# STUDYING NETWORK TIMING WITH PRECISION PACKET DELAY MEASUREMENTS

Lee Cosart R&D, Symmetricom, Inc. 2300 Orchard Parkway San Jose, CA 95131, USA *lcosart@symmetricom.com* 

#### Abstract

As the transmission of telecommunications data is increasingly reliant on new generation packet network transport, new methods of time and frequency transfer are required. While some of these methods are at the physical layer, many involve the network and data link layers and are affected by packet network behavior. Thus, it has become important to develop instrumentation and analysis techniques applicable to the study of packet latency and packet delay variation.

Several of the packet network protocols which have been proposed for precision time and frequency transfer are NTP (Network Time Protocol) and IEEE 1588v2, also known as PTP (Precision Time Protocol). Systems designed for high-performance NTP or PTP contain the necessary elements for precision packet timing measurements, namely a common precision timing reference from a source such as GPS and hardware timestamping capability. Such equipment has been developed over the past several years and has been used to make packet delay measurements, both in the laboratory and in live operating production networks.

The interpretation and analysis of packet timing data benefit from a combination of traditional techniques and new techniques. Certain analysis techniques applicable to traditional time and frequency measurements carry over to packet timing data, while others are of little utility or in some cases benefit from modification. In this latter category is a new metric, minTDEV, based on one of the Allan deviation calculations.

This paper starts with a discussion of the measurement techniques and systems used in the study of packet network timing. Following this introduction, it presents laboratory and production network measurement results from a number of commercial operators and introduces various techniques for the analysis of such results. During this discussion, references are made to compare and contrast these methods of analysis with the techniques applicable to traditional time and frequency measurements, such as those made on atomic oscillators and synchronization systems.

### INTRODUCTION

Just as understanding oscillator and frequency-locked loop behavior is critical to a study of traditional methods of time and frequency transfer, understanding packet network behavior from the standpoint of

timing is critical to a study of packet network time and frequency transfer protocols such as NTP and PTP. When protocols such as these are used, the measurement of packet network timing becomes an essential measurement, along with the direct characterization of system phase and frequency performance made in the traditional way. This is akin to the necessity of measuring environmental parameters such as temperature, while measuring system output phase and frequency if one wishes to study how the system is affected by these conditions. Thus, the study of packet network timing often involves two separate measurements, one traditional phase and frequency measurement of an output signal and another involved with characterizing packet delay and packet delay variation.

Traditional methods of measuring network packet delay have typically employed software methods over wide areas or lab equipment with limited precision, yielding only histograms of the packet delay data. When network timing behavior is not a stationary process, but rather is subject to dynamic conditions and changes over the short term and longer term, a histogram does not suffice; packet delay must be characterized over time. When applications demand precision on the order of microseconds or nanoseconds, packet network timing measurement equipment must have similar capabilities.

### PRECISION PACKET TIMING MEASUREMENT EQUIPMENT

Precision packet timing measurements require two principal components. First, hardware capable of producing precision timestamps is necessary. Second, since packet timing measurements involve timestamping a sequence of probe packets at two or more points in a network and correlating the results, a common, precise time-of-day (TOD) clock is a requirement. Equipment incorporating GPS timing receivers and packet hardware timestamping such as IEEE 1588 grandmaster clocks and NTP timeservers, fulfills these two requirements. Dedicated GPS-referenced network timing probes using specialized probe packets can also be used. These three configurations are shown in Figure 1 below.



Figure 1. PTP (IEEE 1588), NTP, and network timing probe measurement setups.

The IEEE 1588 PTP equipment serves to illustrate a few other points with regards to making these packet timing measurements. Since the goal is to make one-way delay measurements in both directions, there is a packet generation function as well as a packet timestamping probing function at both nodes. This probing function can either be combined with the packet generation function or separated from the packet generation function. A probe combining the two functions can be referred to as an active probe. In contrast, a probe that serves only as a listening device without implementing the PTP protocol can be referred to as a passive probe. The passive probe is shown in Figure 2 and the active probe is shown in Figure 3.



Figure 2. Passive probe packet timing measurement setup.



Figure 3. Active probe packet timing measurement setup.

