TECHNICAL BRIEF

Distributing Timing to EndaceProbes

This technical briefing describes industry best practice for implementing a timing distribution infrastructure for Endace Probe™ deployments[1]. Endace has partnered with Trimble and EndRun Technologies to offer high-quality GPS and CDMA receivers. This paper refers to these products when recommending solutions.

Packet Timestamping

The first thing every EndaceProbe does when it receives a packet on its monitoring-port is to add a timestamp to the packet. The timestamp indicates the exact time the packet arrived at the monitoring interface. Analysis applications use packet timestamps to accurately reconstruct the order and timing of protocol-based conversations, while latency monitoring applications rely on packet timestamps to work out the end-to-end network delays in transaction processing.

Timestamped packets are stored in 'trace files.' Traces consist of time-ordered records, each comprising the original packet plus overhead information such as port number, packet type, and timestamp. There are two file formats for trace files – pcaps and ERF records (Endace Record Format). EndaceProbes can produce and manipulate pcap and ERF files.

A key difference between pcap and ERF formats is the way the timestamp is stored. The ERF timestamp is a single 64-bit fixed-point number, representing seconds since Jan 1, 1970. The high 32 bits contain the integer number of seconds, while the lower 32 bits contain the binary fraction of the second. This allows timestamp resolution of 2–32 seconds, or approximately 233 picoseconds. Pcap timestamps use two 32-bits fields, one for number of seconds, and one for number of microseconds. Therefore, regardless of the accuracy of the timestamping hardware, standard pcap records can only record timestamps to the nearest microsecond (μs).

Currently, Endace hardware generates timestamps with 28 meaningful bits of information (the last 4 bits are always zero), which allows for 7.5ns of resolution[2]. Higher resolution, while theoretically possible, is neither practical nor required for most applications. 7.5ns resolution means you can take any two packets from a trace file generated from a single clock source, compare their timestamps, and know with certainty which packet was transmitted first. That's important because the packets in a trace files are frequently an aggregation of the packets sent in both directions of a link. Often, trace files contain packets captured from both directions of multiple links. Without this high degree of timestamp accuracy, you would not be able to accurately reconstruct the packet sequence.

Measuring Packet Latency

In a trading environment, measuring the time it takes for a packet to get from point A to point B is crucial. The simplest approach is to capture and timestamp packets as they are sent from point A, capture and timestamp packets as they arrive at point B, send the ERF file from the two monitoring points to a single location, look for identical packets in the two files (based on the packets' address, sequence number, and content), and then subtract the timestamp at point B from the timestamp at point A. This approach is fine, but of course only works if the source of timing at points A and B are accurately synchronized.

Measuring the accuracy of the clock at point A and the clock at point B is a matter of relative error (also called clock skew). In practice, the timestamps in most ERF file are ultimately derived from a local clock that has been synchronized to the single worldwide time source (called the Coordinated Universal Time, or UTC).

In monitoring equipment such as the EndaceProbe, the internal hardware clock used to generate timestamps derives its basic timing from a crystal oscillator. If crystal oscillators were perfect you could simply set the clock relative to UTC at time of manufacture. However, variances in materials, temperature, and voltage cause every crystal oscillator to drift. Even within a single second a local clock based on a crystal oscillator will drift far enough for the packet timestamps to be significantly skewed from UTC. The simple

1 The monitoring-ports on all EndaceProbes are based on the Endace DAG card. Therefore, this application note is suitable as a guide for deploying any DAG®-card-based equipment.

2 To put this level of accuracy into context, timestamps generated by software will have mean errors measured in 100s of milliseconds, with maximum errors that can exceed half a second due to context switches in the OS kernel.

solution is to periodically correct the natural drift of the local clock by continuously realigning it to a UTC reference signal distributed to the EndaceProbe by a clock distribution system.

Some monitoring port cards use HW-based timestamping but only correct the drifting local clock once per second. This technique does not account for clock drift within each second, and this drift can be significant. For example, if a +/- 50 ppm (parts-per-million) crystal oscillator is used, which is typical accuracy for these types of crystals, the maximum accumulated error at the end of a one-second period could be +/- 50μs. By using sophisticated DLL (Digital Locked Loop) technology Endace hardware provides continuous clock alignment maintaining +/- 30ns clock accuracy at all times.

