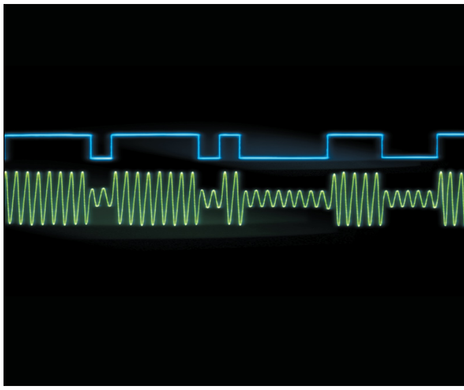
Symmetricom®

Deployment of Precision Time Protocol for Synchronization of GSM and UMTS Base Stations



White Paper

Table of Contents

Deployment of Precision Time Protocol for Synchronization of GSM and UMTS Base Stations

Abstract

This paper is designed to help network engineers, network planners, and network operations understand how to deploy Precision Time Protocol (PTP, standardized as IEEE Standard 1588-2008 [reference1]). PTP is a next generation, packet-based timing protocol targeted for use in asynchronous network infrastructures based on packet transport technologies. This paper specifically focuses on the synchronization requirements for wireless backhaul applications across native Ethernet based networks within the UMTS/GSM mobile wireless environment. It discusses the relevance of PTP within this paradigm, and describes some of the considerations that have to be taken into account for deployment of PTP into such a network. The paper also discusses some of the advantages and limitations of packet-based timing technologies, with specific reference to PTP.

Introduction

This paper is one of a series of White Papers and Application Guidelines produced by Symmetricom as part of an overall Framework for Synchronization and Timing in the Next Generation Network (NGN). These papers are intended to help service provider network architects, planners, and engineers design and deploy stable, robust synchronization and timing architectures to support applications and services that will be deployed on the NGN.

This document specifically addresses the application of the Precision Timing Protocol to the frequency synchronization of GSM and UMTS-FDD base stations (i.e. keeping all base stations running at the same frequency). It covers the synchronization requirements, types of base stations, and guidelines for the deployment of PTP in the GERAN and UTRAN (i.e. the GSM and UMTS radio access networks). The document concentrates on the case where the base station is connected over an Ethernet access technology.

Future documents will address other related applications such as CDMA and CDMA2000 base stations, and WiMAX base stations, all of which require time synchronization in addition to frequency synchronization. They will also address the operation of PTP over alternative access technologies, such as xDSL and GPON.

GSM and UMTS Base Station Synchronization

One of the most common applications currently being cited for packet timing technologies such as PTPv2 is for the synchronization of various wireless telephony and data services, e.g. GSM, UMTS, CDMA, WiMAX etc. These are gradually transitioning from a TDM-based backhaul network to a packet-based network. This white paper examines the synchronization implications of the shift from TDM to packet network, and the consequences for the synchronization requirement.

GERAN and UTRAN Architecture

Figure 1 shows the architecture of the UTRAN (UMTS Transport Radio Access Network). The RNC (Radio Network Controller) is connected to the Node B (UMTS Base station) over an interface called the IuB. The GERAN (GSM/EDGE Radio Access Network) is similar, except that the base station interface is called the Abis.

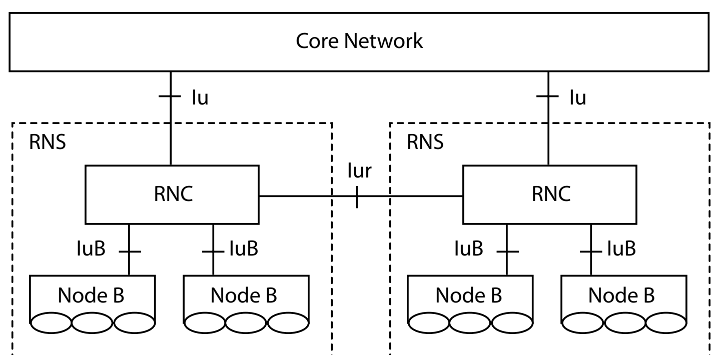


FIG 1 UTRAN Architecture [Fig. 4 from 3GPP TS 25.401, reference 2]

Typically, the Abis and the IuB interfaces have been TDM based, e.g. E1 or T1 interfaces. However, these are increasingly expensive when compared to packet networks, such as Metro Ethernet, or high speed DSL. Secondly, with the increase in data services to mobile devices, TDM based backhaul connections are not sufficiently scalable to keep up with the new bandwidth demands.

A third driver is the deployment of 3rd. generation UMTS Node Bs alongside GSM base stations. If the Node B is going to require a packet interface, it reduces the operating cost to eliminate the TDM connection to the GSM base station and run the GSM backhaul over the packet network.

Synchronization of Base Stations

The problem with eliminating the TDM interface is that this is often used as a source of synchronization for the base station itself. In order to permit correct handover between adjacent base stations in the presence of Doppler shift generated by a moving mobile handset, the RF frequency at a GSM or UMTS base station must be accurate to within 50ppb (parts per billion) of the nominal frequency at all times (3GPP TS 25.402, section 4.2, [3]).

Typically this frequency is derived from the knowledge that the TDM input clock will be traceable back to the wireline carrier's PRC (Primary Reference Clock). Over the long term this makes an extremely accurate reference, better than 1 part in 10^{11}. However, it may vary over the short term, and hence a PLL is used to filter this and ensure that the input to the RF circuits is well within the 50ppb requirement. Typically the output of the PLL will be stable to within 16ppb (according to various base station manufacturers, see also G.8261, Appendix IV.2.3, [5]). To achieve this, the reference input to the base station PLL must also be stable to better than 16ppb over the time constant of the filter.

These requirements are shown in Figure 2. There is no exact specification on the network input frequency other than long term traceability, since that will depend on the manufacturer's implementation of the base station PLL. Some manufacturers filter the network input heavily to ensure the RF frequency is kept stable; certain others rely more on the network input stability.

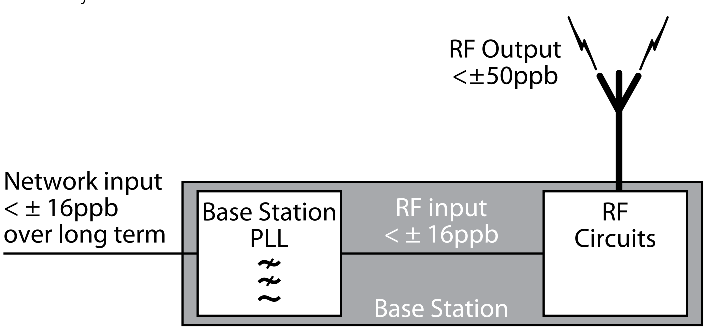


FIG 2 Frequency Accuracy Requirements at the GSM or UMTS Base Station

When the TDM backhaul is replaced by a packet network, the synchronization requirement must be met by some other means. The operation of a packet timing technology such as IEEE1588 PTP [1] is one such possible means.

