Improving Time Accuracy of Servers Across the Facebook Fleet
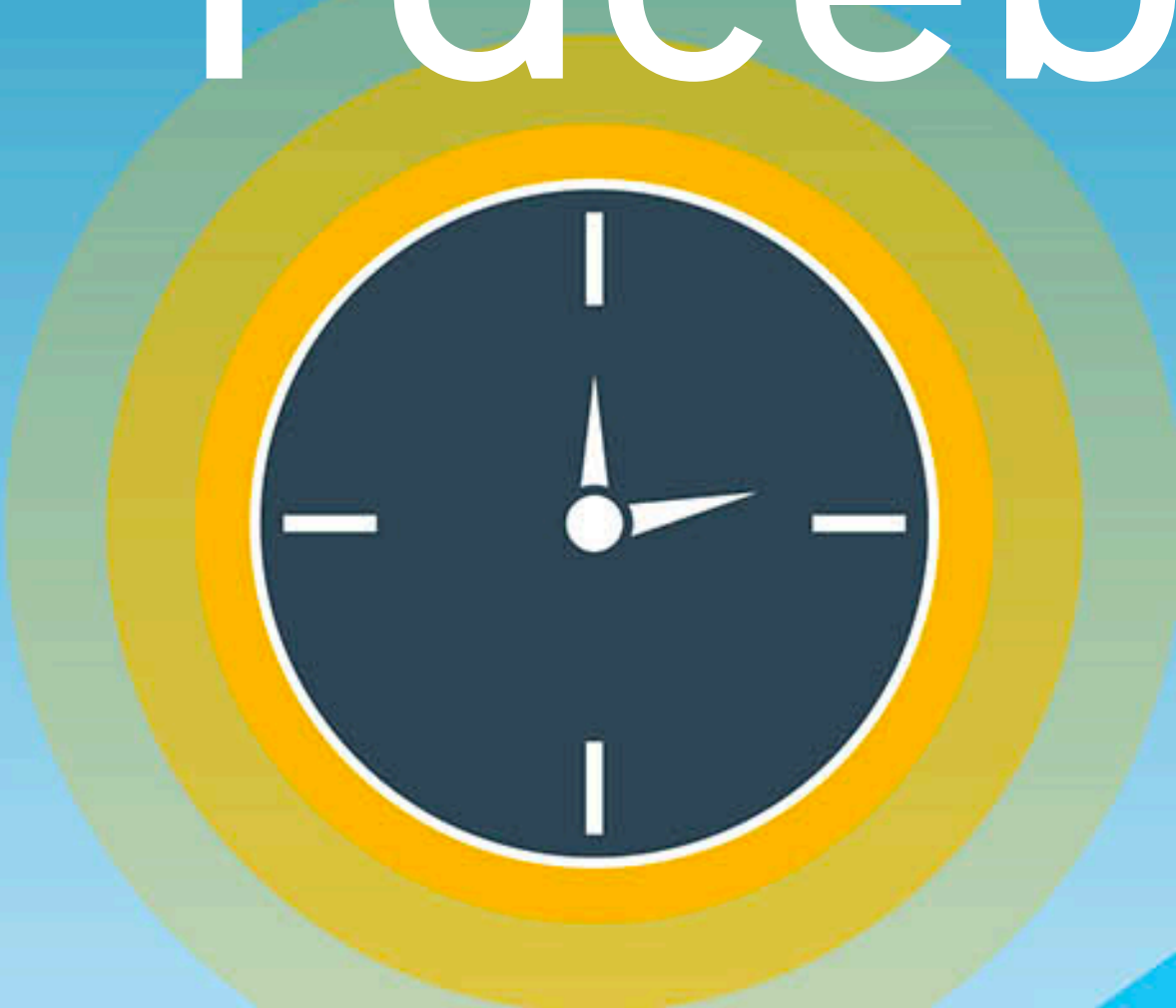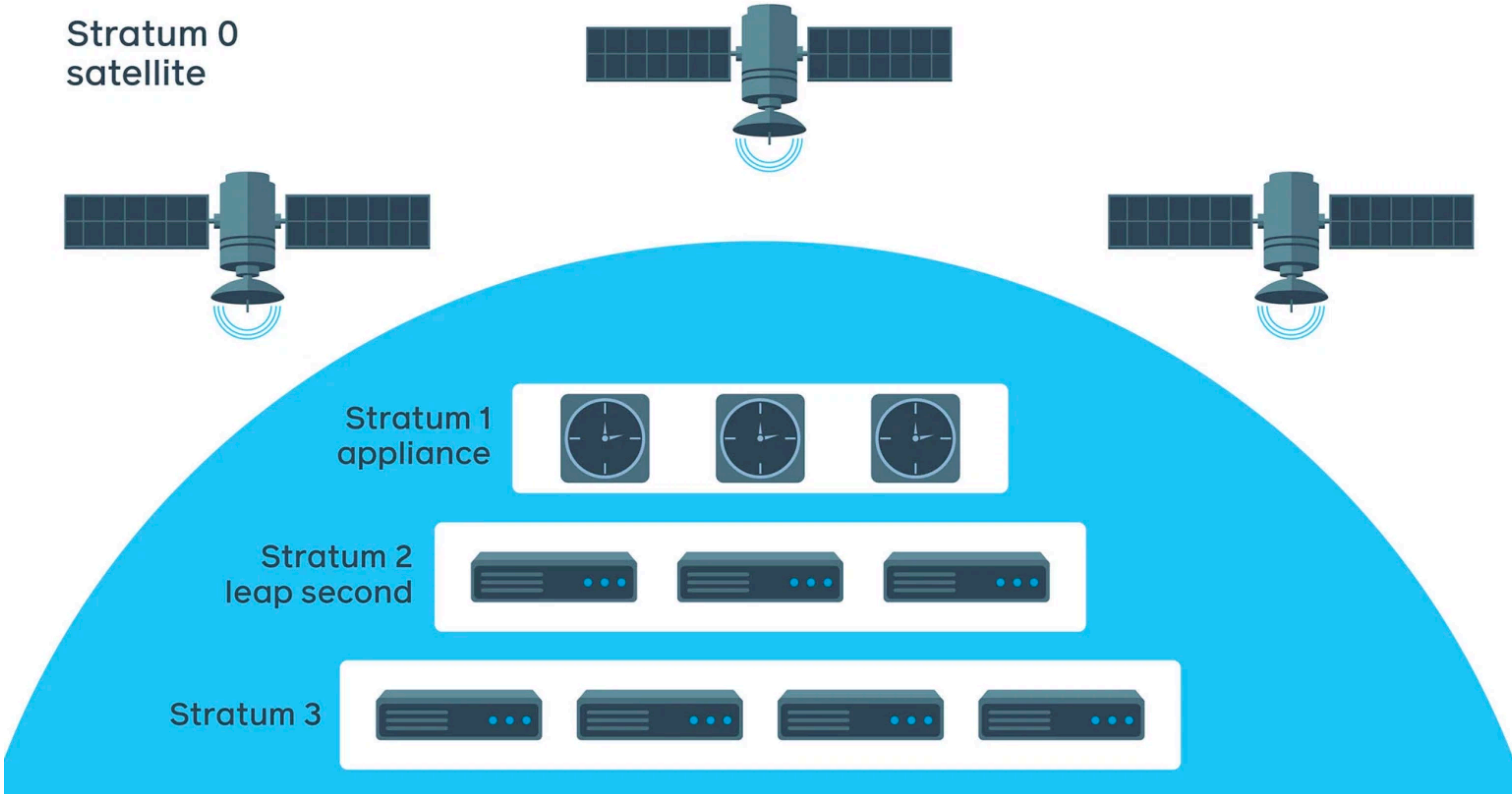
Oleg Obleukhov
Production engineer

FACEBOOK

Typical NTP hierarchy

Stratum 0
satellite

Stratum 1
appliance

Stratum 2
leap second

Stratum 3

Once upon a time

```
1  [user@client ~]# ntpq -p
2  remote              refid           st t when poll reach   delay    offset  jitter
3  ===============================================================================
4  +server1            .FB.             2 u  406 1024  377    0.029     0.149   0.035
5  +server2            .FB.             2 u  548 1024  377    0.078     0.035   0.083
6  *server3            .FB.             2 u  460 1024  377    0.049    -0.185   0.114
