ENT-AN1046

Application Note

IEEE 1588v2 and SyncE—Applications and Operation Using Microsemi's Synchronization Solution

Microsemi
Power Matters.™

Microsemi Corporate Headquarters
One Enterprise, Aliso Viejo,
CA 92656 USA
Within the USA: +1 (800) 713-4113
Outside the USA: +1 (949) 380-6100
Fax: +1 (949) 215-4996
Email: sales.support@microsemi.com
www.microsemi.com

About Microsemi

Microsemi Corporation (Nasdaq: MSCC) offers a comprehensive portfolio of semiconductor and system solutions for aerospace & defense, communications, data center and industrial markets. Products include high-performance and radiation-hardened analog mixed-signal integrated circuits, FPGAs, SoCs and ASICs; power management products; timing and synchronization devices and precise time solutions, setting the world's standard for time; voice processing devices; RF solutions; discrete components; enterprise storage and communication solutions; security technologies and scalable anti-tamper products; Ethernet solutions; Power-over-Ethernet ICs and midspans; as well as custom design capabilities and services. Microsemi is headquartered in Aliso Viejo, California, and has approximately 4,800 employees globally. Learn more at www.microsemi.com.

VPPD-02935

Revision History

The revision history describes the changes that were implemented in the document. The changes are listed by revision, starting with the most current publication.

1.1 Revision 3.0

In revision 3.0 of this document, references to Figure 3 were corrected.

1.2 Revision 2.0

In revision 2.0 of this document, section header 3 was corrected. For more information, see Network Deployment: LTE Equipment for ToD Synchronization.

1.3 Revision 1.0

Revision 1.0 was the first publication of this document.

Contents

Figures

Tables

2 IEEE 1588v2 and SyncE—Applications and Operation Using Microsemi's Synchronization Solution

Synchronous Ethernet (SyncE—ITU-T Rec. G.8261) and 1588 (IEEE 1588-2008 or version 2) are technologies used for distribution of frequency and time of day (ToD).

The architecture and features of SyncE are very similar to those found in SDH/SONET networks. SyncE uses the physical layer (Ethernet interfaces) to distribute frequency from the primary reference clock (PRC) to the slaves. The means for redundancy are provided.

The equipment sync controller selects between received clocks from multiple inputs, filters the selected reference, and uses that as the transmit direction clock.

IEEE 1588v2 defines the precision timing protocol (PTP) at the packet layer, which may be used to distribute frequency and/or ToD (phase).

Slaves lock to grand masters (GM). Each GM issues PTP packets time-stamped with ToD. Each slave estimates the delay between each GM and slave, then adds the delay to the received ToD (the slave also knows ToD). For frequency-only distribution, the GM-to-slave delay need not be known—this is sometimes called the "one-way method." Redundancy is an inherent part of the 1588 PTP.

SyncE and 1588 may both be used in heterogeneous networks to distribute frequency. As mobile networks converge to packet platforms, one major application for SyncE/1588 frequency distribution is to ensure that wireless base stations (BTS, NodeB, and eNodeB) stay synchronized.

Newer generations of mobile technology focus on increasing the data throughput, uplink, and downlink. This requires tighter phase alignment between neighboring towers to facilitate hand-over. 1588 can provide phase alignment where SyncE cannot; however, SyncE can be used to increase the achievable accuracy in the network.

The frequency and phase accuracy requirements for various mobile generations is shown in the following table.

Table 1 Frequency and Phase Requirements for Mobile Technologies

Technology	Frequency Accuracy	Phase Accuracy
GSM	≤ ±50 ppb	N/A
UMTS FDD	≤ ±50 ppb	N/A
UMTS TDD	≤ ±50 ppb	≤ 2.5 µsec
3GPP2 CDMA2000	≤ ±50 ppb	≤ 3 µsec
TD-SCDMA	≤ ±50 ppb	≤ 3 µsec
LTE		Long-term cumulative error ≤ 1.5 µsec
LTE-advanced/MIMO		≤ 0.5 µsec between neighboring towers

Microsemi PHYs were the first to implement SyncE. Microsemi PHYs and Carrier Ethernet switches (CES) were the first to implement hardware accurate 1588 in the telecom environment. Microsemi actively participates in driving the technology through standards body participation.

The Microsemi synchronization solution encompasses copper and optical PHYs, 15 100 Gbps Carrier Ethernet switches and comprehensive software that implements the advanced PTP algorithms and servo-enabling turnkey designs.

IEEE 1588 was originally designed for industrial applications. Microsemi implementations may also be used for this purpose, but this document describes the following telecom applications:

- An application scenario where one telecom operator provides synchronization services to two mobile operators
- Three 1588 deployment scenarios with and without GPS
- The concept of SyncE and 1588 operation
- The Microsemi system solution and its elements
- An example upgrade of a legacy system to become 1588-capable

Accompanying application notes describe details of Microsemi PHYs and switches. Device datasheets provide the highest level of device-specific detail.

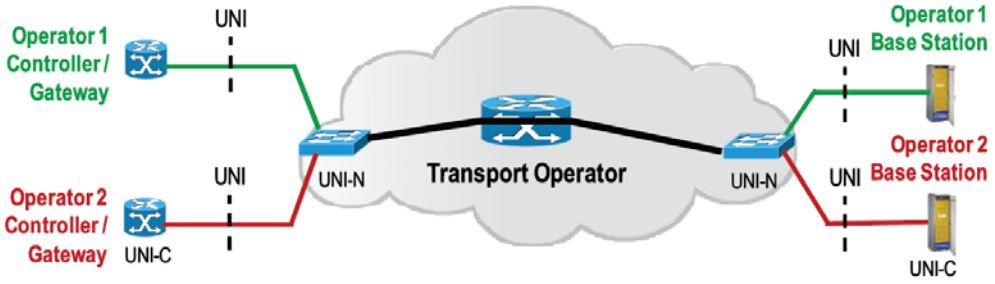
3 Synchronization Service

Synchronization may be delivered as a service from one organization to another within the same company or to a separate company. The following application does not distinguish between the two scenarios, although the wording points to the latter scenario.

The many cell towers or base stations of a mobile operator must be connected with the centralized controllers and gateways. In some cases, the mobile operator also owns this infrastructure, but in many cases, this capacity is leased from a transport operator.

The following illustration shows a metro Ethernet forum (MEF) approach in which the transport operator provides services over well-defined UNIs. The transport network shown may use any combination of technologies, including Ethernet, IP, MPLS, and OTN.

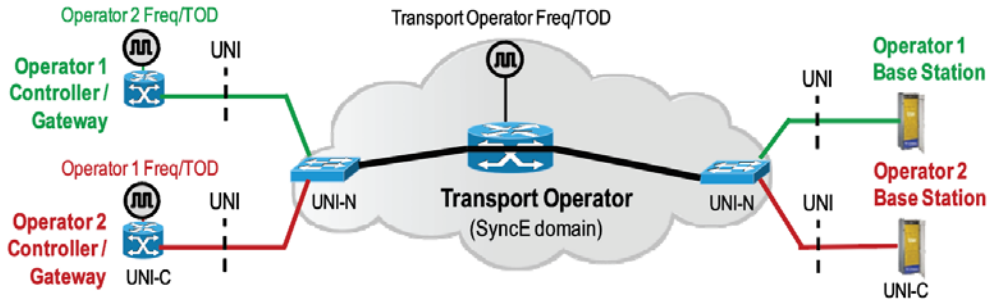
Figure 1 A Transportation Operator Provides Transport Services to Two Mobile Operators



OAM tools are used by both transport and mobile operators. The transport operator monitors connectivity and performance of each path and mobile operator service. Each mobile operator monitors connectivity over the UNI and end-to-end service connectivity and performance. For debugging, the transport operator can proactively detect SLA issues and isolate failures to a node/link, and each mobile operator can isolate failures to the access link or the transport operator service.