Types of Base Stations

There are several types of base stations available to a mobile phone operator. These first is the large "Macro base station" intended to serve a conventional cell. More recently, vendors have introduced "micro" or "pico" base stations. These are intended to enhance coverage on a small campus or inside a large building where signal penetration is weak. Finally, some vendors are now proposing a "femto base station". The object of these is to serve a single house or residential unit, backhauling across the owner's internet connection, and removing the need for a fixed line phone.

Macro Base Station

The key feature of the macro base station is that it requires a high-capacity backhaul, especially with advent of HSDPA (High Speed Download Packet Access) and HSUPA (High Speed Upload Packet Access). Therefore it is a good candidate for a Metro Ethernet backhaul right out to the base station.

The central base station PLL typically uses an ultra-stable DOCXO, filtering the input reference down to ~100μHz (according to some manufacturers). This can remove most frequency transients, relaxing the requirements on the input clock reference.

The frequency accuracy requirement at the RF output is 50ppb or better to permit handoff of calls between sites in the presence of Doppler shift caused by a handset moving at up to 250km/h. This leads to a long-term requirement of 16ppb on the reference interface Macro base stations connected over a Metro Ethernet network are good candidates for the use of PTP or other packet timing technologies to meet the synchronization requirement.

Micro/Pico Base Station

Micro or Pico base stations are typically deployed on a campus or in a building, e.g. at a large enterprise site. If it supports packet access (e.g. HSDPA or HSUPA), a high capacity backhaul is required. Therefore the base station may be connected via Ethernet, although it is more likely to be deployed on a lower-cost access infrastructure, e.g. xDSL, GPON or EPON.

Normally the central base station PLL still uses an ultra-stable DOCXO, although some manufacturers are trying to reduce the PLL and local oscillator cost here. Since on a campus or within a building, handsets are extremely unlikely to be moving at 250km/h, the frequency accuracy can be relaxed to 100ppb at the RF output. This reduces the constraint on the input reference to around 33ppb.

Micro or Pico base stations connected over a Metro Ethernet network are good candidates for the use of PTP. If connected over alternative access infrastructure, such as xDSL or GPON, the performance is degraded because of the characteristics of the physical layer, although the relaxation of the performance requirement does help.

This version of this document concentrates on the operation of PTP over an Ethernet access network. Further characterization of the performance of PTP over alternative access technologies such as xDSL is required, and will be addressed in future versions of this document.

Femto Base Station

These are small devices with a very short reach, deployed within individual houses or residential units. They are operated over consumer-grade access infrastructure, e.g. ADSL.

For these devices, the central base station PLL is typically integrated into the PTP slave itself. The local oscillator is a moderately stable TCXO at best due to cost constraints. Handsets are slow-moving, and if handoff to the local macrocell is not a priority, the frequency accuracy at the RF output can be relaxed to 250ppb, reducing the constraint on the input reference.

Further characterization of the performance of PTP over access technologies such as ADSL is required, and will be addressed in future versions of this document.

Summary Requirements

A summary of the requirements for GSM/UMTS base stations is shown in Table 1.

Base Station Type	Frequency Accuracy		Local Oscillator	Probable Access Network
	RF Output	Reference Interface (long term)		
Macro	±50ppb	±16ppb	DOCXO	Ethernet
Micro	±100ppb	±33ppb	DOCXO or OCXO	xPON or xDSL
Femto	±250ppb	To be established	TCXO	ADSL

TABLE 1 Synchronization Requirements for GSM/UMTS Base Stations

Packet Timing Systems

The Nature of Packet Timing

A conventional timing signal is normally a periodic digital signal, where the edges of the signal are the reference points in time, used to control the timing of operations in a digital circuit. These edges are known as "significant instants" in time. Timing jitter and wander causes these significant instants to vary slightly from their ideal position in time, i.e. they may not occur at precisely equally spaced points in time. A conventional timing signal is shown in Figure 3.

A packet timing signal is similar in concept. While the transmission medium is different (packets on a network as opposed to signals on a wire), the packets still contain significant instants (normally the front edge of the packet), with a defined ideal position in time. The variation of the significant instants around their ideal position is termed "packet delay variation" (PDV). This is shown in Figure 4.

Some packet timing signals may be periodic (e.g. circuit emulation packets containing constant bit rate data), and for these the ideal position in time is implicitly given by the packet rate. Other packet timing signals are not periodic (e.g. PTP or NTP), and for these the ideal position in time is given by a timestamp embedded in the packet data.

Precision Time Protocol (PTP, IEEE1588)

The IEEE1588 Precision Time Protocol (PTP, [1]) enables the accurate distribution of time and frequency over a packet network. It was originally introduced to synchronize networked computer systems by using a master reference time source (or "server") and a protocol by which slave devices can estimate their time offset from the master time reference. It achieves this by sending a series of time-stamped messages between the central time server and the slave devices. Over a suitable, well-designed network, it is capable of achieving time accuracy of better than 1 microsecond, and frequency accuracy of better than 10ppb.

A more in-depth description can be found in the companion paper, "Synchronization for Next Generation Networks: NGN Synchronization and Timing Technology" [9].

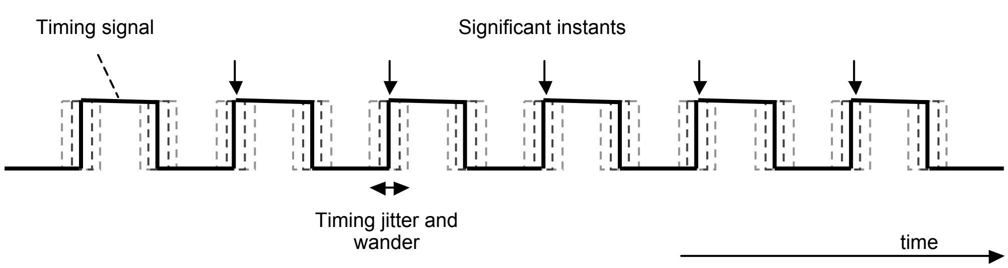


FIG 3 Conventional Timing Signal

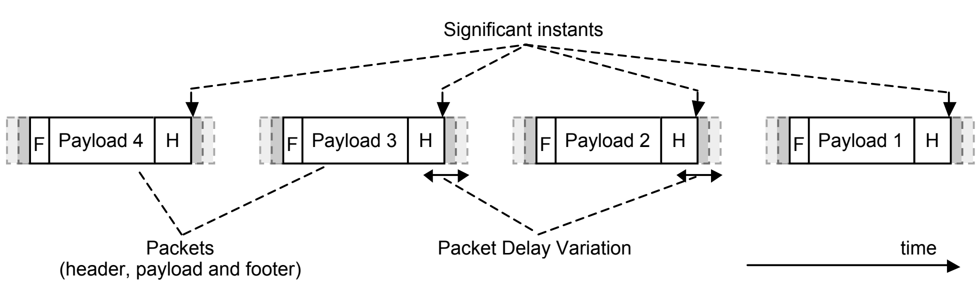


FIG 4 Packet Timing Signal

PTP version 2 (IEEE 1588-2008)

PTP was originally designed for the industrial automation and test and measurement industry, and has been in use there for several years. The standard was ratified in 2002 by the IEEE. In 2005 a new project was started to revise the standard, both to improve performance in its original intended application space, and to allow it to be used in telecom applications. PTP version 2 was approved in early 2008.