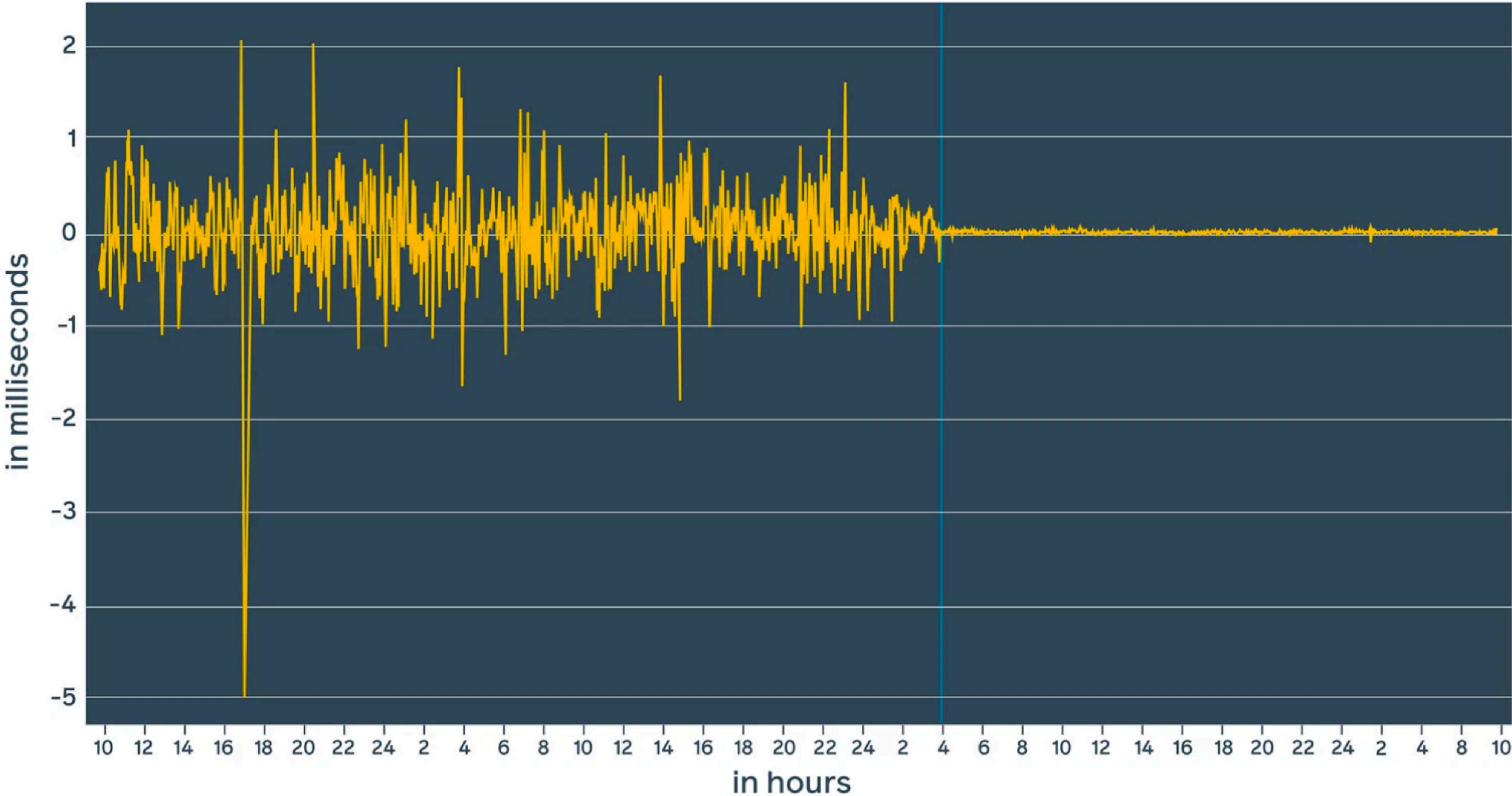```

# Improving precision

# Improving precision

```
1   [user@client ~]# chronyc sources -v
2   210 Number of sources = 19
3   .-- Source mode '^' = server, '=' = peer, '#' = local clock.
4   / .- Source state '*' = current synced, '+' = combined , '-' = not combined,
5 | /   '?' = unreachable, 'x' = time may be in error, '~' = time too variable.
6 || .- xxxx [yyyy] +/- zzzz
7 || Reachability register (octal) -. | xxxx = adjusted offset,
8 || Log2(Polling interval) --. | | yyyy = measured offset,
9 || \ | | zzzz = estimated error.
10 || | | \
11 MS Name/IP address   Stratum Poll Reach LastRx Last sample
12 ===
13 ^+ server1   2   6   377   2   +10us[+10us] +/-   481us
14 ^- server2   2   6   377   0   -4804us[-4804us] +/-   77ms
