CASE STUDY I FACEBOOK

Facebook produces some of the most popular websites and apps in the world. Their technologies empower over a billion people around the globe and a crucial enabler for this is their network of datacenters. While maintaining accurate timing across any network is paramount, Facebook's challenge was to synchronize all their servers across many datacenters with sub-millisecond precision.

Maintaining accurate time within Facebook datacenters

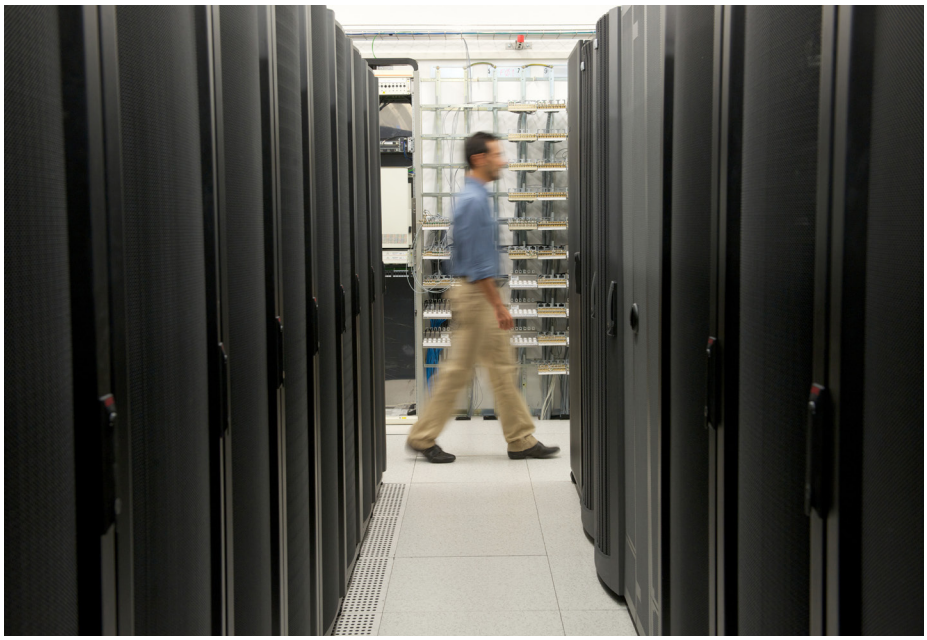
As Facebook's infrastructure has grown, time precision in their systems has become increasingly important. Not only do they need to know the time difference between two random servers in a datacenter with a high level of accuracy, they also have technical requirements for better precision. For example, multi-master databases translate microseconds and even nanoseconds of precision directly into their theoretical throughput. Logging, too, requires precision timing; in order to match logs between nodes of a distributed system, millisecond precision is often required.

To achieve the necessary timing accuracy, Facebook tested Chrony. Chrony is a different implementation of the network time protocol daemon (NTPd) and is able to synchronize system clocks faster – and with better accuracy – than NPTd.

While Chrony promised better accuracy, testing its performance proved to be problematic. Both NTPd and Chrony systems provide estimates as to their timing accuracy. These estimates (and, consequently, inherent inaccuracies) meant that Facebook's engineering team could only be confident that the timing accuracy was within the ±10ms expected for NTPd. In other words, they couldn't properly determine the improvements in accuracy from using Chrony.

Facebook developed a number of different ways to measure the actual time error in their datacenters. Measurement solutions, though, proved impractical to implement at scale, required 1pps outputs that were not always available from devices, and it was often necessary to run coax cables to various locations to make spot checks which again proved impractical. Clearly, another measurement solution was required.

Ideally, Facebook wanted a test solution with a GNSS receiver and a built-in stable atomic clock. They also wanted it to have multiple interfaces (1pps, network etc.) which would allow their engineering team to perform checks throughout the datacenter from a central test instrument using the NTP packet stream. In addition, Facebook wanted to monitor their Time Appliances and have the flexibility to probe at other points throughout the network. This required the ability to remotely operate the test equipment as well as access the live data 24/7 in order to be aware of any QoS trends or issues before they affected system performance. The test solution also had to support IPV6, have vital security features, and as it was to be deployed at scale, had to be a minimum effort solution with limited configuration required.



Facebook's work on synchronizing system clocks using Chrony is documented in their engineering blog at https://engineering.fb.com/2020/03/18/production-engineering/ntp-service/

> *'With Sentinel we are now able to verify that we've improved our timing synchronization accuracy from 10 milliseconds to 100 microseconds.'*
> **Oleg Obleukhov, Production Engineer, Facebook**

Calnex Solution

Facebook identified the following features in Calnex's Sentinel Synchronization Tester that addressed some of their unique challenges:

- Internal Rb clock disciplined to GNSS or Stratum 1 clock
- Supports NTP and PTP
- Measures up to 4 clock signals such as 1pps
- Remote operation for continuous monitoring

Additional features required by Facebook are detailed on the page below.

