napatech

# THE TIME HAS COME FOR GREATER PRECISION

> **"** One of the first challenges of understanding time-stamping and time synchronization is the need to be clear on our concept of time and what it is

# THE TIME HAS COME FOR GREATER PRECISION

Packet networks and the Internet Protocol (IP) are now firmly established as the platforms for new communication services. No longer just for "best effort" data transmission, IP has matured to encompass not just one, but a multitude of packet protocols, algorithms and methods designed to enable IP to support real-time services, such as voice, video and real-time transactions.

These developments have driven the need for more network monitoring, testing and analysis to ensure that networks can support real-time delivery without compromising the quality of service offered to other services. Accurately measuring when events occur in the network at a packet-by-packet level is therefore essential.

With 10 Gbps networks being the standard and 40 Gbps and 100 Gbps underway, the stakes have been significantly raised. High precision time measurements are now essential. At 100 Gbps, there can be as little as 6.7 nanoseconds between packets that need to be analyzed. A nanosecond (ns) is one billionth of a second.

The ability to time-stamp packets with nanosecond precision is therefore essential in understanding what is going on in the network at a packet-by-packet level. Achieving this precision, however, is far from straightforward.

In the face of this challenge, many will claim that microsecond precision is "good enough", but as this paper will show, this is a short-lived assessment. If an accurate analysis of 10 Gbps networks is to be performed, then nothing other than nanosecond precision will do.

This white paper outlines what is required to achieve nanosecond precision time-stamping. The world of precision time-stamping can be foreign to even the most experienced engineers, but this paper will strive to de-mystify some of

the concepts that lie behind time synchronization and time-stamping based on the latest knowledge and solutions available.

Napatech accelerators for network management and security applications will be used as examples throughout the white paper of a current state-of-the-art implementation of many of the concepts discussed.

**WHAT IS TIME?**

It might seem a strange question, but one of the first challenges of understanding time-stamping and time synchronization is the need to be clear on our concept of time and what it is.

For most of us, time is dictated by a clock on the wall, our watch or the time on our PC or mobile. We are aware that each of these time sources can be out of sync, but for our daily needs, minutes-accuracy is sufficient. We hardly ever ask ourselves what the "true" time is.

Time is actually based on atomic seconds. One atomic second is defined as the time it takes for the cesium frequency to oscillate 9,192,631,770 times. Cesium atomic clocks provide the most accurate timekeeping today and provide a stable time-scale known as International Atomic Time (TAI).

Coordinated Universal Time (UTC) is used as the reference for setting all public and private clocks. UTC is based on TAI, but is synchronized with solar time (known as UT1), which is derived by astronomical measurements, to take account of irregularities in the earth's rotation. Because of tidal friction and other effects, the rate of the earth's rotation is not constant. It is therefore necessary to add leap seconds to keep UTC within 0.9 seconds of UT1.

## UTC AND GPS

UTC is used as the standard time for networking and is widely transmitted using radio signals. However, one of the most common UTC time sources used in networking is Global Positioning System (GPS) satellites.

GPS consists of 24 satellites orbiting the earth. Initially used for military purposes (and incidentally still owned by the US military), GPS has also been made available for civilian use. The most familiar application is navigation systems, where triangulation is used to estimate a GPS receiver's location to within an accuracy of meters.

Triangulation uses a minimum of three satellite signals to determine the location of the GPS receiver. Key to this calculation is the measurement of the distance of each satellite from the GPS receiver. GPS satellites all have atomic clocks on board, which are extremely accurate. Each satellite sends a pseudo-random code using radio signals at exactly the same time, which is simultaneously generated by the GPS receiver. By comparing the pseudo-random codes received from each satellite to the receiver's own version, the delays measured can be used to calculate the distance from each satellite.

The pseudo-random code is in essence a time synchronization pulse signal, but for triangulation to provide an accurate calculation of the GPS receiver's location, an accurate time source is needed at all measurement points. However, it would cost up to $100k to include an atomic clock in each GPS receiver.

To circumvent this problem, GPS receivers use a fourth satellite measurement. If the GPS receiver's clock is imperfect, the fourth satellite measurement will not correspond with the triangulation result calculated earlier. This must mean that the GPS receiver's clock is out of sync and must be adjusted. By adjusting the clock, the GPS receiver can now provide a very accurate UTC time. Atomic clocks are synchronized to UTC with an accuracy of +/- 100 ns.

A GPS receiver can thus be used to provide a very accurate UTC time source. For this to work, the receiver's antenna must be positioned where it can receive signals from as many satellites as possible. This usually means on the roof of the building where the systems are to be installed.

A typical example of a GPS receiver for time synchronization is the Trimble Acutime Gold GPS Smart Antenna. This receiver can track up to 8 satellites at a time; the more satellites tracked, the better the accuracy. The antenna uses the satellite signals to calculate its location and adjust its clock. It then generates a Pulse Per Second (PPS) signal, which is synchronized to within 15 ns of GPS/UTC time. A UTC time tag is generated for each pulse. With the PPS signal and the UTC time tag, we have an accurate time reading and a method for synchronizing our clocks to UTC.

## CDMA

One of the disadvantages of using GPS antennae is that in order to receive a strong enough satellite signal, a GPS antenna needs to be deployed on the roof of the building hosting the network. This is not always practical.

One alternative approach is to use a CDMA receiver. CDMA stands for Code Division Multiple Access and is a mobile/cellular telephone technology widely deployed in North America, South America and parts of Asia. One of the distinguishing features of CDMA compared to other mobile telephone technologies is that it uses a GPS receiver at each base station to provide accurate timing for the mobile network. The timing of the network must be synchronized to within 10 microseconds.

The base station broadcasts this time signal derived from GPS to each CDMA telephone or receiver. The advantage of CDMA is that the broadcast signal is based on a much lower frequency, which means stronger signals, which can be received inside of buildings. Thus, CDMA antennae can be installed close to the systems that require time synchronization.

However, CDMA receivers are less accurate than GPS receivers due to the fact that distance information is not available (i.e. it is not possible to determine how far the received signal has traveled and hence the delay induced). If the receiver is located at the border between two CDMA cells, there is also a risk of sporadic cell handover, which in turn may result in a jump in time because of the new distance to the base station. Yet, according to CDMA receiver vendors, such as EndRun Technologies, CDMA receivers still provide an accuracy of ca. 10 microseconds relative to UTC.

## ACCURACY, PRECISION AND RESOLUTION

Before diving deeper into timekeeping we should define the basic terms for specifying the performance of time synchronization and time-stamping. This enables us to be specific and compare capabilities of different technologies and products.

Accuracy: For time-stamping: The average offset between a time stamp and the time of the event measured relative to the time difference. For time synchronization: The average offset (skew) between the synchronized clock and the reference clock.

Precision: For time-stamping: The variation of the offset between a time stamp and the time of the event measured relative to the time difference, compared with the average of all offsets. For time synchronization: The variation of offset (skew) between the synchronized clock and the reference clock, compared with the average of all offsets.

Resolution: The smallest interval between readout updates, determined by the clock sampling frequency.

A high precision value itself does not tell you how well a system performs. The resolution should be smaller or equal to the precision. Otherwise the precision specified may give the impression of a fine grained measurement, while in the real world the time readout is incrementing in larger steps determined by the resolution.