The main changes from version 1 to version 2 are:

- Shorter message formats allowed
- Higher update rates allowed
- Specified transport over more network layer protocols, including UDP/IPv4, UDP/IPv6 and Ethernet
- Defined a security protocol for PTP (experimental)
- Defined the "transparent clock" concept, a means of compensating for the message delay through network elements

The Telecom Profile

Another concept introduced in PTP version 2 is the "PTP Profile". This was born out of the recognition that different applications needed different options and subsets of the full PTP protocol. Therefore it was decided to allow standards organizations or industry bodies to create profiles for a specific application or groups of applications.

The "Telecom Profile" is currently under development by the ITU-T (Study Group 15, Question 13), to define the characteristics required in the telecommunications industry. Symmetricom is leading its development by co-authoring contributions to the relevant groups on the contents of the profile (see [6], [7]).

Issues affecting Timing Performance in Packet Networks

There are several phenomena in packet networks that can affect the performance of packet timing algorithms such as PTP. These include:

- Packet Loss
- Packet Error
- Extended Packet Loss
- Packet Delay
- Packet Delay Variation

Packet Loss

Packet loss is not an issue for packet timing protocols such as PTP, because the slave servo mechanisms integrate over several seconds' worth of data. The timestamps are all relative to a stable time references, not to the previous packet, therefore the slave can simply wait for the next packet to arrive to obtain the information it requires. Therefore the loss of an individual packet or even a group of packets will have little effect on the clock recovery performance.

Packet Error

Bit errors or corruption in the packet normally results in the packet being discarded due to a bad checksum or frame check sequence value. It is extremely unlikely that a packet with one or more bit errors will pass both the CRC check on the frame check sequence value, and the UDP or IP checksum tests. Discarded packets are treated as lost packets, and hence have little effect on the clock as described above.

Even if it does pass these two tests, the error would have to be in the timestamp field to affect the clock recovery process. The servo algorithm is likely to reject a corrupted timestamp value as being outside the expected range. If the errored timestamp does fall within the expected range, it is typically averaged with other timestamps, reducing still further the effect it may have on the clock itself.

Extended Packet Loss

Network outages may give rise to an extended period of packet loss, such as a temporary outage or period of congestion. If this occurs, the clock servo at the slave must go into a holdover mode, as with conventional clock mechanisms when the source is lost. This enables it to ride out the outage until the network is restored.

Path protection mechanisms, such as IP re-routing, should in general cause the path to be restored quickly, enabling the servo to re-lock without any degradation of the clock accuracy.

Packet Delay

The delay through a packet network can be several milliseconds, which is larger than in many traditional synchronization networks. However, this is still small compared to the filter bandwidths typically employed in the slave servos. Therefore, this increased delay has no effect on the accuracy of the clock.

Packet Delay Variation

This is the main issue affecting the accuracy and stability of slave clocks when using packet timing protocols such as PTP. The variation in delay from packet to packet through the network induces noise in the slave's perception of the time at the master. Constant delay would cause a fixed offset, however, variable delay causes a varying estimate of the offset. The performance of the slave is affected by both the magnitude of this variation, and how effective the slave's filter is at removing this noise.

Packet delay variation (PDV) is caused both by the network elements themselves (e.g. switches or routers), the physical network layer, and even the topology of the network. It increases with the size of the network, and tends to be correlated to network load, i.e. if the amount of traffic in the network increases, the delay variation is also likely to increase. A detailed discussion of the different causes of packet delay variation is given in the companion white paper "Synchronization for Next Generation Networks: NGN Synchronization and Timing Technology" [9].

PTP Deployment Guidelines

The deployment guidelines have been split into two sections. "Engineering Guidelines" deals with parameters that can be controlled at the network or system design stage, and need to be considered in advance of deploying a packet timing system. "Operational Guidelines" deals with parameters that can be controlled once the packet timing system is in operation.

Engineering Guidelines

Planning a Packet Synchronization Network

As noted in the previous section, the packet delay variation increases with both the size and loading of the network. Therefore, it is important both to place the PTP grandmaster as near as possible to the slaves, and to keep the loading on the network as low as possible. In practice, this is generally achieved by locating the PTP grandmaster at the edge of the network, as shown in Figure 5:

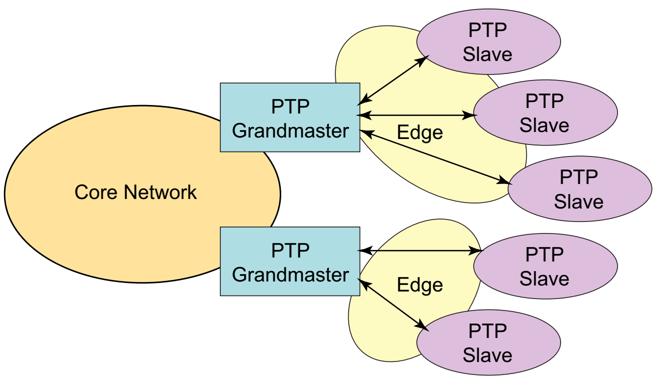


FIG 5: Distribution of Masters to the Edge

As a basic "rule of thumb", Symmetricom recommends that for effective deployment of PTP, the span of the network is limited to no more than four switches or routers, with traffic loads of less than 80%.[1] Longer spans will require either reduced traffic loads or more stable local oscillators (or both). This rule only applies to Ethernet networks; it does not apply when the synchronization link is operated over alternative physical layers such as xDSL or xPON.

1 This is merely a rule-of-thumb for planning purposes, it is not intended to be a guarantee of operation. The purpose is to provide guidance as to appropriate locations for the PTP grandmaster. Since different switches and routers perform differently, these locations may have to be adjusted based on the results from field trials.

The synchronization planning process can be broken down into the following steps:

- Step 1: Identify the PTP slave locations
- Step 2: Identify suitable locations for the PTP grandmaster, as near as possible to the slave locations
- Step 3: Check that the grandmaster has sufficient capacity to serve all the slave devices in the vicinity, and that the "four switch" rule is maintained with the chosen locations, adjusting the grandmaster distribution where appropriate
- Step 4: Field trial and monitoring – measure the performance of both the output clock and the network PDV to verify that the network is suitable and that the chosen locations are correct
- Step 5: Ongoing monitoring of critical or selected links to ensure the synchronization quality is maintained

Measuring Network Suitability

The suitability of the network for distributing packet timing varies with a number of different parameters. These include the types of switches being used, the performance of the slave device, and the stability of the local oscillator. For example, some switches may add more delay variation than others due to the way they have been designed. Therefore it is important to measure the suitability of the network once the initial planning process has identified likely sites for the PTP grandmaster,

A new metric is required to quantify the suitability of a packet network for time and frequency distribution, once the rule of thumb has been used to size the network. Symmetricom has developed a metric called "Minimum Time Deviation" (MinTDEV) for this purpose, and is working with the ITU to standardize the approach (see Appendix V of G.8261 [5] for a full definition). MinTDEV is calculated from a set of packet delay values, and enables the performance of the output clock to be predicted from the packet delay variation of the network. The metric allows a mask to be calculated (similar to the masks developed for the MTIE and TDEV metrics), showing the boundary of acceptable performance of the packet network. The mask is independent of the number or type of switches or network elements, and shows quickly whether the network is fit for purpose.