When aligned to exactly the same reference clock EndaceProbes will stay within +/- 30ns of each other. Which leads to a discussion about the distribution of reference clock information.

UTC Distribution Systems

There are several UTC clock distribution systems available. Which one is best depends on the situation and the accuracy required. The four most common clock distribution systems are GPS, CDMA, NTP, and PTP. Each has its advantages and disadvantages.

GPS is the primary, but not the only source of UTC reference information, and knowing whether the UTC distribution system being used is aligned to GPS UTC or true UTC can be quite important. This briefing assumes that any GPS-derived time-of-day value has been adjusted to true UTC[3].

Global Positioning System (GPS)

GPS uses satellites that have an internal atomic clock, corrected to the UTC, so GPS is typically accurate to within 15ns of UTC. GPS signals carry information that makes highly accurate UTC available more-or-less anywhere on Earth. While GPS offers the greatest precision of all the clock distribution systems, installation of GPS receivers can be challenging due to the need for a wide-angle view of the sky.

Endace offers the Trimble™ Acutime™ Gold GPS receiver[4], a ruggedized, integrated GPS receiver that provides a highly accurate 1 PPS (pulse-per-second) timing signal to the Endace TDS-2/6 or TDS-24 time distribution servers. It is a compact pole-mounted device with all electronics integrated into a small sealed weather-tight unit. Accuracy to UTC is rated as +/- 15ns.

Code Division Multiple Access (CDMA)

In Code Division Multiple Access (CDMA) mobile phone networks, timing signals are distributed as a side effect of establishing communication with handsets. CDMA uses GPS receivers on its towers as the master clock source[5], but subsequent processing and transmission reduces the accuracy to UTC to about 10 microseconds. Because CDMA signals can penetrate buildings, CDMA can be a cost-effective way of obtaining an accurate clock feed. Like GPS, a CMDA receiver is typically installed in each building, with a clock distribution system used to distribute the reference information to the EndaceProbes.

CDMA is being phased out in many locations. CDMA signals also suffer a transmission delay from the cell radio tower to the receiver. Therefore there will be skew from UTC depending on the distance between the transmission tower and the CDMA receiver.

Endace offers the EndRun Technologies Præcis Ct™ CDMA receiver. It is an integrated receiver that provides an accurate 1 PPS timing signal to the Endace TDS-2 or TDS-24 Time Distribution Servers. Accuracy to UTC is rated as +/- 10μs. The EndRun CDMA receiver can be configured to compensate for the propagation delay from the cell tower and the receiver.

Network Time Protocol (NTP)

NTP is the most widely used clock reference across LANs and WANs. It requires no special hardware, and is a simple way for computers to maintain accurate time of day. NTP provides more than adequate accuracy for most enterprise applications but is not an accurate enough UTC clock distribution method for network-monitoring tools measuring latency.

3 Note that the difference between GPS UTC and UTC only impacts the TOD aspect of the timestamp. Any GPS-derived pulse-per-second reference (used to align local clocks) will be aligned to UTC seconds whether or not the GPS is outputting UTC or GPS UTC TOD.

4 The Trimble Gold is ideal for timing applications because it allows the user to configure the receiver to output GPS UTC or true UTC. It also allows the user to enter a timing offset to compensate for the length of coax cable between the antennae and the receiver. Without this compensation, the PPS output of the receiver will be skewed from UTC by the propagation delay of the cable run, which is length dependent, of course.

5 However the time-of-day output by the CDMA receiver is UTC, not GPS UTC.

EndaceProbes use NTP only for coarse date and time-of-day information (year, month, day, hour, minute, second). For higher levels of resolution Endace recommends an accurate UTC timing signal also be provided via a PPS.

IEEE 1588 Precision Time Protocol (PTP)

The IEEE 1588 Precision Time Protocol (PTP) is a standard designed to transmit a clock reference across a LAN or WAN. It uses a combination of multicast and unicast protocols to synchronize multiple devices (called timing slaves) with sub-microsecond precision single reference source (called the timing Grandmaster).