The accuracy is usually measured as a mean and a standard deviation (STD) over a given time period and number of data points. The mean time skew represents a fixed offset, which can be compensated through configuration of the equipment involved or in application software. Assuming a normal distribution of the measured data points, 68.2 percent of the measured data points are within +/- the specified STD. 3-sigma is another commonly used parameter for specifying time synchronization accuracy, and 97.7 percent of the data points are within +/- the specified 3-sigma value.
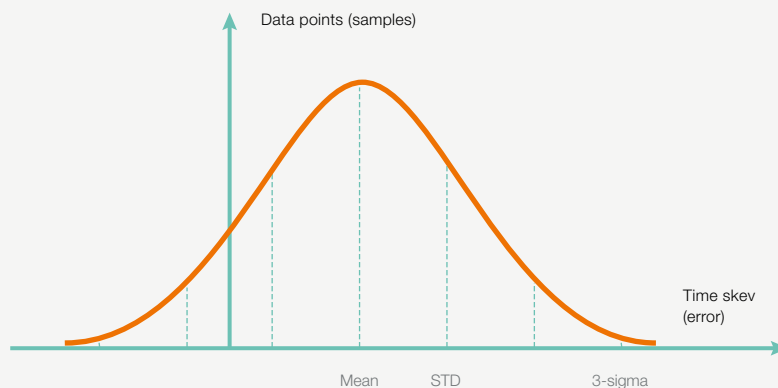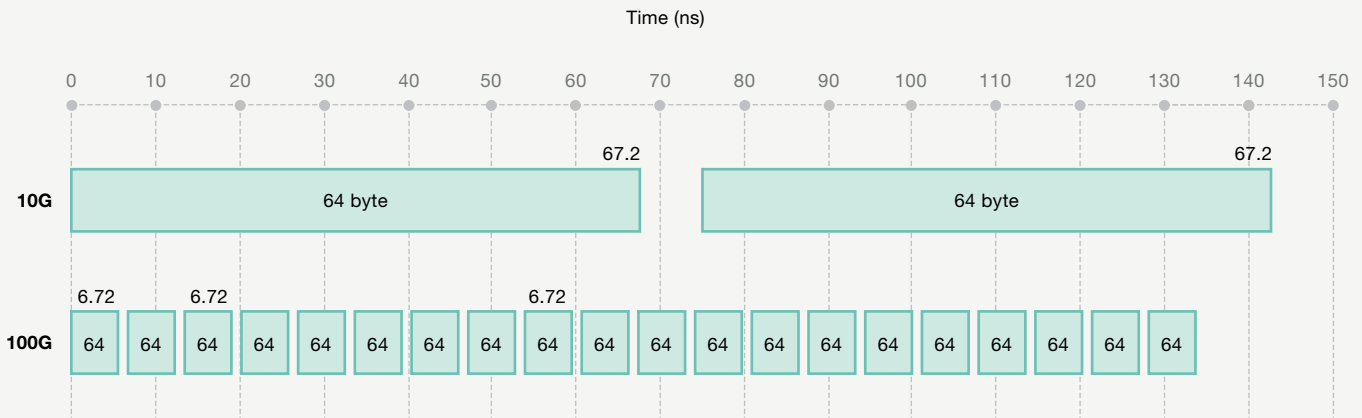
## PACKET TIME-STAMPING

The significance of time stamp precision depends on the measurements applied. For absolute network latency measurements, packet time stamps collected from monitoring equipment at different locations are compared. The precision of such measurements depend on the accuracy of UTC absolute time maintained by the local time stamp clock. For jitter measurements accurate clock rate is the key to high precision, whereas the absolute UTC time is irrelevant.

## TIME STAMP FORMATS

Different time stamp formats are used depending on the application or operating system involved. Many UNIX, Linux and FreeBSD applications refer to the UNIX time format, since this is the built-in time format in these operating systems. Windows applications or other applications over NDIS may use the NDIS time format. The PCAP time stamp format is used with applications such as SNORT, tcpdump and Wireshark.

**FIGURE 1**
Mean, standard deviation and 3-sigma

**FIGURE 2**
Frame size compared to 10 ns clock resolution

• The UNIX format is based on a 10 ns unit, with a
64-bit unit counter. The start time is January 1st 1970.
• The NDIS format is based on a 100 ns unit, with a
64-bit unit counter. The start time is January 1st 1601.
• The PCAP-μs format is based on a 1000 ns (1 μs) unit,
with a 32-bit unit counter in the most significant bits of
the time stamp, and a 32-bit second counter in the least
significant bits of the time stamp. The start time is
January 1st 1970.

Some specialized network adapters support configurable
time stamp formats. For example, Napatech accelerators
can be configured to time stamp in six different formats.
Hardware generated time stamp formats offload the
application software from time format conversion.

**SOFTWARE BASED TIME-STAMPING**
The importance of packet time-stamping in high-speed
networks is driving more vendors to include this functionality
in their products, but there are still many situations where
hardware packet time-stamping is not available thus requiring
time-stamping to be performed in software. The server time
resolution is in best case 100 ns, but due to unpredictable
and variable OS latency the accuracy of this type of time-
stamping is low and typically in the milliseconds range.
For a few applications this is sufficient, but for latency and
jitter measurements in for example financial trading systems,
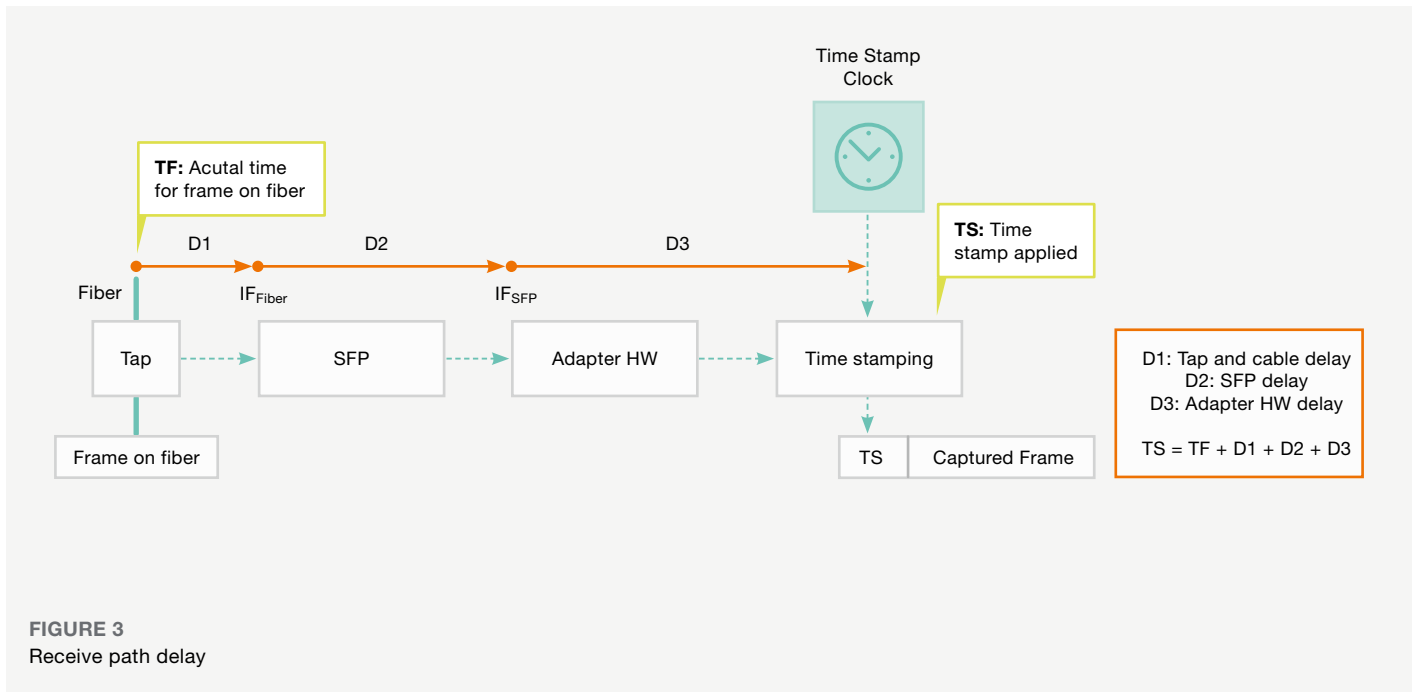nanosecond accuracy is required.

**HARDWARE BASED TIME-STAMPING**
To achieve high precision, time-stamping must be applied as
close to the wire as possible and based on a high resolution
clock. Packet arrival time is typically stored as a time stamp
in an extended header pre-pended to the raw packet
received from the network. The extended packet is passed
on to the host for processing.