The transport operator can run SyncE and IEEE 1588 inside its own network. It may offer a synchronization service to the mobile operators, as shown in the following illustration:

Figure 2 Synchronization Service



Each mobile operator has the following timing options:

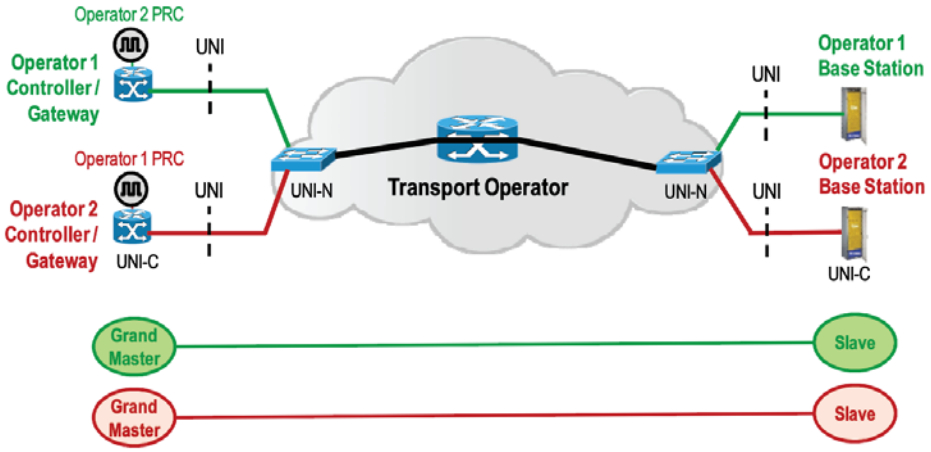
- GPS at base stations
- TDM connection over MPLS PW or legacy equipment

- IEEE 1588 without assistance from transport operator
- IEEE 1588 with transparent or boundary clock service from transport operator
- Use of the transport operator's SyncE or IEEE 1588 (not fully synchronous)

GPS at every base station tends to be expensive and not always possible because it requires an aerial antenna. A TDM connection can be, and is, used for frequency distribution but not ToD. The TDM connection only supplies timing, so the migration to Ethernet makes it less desirable.

The following illustration shows where the transport operator only offers transport services, not synchronization services.

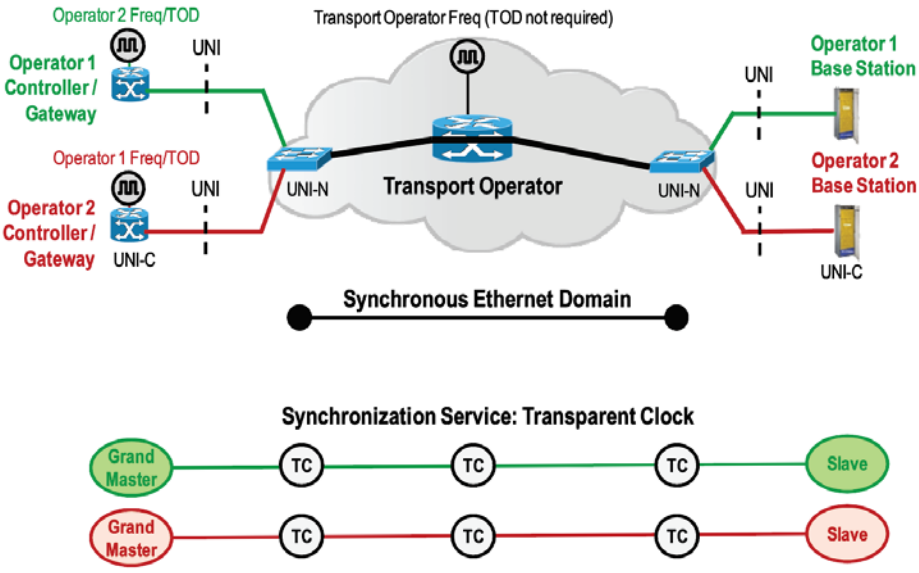
Figure 3 No Synchronization Service



Each mobile operator connects its slaves with its grand master. The transport operator is unaware of the 1588 content and does not provide support to improve the quality or accuracy of the distribution—this is most suited for frequency distribution.

The following illustration shows that the transport operator provides a transparent clock (TC) service to each mobile operator.

Figure 4 Transparent Clock Synchronization Service

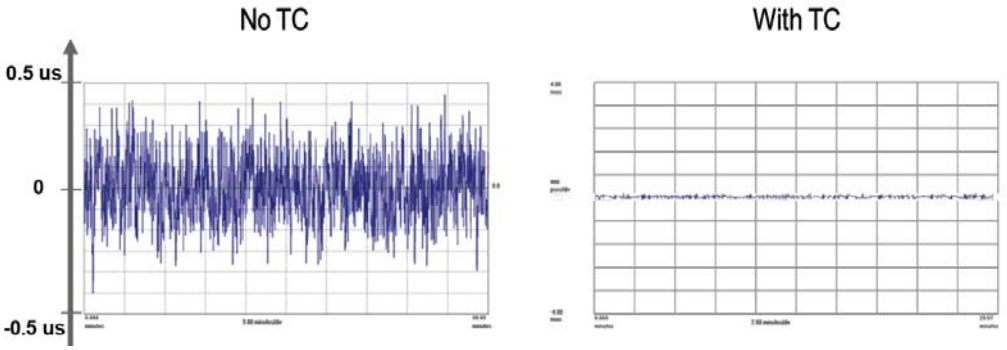


Without TC operation, the packet delay through the network depends on the traffic load. Large packet delay variation (PDV) can degrade performance. A transparent clock measures the transit delay (or residence time) and inserts that information in the correction field of the PTP packet. Accordingly, a fast packet will have a small correction value and a packet going through a highly congested switch will have a large correction value. In the end, the slave can know, packet by packet, what network delay the packet has experienced.

In the TC service, the transport and mobile operator must agree on the type of PTP packet encapsulation. Then, the transport operator can identify the PTP packets, insert residence time correction, and deliver an enhanced service to the mobile operator.

The following illustration shows measurements of the slave clock offset from the master clock within a 30-minute measurement interval. A single unloaded carrier Ethernet switch is placed between the master and the slave clocks. The first measurement is without TC operation, the second includes it.