Having successfully met all their technical requirements, Facebook has recently begun to rollout Sentinel into their datacenters worldwide.

Calnex Sentinel Sync Tester

The Calnex Sentinel Synchronization Tester provided Facebook with a solution that met their requirements providing verification of their data center timing accuracy and alignment. The key features of Sentinel that addressed the challenges that Facebook were facing included:

GNSS and Rubidium Clock

The full performance internal Rubidium clock in Sentinel can be disciplined to either GNSS or a Stratum 1 device. Should the GNSS signal be temporarily interrupted, the internal clock in Sentinel has excellent holdover performance and will maintain accuracy allowing monitoring of system sync performance to continue.

NTP Testing

To meet the requirements of Facebook, Calnex added IPv6 capability so that it can measure both NTP and PTP in IPv4 and IPv6 protocols. Sentinel can also measure multiple NTP packet streams simultaneously.

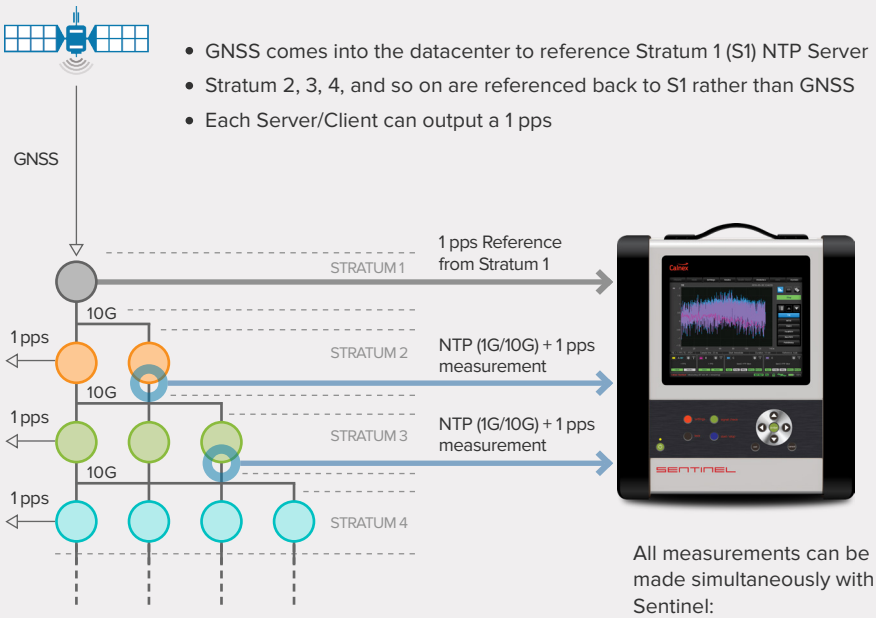
Continuous Measurement and Live Data

Sentinel was able to make measurements for up to several days before the measurement had to be restarted. However, in the Facebook environment, it was important that the measurement should always be running. Sentinel was adapted so that it would continuously acquire measurement data without having to restart the measurement. The currently running measurement data, and that from several days prior, is always available to download.

Security

From the start, Facebook emphasized the importance of security in the instrument. Calnex added additional security features to allow it to be integrated into their network. These included, but were not limited to, password protection, secure HTTP and the ability to wipe any network acquired data from the Sentinel internal memory.

Typical Datacenter Architecture



- GNSS comes into the datacenter to reference Stratum 1 (S1) NTP Server
- Stratum 2, 3, 4, and so on are referenced back to S1 rather than GNSS
- Each Server/Client can output a 1 pps

All measurements can be made simultaneously with Sentinel:

- 2 x packet ports (2 x independent Pseudo NTP Clients or PTP Subordinates)
- 4 x clock ports

API

Working at scale, Sentinel had to integrate into their network management system. Calnex added an API to Sentinel allowing full remote and secure operation of the instrument from the network management software. Through this, Facebook can upload new firmware, check and change the instrument configuration, lock and unlock the instrument screen, check the channel status, start and stop measurements and of course, download the live measurement data.

Easy Installation

Recognising that even a simple task like mounting equipment in a rack quickly becomes a big task when multiplied up by the scale of large datacenters, Calnex shipped Sentinel with a special option whereby the instrument was pre-assembled into the rack mount along with the necessary jumpers. This greatly simplifies the job of getting Sentinel installed and into Facebook's global operations.

Calnex Solutions Ltd
Oracle Campus, Linlithgow
West Lothian EH49 7LR
United Kingdom

t: +44 (0) 1506 671 416
e: info@calnexsol.com

calnexsol.com

CX9037 v1.0