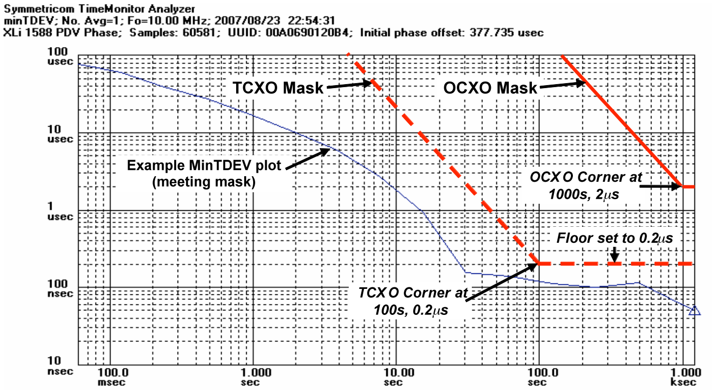


FIG 6 MinTDEV Mask for GSM/UMTS Frequency Accuracy

The MinTDEV mask is appropriate to determine whether a PTP slave is going to be capable of meeting the GSM or UMTS frequency accuracy requirement is shown in Figure 6. Two masks are shown dependent on the local oscillator in the slave, since this affects the time constant of the slave's filter. The derivation of the mask is described more fully in Appendix 1.

Application to Alternative Packet Network Access Technologies

The above discussion has concentrated on an Ethernet-based packet network. However, when the packet timing flow has to go across other types of networks, the physical network layer can affect the characteristics of the packet delay variation in different ways. The companion paper, "Synchronization for Next Generation Networks: NGN Synchronization and Timing Technologys" [9] describes this in more detail.

For these networks such as the various types of DSL links, GPON and EPON, the MinTDEV metric may not be the best metric to use, and an alternative characterization technique may be required. Similarly, the "4 switch" rule only applies over Ethernet networks. At present, Symmetricom does not recommend the operation of packet timing technologies over alternative networks without careful evaluation of the conditions and testing of the devices to be used.

Grandmaster Performance Constraints

The budgeting process described in the Appendix yields a performance constraint on the grandmaster clock of a fractional frequency offset of no more than 0.5ppb. This is satisfied by locking the grandmaster to a primary reference source such as a G.811 or Stratum 1 clock, which have a frequency accuracy of better than 1 part in 10^{11}.

The grandmaster may introduce temporary inaccuracies into this through quantization errors in the timestamps. The slave device must be capable of filtering this noise out, in addition to the network noise. Over a 100s period (e.g. as appropriate for a TCXO), a phase offset of 50ns results in a frequency error of 0.5ppb. This implies that the maximum timestamp error should be less than 50ns. When operating into a slave using a 1mHz filter, the timestamp accuracy can be permitted to be up to 500ns.

Redundancy Strategy

IEEE1588 version 2 describes an algorithm for a slave to determine the best master within its field of view. The slave then chooses this as its "grandmaster", or the master it is going to synchronize to. The algorithm is called the "Best Master Clock Algorithm", and it is dynamic, allowing the slave to switch to an alternative master if the original master fails or is excessively masked by network noise.

However, in some circumstances, operators may choose to define an alternative algorithm, or to manually configure slaves to synchronize to a specific grandmaster device. For example, some operators may not want to give the freedom to slaves to autonomously choose between masters. It may be better to manually configure slaves, and then instruct all slaves to switch to the same alternative master in the event of a failure, rather than potentially having each slave listening to different masters.

No one method can be said to be the better than another. The strategy chosen is dependent on the operator's preference for managing the synchronization network, and needs to be considered at the engineering planning stage. In either approach, the grandmaster must have sufficient capacity to cope with the additional load caused by a fail-over.

Operational Guidelines

Frequency of Timing Messages

The frequency of timing messages can be adjusted dynamically to adapt to changing conditions in the network. The required frequency is dependent on several factors, for example the performance of the slave device, the stability of the slave's local oscillator, and the amount of noise in the network.

In general, doubling the number of timing messages does not double the reach of the network. If possible, it is better to manage the traffic load rather than increase the frequency of timing messages.

As a general guideline, Symmetricom recommends using 64 sync messages per second. The number of delay_request messages required by the slave to fix the time offset is dependent on the slave implementation. Most slaves use the same number of delay_request as sync messages, and Symmetricom recommends this setting. A few slaves make primary use of the sync messages for achieving frequency lock, and then use a much reduced number of delay_request messages solely to fix the time offset. This reduced setting is not recommended unless specifically advised by the slave device manufacturer.

Quality of Service (QoS)

Carrier-class switches and routers are often designed with many options for addressing quality of service. These may include priority management, bandwidth reservation, load balancing, traffic policing and shaping, etc. These tools are invaluable in order to differentiate between traffic classes, and to optimize performance for each traffic class, especially in situations where limited bandwidth is available.

However, some care must be taken in the application of QoS techniques. Some of the more complex schemes may impede the raw performance of a device in order to improve the overall performance of a network. A simple analogy is that of road traffic control – these even out the delays to road traffic across all users, but they do that by causing additional delay to some road users while other users are given a turn, for example at a traffic light controlled intersection. In doing so, the overall capacity of the road network is somewhat reduced from the maximum possible.

The simplest, "low-touch" schemes such as strict priority (SP) are generally most appropriate for PTP traffic. Complex, "high-touch" schemes consume more computational resources within the network element, which can cause queuing or delay while waiting for these resources. The actual scheme used to manage particular traffic classes is dependent on the implementation of the network element.

Symmetricom recommends setting the following QoS levels for PTP traffic:

- Diffserve Expedited Forwarding (EF) class
- IEEE 802.1 p-bit marking of 5 or above
- UMTS conversational class (see 3GPP TS23.107 [4], normally mapped to a "p-bit" of 5)

In particular, it is best to place the PTP traffic in a different traffic class to bandwidth-intensive applications such as video streaming.

Unicast vs. Multicast

PTP was originally intended as a multicast protocol. There seemed little point in sending individual streams of sync messages to each slave, so a multicast model appeared to be more efficient. However, in the telecommmunications network, this is not such an obvious choice. Unlike the closed, controlled, single purpose industrial networks that PTP was originally designed for, telecommunications networks have to handle data from all sorts of different applications, and the use of unicast messages is more appropriate.