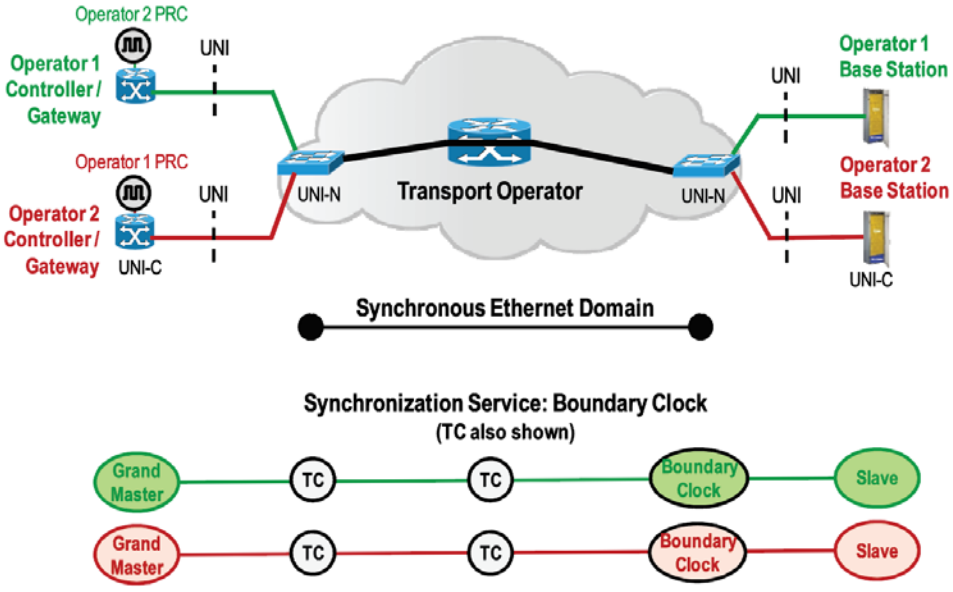
Figure 5 Measurements of Slave Clock Offset from Master Clock



The variation in offset shifts from ±500 nsec to ±5 nsec; therefore, slave tracking is dramatically improved with TC operation.

The following illustration shows that the transport operator also supplies a boundary clock (BC) function to each mobile operator (TC is still in operation).

Figure 6 Boundary Clock Synchronization Service

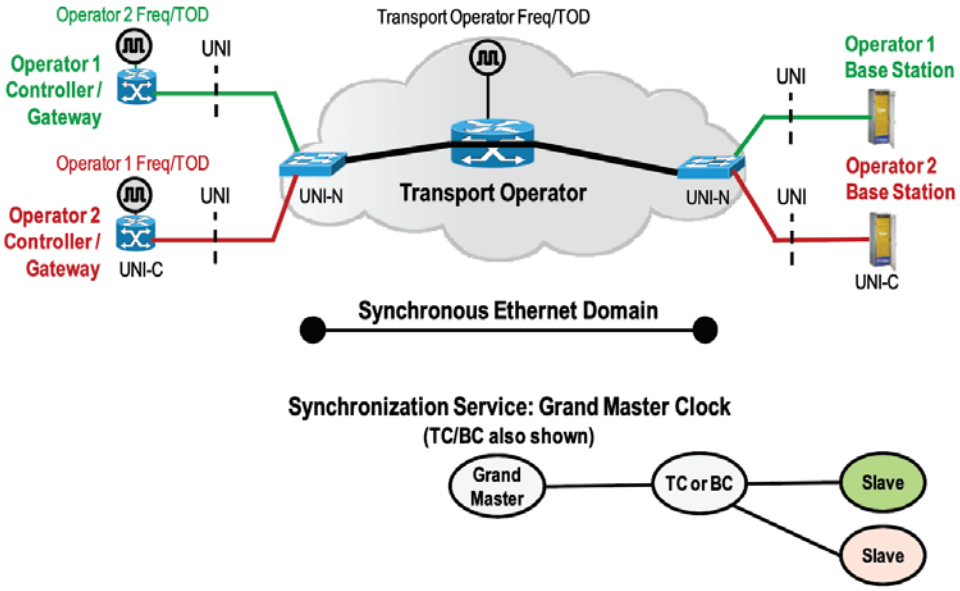


The BC provides increased scalability by imposing hierarchy in the network—the GM connects to a number of BCs and each BC connects to a number of slaves. The BC acts as a slave to the GM and a master to the slave-only clocks.

One network element may support only one BC. The Microsemi solution enables multiple BC instances in one network element. This could be used for the previous case where the transport operator's equipment could only support one BC, in addition to its own, for each mobile operator. Each BC would operate in its own ToD domain, the quality of which depends on the transport operator's quality.

The following illustration shows that the transport operator provides its own ToD for the mobile operators to use.

Figure 7 IEEE 1588 Master Clock Synchronization Service



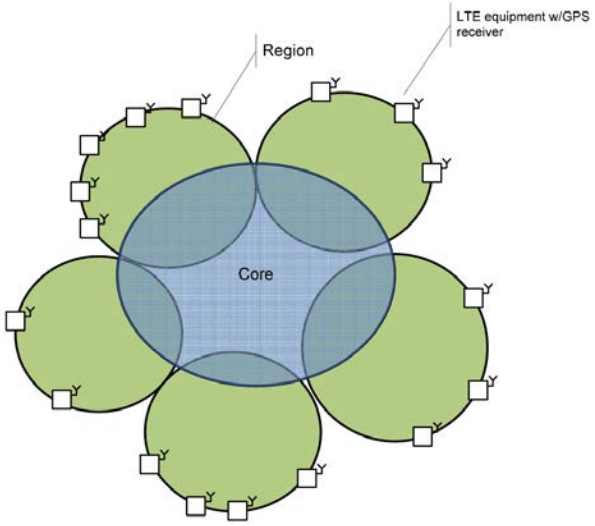
This service results in high-quality timing that is not fully synchronous to the mobile operator's timing. This service is also easier to deploy for the transport operator but makes the mobile operators more dependent, resulting in ToD "islands" within their network. However, the transport and mobile operators' ToD are all locked and calibrated to UTC such that the difference, in practice, is below required accuracy.

4 Network Deployment: LTE Equipment for ToD Synchronization

The following three illustrations show how LTE equipment is used for ToD synchronization using GPS, GPS with 1588 backup, and 1588. All these illustrations also show a core network connecting a number of regions.

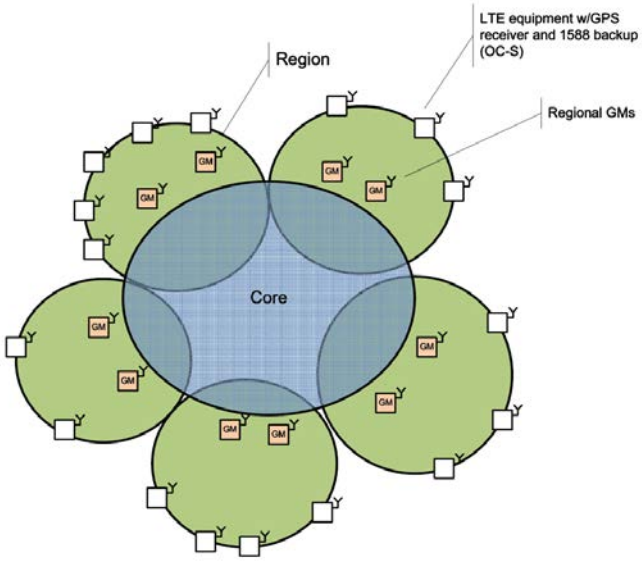
In the following illustration, the network operator relies completely on GPS. This strategy is beyond the scope of this document.