15 ^* server3   2   6   377   59   -42us[-46us] +/-   312us
16 ^+ server4   3   6   377   60   +11ns[-3913ns] +/-   193us
```
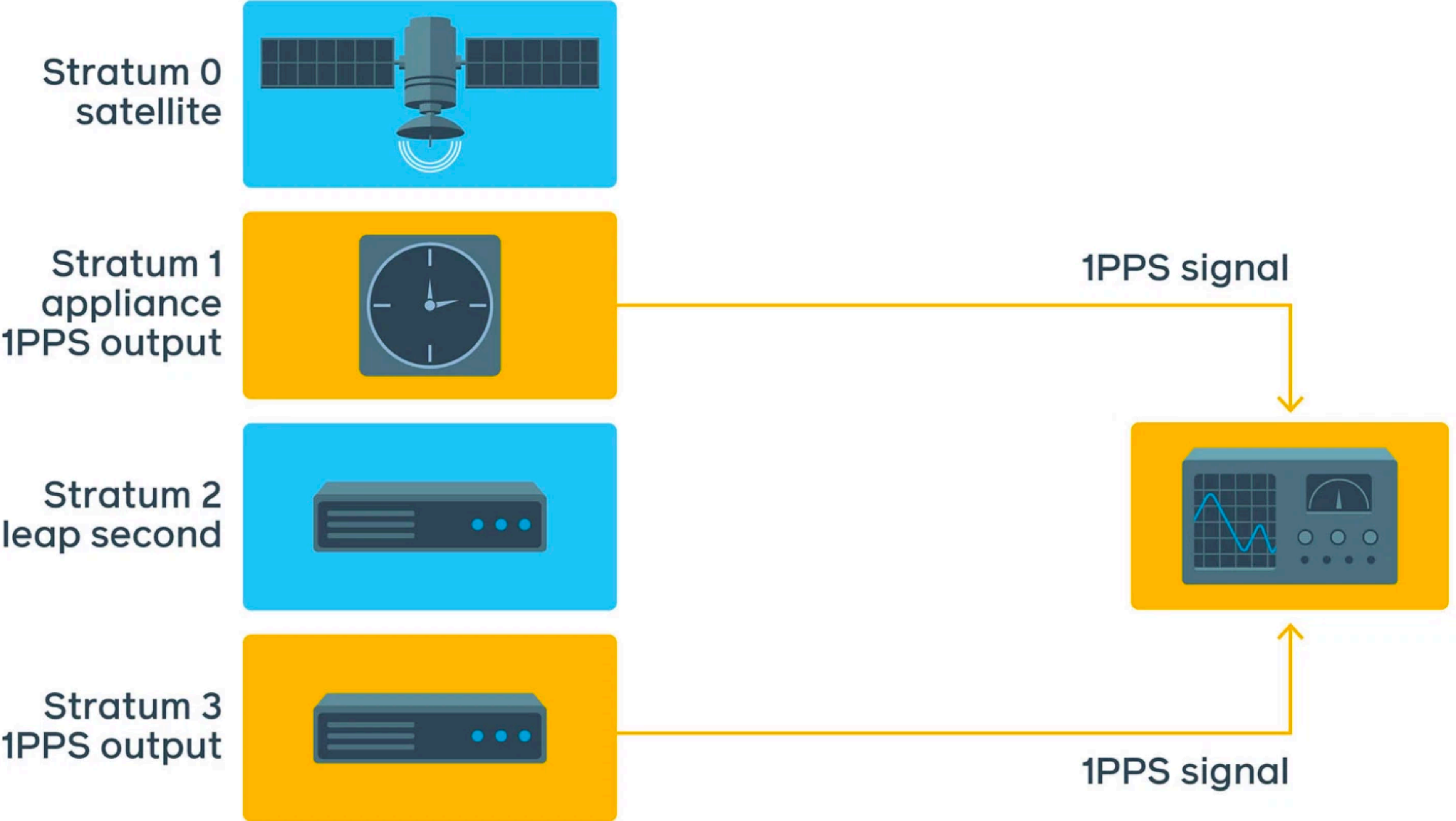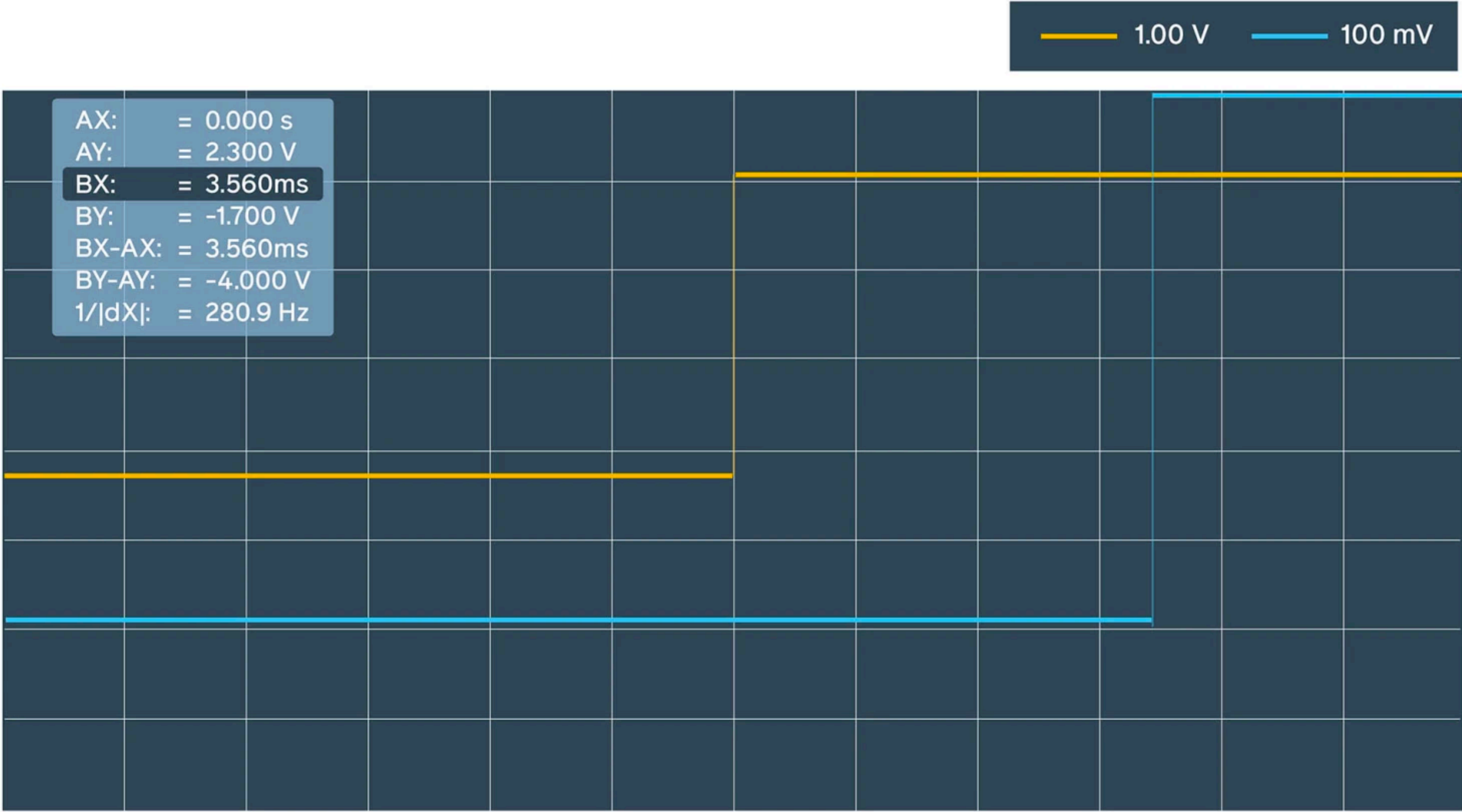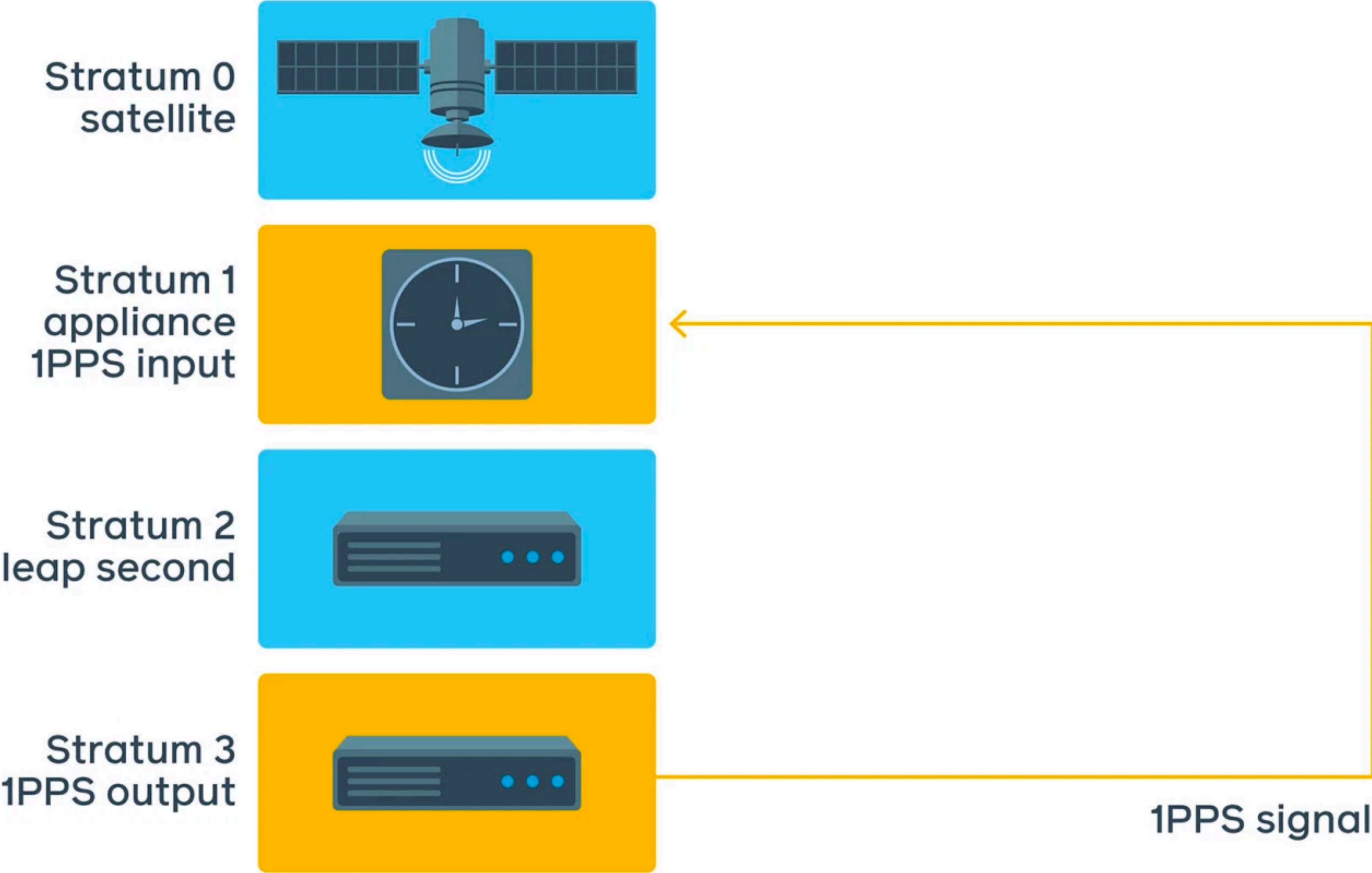
# Improving precision

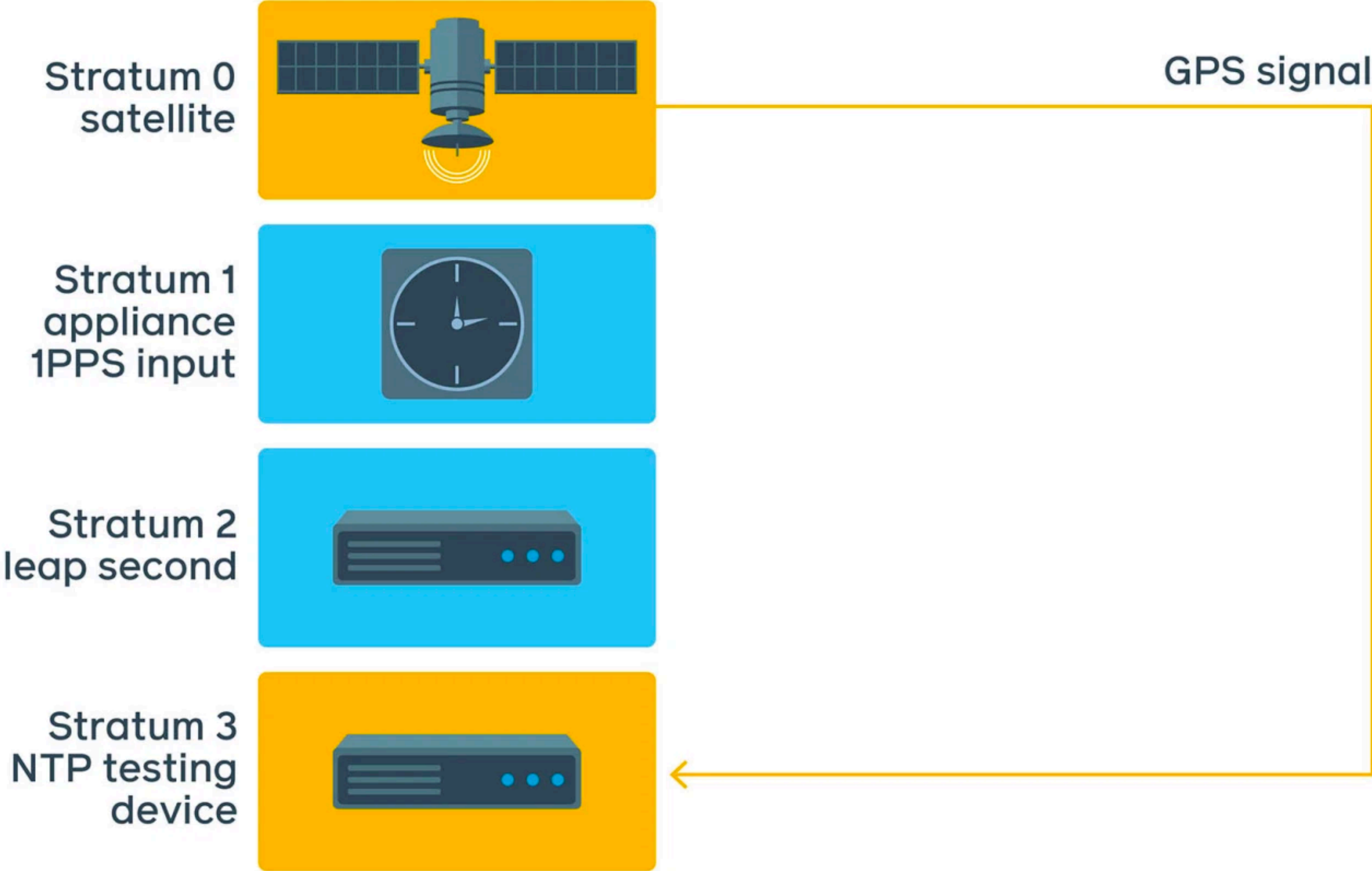# Validating improvements

# Validating improvements. 1PPS



Stratum 0
satellite

Stratum 1
appliance
1PPS output

1PPS signal

Stratum 2
leap second

Stratum 3
1PPS output

1PPS signal

# Validating improvements. 1PPS

# Validating improvements. Dedicated testing device

**Stratum 0 satellite**

**Stratum 1 appliance 1PPS input**

**Stratum 2 leap second**

**Stratum 3 1PPS output**

1PPS signal