Packet arrival time may be measured at the beginning or the
end of the packet received. The preferred method depends
on the purpose of the measurements. It is always possible
to calculate the alternative time stamp from the size of the
packet and the line speed. For correct correlation of packet
time stamps, the same configuration must be used for all
equipment in a system.

Time-stamping at the start of frame reception is most
convenient when analyzing packets at the link level. Time-
stamping at the end of frame reception is useful when
analyzing data from multiple links at the application level.

With this in mind, a monitoring system should process the
packets in the same order as seen from a real application
running on a host connected to the network. In a real system,
data is delivered to the application after reception of the
complete packet. Time stamps at the end of packets
therefore ensure that packets received on different ports
can be sorted based on time stamps and delivered in correct
order to the application.

**FIGURE 3**
Receive path delay

The resolution of the time stamp clock determines how accurate we can measure packet arrival, latency and jitter. At a minimum, it should be possible to identify each packet on a link by a unique time stamp. This requires that the time stamp resolution must be less than the time it takes to transmit a minimum-size Ethernet frame of 64 bytes.

Including a preamble of 8 bytes and the minimum inter-frame gap of 12 bytes, a frame can be transmitted every 672 bit times. Transmission of the minimum-size packet takes 67.2 ns on a 10 Gbps link, 16.8 ns at 40 Gbps and 6.72 ns at 100 Gbps. For 10 Gbps a resolution of 10 ns is sufficient, but for 100 Gbps an even better resolution is needed to distinguish packets. This is illustrated in Figure 2.

As can be seen for 10 Gbps, the time-stamping precision needs to be lower than 60 ns to ensure a unique time-stamp for each packet. A 1 microsecond precision, on the other hand, would only be useful for applications that needed to analyze packets of 1280 bytes and above. However, with IP networks now being used for voice and other time-sensitive applications, the use of smaller packet sizes in transmission is more common requiring network monitoring and analysis to not only support lower byte sizes than 1280, but all packet sizes, thus requiring nanosecond precision time-stamping.

Specialized network adapters, like Napatech accelerators, provide on-the-fly time-stamping at the MAC level with a precision of 10 ns and resolution of 4 ns. A local time stamp clock is used as the reference for time-stamping. At this level of precision, every nanosecond counts. So what are the components contributing to the accuracy of time-stamping?

**RECEIVE PATH DELAY**
We need to look at the path from the physical wire to the location in the monitoring network adapter where time stamps are applied.

Figure 3 illustrates the receive path and the related delay components. The network adapter is typically attached to the network links via an optical tap and cables connected to SFP+ modules on the network adapter. From the SFP+ host interface the monitored frames are transferred via adapter hardware to the time-stamping logic. All the components on the way from the wire to the time-stamping logic introduce a delay to the time stamp compared to the actual time of the packet on the wire. The time-stamping logic may be located in the PHY or later in the processing pipe. Time-stamping close to the wire, for instance in the PHY, reduces the total delay. The stability of the reference clock also contributes to the accuracy of the time stamp.

Each of these components contributes with fixed and variable delay. The variable delay is measured as standard deviation (STD). The receive delay depends highly on the hardware and SFP models used.

The fixed delay can be compensated, whereas the variable delay represents an error which determines the accuracy of the time stamp function. The fixed delay is usually a significant order of magnitude higher than the variable delay, which makes it possible to configure compensation for the fixed delay contribution and maintain precision. For example, the D1 component contributed by the Tap and Cable per meter typically contributes a 5 ns fixed delay with a standard deviation of 0.1 ns.

The components discussed so far relate to the precision of a standalone system, where measurements on a single link or links on the same network adapter relative to a local clock are sufficient. If correlated measurements between different chassis or locations are needed, then the systems involved must be synchronized, and the synchronization introduces another contribution to the overall accuracy of time-stamping. We will look at the overall accuracy of complete systems later on.

It's important to note that when measuring relative timing between packets on a single link, the fixed delay does not influence the results. Considering fixed delay contributions

is only relevant when measuring absolute timing of packets or when correlating packets between links attached to different synchronized systems.

## MEASURING TIME STAMP ACCURACY
How do we measure the accuracy of time-stamping? When operating in the nanosecond precision range, verification of the actual accuracy is not trivial. We will describe two different measurements.

• Relative time stamp accuracy: How accurate can time between frames be measured?

• Absolute time stamp accuracy: How close is the recorded time stamp to the actual time of the frame on the wire?

Figure 4 shows a setup for relative time stamp accuracy measurement. A traffic generator with a high precision and stable transmit clock is used to generate a frame sequence at a well-defined rate and with a specific inter-frame gap. The delta time between time stamps in captured frames should correspond to the rate of generated traffic. With appropriate traffic load this test setup generates millions of data points in a few minutes. The data is analyzed to calculate and report the accuracy of relative time-stamping.

**FIGURE 4**
Test setup for relative time stamping accuracy



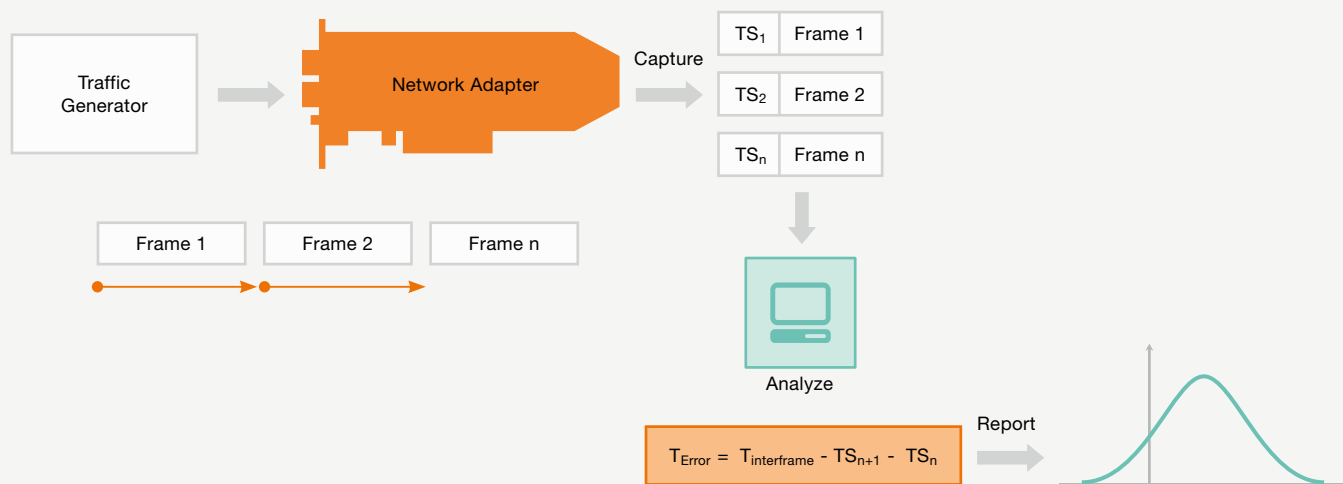$$T_{Error} = T_{interframe} - TS_{n+1} - TS_{n}$$

Figure 5 shows a setup for absolute time stamp accuracy measurement. A time server with high precision and a stable clock provides the reference clock for the test setup. Time server PPS outputs are connected to the traffic generator, the oscilloscope and the network adapter under test. The network adapter is synchronized to the PPS pulse.

A traffic generator with high precision and a stable transmit clock is used to generate a single packet once a second. The traffic generator is configured to transmit a frame shortly after every PPS pulse, but the transmit timing of this frame is not critical. The oscilloscope is connected to the SFP host connector electrical interface IFSFP on the network adapter. The oscilloscope measures the delay TPPS_SFP from when the PPS pulse is sent until the frame is available on IFSFP.