The relative simplicity of the active probe setup shows an advantage it has compared to the passive probe. Using a passive probe requires insertion of a hub or network tap and an additional piece of equipment, an

IEEE 1588 slave to initiate and maintain the PTP protocol. It is important to point out that a 1588 slave cannot be used as a precision probe, since it takes its timing from the very network the equipment is set up to characterize and is subject to the packet delay variation in the network. The 1588 grandmaster, on the other hand, takes its reference from GPS, and so contains the necessary components for precision packet timestamping. In both cases, illustrated in Figure 2 and Figure 3 above, the 1588 grandmaster module is serving as an active probe of sorts, since it both generates a packet stream and times the packets.

The setups in Figure 2 and Figure 3, in contrast to the PTP setup in Figure 1, incorporate two-node measurement capability into a single shelf. Such a setup cannot be used for wide-area network measurements, but is ideal for measurements in a lab where both sides of a test network are co-located. An advantage of such a lab setup is that the two modules share the same reference, since they are in the same shelf. In this common shelf setup, the GPS reference could be replaced with a frequency reference, such as a 10 MHz lab reference from a cesium clock; in either case, the two modules will share the same time reference.

While measuring an oscillator or output signal from a timing system is a single-node measurement, measuring packet delay requires correlating measurement data taken at two nodes. The timing system output signal measurement involves timestamping signal edges to derive phase deviation samples. As indicated in the discussion above, the packet delay measurement requires timestamping a stream of packets at two different network nodes. The data from the two nodes are combined to form a packet delay sequence. To ascertain the procedure for doing so, it is instructive to view a short sample of packet timestamp data. Such a sample, taken using IEEE 1588 equipment, is shown below in Figure 4.

```
Source File:

SEQ: 01195 UUID: 00A069012FB9 UTC: DATE 2008:124:01:48:56 NSEC 0650092776

SEQ: 23238 UUID: 000055010016 UTC: DATE 2008:124:01:48:56 NSEC 0772791251

SEQ: 23239 UUID: 00A069012FB9 UTC: DATE 2008:124:01:48:57 NSEC 0742766301

SEQ: 01196 UUID: 00A069012FB9 UTC: DATE 2008:124:01:48:58 NSEC 0649898076

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:49:00 NSEC 0372820611

SEQ: 01197 UUID: 00A069012FB9 UTC: DATE 2008:124:01:49:00 NSEC 0649723496

Destination File:

SEQ: 01195 UUID: 00A069012FB9 UTC: DATE 2008:124:01:48:56 NSEC 0650356493

SEQ: 23238 UUID: 000055010016 UTC: DATE 2008:124:01:48:56 NSEC 0772511963

SEQ: 23239 UUID: 000055010016 UTC: DATE 2008:124:01:48:56 NSEC 0742522643

SEQ: 23239 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 01196 UUID: 00A069012FB9 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 23240 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 01197 UUID: 000055010016 UTC: DATE 2008:124:01:48:58 NSEC 0650169943

SEQ: 01197 UUID: 000055010016 UTC: DATE 2008:124:01:49:00 NSEC 0372557873

SEQ: 01197 UUID: 000055010016 UTC: DATE 2008:124:01:49:00 NSEC 049977513
```

Figure 4. Active probe packet timing measurement setup.

This data sample in Figure 4 contains forward and reverse packet flows timestamped at two locations, "source" and "destination." The individual flows are distinguished using the "UUID" field, which is something akin to the MAC address of the originating module. Individual packets are identified using the "SEQ" field. The packet delay for a particular packet is calculated by computing the difference between the "source" and "destination" date/time stamps. In most cases, the date stamp is the same, and the difference is simply the difference between the two "NSEC" field values. The placement of the

timestamp difference samples is also based on the combination of the "DATE" and "NSEC" fields. The sample interval between packet 01195 and packet 01196 is seen to be 2 seconds.

As discussed above, the probe packets can be PTP, NTP, or some other kind of probe packets. Confidence that a particular packet measurement system is performing well is heightened if results from an independent packet measurement system utilizing different techniques yields similar results. In Figure 5 below, two separate pairs of measurements made using NTP and PTP equipment on two different networks are shown. The results are shown as histograms of the packet delay which reveal the similarity of the NTP and PTP measurement equipment results.