The Grandmaster is typically synchronized to a high-quality UTC reference source such as a GPS or CDMA receiver.

Each slave continuously synchronizes its clock by exchanging timestamped messages with the Grandmaster. The protocol is based on the assumption that the packet transmission delay for packets travelling from the master to the slave are identical to the delay of packets heading from slave to master. The Grandmaster periodically broadcasts (multicasts) a timestamped sync message (and optional follow-up message) every two seconds or so. Also at regular intervals (every 20s or so), each slave and the Grandmaster exchange timestamped messages directly using unicast UDP/IP. The slaves use the arrival time of messages and their internal timestamps to estimate the packet delay between master and slave. The slave then uses this delay as an offset to synchronize itself to the master's timestamps.

Implementations typically use hardware packet timestamps to improve accuracy over software-based systems such as NTP. PTP is less accurate than using an external reference such as GPS, but is more accurate than NTP when used over a dedicated LAN designed to support constant, symmetric packet transmission times. The PTP protocol does work over LANs or WANs carrying other traffic, but clock accuracy declines depending on how much the packet latency changes under heavy traffic load.

PTP-based timing distribution equipment can be deployed in much the same way as GPS or CDMA receivers. A PTP slave or master is required in each building to provide the EndaceProbe with a PPS synchronization signal.

Connecting a UTC reference to EndaceProbes: PPS and TOD Distribution

To ensure that the local clock on the EndaceProbe is continually synchronized to an accurate external reference clock, EndaceProbes are designed to accept a PPS hardware reference signal.

The required PPS reference is a binary signal compliant with the RS-422 electrical standard. Using RS-422 provides the EndaceProbe with an accurate timing pulse over cables up to 328 yards (300 meters) in length[6]. The PPS signal provides an extremely precise and accurate signal edge (transition) aligned to the one-second boundary of UTC. The monitoring port hardware uses this PPS transition to align the one-second boundary of its internal clock to within +/- 30ns of the transition.

Unlike the PPS reference, the UTC-aligned TOD provided by NTP does not require the separate physical distribution of a timing reference. Rather, the EndaceProbe's management Ethernet port need only reside on a LAN/WAN equipped with an accessible NTP server, which is a common configuration.

Most GPS and CDMA receivers provide a PPS output synchronised to UTC. Various PTP products also support PPS outputs, as discussed previously.