There are several reasons why the use of unicast can increase performance

- **Packet Replication –** when a packet is multicast through a network, it needs to be replicated at each network element where it exits on multiple ports. This replication process takes time, and may add to the delay variation experienced by the packet in its journey from server to slave.
- **Priority –** in a telecommunications network environment, multicast traffic may often be blocked in the upstream direction (e.g. for security or operational reasons). This prevents the use of PTP delay_request messages when operated in multicast. In the downstream direction, the increase in multicast traffic for applications such as broadcast video streaming means that the amount of bandwidth and priority allocated to these traffic types is often limited to avoid bringing down the rest of the network.
- **Slave resource limitation –** with a multicast model, every message transmitted has to be examined by every device in the multicast group. This means that the slaves end up listening to all the delay_request and delay_response messages produced by or for other slaves, leading to the slave's processor being saturated by passing messages up the protocol stack that it then throws away.
  Slaves are supposed to be the lowest-cost elements of the synchronization eco-system, and hence it is important to minimize the amount of processing power required. While this can be solved by adopting multicast solely for sync messages, or building hardware into the slave to throw these messages away before they reach the protocol stack, both of these increase the complexity of either the slave or the overall system.

Therefore, Symmetricom recommends that unicast transmission be used at all times for telecommunications applications.

Network Performance Metrics

Symmetricom is leading the development of new metrics to quantify the performance of the network, such as Minimum Time Deviation (MinTDEV). Such metrics need to be continuously monitored, and the operator needs to know how to control the network performance to maintain them within the budgeted performance.

At present, the primary means of controlling the MinTDEV is to reduce the amount of traffic within the network using admission control on the non-PTP traffic to manage the load. Above 80% traffic loads switch and router performance can degrade significantly, and becomes very dependent on implementation. Therefore the network should always be operated below this "knee point".

Conclusion

PTP Application

With the migration from TDM to NGN, the challenge for the network operator, the network planning engineers, and the network element vendors, is to be able to provision a packet-based frequency delivery of the quality, accuracy, and consistency that enables time and delay sensitive applications to be delivered with equal or better quality than those available today. To meet such stringent requirements these services have to be delivered with carrier class availability and reliability, and with rich measurement, diagnostic, and management features that fit the operational model already established in service provider operations centers.

Packet-based networking is now entering a new phase; best-effort data is no longer the only service offered, and high QoS is now considered fundamental to the operation of robust services and applications. The delivery of synchronization and time using a packet protocol such as PTP is such a service. Engineering PTP will significantly change the way that service providers deploy and manage both frequency syntonization (synchronization) and time services.

Mobile wireless operators and vendors of wireless network elements such as base stations (Node B or BTS) and the Radio Network Controllers (RNC) are in the forefront of investigations into PTP because of the compelling economics of a move away from E1/T1 TDM transport to packet-based transport for wireless backhaul. The move to Ethernet transport is also a catalyst for change in the synchronization instances that enable networks to deliver real time and mobility services. The objective of this document therefore has been to examine the application of PTP to GSM and UMTS wireless backhaul, the most cost sensitive part of the Mobile Network, and the domain in which PTP plays a critical role.

Key Considerations for Deployment of PTP

Several conditions must exist for PTP to be deployed as the synchronization technology. The first is that the underlying network is built on a packet-based transport technology, such as native Ethernet, without embedded synchronization or time services available. Secondly, deployment of PTP requires that the access nodes or end stations should be enabled with PTP slave functionality. A third condition is that the network is to provide mission critical real-time services and applications; mobile network services such as wireless backhaul impose stringent frequency and time requirements on the underlying transport as we have seen above. Finally, the underlying network architecture is also important in order to determine the placement of the PTP grandmaster Clock (GM) servers.

The location and distribution model of the GM is critical for the overall accuracy, consistency, and cost of the synchronization service and the choice made here can seriously impact the performance of the network. To determine exactly how and where GM servers are deployed, the network planner has to take into factors such as potential network load, congestion – especially at the aggregation points of the network – and the performance of the individual network elements. The Minimum Time Deviation (MinTDEV) analysis shown in this document enables an explicit evaluation of the tolerable noise budget on the access and aggregation links. This is a major benefit in determining the placement of the GM servers with reference to this overall network performance.

To avoid excess accumulation of delay or packet jitter in propagation of PTP, it is advisable to deploy the grandmaster servers as close as possible to the edge devices. However, the actual hop count will be determined by factors other than just the underlying transport technology. In addition to the overall network load and congestion state these include the efficiency of the network elements on the packet path, the stability of the local oscillator on the final access device, and finally the quality of the PTP slave servo algorithms.

Grandmaster Clock server placement will also be affected by innate scale factors such as the CPU performance limitations of the grandmaster servers, and reliability considerations. For the latter, the critical importance of synchronization and timing services implies that the network architect will require redundant grandmaster each of which is able to provide consistent PTP service to the slaves. Cautious engineering will ensure the GM is deployed at the point where it is most effective and least risk, most probably at the aggregation point nearest the fan out to the end stations. Thus to ensure carrier class availability of the PTP GM the network planner must carefully evaluate the number of active slaves per server under various failure conditions (capacity planning), the redundancy architecture of the servers, and the quality and type of network element in terms of impact on propagation of PTP.

To understand how PTP will add value in this environment, it is important for network planning, operations, service engineering, and cost-benefit analysis, to have an evaluation tool that will enable the operator to predict the performance of the network whatever the underlying transport mode. The deployment of packet-based networks and the migration to NGN has therefore introduced a need to analyze the performance of synchronization and timing instances in a network in a different way from the methodologies used in TDM networks.

New Metrics

The analytical method presented in this document, Minimum Time Deviation, is a new and highly effective way of analyzing the service capability of a network by looking at the underlying noise budget and its impact on the transport of a synchronization or time service. Although the focus of this document is to discuss packet-based wireless backhaul for GSM/UMTS over native Ethernet transport, the technique can be applied to every network element, to every transport technology, and in every network domain. It introduces an innovative and fundamental parameter that enables the operator to characterize a network in terms of applications and services supported by the deployed synchronization architecture. Moreover, MinTDEV is not just applicable to packet networking it is equally applicable to TDM transport and enables comparisons between TDM and packet-based access in environments such as wireless backhaul.

About Symmetricom and PTP

Symmetricom is a global leader in innovating, architecting, and delivering synchronization and time solutions to the networking industry. Symmetricom PTP solutions exploit and rely on this well established leadership and on the capabilities of the Symmetricom carrier-class SSU platforms, TimeHub and SSU 2000, and the TimePictra suite of management tools. The Symmetricom carrier-class PTP blade can be simply deployed into any existing Symmetricom SSU, and immediately begin to provide a rich suite of PTP grandmaster services. Redundancy of a mission critical service is a fundamental operational parameter and is enabled in Symmetricom SSUs by the simple deployment of a second PTP card to enable redundancy at card, link, and port level. Integration of the management layer of Symmetricom PTP cards is seamless and provides a rich feature set enabling the service provider to ensure rapid service provision to the PTP slaves installed on the network.

As the industry leader in this domain, Symmetricom has taken the initiative to deliver a suite of advanced synchronization solutions, which includes PTP, to the networking industry under the Framework for Synchronization and Timing in Next Generation Networks. The Framework not only outlines a methodology for analyzing and understanding the different time and synchronization technologies available, including legacy (TDM), NGN physical layer, and packet-based implementations, but it also determines the inter-working scenarios for these different technologies.