Figure 5. Comparison of IEEE 1588 PTP and NTP measurement equipment.

As a final point before moving away from this discussion of packet measurement equipment, it should be noted that often the motivation for characterizing packet network timing behavior arises from the interest in understanding the background conditions for a study of time and frequency transfer performance of packet timing equipment. For example, characterizing the performance of an IEEE 1588 slave requires measuring output signal phase with a separate measurement of packet network timing performance, providing a view into underlying conditions and often insight into the slave performance. Figure 6 below shows an example of a measurement setup where both slave performance and network performance are studied.



Figure 6. Characterizing IEEE 1588 slave performance and network timing performance.

## PACKET TIMING MEASUREMENT ANALYSIS

While much of the analysis applicable to clock measurement data carries over to packet delay data analysis, some of the techniques do not. In addition, new techniques of analysis provide insight into packet delay data behavior. The following list contrasts techniques applied to clock data with some of those applied to packet data:

- Clock data analysis
  - o Phase
  - o Frequency
  - o Phase noise
  - o MTIE
  - o ADEV/MDEV/TDEV
- Packet data analysis
  - Phase (packet delay sequence)
  - o Phase noise
  - o Histogram/PDF(probability density function) & statistics
  - Running statistics
  - o TDEV/minTDEV/bandTDEV.

These new packet TDEV-based metrics and other new metrics are under study at the ITU-T Q13/SG15 synchronization experts group. The minTDEV metric has references in the latest G.8261 draft. Examples of some of these approaches to packet data analysis are shown in packet timing measurement examples in the subsequent discussion, namely examples in the Lab Measurements and Live Network Measurements sections below.

One of the differences between a clock signal measurement and a packet delay measurement manifests itself in the phase data, particularly when longer term measurements are made. Packet delay phase sequences are typically dominated by short-term variations, and a plot of packet delay phase often appears as a band of data. Often a scatter plot view reveals aspects of the data not seen when the data points are connected. Such is the case for the example shown in Figure 7.



Figure 7. Packet delay sequence with points connected and as scatter plot.

### LAB PACKET TIMING MEASUREMENTS

Before studying live networks and more complicated laboratory network configurations, it is instructive to characterize various network devices individually. Figure 8 shows phase, MTIE, and TDEV for a crossover cable, a hub, an enterprise switch, and an enterprise router.

Figure 9 shows packet delay statistics for the same four devices as well as for a multilayer switch. With increased device complexity come increased delay and spread. Multilayer switches and wire speed routers incorporate both layer 2 and layer 3 transport functionality and can support routing functions. But their performance is more like a switch than an enterprise router, which is software based. The router measurements in Figure 8 and Figure 9 were made on an enterprise router.

Figure 9 also includes statistics for two enterprise routers connected in series. Mean delay and packet delay variation (PDV), as represented by standard deviation, increase by a factor of two. The central limit theorem applies, with the two router distribution approaching a Gaussian shape rather than the distorted flat-top shape for the single router. Note that the y-axis is on a logarithmic scale, so the classic bell shaped curve becomes a parabola.



Figure 8. Crossover cable, hub, switch, and router packet delay phase, MTIE, TDEV.



Figure 9. Crossover cable, hub, switch, multilayer switch, and router packet delay statistics.

If network conditions are not static, then tracking a statistic over time can be an effective way of studying a packet timing measurement. An example is shown in Figure 10 below.



Figure 10. Tracking packet delay statistics over time.

Traffic loading is a critical factor in packet network timing. Any laboratory study attempting to simulate real network conditions requires some means of traffic generation. Real networks do not operate under static loads, and dynamic changes in loading levels often result in dynamic changes in packet delay. Figure 11 describes a lab setup which includes a traffic generator.