With reference to Figure 3, assuming that D1 and D2 are known and fixed values, we can calculate the time of the frame on the fiber relative to the last second tick:

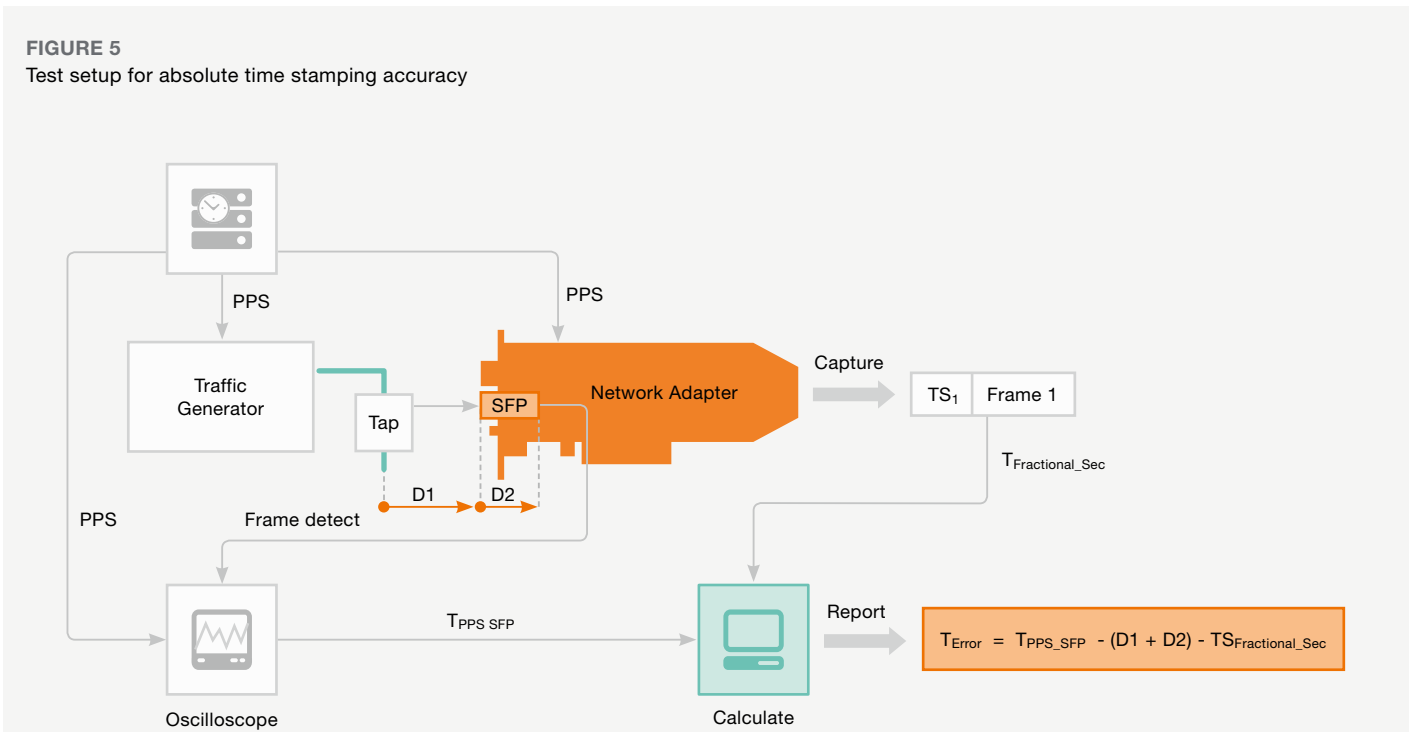$$T_{Relative\_Fiber} = T_{PPS\_SFP} - (D1 + D2)$$

Seconds and fractional seconds can be retrieved from the time stamp. The fractional seconds in the time stamp represents the offset from the second tick as recorded by the network adapter. Since the network adapter time stamp clock is synchronized to the PPS and the second tick, we can compare the measured relative time offset TRelative_Fiber with the fractional seconds TSFractional_Sec in the time stamp of the captured frame. The difference between the two represents the error, which determines the accuracy of the time stamp:

$$T_{Error} = T_{Relative\_Fiber} - TS_{Fractional\_Sec}$$
$$= T_{PPS\_SFP} - (D1 + D2) - TS_{Fractional\_Sec}$$

If the network adapter does not have support for receive path delay compensation, then TError includes a fixed delay contribution from D1, D2 and D3. Cable lengths must be equal for the PPS connections to the network adapter and the oscilloscope to ensure correct measurements.

So far we have looked at measurement for a single frame. Automated testing over multiple frames requires correlation between measurements sampled by the oscilloscope and frames captured by the adapter. This can be achieved if the oscilloscope and the network adapter are synchronized to time of day provided by the time server.

**FIGURE 5**
Test setup for absolute time stamping accuracy

## CLOCK SYNCHRONIZATION AND TIME DISTRIBUTION

Correlated measurements between different chassis or locations require synchronization of clocks. For some use cases the clocks must be synchronized to absolute time, while other use cases only need frequency synchronization.

At this point it is useful to define some terms used for different aspects of clock synchronization.

**Time of Day:** Time of Day (TOD) represents the year, month, hour, minute and second.

**Phase:** Phase represents the exact time when the second increments – the second tick. A PPS signal is typically used to indicate the phase.

**Frequency:** The frequency represents the speed of the clock. This can also be translated into clock increments per time unit.

## TIME OF DAY AND PHASE SYNCHRONIZATION

Time synchronization is based on comparing the clock being synchronized (slave) with a reference clock (master) and adjusting the slave clock accordingly. The time error is widely referred to as the clock skew. If the slave clock is behind, then the slave clock rate is increased to catch up with the master. If the slave clock is ahead, then the slave clock rate is decreased to align with the master. Clock skew measurement and clock adjustment is reiterated until the slave clock is synchronized with the master.

Some specialized network adapters, like Napatech accelerators, have built-in algorithms which can do very fine clock rate adjustment, resulting in accuracy better than +/- 10ns STD compared to the reference signal (i.e. PPS).

If the reference signal is lost, the slave runs its clock based on the local oscillator. Over time the slave clock will drift away from the reference time. When the reference signal is regained, the clock skew can be several seconds, and the adapter algorithm ensures that the clocks are smoothly synchronized without any negative or positive jumps in time.

If only a phase reference, such as PPS, is available, then the slave clock is synchronized to the second tick only.

## FREQUENCY SYNCHRONIZATION

Frequency synchronization locks the slave clock rate to the master clock rate. PPS, 10MPPS (10 million pulses per second) and SyncE (Synchronous Ethernet) are examples of technologies which can be used for frequency synchronization. Frequency synchronization is sufficient in situations where only relative timing is of importance. Packet jitter measurement is one example.

## MEASURING SYNCHRONIZATION ACCURACY

Comparing PPS outputs from both master and slave clocks is a simple and commonly used method for measuring synchronization accuracy.

Figure 6 shows a test setup for measuring the accuracy of the network adapter time stamp clock, which is synchronized to an IEEE 1588 / PTP grandmaster. An oscilloscope is configured to sample and store the two PPS signals and the system runs for typically 24 hours, generating 86400 (24 x 60 x 60) data points. Based on this data we can calculate the mean and STD offset between the second ticks of the two clocks. Cable lengths must be equal for the 2 PPS outputs to ensure correct measurements.

## TIME SYNCHRONIZATION CABLE DELAY COMPENSATION

Some equipment, like Napatech accelerators, provides the ability to configure for time sync cable delay compensation. This configuration can be used when a PPS time reference is connected.

IEEE 1588 / PTP has built-in delay compensation, which eliminates the need for network delay compensation.