Symmetricom has taken a leading role in the development of new NGN metrics and analytical tools such as MinTDEV that will allow network planners and synchronization experts to drive coherent network synchronization into NGN systems.

Symmetricom PTP is a leading-edge best-in-class implementation of this new standard for synchronization. It leverages Symmetricom's well established experience in this domain, and adds new and vital features that facilitate and enhance NGN services.

Appendix 1 Derivation of the Minimum Time Deviation Mask

As described above, the principal effect in packet networks that affects the performance of packet timing protocols such as PTP is variation of packet delay. As each network element and each network segment introduce some variation to the delay experienced by a flow of packets, the main way to control this effect is to limit the span of the network over which the protocol is deployed, and the amount of traffic in that network.

Empirical Behavior

A graphical way to show how the size of the network and the amount of traffic affects the stability of the packet timing slave is given in Figure 7, which illustrates the "operational area" for a network. This graph was based on both characterization of the network PDV and the performance of a particular slave device, measured across a network with varying numbers of switches and traffic loads.

From this graph, a basic "rule of thumb" can be developed. Symmetricom recommends that for effective deployment of PTP, the span of the network is limited to no more than four switches or routers, with traffic loads of less than 80%. Longer spans will require either reduced traffic loads or more stable local oscillators (or both).

The problem with this "rule of thumb" type of approach is that the operational area varies with a number of different parameters. These include the types of switches being used, the performance of the slave device, and the stability of the local oscillator. There is no way of calculating where the boundary might be, other than by empirical means through observation and measurement. Even these measurements are only valid for the network it is tested over.

Therefore a new metric is required to quantify the suitability of a network for time and frequency distribution once the network has been sized using the rule of thumb developed from the empirical data.

Characterization of PDV

Symmetricom has developed a metric to characterize the delay variation of a packet network called "Minimum Time Deviation" (see Appendix V of G.8261, [5]). In essence, most slave servo algorithms make use of the fact that the fastest packets traverse the network at an approximately constant rate (an observation made in the development of the NTP specification back in 1989, see RFC1129 [8]).

For example, Figure 8 shows histograms of packet delay measured through a 10-switch network at different loads. The highlighted peaks show those packets that traverse the network without being queued at any of the switches. This is

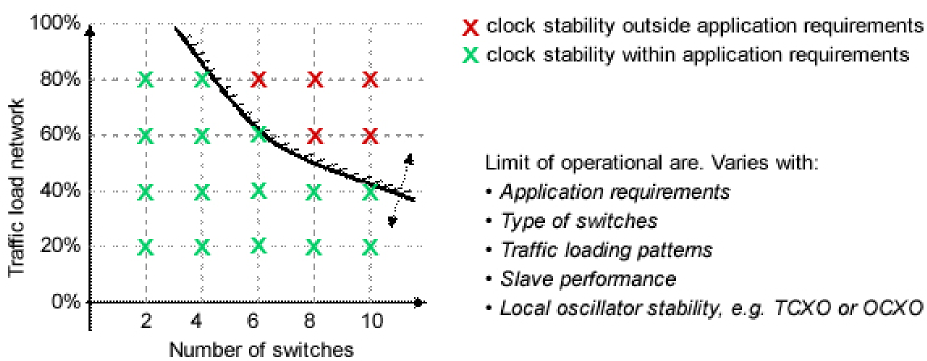


FIG 7 Operational Area for GSM/UMTS over Symmetricom Test Network

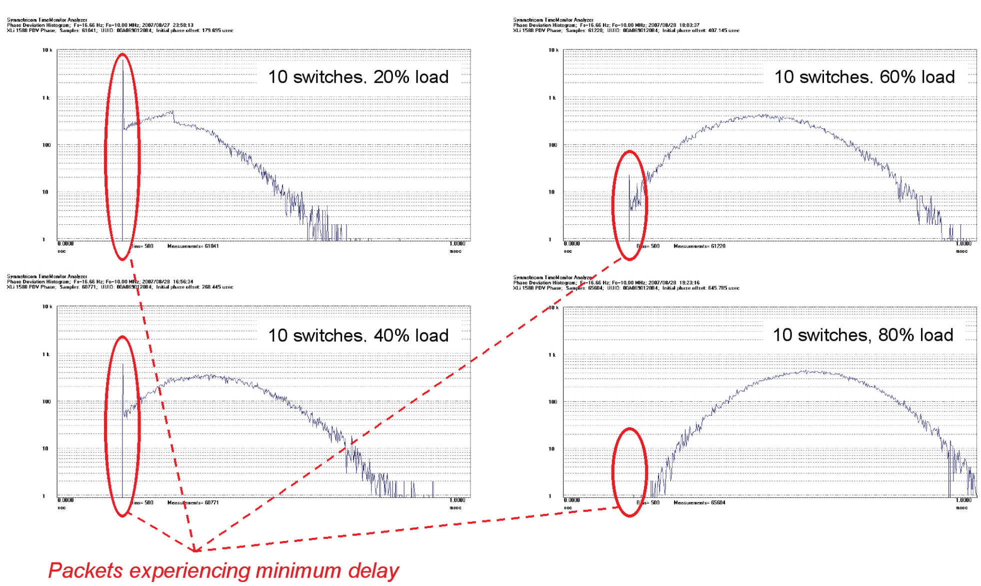


FIG 8 Packet Delay Histograms at Different Traffic Loads

the fastest that any packet can travel through the network. As the load increases, the probability of being queued at one or more switches increases, but the minimum packet delay through the network remains constant, at least up to the point where the "minimum peak" disappears altogether (i.e. where the probability of not being queued at any of the switches becomes vanishingly small). This feature can be used in a packet selection process to discard the packets that are going to cause the biggest errors in the time calculation.

Minimum Time Deviation is a measure of both how constant this minimum delay is, and the frequency of occurrence of packets experiencing the minimum delay. It measures the minimum delay over a series of three consecutive time intervals, and calculates the variation between the minimum delay values between these intervals. The time intervals are progressively widened, giving an idea of how long the wait is for a minimum delayed packet. The actual calculation is derived from the Allan deviation used for characterizing oscillator stability, and will be described in more detail in a forthcoming white paper from Symmetricom.

An example of a Minimum Time Deviation plot is shown in Figure 9. A mask has been drawn on the diagram – provided the plot is below the mask, there is enough information in the packet timing messages for the slave to be able to produce a stable clock. The horizontal section of the mask is derived from the maximum phase or time deviation permitted by the application, while the diagonal section represents the filter characteristic. They intersect at the corner frequency for the filter, which is determined by the stability of the local oscillator.

Such masks provide a means to quantify the network performance independent of the number of switches or network elements. For example, some switches may add more delay variation than others due to the way they have been designed. Software-based switches or routers are generally more variable than hardware-based devices, and hence the hop count for these devices might have to be reduced. However, the Minimum Time Deviation performance is independent of this, and shows quickly whether the network is fit for purpose.