When a multilayer switch is subjected to various traffic loads, the packet delay distribution shows increased spread as more and more packets are subjected to greater delay. Still, many packets avoid the increased queuing delays and traverse the device in times comparable to the unloaded switch. The result is a single-sided distribution with greater spread as traffic levels increase. This is shown in Figure 12 below. The second characteristic to note is the bimodal or multi-modal nature of the distribution, particularly apparent at lower traffic loads. Figure 13 shows the elevation of TDEV curves with increased traffic for the same set of measurements.



Figure 11. Lab network configurations with traffic generation.



Figure 12. Statistics for multilayer switch with traffic.



Figure 13. TDEV for multilayer switch with traffic.

If the TDEV calculation is adjusted to include a selection algorithm, packet delays at or around the minimum packet delay floor can be favored, and the resulting modified TDEV shows similar noise levels at varying degrees of traffic. Servo designs must include a similar selection algorithm in order to produce a slave clock with good performance, so this minTDEV algorithm is an indicator of how well packet selection might work under various network scenarios. Figure 14 shows minTDEV analysis on measurements made on a switch with traffic levels ranging from 0% (no load) to 50%. The curves converge for tau values greater than 100 seconds. The 0% and 5% curves are virtually indistinguishable. Figure 15, which includes both TDEV and minTDEV curves for one value of traffic loading, shows the potential noise reduction as a result of packet selection.



Figure 14. minTDEV for multilayer switch with traffic.



Figure 15. TDEV vs. minTDEV for multilayer switch with traffic.

A further generalization of the selection aspect of minTDEV is to include a selected band of the data, producing a metric called bandTDEV. To define bandTDEV, it is first necessary to represent the sorted phase data. Let "x" represent this sorted phase sequence from minimum to maximum over the range  $i \le j \le i+n-1$ . Next, it is necessary to represent the indices which are themselves set based on the selection of two percentile levels. Let "a" and "b" represent indices for the two selected percentile levels. The averaging is then applied to the "x" variable indexed by "a" and "b". The number of averaged points "m" is related to "a" and "b": m=b-a+1. A special case of bandTDEV with "a" set to 0 is percentileTDEV. The TDEV, minTDEV, and bandTDEV calculations are compared in Figure 16.

**TDEV** 
$$\boldsymbol{\sigma}_{x}(\boldsymbol{\tau}) = TDEV(\boldsymbol{\tau}) = \sqrt{\frac{1}{6} \left\langle \left[\frac{1}{n} \sum_{i=1}^{n} x_{i+2n} - 2\frac{1}{n} \sum_{i=1}^{n} x_{i+n} + \frac{1}{n} \sum_{i=1}^{n} x_{i}\right]^{2} \right\rangle}$$

$$\min \mathsf{TDEV} \qquad \sigma_{x_{\min}}(\tau) = \min TDEV(\tau) = \sqrt{\frac{1}{6} \left\langle \left[ x_{\min}(i+2n) - 2x_{\min}(i+n) + x_{\min}(i) \right]^2 \right\rangle} \quad x_{\min}(i) = \min \left[ x_j \right] for(i \le j \le i+n-1)$$

**bandTDEV** 
$$\sigma_{x\_band}(\tau) = bandTDEV(\tau) = \sqrt{\frac{1}{6} \left\langle \left[ x'_{band\_mean}(i+2n) - 2x'_{band\_mean}(i+n) + x'_{band\_mean}(i) \right]^2 \right\rangle} \quad x'_{band\_mean}(i) = \frac{1}{m} \sum_{j=a}^{b} x'_{j+i}$$

Figure 16. TDEV, minTDEV and bandTDEV calculations.

The bandTDEV calculation reduces to minTDEV, TDEV, and to percentileTDEV, in certain special cases.

- TDEV is bandTDEV(0.0 to 1.0)
- minTDEV is bandTDEV(0.0 to 0.0)
- percentileTDEV is bandTDEV(0.0 to B) with B between 0.0 and 1.0.

An example bandTDEV calculation for a measurement of an enterprise router along with the packet delay sequence phase for the same measurement is shown in Figure 17 and Figure 18. The packet delay sequence phase is shown both as a plot with the data points connected and as a scatter plot, with the latter revealing the banding. The bandTDEV calculation is seen to reduce noise in comparison to TDEV over most of the range of the plot.



Figure 17. Packet delay sequence phase for an enterprise router.