# Validating improvements. Dedicated testing device



**Stratum 0 satellite**

**Stratum 1 appliance 1PPS input**

**Stratum 2 leap second**

**Stratum 3 NTP testing device**

GPS signal

# Validating improvements. Dedicated testing device
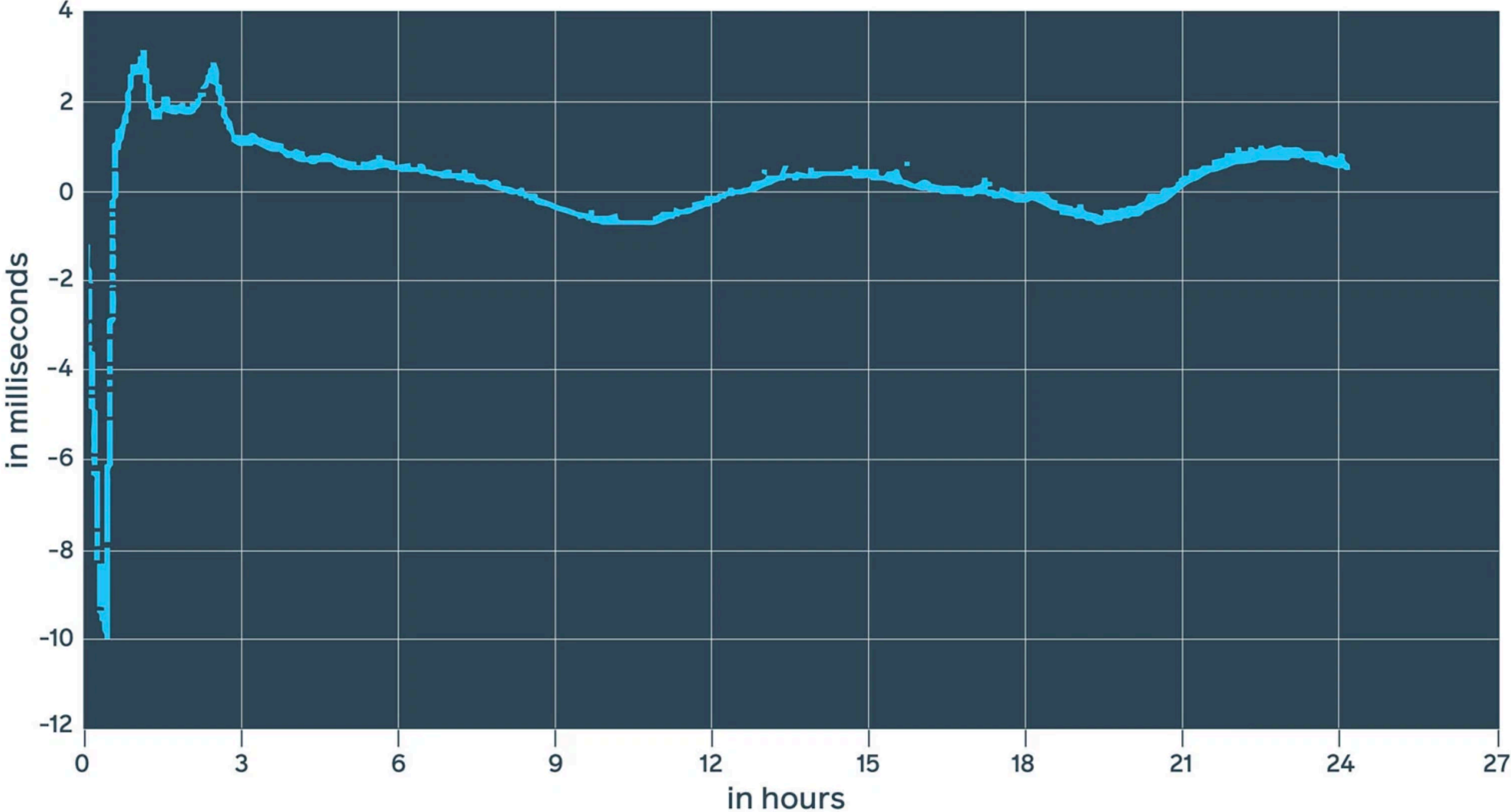
# Validating improvements. Dedicated testing device

# Improving precision

```
1 [user@client ~]# chronyc sources -v
2 210 Number of sources = 19
3   .-- Source mode '^' = server, '=' = peer, '#' = local clock.
4   / .- Source state '*' = current synced, '+' = combined , '-' = not combined,
5 | /   '?' = unreachable, 'x' = time may be in error, '~' = time too variable.
6 || .- xxxx [yyyy] +/- zzzz
7 || Reachability register (octal) -. | xxxx = adjusted offset,
8 || Log2(Polling interval) --. | | yyyy = measured offset,
9 || \ | | zzzz = estimated error.
