine)

<https://www.youtube.com/watch?v=nvlt0dA7rsQ>

The Future (Beyond Accuracy)

APIs Beyond “Now()”

Authenticated Time Source

Resilient Timing Sources / Paths

Explicit Time Accuracy / Jitter Bounds

Explicit Consistency / Availability Tradeoffs

Non-Newtonian Timekeeping

https://www.cisa.gov/sites/default/files/2023-02/Technical-Level_Resilient_Timing_Overview-CISA_Fact_Sheet_508C.pdf



CISA
CYBER+INFRASTRUCTURE

DEFEND TODAY. SECURE TOMORROW.

TIME – THE INVISIBLE UTILITY



WHY IS TIME IMPORTANT?

Time is critical to certain services used within most organizations, yet many organizations are unaware of their dependence on time, the source of their time, or the existence of a world time standard.¹ As systems grow in complexity, becoming global and mobile, access to resilient, accurate, and precise time is a necessity in both the private and public sectors worldwide. Without accurate and resilient time, critical functions and services can become unreliable, inaccurate, or unavailable.



SECTORS AND INDUSTRIES DEPENDENT ON TIME

Communications	Transportation	Power Grid	Finance	Security	IT
Telecommunication	Aviation	Frequency Monitoring	Regulatory Requirements	Cryptography	Smart Devices
Cloud Operations	Maritime	Multi-rate Billing	ATM Networks	Access Control	Incident Investigations
Internet of Things (IoT)	Pipelines Rail	Fault Detection		Forensics Surveillance	



WHY SHOULD YOU BE CONCERNED ABOUT TIME NOW?

GPS has become the *de facto* time standard for many commercial users because of its relatively low cost and ubiquitous availability. In 1997, the President's Commission on Critical Infrastructure Protection

Summary

1. The Last Centimeters Problem is “solved”

Implement, adopt, and deploy PTM

2. Software applications need more information

Today, they get only

now ()

As we increase our TRUST in the TIME, let's not forget to inform the application

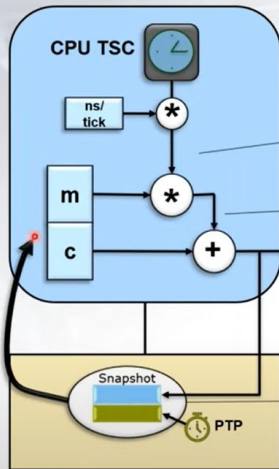
Thank You!

Software Access to “Now”

OCPTAP CALL: “Precision Time in the Last Centimeters with PCIe PTM: A Deeper Dive” August 23, 2023



CPU COUNTER → SYNCHRONIZED TIME



Time “now” (from a Linux application)

- `clock_gettime(CLOCK_MONOTONIC_RAW, &now);`
 - Returns TSC scaled to nominal nanoseconds
- `clock_gettime(CLOCK_MONOTONIC, &now);`
 - Returns TSC scaled to track TAI, in nanoseconds

- `clock_gettime(CLOCK_REALTIME, &now);`
 - Returns `CLOCK_MONOTONIC + (now-1/1/1970)` [incl. leap seconds]

Cross-Timestamp “Snapshot”

- `ioctl(phc_fd, PTP_SYS_OFFSET[_PRECISE])`
 - Returns `(CLOCK_REALTIME, PTP_TIME)`
 - Used to nudge `m` up or down over time

Piecewise-Linear Clock Model: $y=mx+c$

<https://youtu.be/9OILFLV-Sfc?si=ohnZ9U07JT7eoYN6&t=1211>

PCIe PTM: How it works — between chips / chiplets



PCIe PTM IN ACTION

At $t1'$, snapshot PTP Counter

Soon after $t4'$, use $t1$, $t4$, $t2'$, $t3-t2$ to compute:

- $LD = LinkDelay = \frac{[(t4-t1)-(t3-t2)]}{2}$
- **PTM Root time @ $t1'$** = $t2' - LinkDelay$

Return the Cross-Timestamp:

- (PTM Root Time @ $t1'$, PTP Counter @ $t1'$)

Repeat as requested (e.g. by software)

Note: All PTM timestamps are in units of nanoseconds
Note: Clock discontinuities are allowed at any time in the Upstream Port except between $t1 \rightarrow t4$, $t1' \rightarrow t4'$, etc.

intel | OCP TAP / 23 August 2023 / OCP TAP / K. Stanton | 12

<https://youtu.be/9OILFLV-Sfc?si=6n4kS-W7645JfwJd&t=1975>