Figure 8 GPS



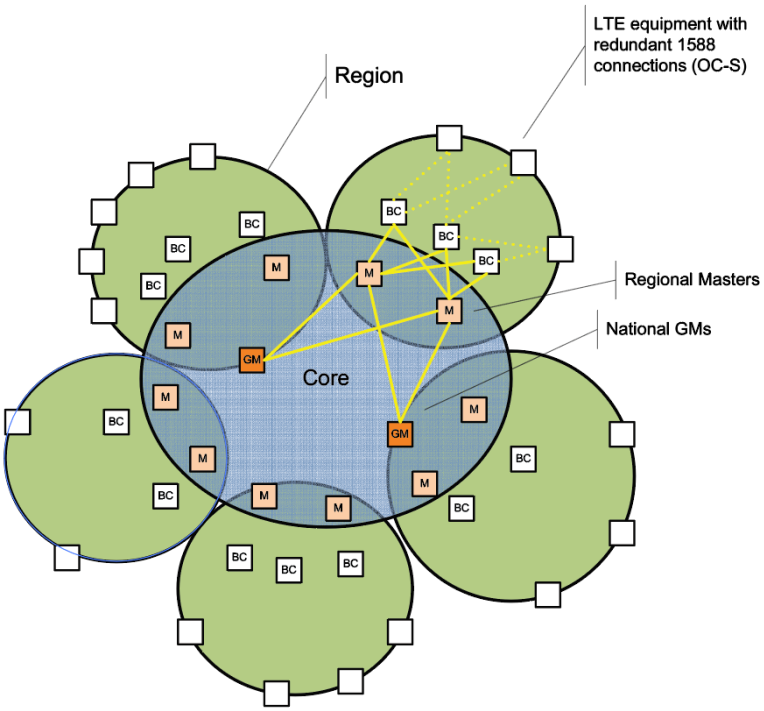
In the following illustration, the GPS signal is jammed or the GPS system falls out because the sky view is obstructed. Here, the operator has a backup in the form of a 1588 distribution network.

Figure 9 GPS with 1588 Backup



Many operators strategically decide not to rely on GPS at all. In the following illustration, GPS is abandoned and a redundant 1588 ToD distribution is chosen instead. Two national/central grand masters, such as a clock calibrated to UTC, are used.

Figure 10 1588

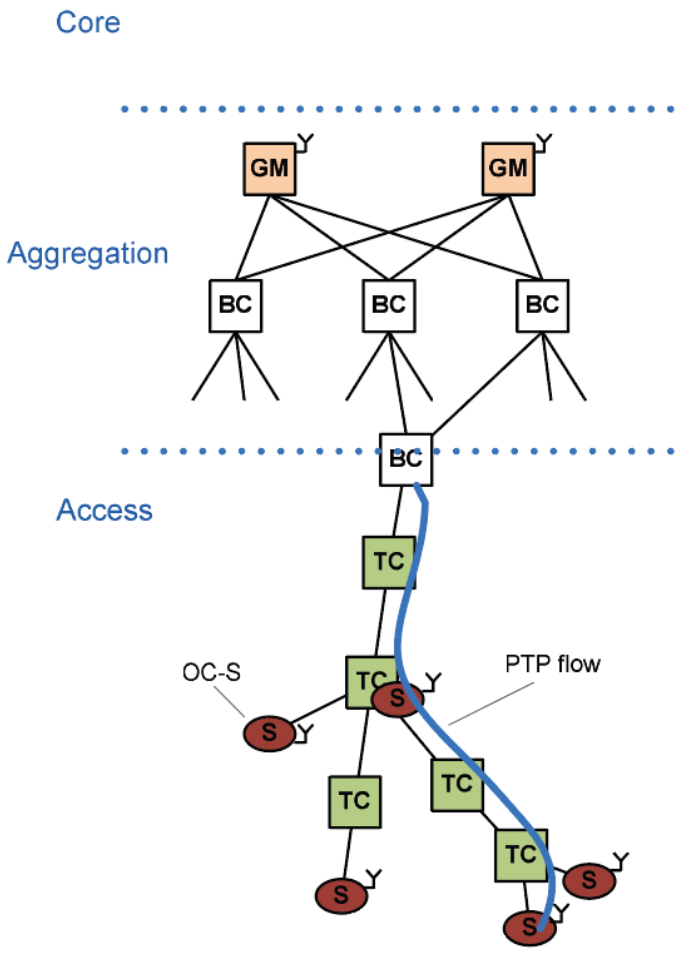


In the 1588-only case, there PTP flows need to be transported from the GMs to the regional masters over the core. This would typically require MPLS encapsulations. More generally, MPLS(-TP) will likely be used by some operators for regional networks, which requires MPLS encapsulations. Other operators prefer simpler Ethernet or L3 distribution networks, particularly in access networks.

For existing IP/MPLS networks, the technique is to put a BC at each IP/MPLS hop. This fits well with traditional L3 networking but is very expensive for 1588. In terms of cost, a BC is more expensive than a TC. In terms of performance, in a network with two LERs connected via at least one LSR, the LSR can implement a TC, but not a BC (this requires that the LSP is terminated). The BCs would thus be interconnected over a 1588-unaware network, which increases PDV and degrades performance.

The following illustration shows in more detail the deployment of clocks in the case of GPS with 1588 backup.

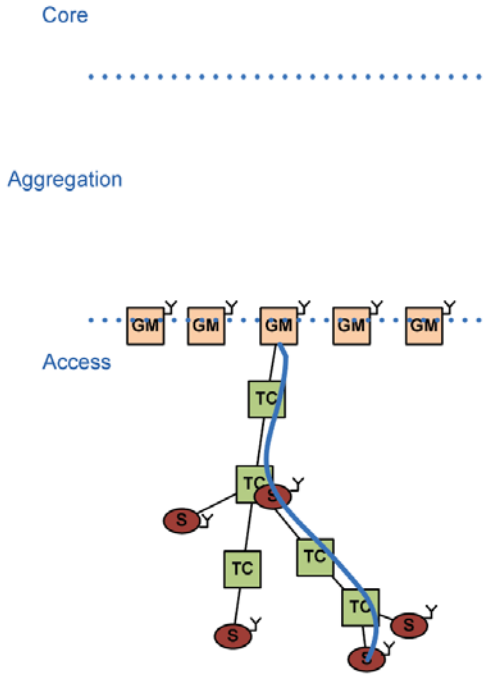
Figure 11 GPS with 1588 Backup (detailed)



The GMs reference GPS or other sources with redundant distribution to the border BCs. This is most likely dominated by BCs, whereas the access network will benefit from the lower cost of TCs. A linear access topology is shown in this example but could as well use ring topology.