10 || | | \
11 MS Name/IP address   Stratum Poll Reach LastRx Last sample
12 ===
13 ^+ server1   2   6   377   2   +10us[+10us] +/-   481us
14 ^- server2   2   6   377   0   -4804us[-4804us] +/-   77ms
15 ^* server3   2   6   377   59   -42us[-46us] +/-   312us
16 ^+ server4   3   6   377   60   +11ns[-3913ns] +/-   193us
```
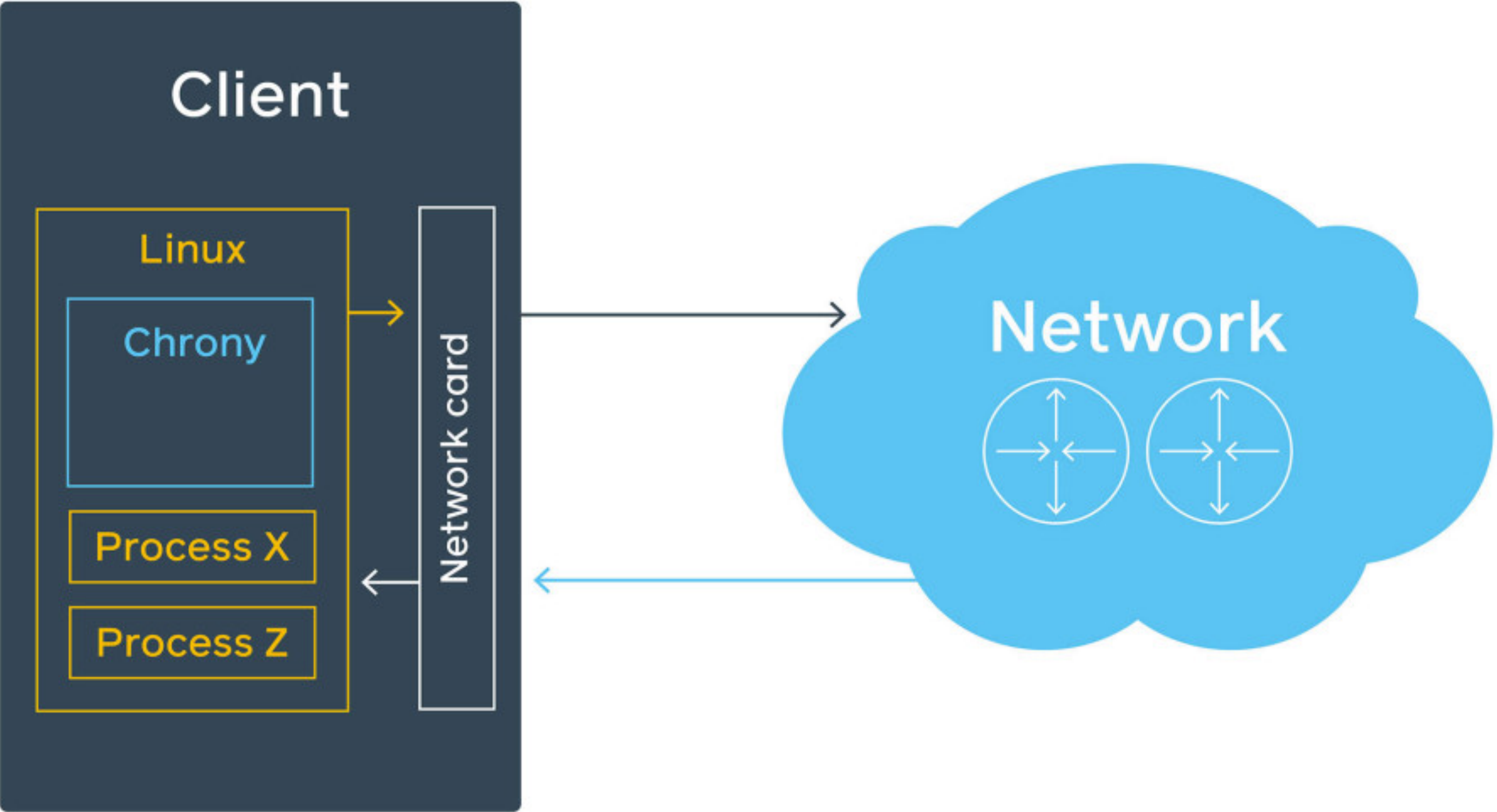
# Validating improvements. Timestamps
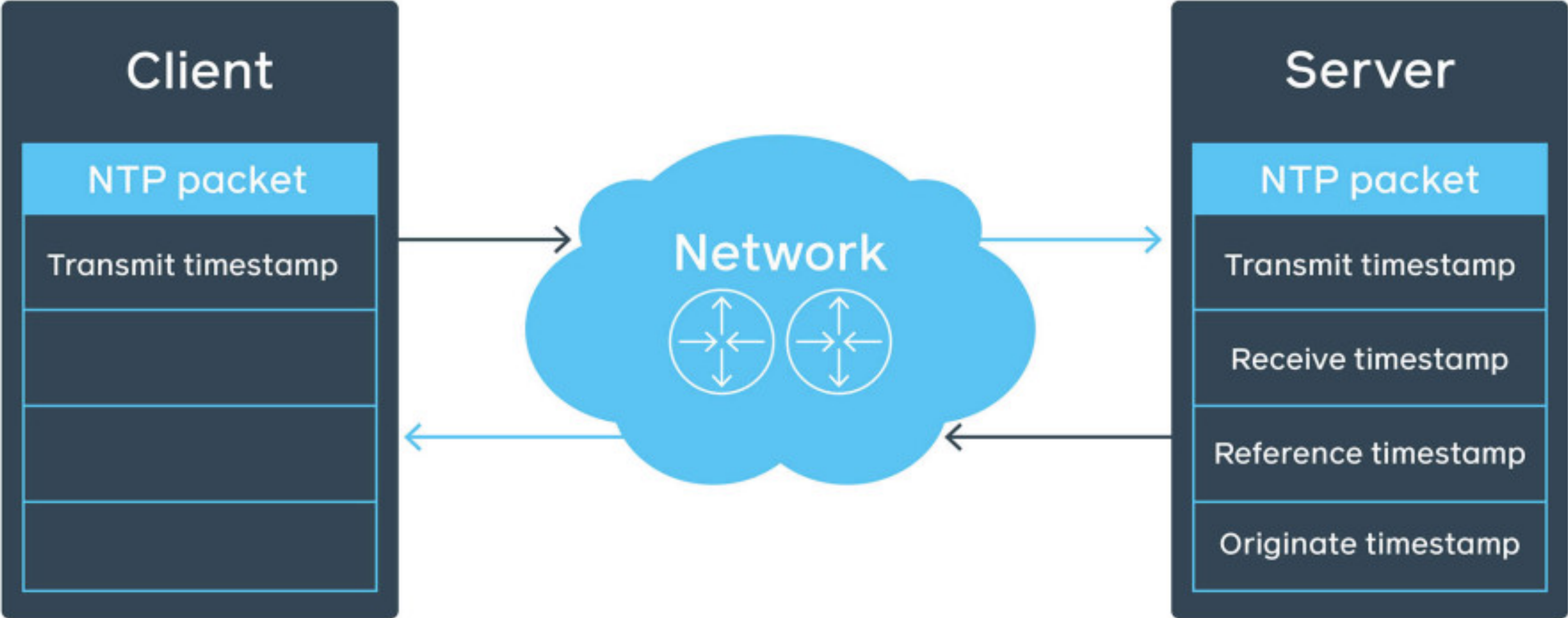