Figure 18. TDEV and bandTDEV for an enterprise router.

### EFFECTS OF DIFFERENT APPROACHES TO TRAFFIC GENERATION ON PDV

Many aspects of traffic generation, from the choice of equipment to the way the equipment is configured, can have a great impact on packet delay variation and by extension on the performance of devices timing from the packet flows, such as IEEE 1588 slaves. A five-node network comprised of carrier-class switches was set up with two ports of traffic generation into each switch in the forward direction and one port of traffic generation in the reverse direction. The setup for the reverse direction (not shown) was kept constant, while setup parameters were varied in the forward direction. Packet delay variation was measured using IEEE 1588 packets as probe packets at a rate of 64 Hz. The setup is depicted in Figure 19.



Figure 19. Network including five carrier-class switches with traffic generation.

Traffic was generated in the forward direction, distributed across two ports, so that each switch was subject to 40% bandwidth utilization. In all cases, traffic was generated as bursts. The frame size, burst length, gap between bursts (interburst gap), and distribution of load between the two ports were the parameters adjusted. The five distinct scenarios are shown in Figure 20.

	<b>Total BW</b>	BW Port 1/2	Frame Size	<b>Burst Length</b>	Interburst Gap
1	40%	20% / 20%	Different	200000	Same
2	40%	20% / 20%	Same	200000	Same
3	40%	20% / 20%	Same	50000	Same
4	40%	20% / 20%	Same	50000	Different
5	40%	21%/19%	Same	50000	Different

Figure 20. Traffic generation parameters for five setup scenarios.

The packet delay sequence phase measurement for the full test including all five scenarios is shown in Figure 21. The five are clearly distinguishable, with each exhibiting certain signature features. Most striking is the floor with change in parameters, particularly the elevated floor in scenario number two. The periodicity present in some of them is apparent. The greatest amount of randomness appears to be in the fifth scenario, indicating that for this particular traffic generator, splitting the load into nearly equal, but not exactly equal, parts helps to reduce artificial periodicities.



Figure 21. Packet delay sequence phase measurement for five traffic generator setup scenarios.

### LIVE NETWORK PACKET TIMING MEASUREMENTS

#### **PRODUCTION NETWORK WITH DSL ACCESS**

Over a period of several weeks, a live production network with DSL access (Figure 22) was studied using 1588 packet delay measurement equipment. With a master at one end and a hub-connected slave and probe at the other end, one-way packet delay measurements were made in both directions. This was of particular importance in this case, since DSL leads to asymmetrical packet delays (see Figure 23). An important capability made use of here and for other live network measurements, new to 1588v2 and not present in 1588v1, is the option for unicast transport. For traversing DSLAM's and DSL modems, unicast rather than multicast transport is essential.



Figure 22. Production network spanning 500 km with DSL access.

Plots of packet delay versus time show periodic variations of traffic. A traffic level increase is seen as increased PDV occurring approximately every 30 minutes (see Figure 24). Such network activities as the regular collection/download of SNMP data can account for this. As another example, it was noted in other measurements made on this network (not shown here) that FTP transfer of large files had a great effect on PDV. Running the phase power spectral density calculation (the lower plot in Figure 24) clearly shows the periodicity as well. It is a useful tool, particularly when the phase plot doesn't show the periodic nature of the data as clearly as it does here.



Figure 23. Asymmetrical packet delay with DSL.



Figure 24. Packet latency (20 hr) and phase power spectral density analysis show periodicities.

Over this same network, a study with network probes capable of using arbitrary length probe packets was used to determine the effect of packet size on packet delay. As might be expected, short packets traverse the network more quickly than longer packets. This is shown in Figure 25.



Figure 25. Short packets (blue) vs. long packets (red).

#### **METRO ETHERNET NETWORK**

A second live production network studied for its packet delay characteristics was configured for layer 2 transport. This network is shown in Figure 26.



Figure 26. Metro Ethernet network with layer 2 transport.