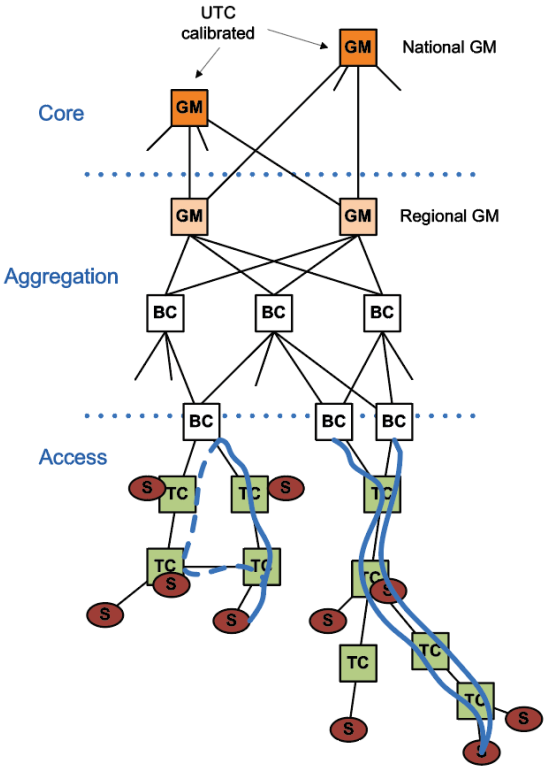
Figure 12 shows the same case, but now the GMs are moved much closer to the access border to avoid the aggregation layer distribution and rely even more on GPS as a ToD source.

Figure 12 GPS with 1588 Backup—GMs Closer to Access Border



The following illustration shows a 1588 distribution network with full redundancy to the border BCs. The access network can also benefit from redundancy if rings are utilized, whereas linear topologies can only be protected by 1+1 link redundancy.

Figure 13 1588 with Redundant Distribution to Border BCs



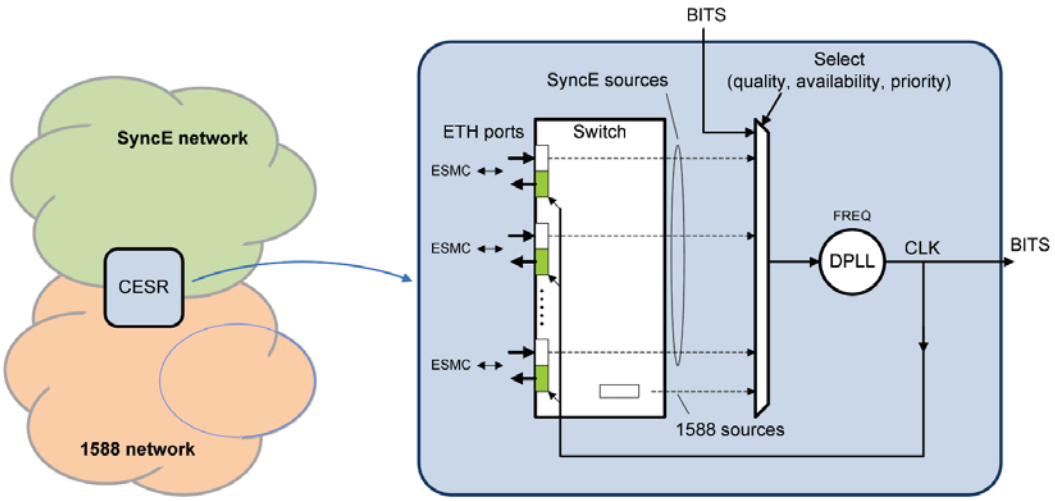
5 SyncE and IEEE 1588 Operation

This section describes frequency and time of day synchronization.

5.1 Frequency Synchronization

The following illustration shows a generic frequency distribution that supports a network.

Figure 14 Frequency Distribution



This network is sub-divided into two domains, one supporting SyncE, the other supporting 1588. In general, this includes support for both or only one of these technologies.

The middle of this network shows a carrier Ethernet switch router (CESR), a network element. This represents the multitude of equipment types that also use SyncE and/or 1588 to support frequency distribution.

The synchronization hierarchy is a tree structure with options for choosing alternative sources, thus providing self-healing architectures.

The CESR is shown in more detail on the right, focusing only on the functional blocks relevant for synchronization.

For frequency synchronization, three sources are principally available:

- The clock recovered from each Ethernet port (Ethernet may be embedded in an OTN signal, for instance)
- 1588 PTP flows (where this device acts as a slave)
- External sync ports (such as E1/T1 used in SDH/SONET networks and called BITS in ANSI terminology, or T3/T4 in ITU-T when used inside central telecom offices)

The source selection may be configured manually or determined by a selection algorithm. Input criteria include source state (up/down), source quality (one source is closer to the primary reference clock (PRC) than another) and assigned priority.

The oscillator locks to the sync source though a digital PLL, which tracks slow variations of the source and filter high frequency jitter, and the result is used in the equipment for timing the transmit ports.

The basic operation of this synchronous Ethernet equipment clock (EEC) is specified in ITU-T Rec. G.8261, G.8262, and G.8264. The ITU-T telecom profile of IEEE 1588v2, which specifies choices made in order to create interoperable frequency synchronization, is specified in G.8265.1.

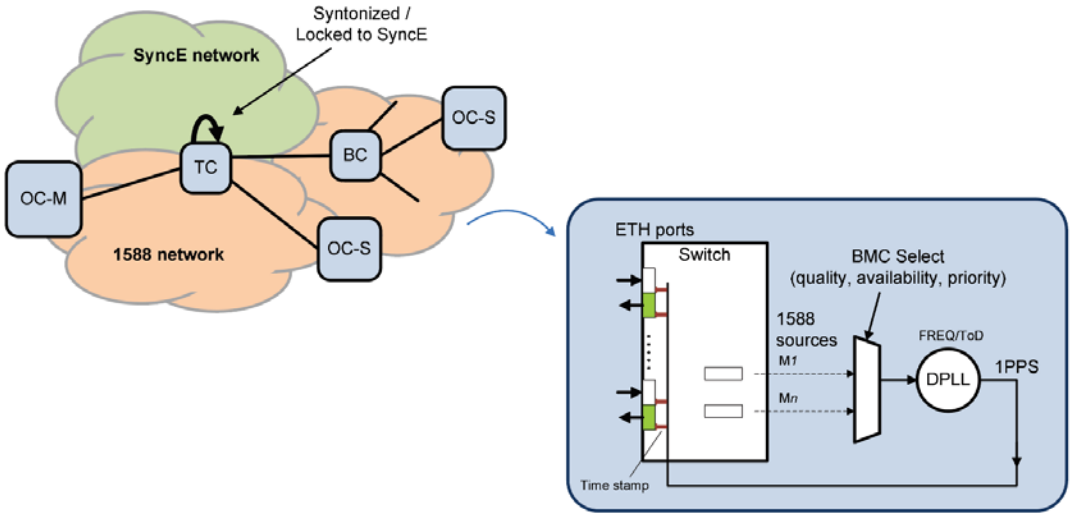
Note: The Ethernet synchronization status message (SSM), which carries the information about the quality and status of the sync source, is carried in a slow protocol PDU in an Ethernet synchronization messaging channel (ESMC).

Microsemi PHYs and CE switches support extracting the individually received clocks and using an externally provided transmit clock.

5.2 Time of Day Synchronization

The following illustration shows a generic network supporting ToD distribution.

Figure 15 ToD Synchronization



This network is subdivided into two domains: one supports 1588 and the other supports SyncE. 1588 is necessary, but SyncE is optional supports ToD distribution by enhancing network stability.