## TIME SYNCHRONIZATION AND DISTRIBUTION OVERVIEW

Here we will discuss commonly used technologies for time and clock synchronization and distribution. Synchronization features, accuracy, deployment and cost are parameters to consider. The table gives an overview of synchronization features and deployment for the different technologies.

| Technology | TOD | Phase | Frequency | Cabling | Topology |
|---|---|---|---|---|---|
| IEEE 1588/PTP | X | X | X | UTP | LAN/WAN |
| NTP | X | X | X | UTP | LAN/WAN |
| IRIG-B | X | X | X | Coax, UTP | Point-Point |
| PPS | | X | X | Coax | Point-Point |
| 10 MPPS | | | X | Coax | Point-Point |
| SyncE | | | X | UTP | LAN |

TOD: Time of Day

The reference in Figure 7 gives an overview of the topologies for different synchronization technologies. A time server can be synchronized directly to GPS/UTC via a built-in GPS receiver or indirectly via a built-in CDMA receiver which receives time information from a base station. The CDMA base station is synchronized to GPS/UTC. Many time servers have an internal atomic clock reference, which keeps the time when GPS or CDMA is unavailable.

### PPS / 10 MMPS SYNCHRONIZATION
Most time servers and other time equipment have at least one PPS input/output for synchronization with other devices or for measurement and calibration. The PPS signal provides

the second tick, which can be used in combination with TOD information for accurate synchronization to a reference UTC clock. The PPS signal itself can also be used for frequency synchronization.

In addition to the PPS input/output, many time servers and other time equipment also have a 10MPPS input and/or output for synchronization with other devices or for measurement and calibration. The higher frequency 10MPPS signal provides much better frequency synchronization and stability than the PPS signal. The PPS signal enables clock adjustment once a second, whereas the 10MPPS provides a much more continuous and fine-grained control of the slave clock, resulting in higher accuracy synchronization.
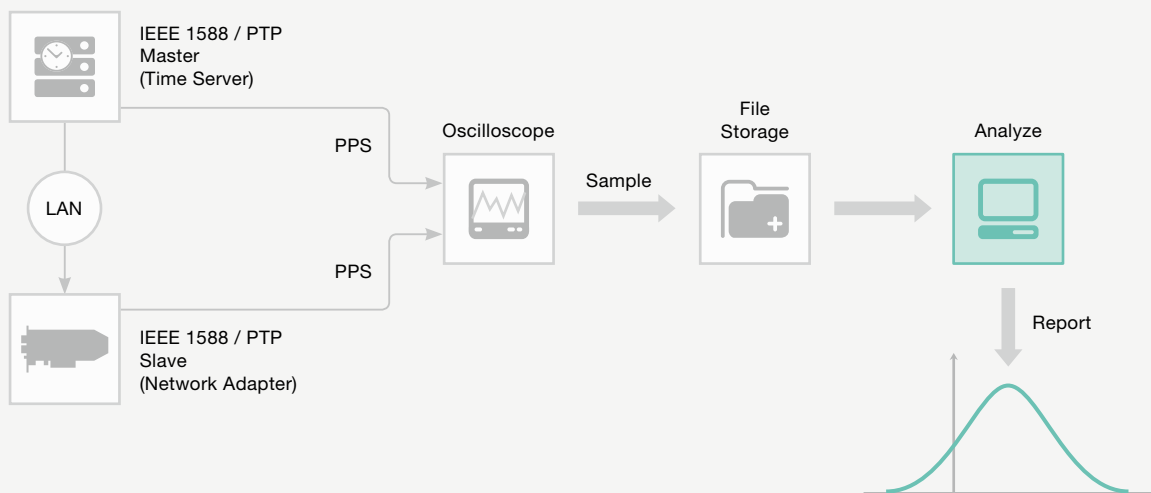
### SYNCHRONIZATION WITH PPS AND SERIAL INTERFACES
Many time servers provide a PPS output and a serial interface for time string output. The PPS signal is typically carried over a coax cable and a separate cable is used for the serial communication. Data servers or other equipment connected to these outputs can synchronize to the UTC time reference by reading the time of day from the time string output and synchronizing the second tick to the PPS signal.
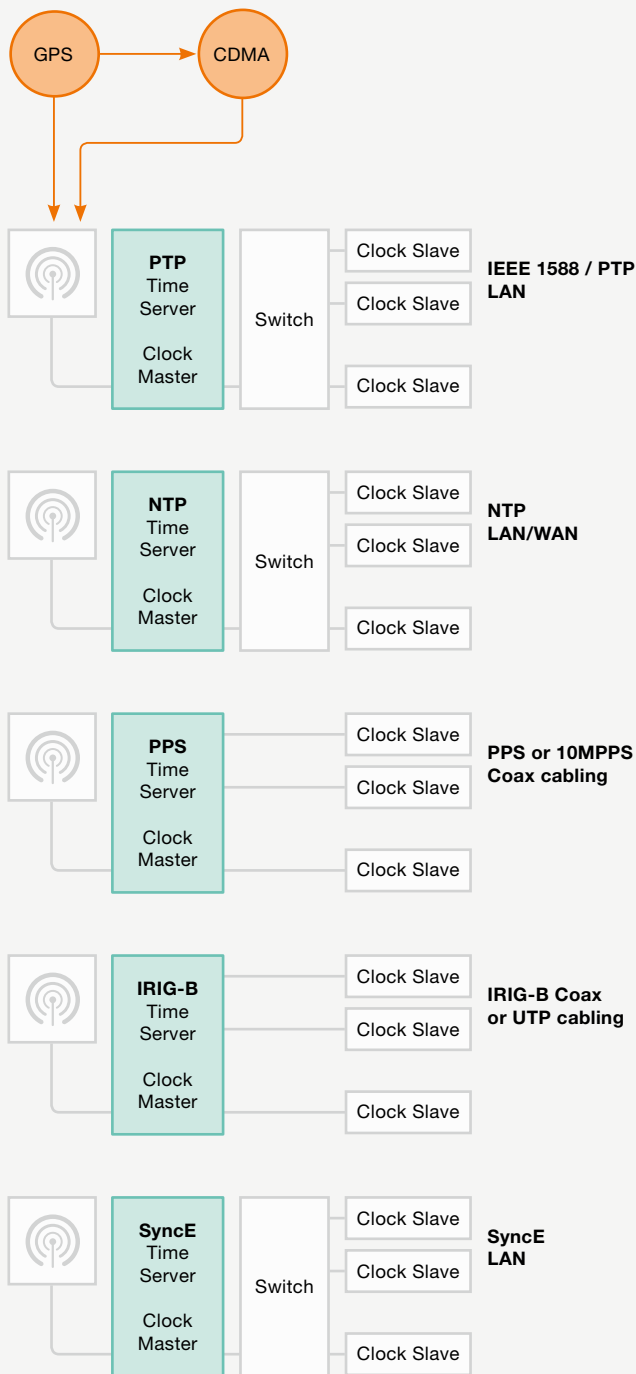
### NTP
Network Time Protocol (NTP) is probably the most widely used method for distributing time information in packet networks. It is a network protocol based on UDP that runs

**FIGURE 6**
Time synchronization test setup

**FIGURE 7**
Time distribution technology overview

on IP networks. As such, it is a software solution based on a client/server principle where an NTP time server can be polled by clients at regular intervals (typically 15 minutes) to synchronize their UTC time. A client can be any networking device running on an IP network with NTP client software installed.

In order to minimize the load on NTP time servers, a hierarchical NTP time server architecture has been defined. Each hierarchical layer is referred to as a stratum. Strata are numbered depending on how far away they are from the primary UTC clock source. A stratum-1 time source/server obtains its UTC time from GPS, CDMA or similar precise clock source. A stratum-2 server receives its UTC time from a stratum-1 server and so on.

The accuracy of the NTP time provided depends on the latency of the network and the devices between the time server and the client. In a Local Area Network (LAN), the accuracy of the UTC time received is usually between ½ and 2 milliseconds. In a Wide Area Network (WAN), the time synchronization accuracy is usually between 10 and 100 milliseconds. NTP time can also be obtained over the Internet from remote time servers, but the significant and unpredictable delays introduced in the Internet make this a very imprecise method.