# Validating improvements. Timestamps

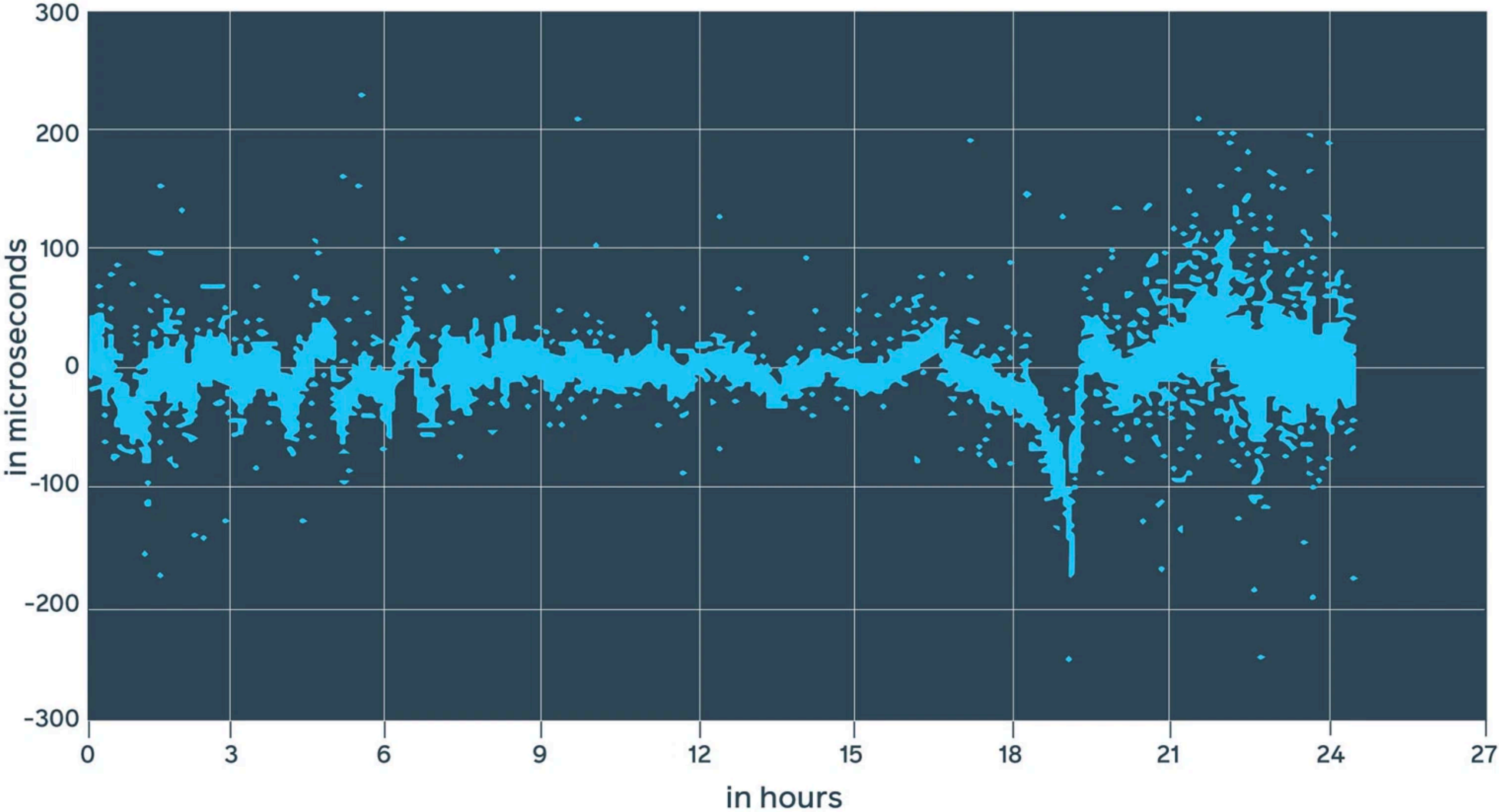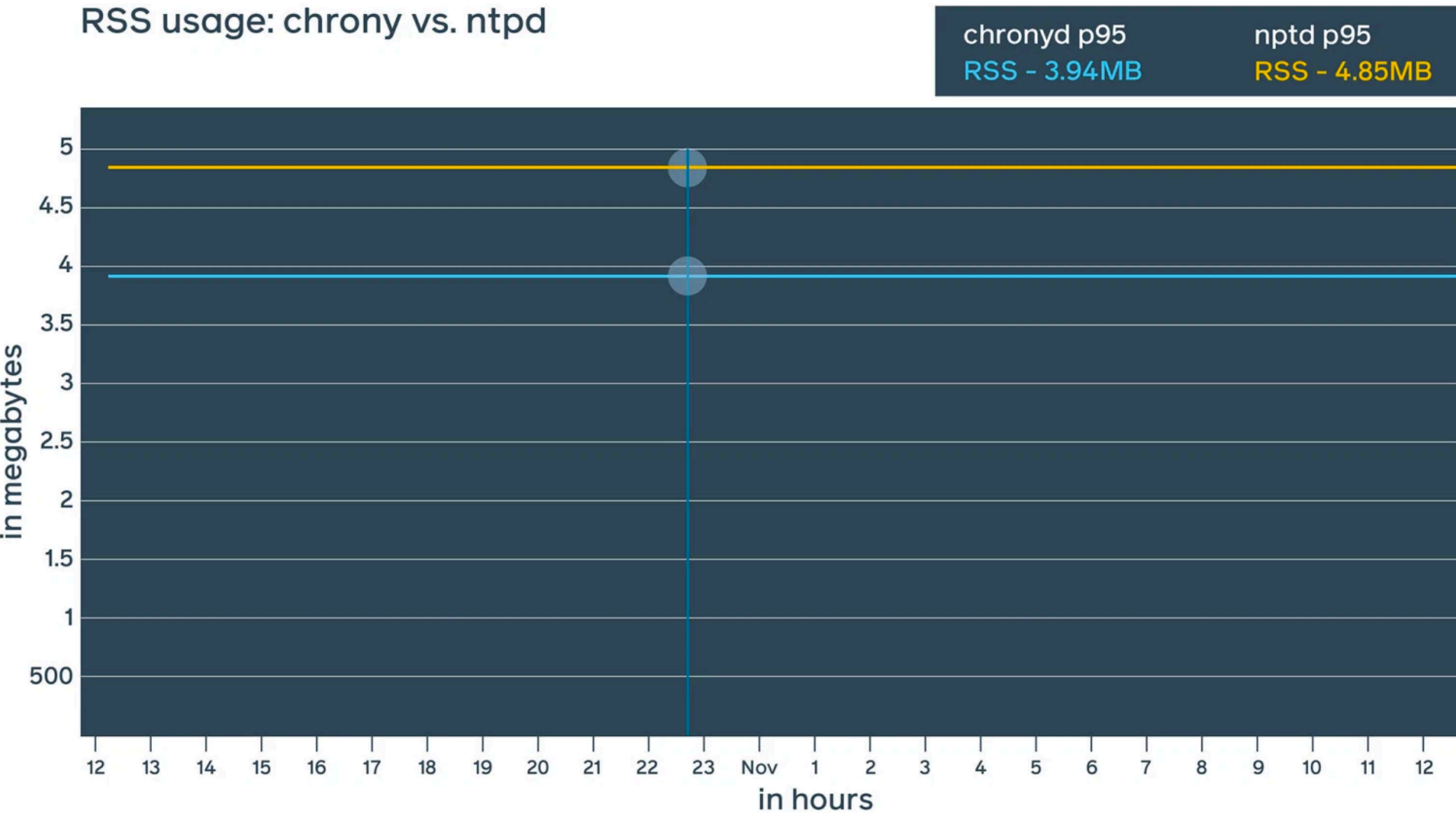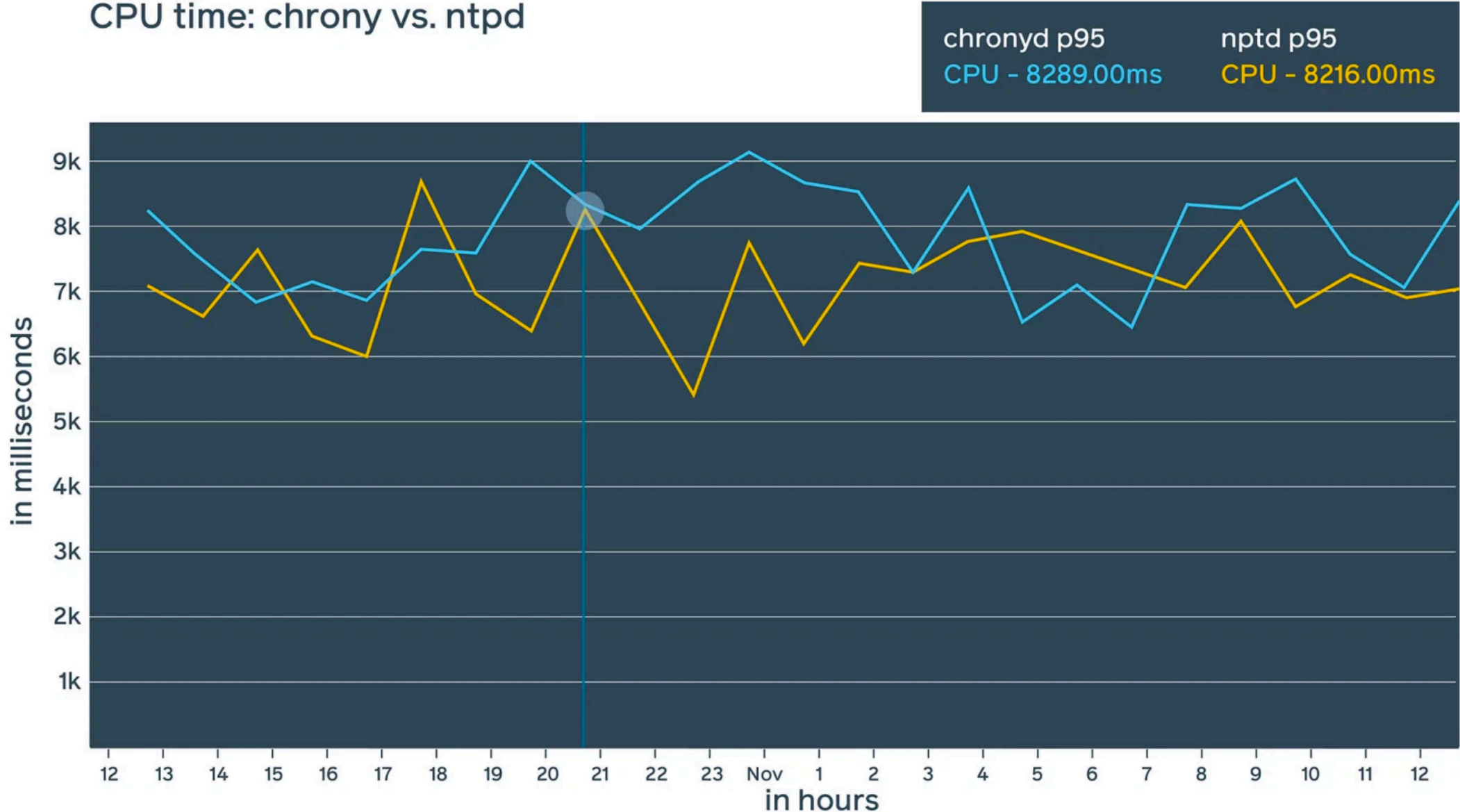# Validating improvements. Timestamps

```
1   ethtool -T eth0
2   Time stamping parameters for eth0:
3   Capabilities:
4       hardware-transmit       (SOF_TIMESTAMPING_TX_HARDWARE)
5       hardware-receive        (SOF_TIMESTAMPING_RX_HARDWARE)
6       hardware-raw-clock      (SOF_TIMESTAMPING_RAW_HARDWARE)
```

Validating improvements. Dedicated testing device