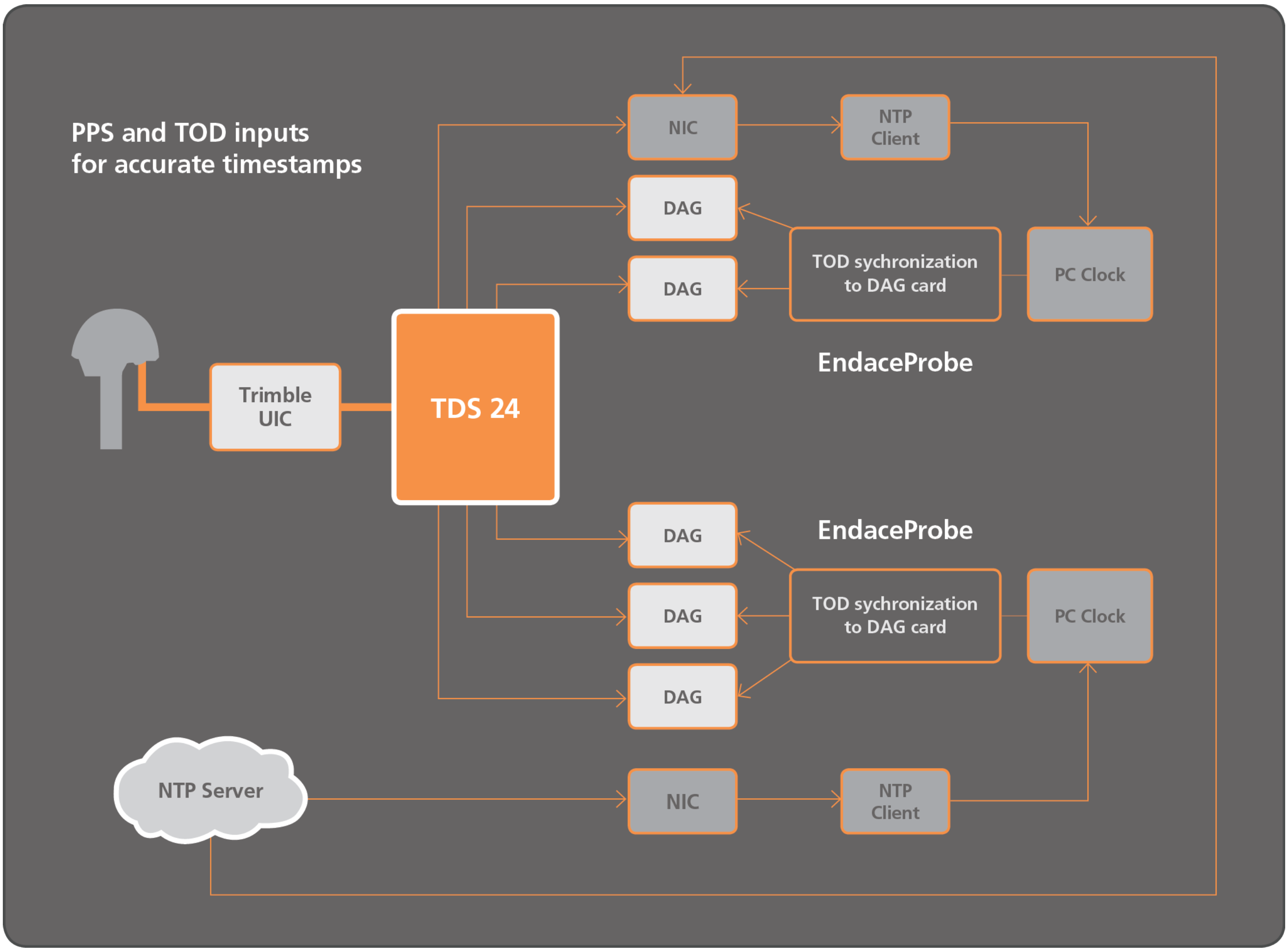
Once you have a GPS, CDMA, or PTP receiver (or any other source of UTC derived RS-422 PPS signal) in the building, the next step is to replicate and distribute the PPS signal to all monitoring port cards on all the EndaceProbes being synchronized. The monitoring ports on EndaceProbes are grouped into one, two, or four-port DAG cards, and a single EndaceProbe can have more than one card. All ports on a card are timestamped by the clock hardware on that card, and each monitoring card in the EndaceProbe must be individually synchronized via its own PPS signal.

Endace has developed the TDS (Time Distribution System) product family to provide building and campus distribution of RS-422 PPS signals. All TDS products use standard cat 5 cabling and RJ-45 connectors to distribute the RS-422 PPS signals, however they are not Ethernet devices and do not use IP. Input to output port latency has been balanced to minimize timing skew between outputs.

6 For those familiar with the RS-422 electrical standard a 328 yard restriction will seem surprising since the RS-422 standard states cable lengths of 1312 yards (1200 meters) are supported. However, the PPS signal is a binary signal, and the longer the cable length the more slewed the signal transition becomes. The clock synchronization circuitry requires a clean edge (a fast slew rate) to maintain accurate alignment, which is why we recommend restricting the PPS cable to 328 yards (300 metres), which restricts additional clock skew to +/-2 ns. Longer cables will work, but timing accuracy is reduced.