NTP time-stamping is based on 4 time stamps. The NTP client time-stamps the NTP request packet before sending it to the NTP server. The NTP server also stamps the received NTP request packet. Before sending a reply, the NTP server time-stamps the NTP reply packet, which is in turn time-stamped by the NTP client upon reception. With these 4 time stamps, it is possible for the NTP client to compensate for network delays and calculate the adjustment that needs to be applied to its own clock in order to synchronize with UTC time. All time stamps are generated by software and this limits the possible delay compensation accuracy.

NTP is adequate for many applications, but for applications requiring high-precision time-stamping (such as networking monitoring and latency measurement), millisecond accuracy is not enough.

## IRIG-B TIME CODES

The Inter-Range Instrumentation Group code B (IRIG-B) time code is a variant of the IRIG time code standards, based on a master-slave hierarchy where time synchronization information is transmitted in words and frames.

IRIG-B has been broadly deployed for military and industrial uses, including electricity supply. It has also been used for time synchronization in financial trading. However, the general opinion is that IRIG-B is a legacy technology. The accuracy of IRIG-B time codes is not sufficient for high precision applications, and it has been replaced by many of the other technologies mentioned here.

## IEEE 1588 / PTP

An alternative IP network distribution mechanism is provided by IEEE 1588, commonly known as Precision Time Protocol (PTP). Standardized first by the IEEE in 2002, IEEE 1588 (which we will hereafter refer to as PTP) uses IP multicast communication (usually over an Ethernet network) to distribute time synchronization signals to network devices.

PTP uses many of the principles of NTP. It is also based on a client/server principle, except in PTP terms, servers and clients are referred to as masters and slaves respectively. As with NTP, PTP also uses the concept of strata, with a stratum-1 clock source indicating a high-precision UTC time source (for example GPS).

However, instead of a strict hierarchical architecture, PTP operates by self-selection of the best clock source. The best master clock algorithm is used to determine which of the available clock sources should be designated as the master clock. Slaves synchronize their clocks with the master clock by reading the transmit time stamp in PTP event messages combined with the received time stamps for these messages.

The master sends a Sync message with the precise time stamp (t1) of when it was sent. The slave then time-stamps the Sync message upon receipt (t2) and uses these two time stamps to calculate the actual offset from the master time. For this to be accurate, the propagation time used by the Sync message, traveling the network from the master to the slave, must be subtracted.

This propagation time is called the mean path delay and is calculated by using the Sync transmit and receive time stamps and two additional time stamps received from a Delay request/response message pair, namely the time stamp when the Delay request message is sent from the slave (t3) and the time stamp when the master receives the Delay request message (t4). This time stamp is then sent back to the slave by using a Delay response message.

With this information, the mean path delay may be calculated as follows:

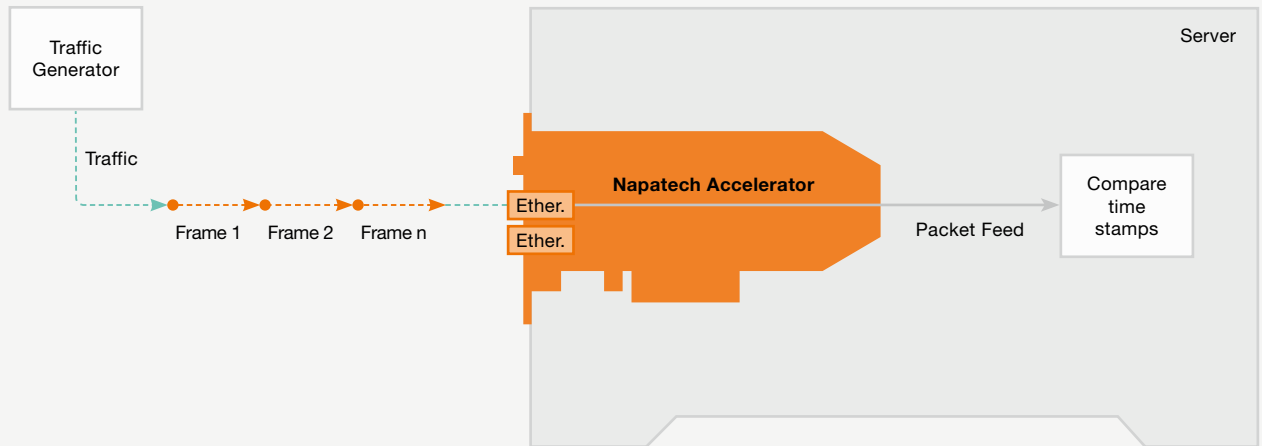$$\text{Mean path delay} = [(t2-t1)+(t4-t3)]/2$$

Then the corrected offset from master may be calculated as:

$$\text{Offset from master} = t2 - t1 - \text{mean path delay}$$

This offset is then used to adjust the slave clock to keep its time-of-day and phase synchronized.

PTP can use hardware for time synchronization, which provides greater accuracy than NTP. A typical PTP network adapter provides hardware time synchronization accuracy of 100 ns between a master and slave. In an ideal environment with two directly connected adapters, the accuracy may be better than 10 ns.

However, the overall accuracy of PTP is still dependent on the underlying network, because PTP is very sensitive to packet delay variations in the network. The reason for this is the method used to find the mean path delay in the PTP algorithm. To get a higher accuracy than, for example, NTP (better than 1 microsecond), the Sync message propagation time from master to slave, and the Delay request message propagation time from slave to master, must be constant. If there is variation in these delays, the clock synchronization accuracy will suffer accordingly.

**FIGURE 8**
Single link – Single accelerator

In PTP aware networks, all network devices, such as switches and routers, are able to recognize PTP event messages, and modify them with the measured residence time (i.e. the amount of time that a packet has spent in the device). The device then adds the measured residence time to a PTP header correction field that the IEEE 1588 v2 protocol provides for exactly that purpose. Such PTP aware network devices are also known as transparent clocks.

In PTP unaware networks, network devices are not able to recognize PTP event messages, and therefore they are not able to add packet residence time to the correction field of the PTP header. This introduces variation in the packet residence time of these devices, and that variation will increase with the network load on the devices. In such PTP unaware environments, the advantage gained by PTP's hardware-assisted time- stamping feature is lost because of the larger network variation delays.