The network elements are all shown as clocks (OC, TC, and BC). In most cases these reside in larger equipment (CESR, CES, POTP, and so on).

The synchronization hierarchy is a tree structure with options for choosing alternative sources, thus providing self-healing architectures.

The equipment is shown in more detail on the right, focusing only on the functional blocks relevant for ToD synchronization. Although the equipment is drawn as a slave, it may support slave, master, and boundary clock functions as well as TC operation.

The 1588 master clock is selected using the best master clock (BMC) algorithm. The result is output for local time stamping and for other equipment.

The sampling system (time stamping), the packet filtering, and subsequent adjustments to the ToD all counter-constitute a PLL that compensates for oscillator wander and tracks both wander in the master and slow delay changes.

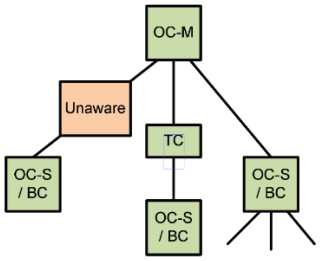
The operation of the various clock types is specified in IEEE 1588-2008. ITU-T specifies the PTP profiles for using PTP within telecom networks.

The Microsemi solution provides deep-level identification of PTP packets and precision time stamping, the PTP protocol stack, and software implementation for OC, BC, and TC operation.

5.2.1 Terms and Definitions

The following illustration shows a simple 1588 clock hierarchy.

Figure 16 1588 Clock Hierarchy



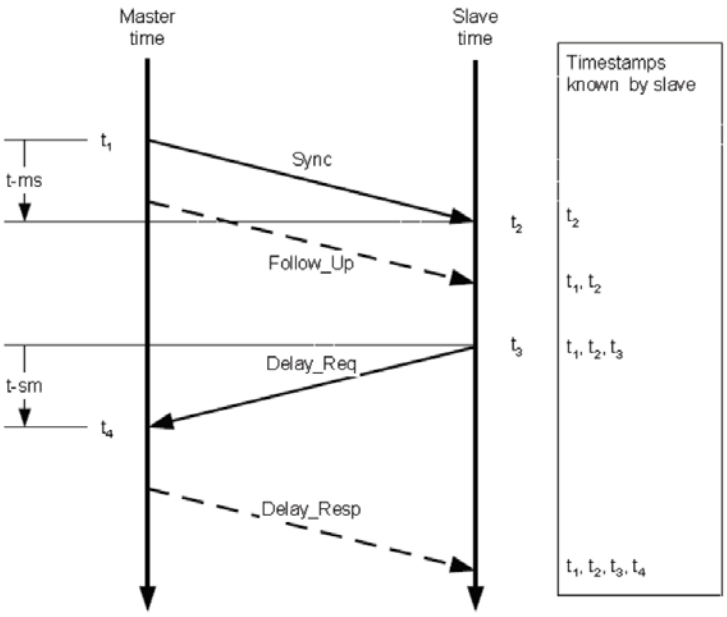
The ordinary clock (OC) is shown at the top. It can operate as a master (OC-M) or a slave (OC-S). In ITU-T terms, packed equipment clock can operate as a master (PEC-M) or a slave (PEC-S). That being said, the IEEE terms will be used in this document.

If the master clock is the ultimate source of time within a timing domain, the clock is also called a grand master clock (GM).

This is a PTP-unaware network, as the slave may connect to the master through a network that is not aware of 1588 PTP traffic. This network cannot reduce packet delay variation (PDV), but it will add PDV.

The following illustration shows how a basic PTP message is transmitted and received.

Figure 17 Basic PTP Message



Sync messages holding the ToD are issued by the master in an end-to-end PTP network (with unaware or aware elements between master and slave). The slave needs to know the delay between master (or BC) and slave. This is derived by issuing Delay_Request and receiving Delay_Response packets from which the slave can estimate the one-way delay. This is shown in the previous illustration.

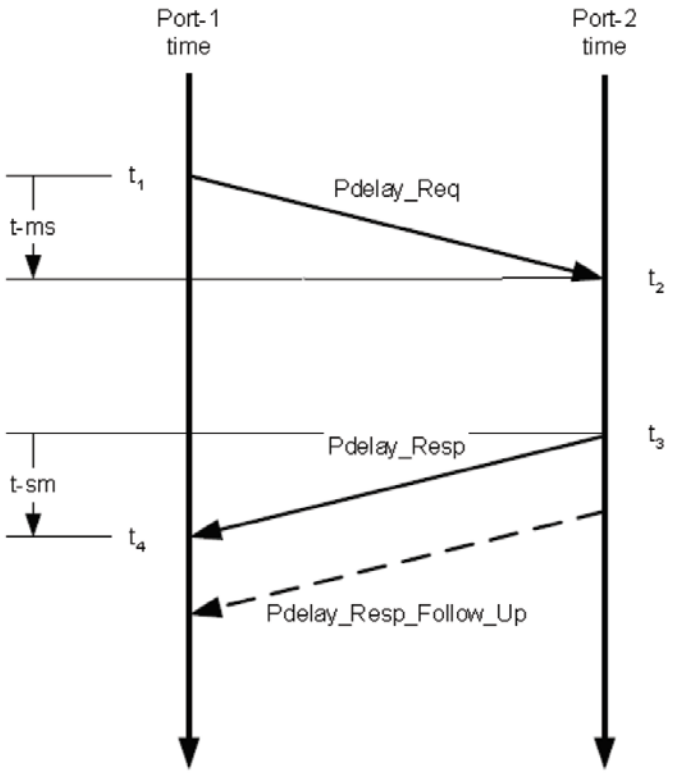
PHYs and switches may be able to time stamp on egress and insert the value in the packet before it leaves the device. This is called 1-step operation. If this is not possible, the time stamp must be carried in a second packet (a Follow_up packet). This is called 2-step operation. This is also shown in the previous illustration.

A transparent clock (TC) is a clock, or a network element, that can identify relevant PTP packets, time stamp them, and correct for the residence time when a PTP packet, on its way between master and slave, passes it. The following lists two types of transparent clocks:

End-to-end TC (E2E TC): each TC only corrects for the residence time measured and the slave still needs to measure the master-to-slave delay.

Peer-to-peer TC (P2P TC): each clock, or network element, measures the link delay between two neighboring elements. When passing a P2P TC, the Sync packet is corrected for both residence time and up-stream link delay. This measurement process is shown in the following illustration.

Figure 18 Path Delay Measurements



Additionally, asymmetric delay can be compensated for if known.

A TC is quite tolerant to the quality of the clock/oscillator used for time stamping. The change in local phase at ingress and egress can be considered constant during the residence time. Furthermore, it need not know the ToD for TC operation. However, for improved accuracy, either a SyncE-derived clock or a syntonized clock should be used. The syntonized clock is derived from the

incoming PTP packet flow, which is copied to a slave function that derives a high-quality frequency reference for the time stamping equipment.

A boundary clock (BC) synchronizes to a master and acts as a master to slaves in order to maintain the ToD for the timing domain. Each master must maintain a connection with all of its slaves, which distributes the network load and enhances the network scale. Inserted in the branching points of the distribution tree, the master must maintain fewer connections.