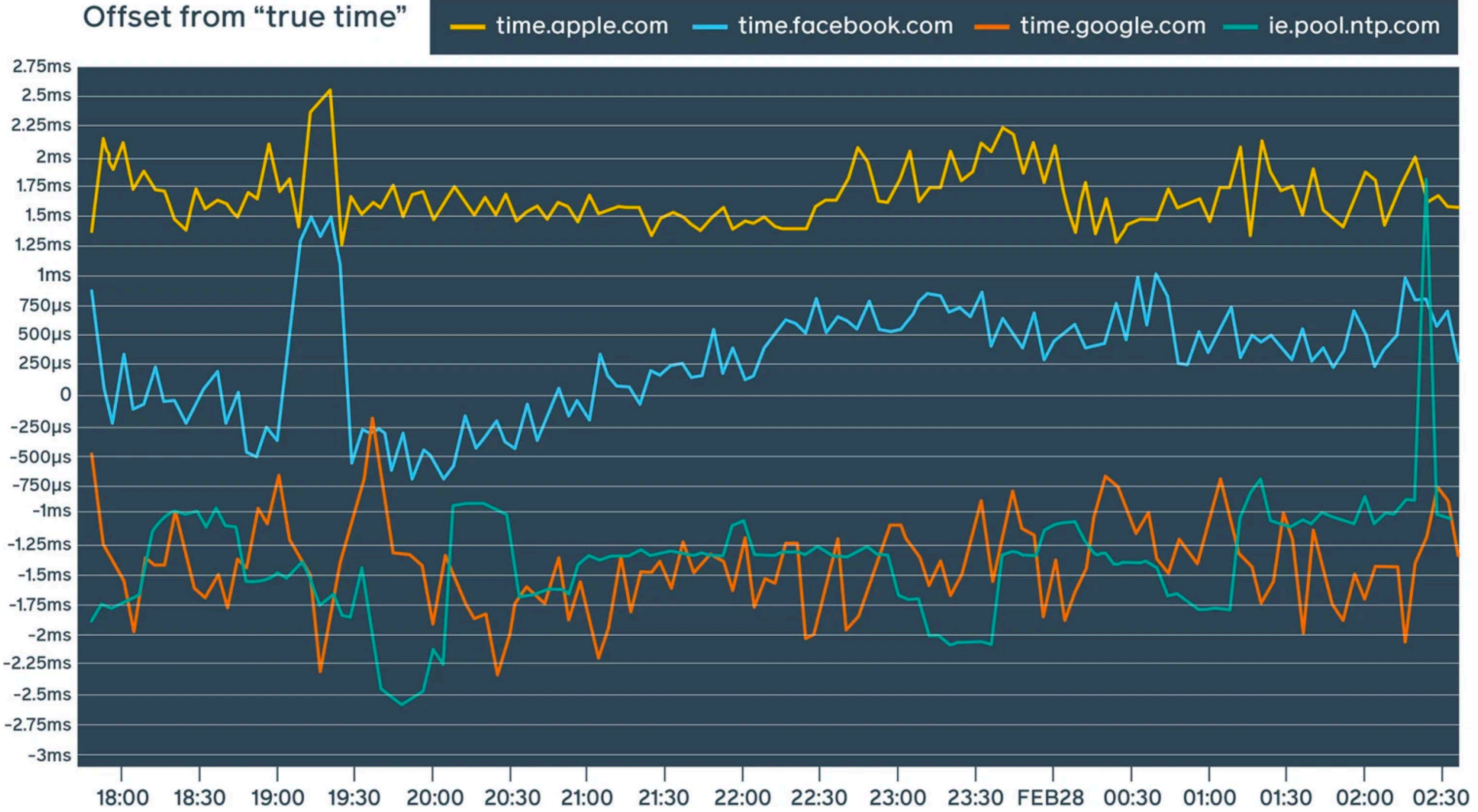
Validating improvements

CPU time: chrony vs. ntpd

chronyd p95	nptd p95
CPU – 8289.00ms	CPU – 8216.00ms

y-axis: in milliseconds (1k–9k); x-axis: in hours (12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, Nov, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)

RSS usage: chrony vs. ntpd

chronyd p95	nptd p95
RSS – 3.94MB	RSS – 4.85MB

y-axis: in megabytes (500–5); x-axis: in hours (12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, Nov, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)