Comparison of the delays seen in the DSL network (22 to 40 msec; see Figure 24) with those seen here (120 to 700  $\mu$ sec; see the first plot in Figure 27) show this network to be considerably faster. During a 24-hour period, network conditions were seen to change, with maximum traffic occurring during business hours, between 8AM and 6PM (Figure 27). The lower plot in Figure 27 shows standard deviation over time and more clearly shows the decrease of traffic as the night progresses.



Figure 27. Packet delay vs. time and standard deviation vs. time for live production IP network over a 24-hour period.



Figure 28. IEEE 1588 slave performance during a 24-hour period.

During this same 24-hour period, the IEEE 1588 slave clock was also measured. It proved to cope with the changes in network conditions very well. While the effects of the increased traffic load between 8AM and 6PM were visible (see the top plot in Figure 28), the 1588 slave clock achieved levels of performance nearly equal to those required of primary reference clocks [3]. Therefore, the 1588 slave clock meets ITU-T packet network performance requirements (see the bottom plot in Figure 28) by a large margin [4]. Note that the requirement is defined in terms of MTIE; while the utility of the MTIE calculation is reduced in the case of direct application to packet delay data, MTIE analysis is useful for the analysis of clocks, especially those driven with a servomechanism.

### COMPANY LAN

An enterprise network shows greater packet latency and far greater packet delay variation than the layer 2 production network just discussed. The 6-day measurement (see Figure 29) shows packet delays ranging from 335 µsec to 5.08 msec. As might be expected, clear diurnal cycles are seen, with peak traffic increasing as the work week progresses, traffic levels relatively quiet over the weekend, and for all days increased traffic starting at 8AM. The distribution is one-sided, with a long tail to the right (see the bottom plot in Figure 29).



Figure 29. Enterprise LAN packet delay sequence and histogram.

### **PUBLIC INTERNET**

A WAN measurement shows the basic characteristics of the LAN measurement except the latency and packet delay variation increases by several orders of magnitude. As shown in Figure 30, a 1588 grandmaster clock was placed in Texas and a 1588 slave connected to a 1588 probe with a hub was placed in California.

A measurement lasting four days shows diurnal cycles of traffic load during the work week as the LAN measurement did, but unlike the LAN measurement, any load changes that might have occurred over the

weekend are difficult to decipher (Figure 31). Packet delays ranging from 29 to 471 msec were observed; recall that packet delays ranged from 0.335 msec to 5.08 msec for the 6-day LAN measurement (Figure 19).



Figure 30. Public internet WAN IEEE 1588 packet delay measurement setup.



#### Packet delay: 29 to 471 msec

Figure 31. Public internet WAN packet delay sequence and histogram.

The WAN packet delay distribution, like the LAN distribution, is a one-sided distribution with a long tail, but interestingly, as the zoomed view shows clearly, there are local peaks near to the floor (see the lower plot in Figure 31).

### REFERENCES

- [1] L. Cosart, 2007, "Precision Packet Delay Measurements Using IEEE 1588v2," in Proceedings of the 2007 International IEEE Symposium on Precision Clock Synchronization (ISPCS) for Measurement, Control and Communication, 1-3 October 2007, Vienna, Austria, pp. 85-91.
- [2] ATIS-PP-0900101.2006, "Synchronization Interface Standard," Annexes C and D, 2006.
- [3] ITU-T Recommendation G.811, "Timing Characteristics of Primary Reference Clock," September 1997.
- [4] ITU-T Recommendation G.8261, "Timing and Synchronization Aspects of IP Networks," May 2006.
- [5] G. Zampetti, 2007, "Definition of Minimum TDEV (minTDEV)," Contribution #401, Contribution to ITU-T Q13/SG15, Geneva, Switzerland, June 2007.
- [6] L. Cosart and G. Zampetti, 2008, "*Definition of Band TDEV (bandTDEV)*," WD 68, Contribution to ITU-T Q13/SG15, Rome, Italy, September 2007.
- [7] L. Cosart, 2008, "Active Packet Measurement Probe," WD 67, Contribution to ITU-T Q13/SG15, Rome, Italy, September 2007.
- [8] L. Cosart, 2008, "*Effects of Different Approaches to Traffic Generation on PDV*," WD 69, Contribution to ITU-T Q13/SG15, Rome, Italy, September 2007.