Applying BCs in a network also reduces jitter. For this purpose, TCs and BCs may be seen as alternatives. In larger network, both have their purpose; however, for the sole purpose of jitter reduction, TCs may prove more cost-optimal by providing lower overall network cost for 1588 deployment.

IEEE 1588 specifies a best master clock (BMC) algorithm for when a slave in a network with redundant masters/GMs must select and be able to change primary reference (in case the first one fails). To select a new master, the slave first aligns to the new master and the new delay between them. This takes time, but the slave must hold the ToD. The longer the holding period, the more difficult it becomes to maintain the correct ToD.

In the telecom profile for frequency distribution, using 1588 ITU-T has expanded upon the basic BMC algorithm by instantiating one slave per master. This following method, V-BMC, has been generalized by Microsemi. Here, the slave connected with the currently primary GM is the one that determines local understanding of ToD. This makes the switch-over time minimized because the standby slave is already active.

6 Microsemi's Synchronization Solutions

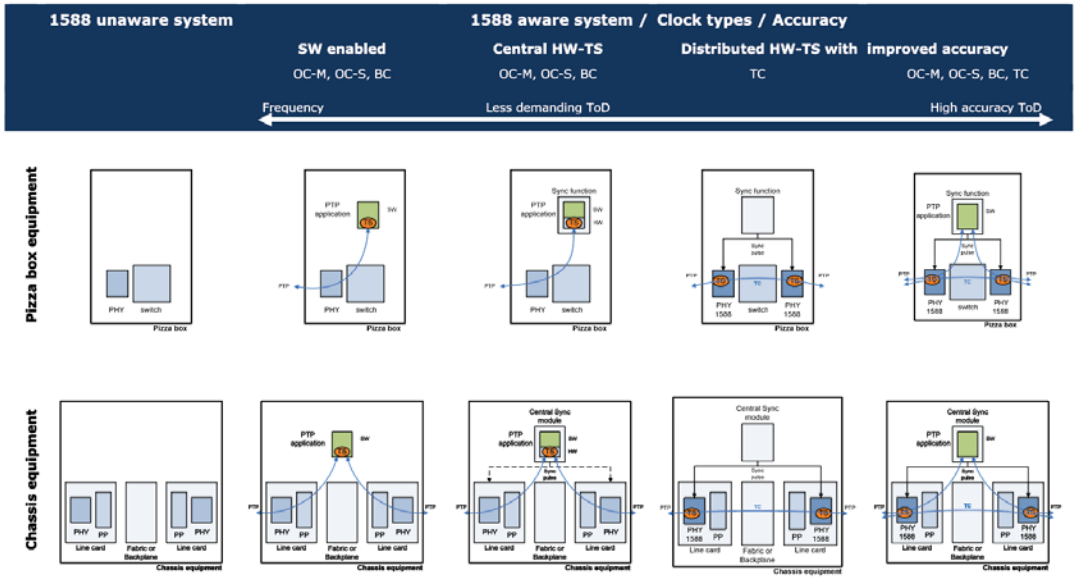
Microsemi's synchronization solutions that support frequency and ToD distribution using both SyncE and IEEE 1588 can be used in both new designs and upgrades of existing platforms. This section discusses and shows examples of legacy system upgrades using Microsemi 1588 PHYs, switches, and PTP application software.

The equipment in this section is generically separated into pizza boxes, or small systems with one module, and chassis-based equipment, or a large system with many modules. In practice, a small system can be a hybrid between a pizza box and a chassis system.

6.1 Upgrading Legacy Systems

The following illustration shows options for upgrading an existing system, which mostly depend on the system architecture and intended application.

Figure 19 Upgrade Options for Legacy Systems to Become 1588-Aware



The existing 1588-unaware system includes 1588-unaware PHYs and switches/packet processors on the main PCB in a pizza box or the line cards in a chassis system. The chassis system also includes a backplane or fabric modules to connect the line cards. Both system types may not even synchronize the line transmitters to an external reference clock (this is more likely on a pizza box than in a chassis system).

Software enabling is the simplest way to upgrading a system to become 1588-aware. In this case, the PTP application is added to the software complex that needs software time stamping. Because of the relatively imprecise time stamping, this upgrade is primarily an option if only frequency distribution using 1588 is wanted. Ordinary clock, master, slave, and boundary clock functionality is possible in this case, as all are supported by Microsemi software.

Better time stamping accuracy is possible with central hardware time stamping. This upgrade is relatively simple and has found its way into several systems. Typically, the original architecture includes some form of internal frequency/frame distribution, but the PDV through the PHY + packet

processor + possibly fabric is not taken into account. Therefore, this solution is well-suited for 1588 frequency distribution and possibly less-demanding ToD distribution networks.

Highest accuracy is achieved using distributed hardware time stamping. This is accomplished by upgrading existing boards with Microsemi 1588-PHYs and designing new boards to existing systems with Microsemi 1588-aware switches and PHYs. Hardware-accurate time stamping is performed at the line interface of the system eliminating the influence of internally-generated PDV.

If only transparent clock operation is needed, OC/BC operation and the associated, relatively expensive OCXO can be excluded. This is especially relevant for pizza boxes that operate at the very edge and near mobile antennas.

Supporting OC, BC, and TC requires both distributed hardware time stamping and centralized PTP application software. These systems are typically deployed in the higher access and aggregation networks.

Note: A 1588-unaware chassis system can also be upgraded with a line card that contains the complete 1588 solution.

6.2 Microsemi Synchronization Solution Architecture

This section describes the Microsemi synchronization solution architecture using Microsemi switches, PHYs, and software. The following concepts apply to system and/or line card upgrades, although they are targeted at new designs.

The following two illustrations, Figure 20 and Figure 21, show pizza box architecture.