EndaceProbes allow each monitoring port card to be pre-programmed with a fixed offset to compensate for latency added by the PPS distribution system. To determine each EndaceProbe's offset from UTC, the following steps are recommended:

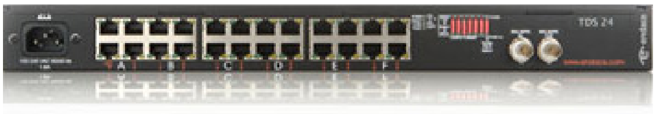
1. Determine the latency from UTC added by the GPS receiver, the CDMA receiver, or the PTP master/slave device. For GPS the coax from the roof antennae to the receiver will add approximately 5ns per meter, although receivers such as the Trimble Gold allow this to be calibrated out right at the receiver. For CDMA, estimate the distance to the CDMA cell tower and calculate latency at about 3 μs/km unless the receiver (such as the EndRun) has been calibrated to factor this out. For PTP the master/slave specification should be consulted.
2. Determine the number and type of TDS products needed to distribute the PPS reference to each monitoring port card (i.e. each DAG card) in each EndaceProbe. Each TDS device in the path adds a fixed latency from PPS in to PPS out, and there will be a small variance in latency depending on operating conditions and which output port is used. See the TDS product discussion below for details.
3. Estimate the cable length for each path. Cat 5 cabling propagation delay depends on the cable type, but



typically adds between 5.1ns and 5.2ns latency per metre of cable length. For simplicity, lengthy cat 5 cable runs going to a single location should be kept to equal lengths so all cards in the EndaceProbe can be programmed with the same timing offset.

4. For each monitoring card, total the overall path delay determined in steps 1 to 3.
5. Use the EndaceProbes management interface to enter the skew offset value (in ns) for each card.

TDS 24



The Endace TDS 24 is a 1U Time Distribution Server capable of providing low latency and accurate timing information to 24 Endace network monitoring cards. TDS 24 modules can be daisychained (stacked) two or more times, providing timing information to hundreds or even thousands of devices[7]. The TDS 24 accepts and distributes RS-422 PPS and IRIG-B00x (IRIG-B DCLS) timing signals[8]. It can convert an IRIG-B input signal to RS-422 PPS output signals, which greatly expands the variety of UTC distribution products that are compatible with EndaceProbe deployments.

The TDS 24 introduces 100ns of latency from PPS in to PPS out, +/-25ns depending on the output port selected.

TDS 2 and TDS 6

The TDS 2 is a small, modular PPS distribution unit compatible with the GPS and CDMA receivers sold by Endace. A 1U rack mount kit is available to house up to three modules. The TDS 2 PPS input connects to the GPS or CDMA receiver's PPS output.

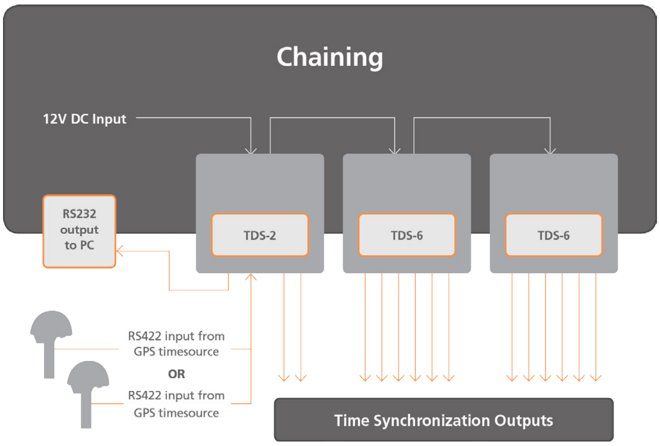
The TDS 2 outputs two RS-422 PPS signals, and two chaining ports for connecting to the TDS 6. When more than two PPS outputs are required, the TDS 6 extender module can be used. The TDS 6 connects to the chaining port on the TDS 2 and replicates the PPS to six RS-422 PPS outputs, and one chaining port. TDS 2/6 combination can be daisy-chained to four levels (one TDS 2 and three levels of TDS 6) without compromising timing accuracy. In this way a maximum of 38 DAG cards may be synchronized from a single source using one TDS 2 and six TDS 6 modules. Note that a chained configuration is no longer required due to the availability of the TDS 24.

TDS 2 GPS input to either PPS output: 140ns +55/-50ns of delay.

TDS 2 CDMA input to either PPS output: 240ns +30/-20ns of delay.