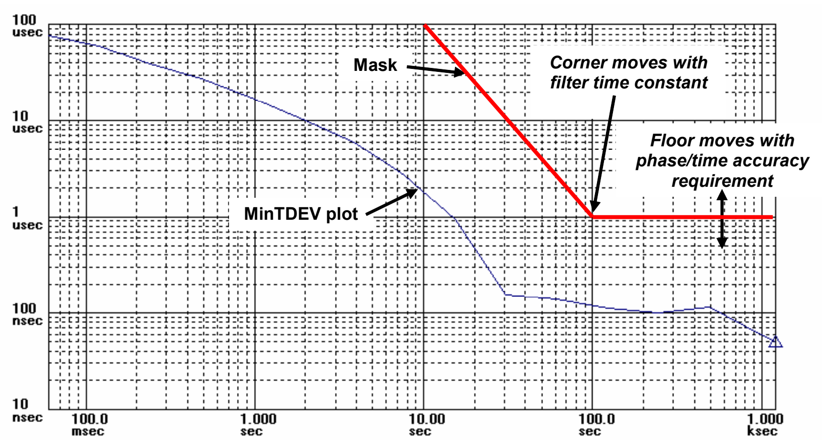


FIG 9 Example of a Minimum Time Deviation Mask

Allocating Performance Budgets

In packet timing protocols such as PTP, each of the elements of the system contribute noise that may degrade the quality of the output clock, as shown in Figure 10. For example, the grandmaster clock converts the reference clock into a series of timestamps carried in message packets. These timestamps will have a small but measurable inaccuracy, caused by quantization processes in the grandmaster. The network contributes its own noise, primarily in the form of variation in delay of the packets carrying the timestamps. The end equipment will also contribute noise as it regenerates the original clock from the packet flow, in particular from the local oscillator.

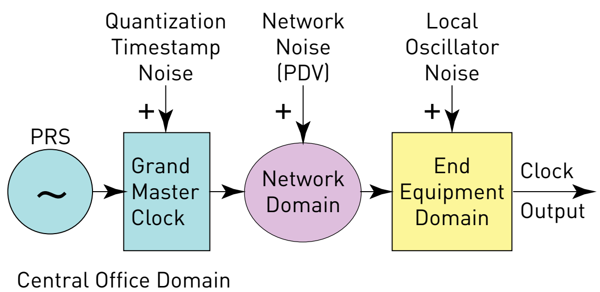


FIG 10 Noise contribution of PTP Elements

In order to determine the performance requirements for each of these different elements, some kind of budgeting process must be used to break down the overall application performance requirements. For example, the application requirements could be divided using a pyramid approach.

This budgeting process can be applied to the GSM/UMTS base station synchronization requirement in the following way. Firstly, the 16ppb frequency accuracy figure is a peak requirement. It must therefore be de-rated to RMS since many of the measurement metrics such as TDEV and MinTDEV are RMS-based. In order to allow headroom for transient fluctuations, the peak requirement must be de-rated by at least 3, leaving approximately 5ppb.

Secondly, the de-rated figure can be apportioned between the different network elements using the pyramid system as shown in Figure 11:

- 10% to the grandmaster, i.e. 0.5 ppb,
- 40% to the packet network, i.e. 2.0 ppb,
- 50% to the end equipment, i.e. 2.5 ppb

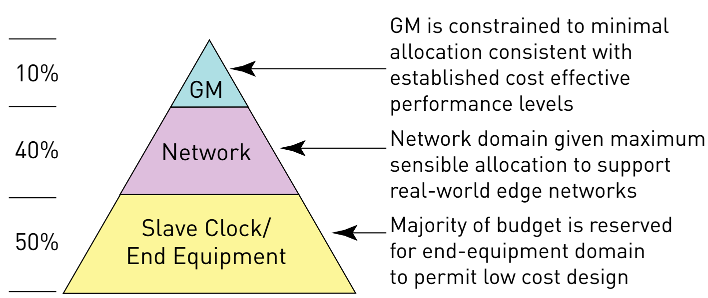


FIG 11 Pyramid Noise Budget Allocation

Derivation of a MinTDEV Mask for GSM/UMTS Operation

The first task is to set the filter corner. Here the decision is based on the characteristics of the local oscillator. The filter must be narrow enough to be able to effectively filter network noise, but wide enough to allow for effective compensation of oscillator noise.

For a good quality temperature-compensated crystal oscillator (TCXO), experience suggests that the corner frequency can be in the region of 10mHz without coupling too much oscillator noise in the output. For an ovenized oscillator (OCXO), the passband can be narrowed to 1mHz. This suggests that the corner should be at 100s for a TCXO, or 1000s for an OCXO.

The second task is to set the horizontal "floor" of the mask. For GSM and UMTS operation there is a frequency accuracy requirement rather than a phase or time accuracy limit. Therefore the approach is to calculate how much phase wander would be generated at the maximum frequency offset over the bandwidth of the filter. The budgeting calculation described earlier indicated that a noise budget of around 2ppb could be allocated to the network domain. This equates to a phase movement of 0.2µs over a 100s period, or 2µs over a 1000s period.

The result is the pair of MinTDEV masks shown in Figure 12:

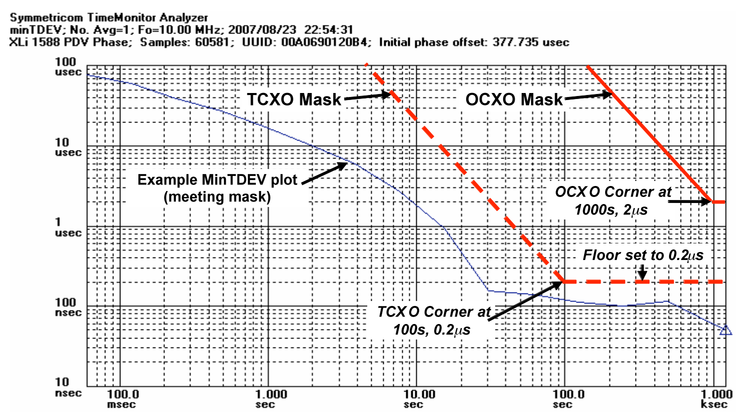


FIG 12 MinTDEV Mask for GSM/UMTS Frequency Accuracy

Comparison to Measured PDV Results

Symmetricom has characterized the packet delay of switched Ethernet networks, covering all the points illustrated in Figure 7 on page 9. The "baseline configuration" used was a collection of Netgear FS108 100BaseT full duplex switches, tested in a configuration very similar to that suggested in ITU-T Recommendation G.8261, Appendix VI, without any priority or other QoS techniques applied.

This section examines the characteristics of this network, and how the MinTDEV performance compares to the mask in Figure 12.

Figure 13 shows the MinTDEV plots for 6 switches at 0%, 20%, 40%, 60% and 80% traffic loads. As can be seen, the 80% plot is well outside the TCXO mask, while the 60% plot is borderline – it meets the floor of the mask, but is just outside the slope. This indicates that the output clock should be within specification up to about 60% load on this network. However, if an OCXO is used with a 1mHz smoothing characteristic, the clock should be within specification at 80% load too.