Public service

Public service

Public service

Offset from "true time"

Legend: time.apple.com · time.facebook.com · time.google.com · ie.pool.ntp.com

https://engineering.fb.com/production-engineering/ntp-service/
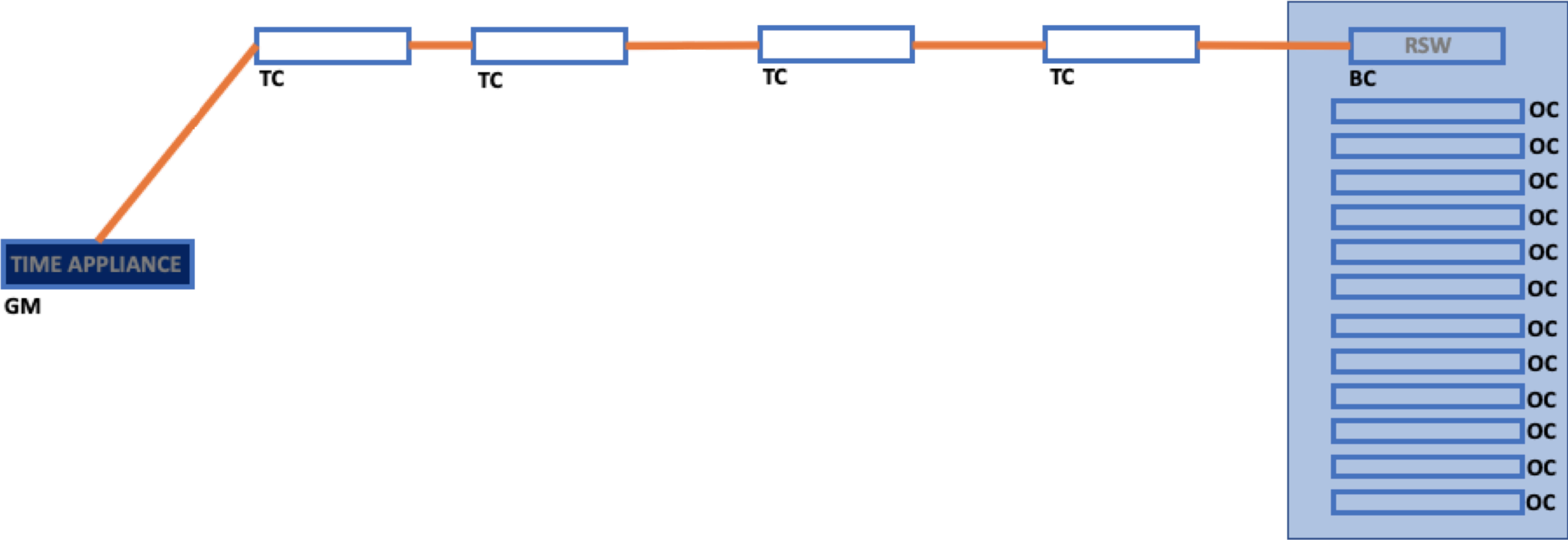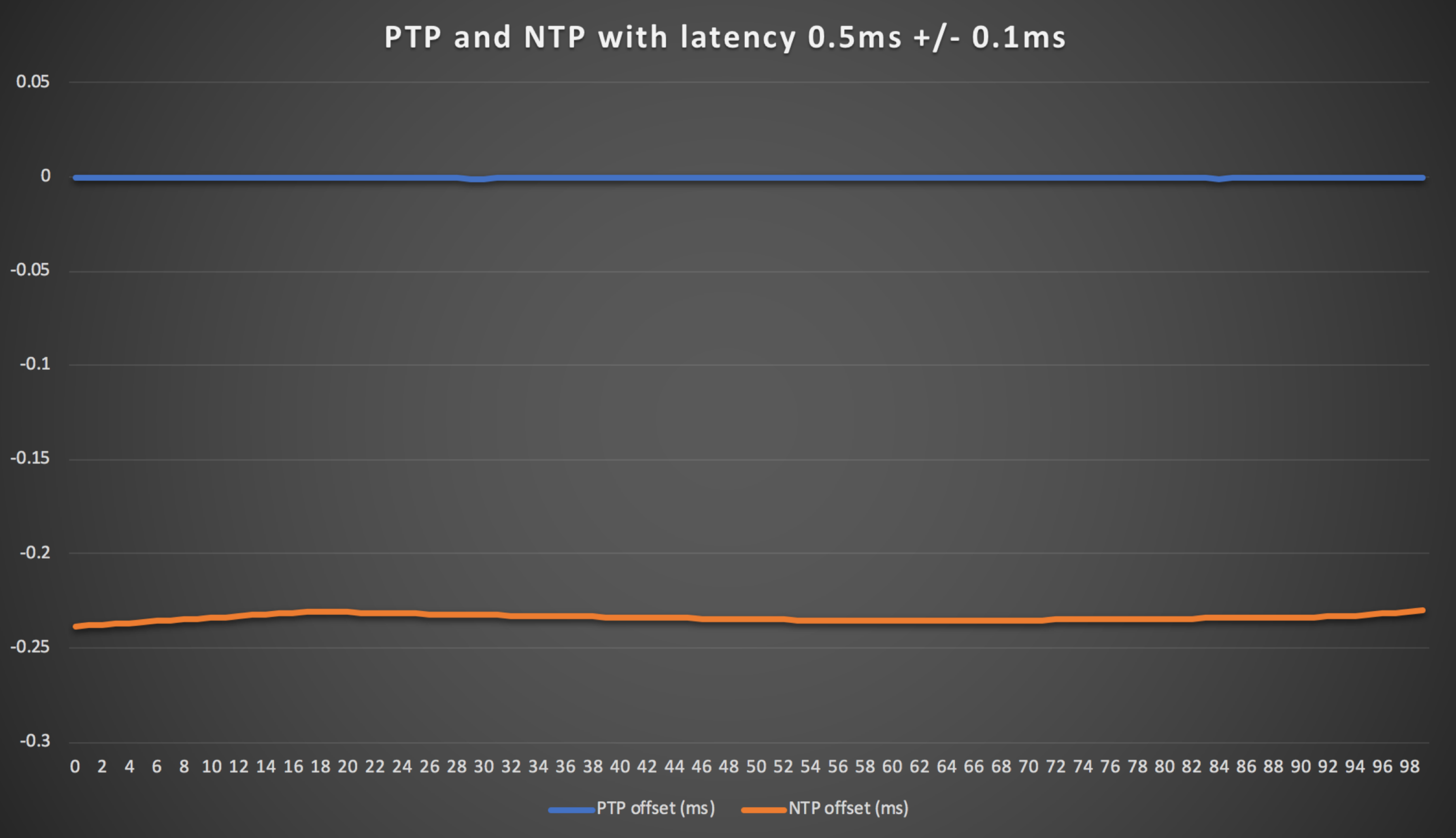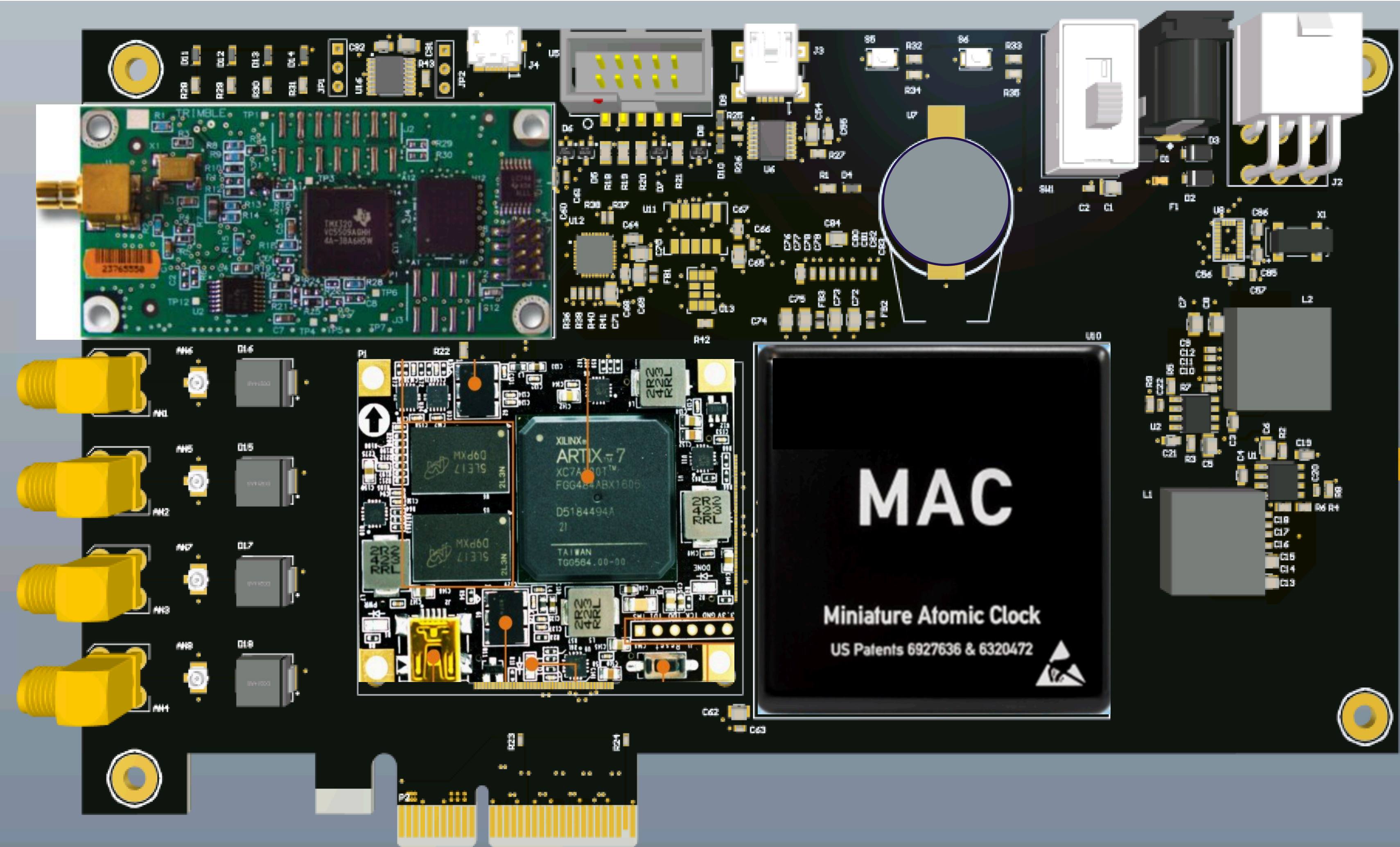
https://github.com/facebookincubator/ntp

What's next

PTP

PTP

Open Compute Time Appliance Project

Thank you