TDS 6: PPS input to any PPS or chain output: 140ns +55/-50ns of delay.

Interfacing to legacy time distribution systems with TIC 1



Data centers and telco central offices often have legacy UTC timing distribution systems in place to provide equipment with UTC or various traditional telecommunication timing references. Often this equipment provides PPS outputs, but frequently the physical interface is a BNC connector rather than the 9-pin RS-422 connector used by the TDS 2 module.

[7] For example, a simple two-layer distribution system comprised of 25 TDS 24 modules can provide 576 PPS outputs.

[8] EndaceProbes use a PPS timing reference only. Certain Endace DAG cards accept an IRIG-B timing reference as well as PPS. Please consult the DAG card data sheets to determine which cards are IRIG-B enabled.

There are two options when interfacing to legacy time distribution equipment:

1. Use the TDS 24, which has a BNC input port.
2. Use Endace's low-cost module called the TIC 1 (Time Interface Converter), a simple BNC to 9-pin RS-422 convertor. The TIC 1 converts the PPS interface from the legacy equipment to what is needed by the TDS 2.

TIC 1 coax input to RS-422 DB-9 output: 130ns +110/-90ns delay added.

Once the legacy PPS distribution system has been implemented the end-to-end propagation delay on each PPS path can be measured and estimated, and this offset programmed into the Endace monitoring ports, as discussed above.

Deployment Examples

Figure 1

UTC reference available per building

This is the ideal scenario, where each location can be equipped with a good UTC reference, either via GPS, CDMA, or PTP receivers. Within each building, the TDS 24 is used to distribute as many PPS reference signals as needed to each EndaceProbe. EndaceProbes with a single monitoring port card (i.e. DAG® card) only require a single PPS reference. EndaceProbes with two monitoring port cards require two (hopefully close to identical) PPS references; EndaceProbes with three cards require three references, etc.