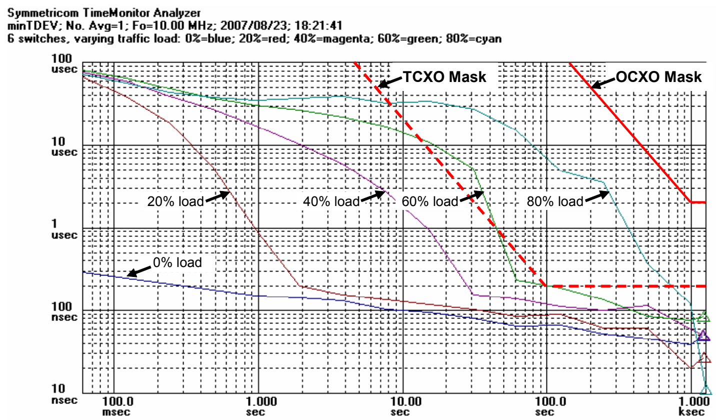


FIG 13 MinTDEV plots for 6 Switches with varying traffic loads

Figure 14 shows the MinTDEV plots for 8 switches at the same traffic loads. This time, the 60% plot is well over the mask, indicating that the output clock may be outside specification at loads above 40-50%. At 80% load, the plot never falls to a floor, indicating that the minimum delayed packets are extremely rare. Even with an OCXO, it is likely that the recovered clock will be outside of the required specification.

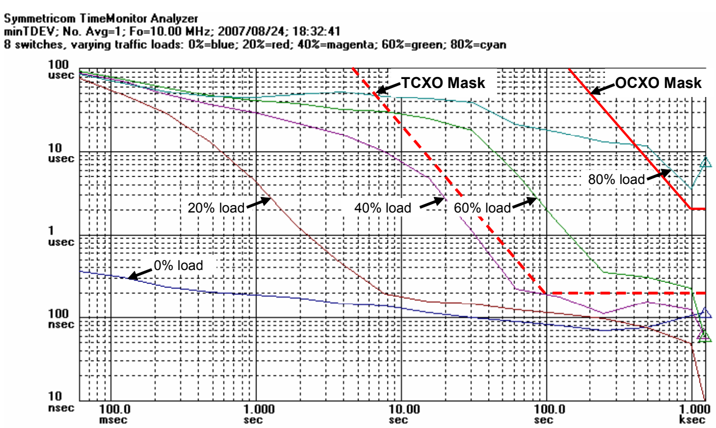


FIG 14 MinTDEV plots for 8 Switches with varying traffic loads

Figure 15 shows the MinTDEV plots for 10 switches at the same traffic loads. Again the 60% plot is well over the mask, and now the 40% plot is borderline, just starting to breach the mask in places. Using an OCXO, brings the 40% and 60% plots below the mask, but as before, at 80% load the plot has no floor, indicating that the recovered clock will be outside of specification with whatever local oscillator is chosen.

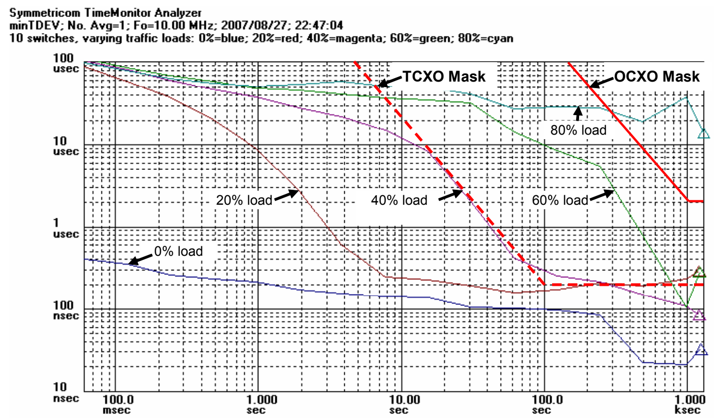


FIG 15 MinTDEV plots for 10 Switches with varying traffic load

References

1. *"IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems"*, IEEE1588-2008
2. *"UTRAN Overall Description"*, 3GPP TS 25.401, September 2007
3. *"Synchronization in UTRAN – Stage 2"*, 3GPP TS 25.402, September 2007
4. *"Quality of Service (QoS) Concept and Architecture"*, 3GPP TS 23.107, September 2007
5. *"Timing and Synchronization Aspects in Packet Networks"*, ITU-T Recommendation G.8261, 2008
6. T. Farley, D. Tonks, A. Pietiläinen, *"IEEE-1588v2 Telecom Profile for Constrained Networks"*, COM15-C399-E, Contribution to ITU-T Study Group 15, Question 13, May 2007
7. T. Frost, *"Template for a Telecommunications Profile for IEEE1588 version 2"*, COM15-C641-E, Contribution to ITU-T Study Group 15, Question 13, January 2008
8. D. Mills, *"Internet Time Synchronization: The Network Time Protocol"*, RFC1129, IETF, October 1989
9. *"Synchronization for Next Generation Networks: NGN Synchronization and Timing Technology"*, Symmetricom white paper, May 2008

Abbreviations and Definitions

3GPP	3rd. Generation Project Partnership (the standards body responsible for defining UMTS)
ADSL	Asymmetric Digital Subscriber Line
CDMA	Code Division Multiplexed Access
CRC	Cyclic Redundancy Check
DOCXO	Double Oven Compensated Crystal Oscillator
DSL	Digital Subscriber Line
EDGE	Enhanced Data Rates for GSM Evolution
EPON	Ethernet Passive Optical Network
GERAN	GSM-EDGE Radio Access Network
GM	Grandmaster clock server
GPON	Gigabit Passive Optical Network
GSM	Global System for Mobile Communications
HSDPA	High Speed Download Packet Access
HSUPA	High Speed Upload Packet Access
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol (e.g. IPv4 – Internet Protocol version 4; IPv6 – Internet Protocol version 6)
ITU-T	International Telecommunications Union – Telecommunications Standards Bureau
MinTDEV	Minimum Time Deviation
NGN	Next Generation Network
Node B	UMTS radio base station
NTP	Network Time Protocol
OCXO	Oven Compensated Crystal Oscillator
PDV	Packet Delay Variation
PLL	Phase Locked Loop
PRC	Primary Reference Clock
PTP	Precision Time Protocol
QoS	Quality of Service
RF	Radio Frequency
RMS	Root Mean Square
RNC	Radio Network Controller
RNS	Radio Network Subsystem
SHDSL	Symmetric High Speed Digital Subscriber Line
SP	Strict Priority
TCXO	Temperature Compensated Crystal Oscillator
TDEV	Time Deviation
TDM	Time Division Multiplexing
UDP	User Datagram Protocol
UMTS	Universal Mobile Telephony Service
UMTS-FDD	Universal Mobile Telephony Service - Frequency Division Duplexing
UTRAN	UMTS Transport Radio Access Network
VDSL	Very high speed Digital Subscriber Line
WiMAX	Worldwide Interoperability for Microwave Access
xDSL	Digital Subscriber Line (of various types, e.g. ADSL, VDSL, SHDSL)

SYMMETRICOM, INC.
2300 Orchard Parkway
San Jose, California
95131-1017
tel: 408.433.0910
fax: 408.428.7896
info@symmetricom.com
www.symmetricom.com