To improve PTP performance in PTP unaware networks, a PTP filter may be applied. Such a filter will typically be capable of improving the time synchronization performance up to 60 times, depending on the actual network traffic and configuration.
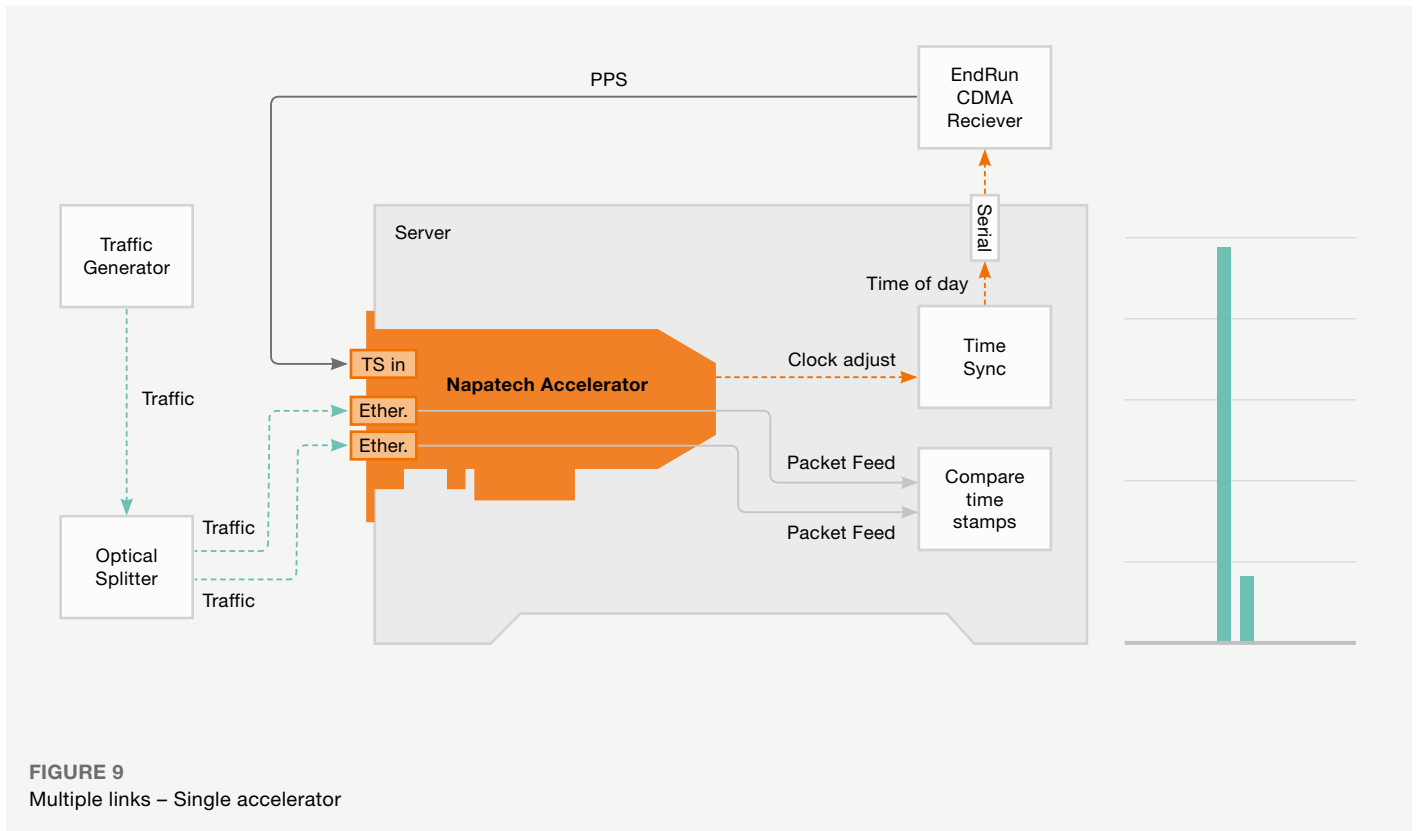
It is possible to create mixed PTP aware and PTP unaware networks, where some PTP slaves that do not require great precision communicate through PTP unaware network devices, while other PTP slaves that do require great precision communicate through PTP aware network devices. However, if just a single point between the PTP master and a PTP slave is PTP unaware, any precision will be lost.

**SYNCHRONOUS ETHERNET**
Synchronous Ethernet (SyncE) is typically used in communication provider networks, to carry timing information. Only the frequency is transported over SyncE, and it can be used as a clock reference for network monitoring devices. SyncE carries only the frequency of the master clock, but this information can be used in combination with PTP to provide accurate synchronization to frequency and TOD.

**NAPATECH TIME SYNCHRONIZATION (NT-TS)**
Napatech Time Synchronization (NT-TS) is the Napatech proprietary time synchronization protocol used for synchronization between Napatech accelerators and time synchronization equipment. NT-TS carries a PPS signal and time of day information that allows time synchronization

**FIGURE 9**
Multiple links – Single accelerator

to an accuracy of +/- 6 ns. Each Napatech accelerator is capable of receiving the NT-TS signal on external and internal connectors as well as transmitting the NT-TS signal in daisy chain fashion to other accelerators or time synchronization equipment.

### ENSURING PRECISION – USE CASES (PUTTING IT ALL TOGETHER)

So, what happens to the accuracy and precision of measurements when applying real world use cases? We will look at some typical use cases.

### SINGLE LINK – SINGLE ACCELERATOR

Let us first look at Figure 8, a simple test setup where we measure the relative time stamp accuracy on a single port. A traffic generator with a high precision and stable transmit clock is used to generate a frame sequence at a well-defined rate and with a specific inter-frame gap. The delta time between time stamps in captured frames should correspond to the rate of generated traffic.
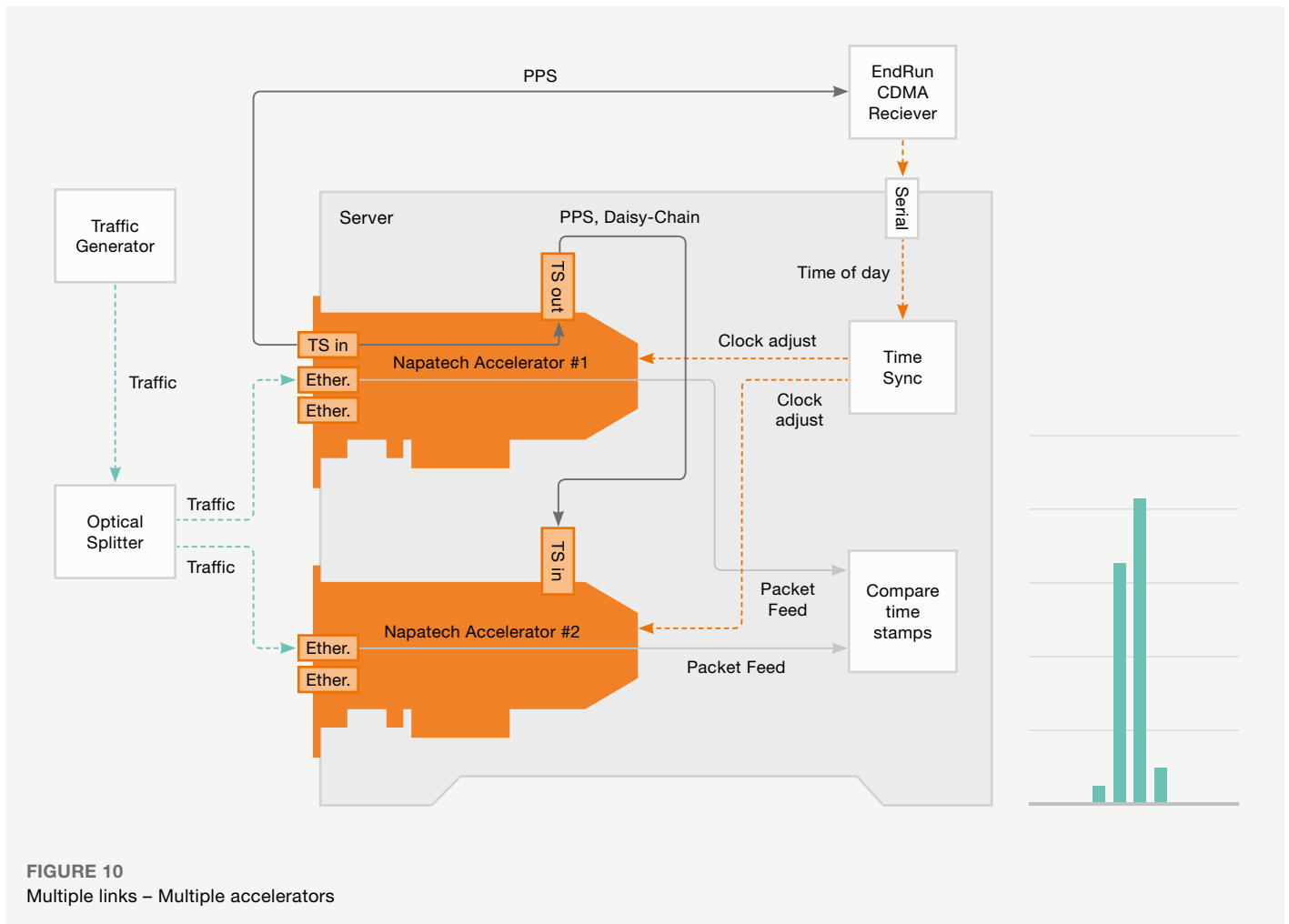
Results are presented as histograms with the emphasis on the standard deviation rather than absolute figures. This is

because the absolute figures will vary from accelerator model to accelerator model due to the influence of different receive path delays and other factors. Many of these are fixed in nature, and it is possible to compensate for them as described earlier in this paper. What is interesting is to understand the variance in measurements, as this cannot be compensated, and will ultimately determine how accurate your timekeeping will be.