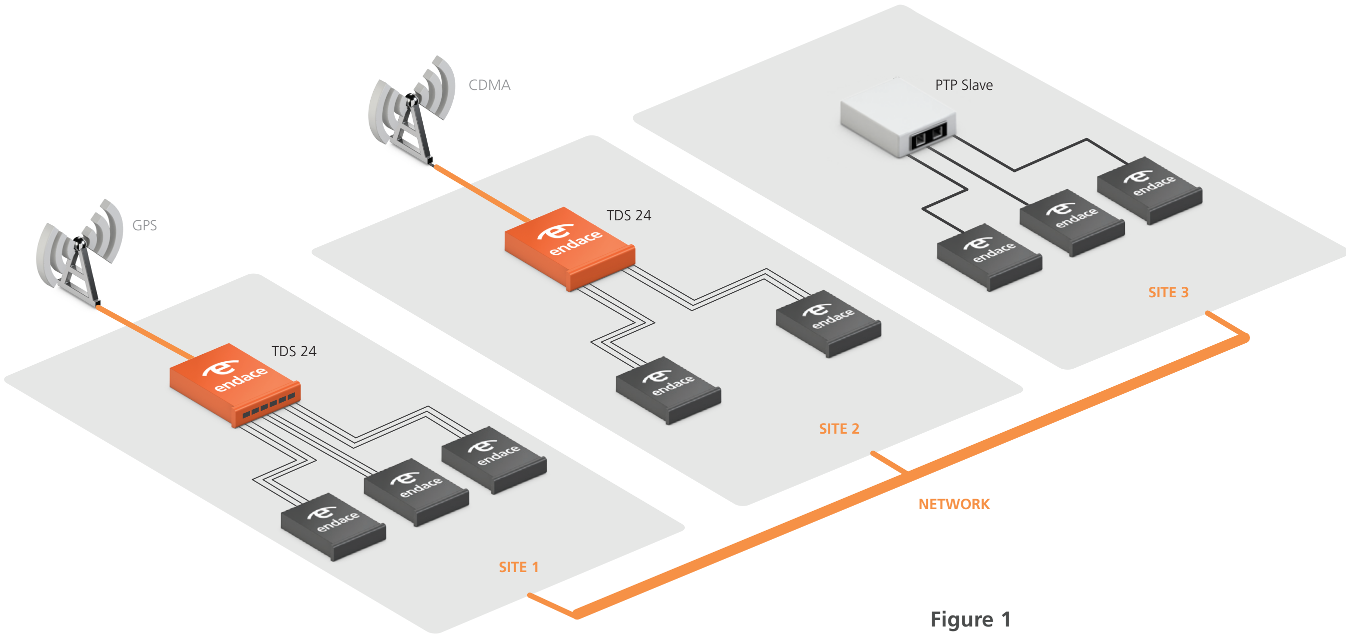


Figure 1
With three sites, one with GPS receiver, one with CDMA receiver, one with PTP slave and TDS-24 in all locations; each DAG card within each probe requires its own PPS signal.

Figure 2
PTP

This scenario covers multiple EndaceProbes spread over multiple sites where GPS or CDMA is not available but a high-quality LAN suitable to PTP-based UTC distribution is available at all sites. Note, at least one location must have a high-quality UTC reference available, and that location will naturally become the PTP Grandmaster for this deployment.

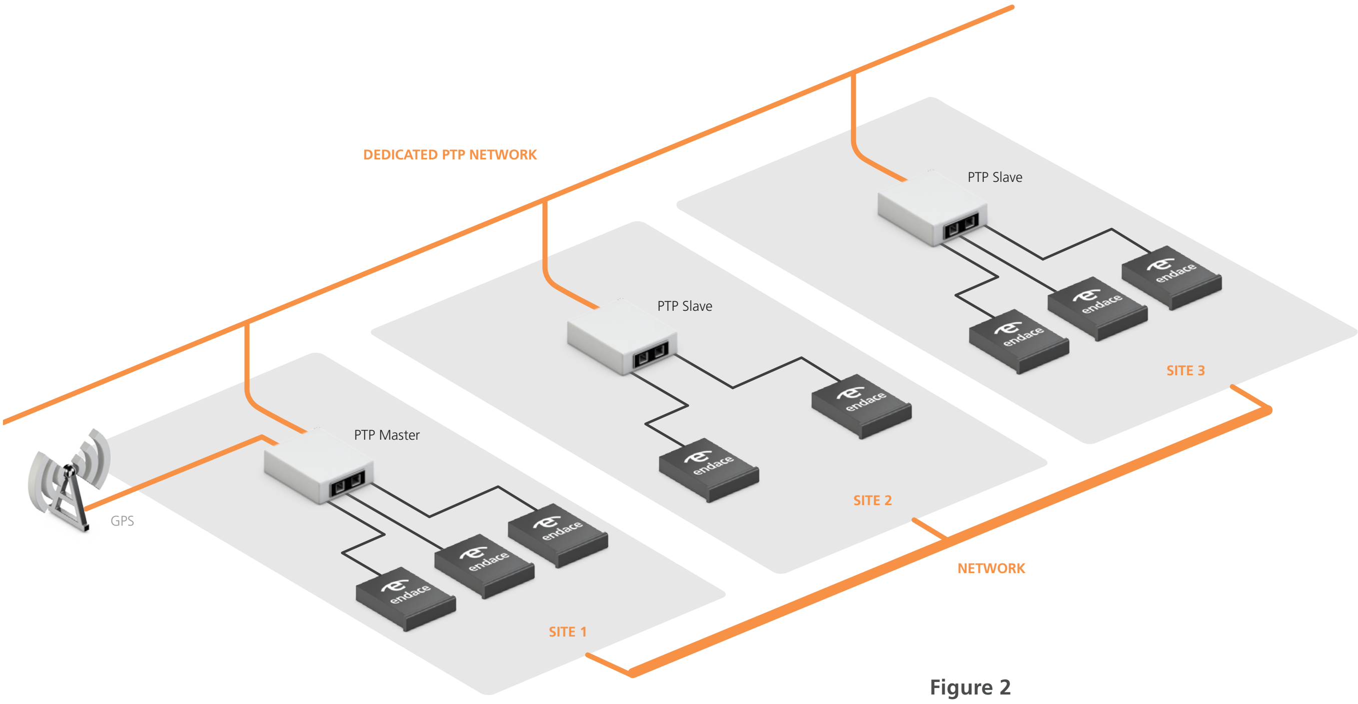


Figure 2
With three sites, one with GPS receiver and PTP Grandmaster, the rest with PTP slave and TDS-24 in all locations. Each DAG card within each EndaceProbe requires its own PPS signal.

Example 2

PTP

This scenario covers multiple Probes spread over multiple sites where GPS or CDMA is not available but a high-quality LAN suitable to PTP based UTC distribution is available at all sites. Note, at least one location must have a high-quality UTC reference available, and that location will naturally become the PTP Grandmaster for this deployment.

Example 3

Third party time distribution server

This scenario shows a single location (perhaps a telco central office) with a time distribution server already installed. Here the TIC module is used to convert the time server's BNC-based PPS signal into the DB-9 RS-422 interface used by the TDS 2. The TDS 2 provides two PPS references to a dual card Endace Probe. Note that if more Probes where to be synchronized the TDS 24, with its built-in BNC PPS connector, would be a better choice than the TIC plus TDS 2 combination shown here.

Example 3

Third-party time distribution server

This scenario shows a single location (perhaps a telco central office) with a time distribution server already installed. Here the TIC module is used to convert the time server's BNC based PPS signal into the DB-9 RS-422 interface used by the TDS 2. The TDS 2 provides two PPS references to a dual card EndaceProbe. Note that if more EndaceProbes where to be synchronized the TDS 24, with its built-in BNC PPS connector, would be a better choice than the TIC plus TDS 2 combination shown here.

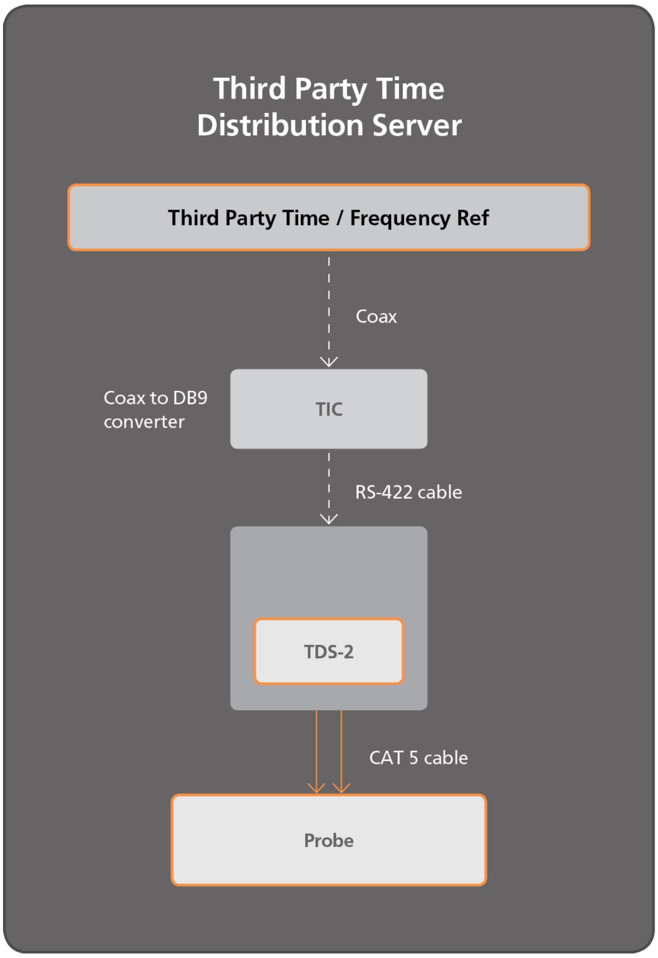


Figure 3
With one site with a generic 'time distribution server' synchronized to GPS. TIC used to convert to RS-422 and connected to TDS 2, then connected to dual card EndaceProbe.

For more information on Endace products visit: **endace.com**
For enquiries email: **enquiries@endace.com**