### MULTIPLE LINK MEASUREMENTS AT ONE LOCATION – SINGLE ACCELERATOR

Let us now look at a system with two monitoring ports on a single accelerator.

The purpose of this example is to demonstrate the relative accuracy between time stamps for frames captured on two ports on the same accelerator. In this case frames received on the two ports are time stamped using the same local reference clock. Clock synchronization accuracy does not influence the accuracy of the relative measurements. Packets received on ports on the same accelerator are handled in parallel and time stamped at line rate.

**FIGURE 10**
Multiple links – Multiple accelerators

When comparing packet time stamps from ports on the same accelerator, it is important to use the same type of Ethernet transceivers and network cable lengths for all ports. Each meter of network cable adds 5 ns of delay. The delay difference between different types of transceivers may be more than 20 ns.

Figure 9 shows a test setup with a time reference connected to one Napatech accelerator installed in a server. Traffic is generated from the traffic generator and distributed to two ports on the accelerator via an optical splitter. The cables and transceivers are the same for both connections. Time stamps are compared on the received packet streams. Please refer to the data sheet Time Precision Performance, available on the Napatech website, for exact results from this and other test setups performed with Napatech accelerators.
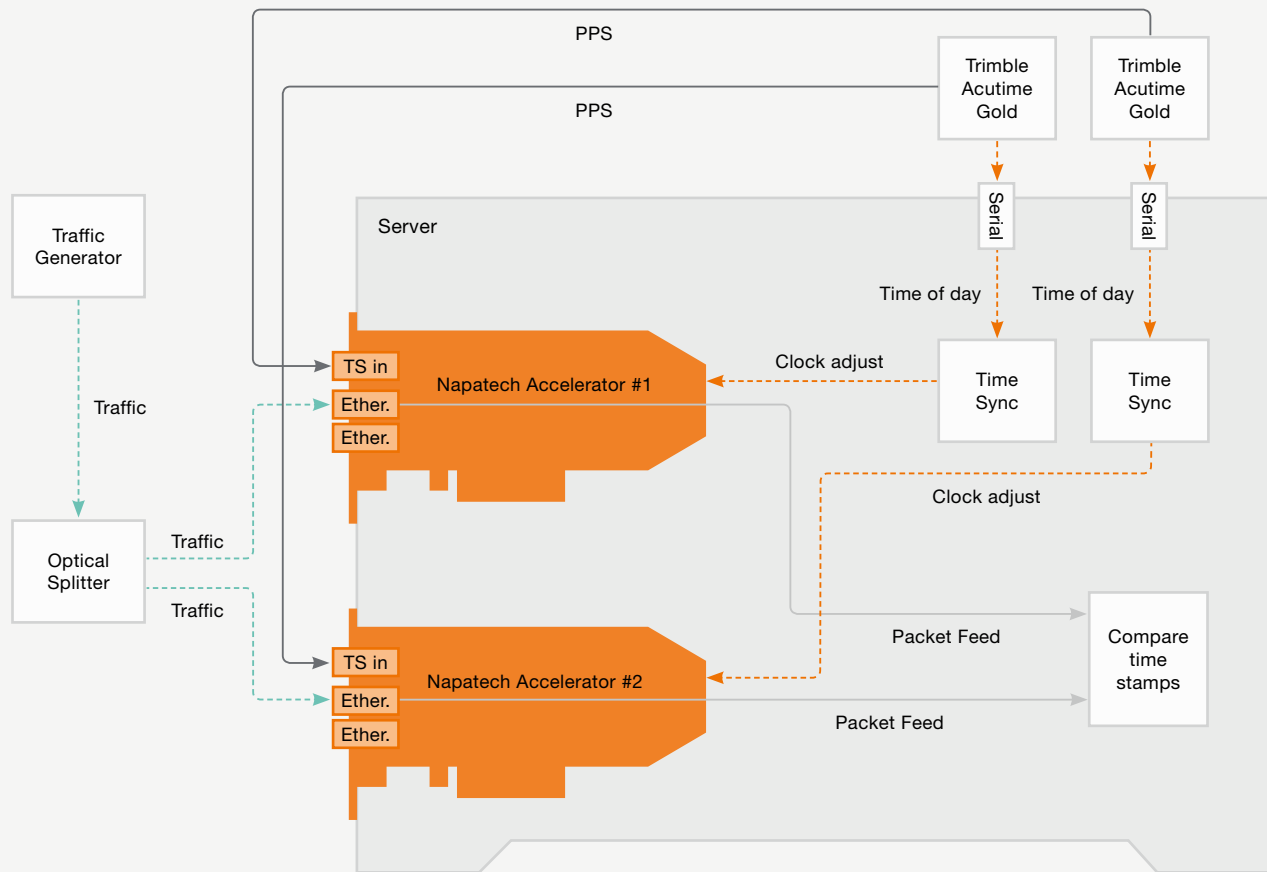
## MULTIPLE LINK MEASUREMENTS AT ONE LOCATION – MULTIPLE ACCELERATORS

Now, let us look at a system where all monitoring ports are at the same location, but on different accelerators within the same chassis.

When accelerators are synchronized to a common reference clock, the precision between time stamps in packets from different accelerators is determined by the accelerator time stamp clock precision and the accelerator internal time stamp precision for each accelerator.

Figure 10 shows a test setup with a common time reference connected to two Napatech accelerators installed in the same server. Traffic is generated from the traffic generator and distributed to one port on each accelerator via an optical splitter. The cables and transceivers are the same for both connections. Time stamps are compared on the received packet streams. Since the same GPS receiver and PPS time

**FIGURE 11**
Measurements across different locations

reference is used for both accelerators, the precision of the GPS receiver does not influence the results.

**MEASUREMENTS ACROSS DIFFERENT LOCATIONS**
Measurements across different locations are more complex, and calculation of the total system time synchronization accuracy involves all the components and cables used to establish the synchronized system. As outlined previously, precision of simple accelerator to accelerator synchronization is straightforward to calculate. If the system comprises several locations and separate GPS receivers, the calculations are more complex. If CDMA or IEEE 1588 / PTP  is used for time distribution in all or parts of the system, the precision of these technologies adds to the overall precision of the system.

Figure 11 shows a test setup with separate GPS time references (in this case Trimble® Acutime™ Gold GPS receivers) connected to two Napatech accelerators installed in the same server. This example applies to a real usage scenario, since the GPS receivers could be located anywhere in the world with GPS coverage. Traffic is generated from the traffic generator and distributed to one port on each accelerator via an optical splitter. The cables and transceivers are the same for both connections. Time stamps are compared on the received packet streams. Since separate GPS receivers and PPS time references are used for the accelerators, the precision of the GPS receivers and the GPS satellites are inherited in the test results.

**THE TIME HAS COME FOR GREATER PRECISION**
Time synchronization and precision time-stamping are more important than ever as we move to higher data rates, such as 40 Gbps and 100 Gbps. Precision is required in not only the technology, but also the terminology and how we understand the various elements and their interaction. Understanding what we mean when we refer to time and accuracy are important in understanding what can be realistically achieved.

The precision and accuracy achievable is relative to the approach taken and the final needs of the application. A GPS antenna is a more accurate time source than NTP or CDMA, but the simple deployment of NTP or CDMA make these options attractive for many applications where the accuracy is good enough for purpose. These are the considerations that must be taken into account when determining the right time synchronization solution for the given application.

Nevertheless, the methods and technologies described here show that highly precise time synchronization and time-stamping are achievable. The tools are available and the technology is in place to support even the most demanding applications.

## COMPANY PROFILE

Napatech is the world leader in accelerating network management and security applications. As data volume and complexity grow, the performance of these applications needs to stay ahead of the speed of networks in order to do their jobs. We make this possible, for even the most demanding financial, telecom, corporate and government networks.

Now and in the future, we enable our customers' applications to run faster than the networks they need to manage and protect.

**Napatech. SMARTER DATA DELIVERY**

SMARTER
DATA DELIVERY