Figure 20 Pizza Box Architecture with Microsemi 1588 PHYs and Switches

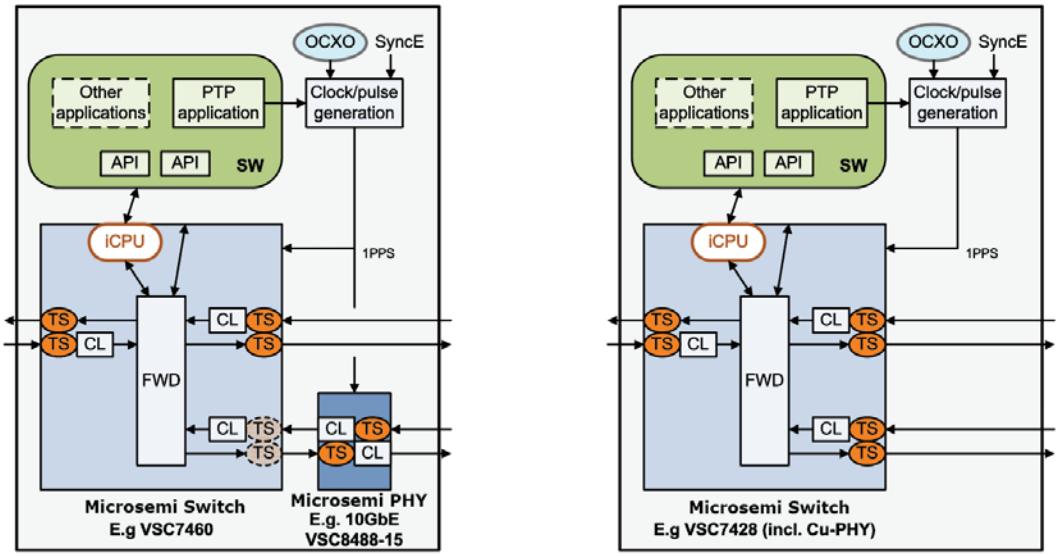
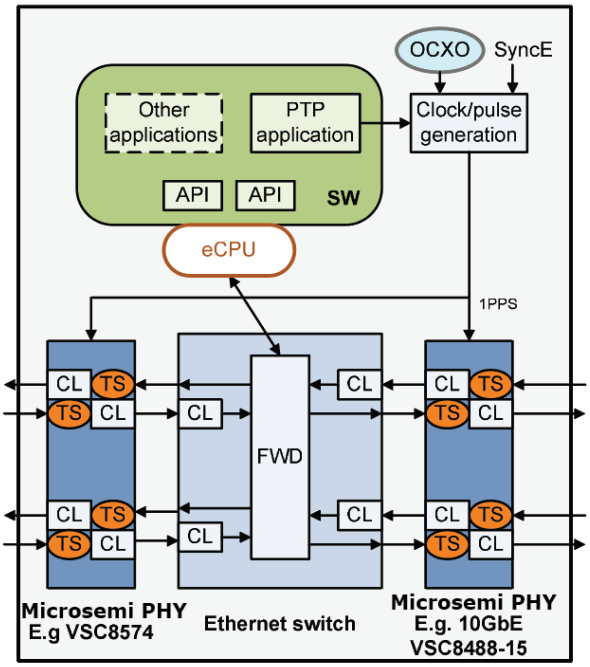


Figure 21 Pizza Box Upgrade with Microsemi 1588-PHYs



In Figure 21, the 1 GbE and 10 GbE PHYs perform the PTP processing because the switch is 1588-unaware.

The assumptions for the pizza box upgrades are:

- With Microsemi switches, all 1588 software runs on the integrated CPU (the iCPU). This is most cost-efficient; however, an external CPU is preferred in some situations.
- There is one CE switch where some ports directly attached to optics modules and others attach to a PHY (as seen on the left hand side of Figure 20, with a combination of 1 GbE optical and 10 GbE optical ports). There is another CE switch where PHYs are integrated with the switch, as in the VSC7428 (as seen on the right hand side of Figure 20, 1 GbE Cu and optical ports). The PHY time stamps are more accurate in the former case where the external links connect with a PHY, so the switch time stamps in that case are ignored.

Microsemi switch examples are the VSC7460 (Jaguar), VSC7462 (Lynx), VSC7428 (Caracal), and VSC7418 (Serval). Relevant Microsemi PHY examples are the VSC8574 (quad 1 GbE Cu-PHY), VSC8488-15 (dual 10 GbE), VSC8492-2/8494-2 (10 GbE with OTN).

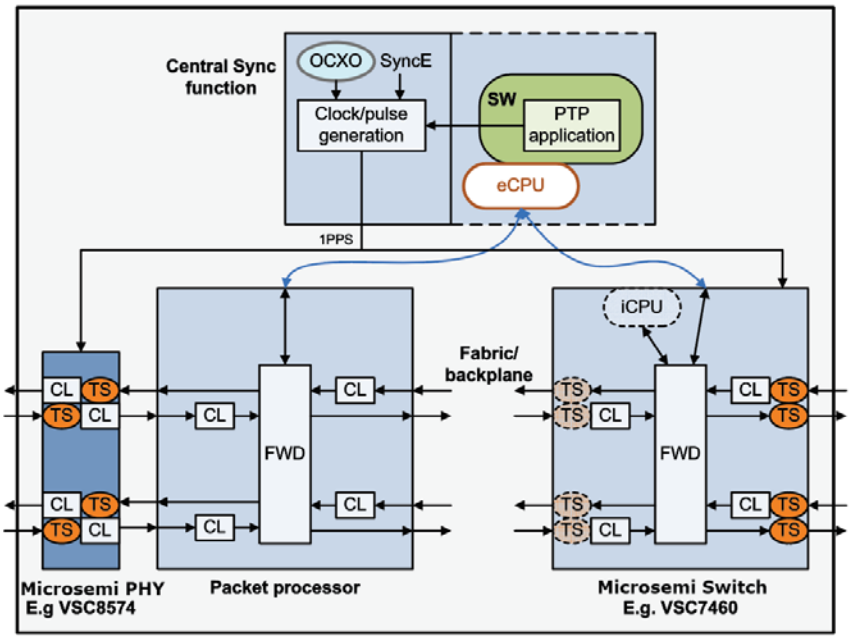
The architecture works as follows:

- Arrival time is recorded for frames arriving on external links at the ports of the switch or the PHY. If the classifier (CL) detects that a frame is indeed an event PTP frame, the time stamp is inserted in the frame or transported as metadata to the iCPU, a CPU-port, or the egress port. The PHY classifier searches for PTP frames to time stamp. The switch classifier also determines to which egress port a frame must be sent.
- At the egress port, the departure time is recorded and can be read by the CPU (E2E Delay request—t3), inserted in the packet as OriginTime (master operation), or the difference between departure time and arrival time is added to the correction field (TC operation). Also, asymmetry and link delay for P2P TC is added.

- In TC operation, the time stampers (TS) just need a common 1 pulse per second (PPS) reference based on a standard crystal oscillator. The 1 PPS signal does not need to be 1588 ToD-aligned. If TC is the only function required, the time stampers do not need align with ToD, but they need to agree on the local time of day. Furthermore, E2E solutions not involving PTP software are possible.
    - On the other hand, it is possible to syntonize the frequency of one TC domain by copying PTP frames to a local slave.
- OC/BC masters transmit sync messages whose OriginTime is inserted/updated at the egress port. Master ports must also participate in OCS delay measurement requests.
- For OC/BC slave operation, the PTP application receives PTP messages from the switch or PHY ports. The filter algorithm continuously estimates the delay from master to slave based on the most appropriate frames (the ones with lowest delay typically provide a better estimate). The result is used to adjust the 1 PPS signal directly or by adjusting the ToD on the ports.
- OC/BC slaves must be able to select between multiple masters. This is included in the 1588 PTP master selection algorithm (BMC). The BMC can have long hold-over periods when shifting to another master (for instance, due to a link failure). ITU has addressed this problem for frequency synchronization, and Microsemi has extended this for ToD synchronization (V-BMC). For more information, see section 5.2.1.
- For OC/BC operation, a stable, local clock is required. If the oscillator wander and drift is high, it is directly viewable on the resulting ToD uncertainty. The slave filter algorithm does not, so to speak, have stable ground on which to measure the master-slave delay. Ovenized crystal oscillators (OCXO) are stable across temperature and the only highly stable option. Other options include locking the local oscillator to the SyncE signal or a Stratum reference, if available.
- 2-step operation can also be supported.

The following illustration shows a chassis upgrade with Microsemi 1588 PHYs on one line card and a 1588 switch on another line card. It assumes a centralized synchronization architecture.

Figure 22 Chassis Architecture



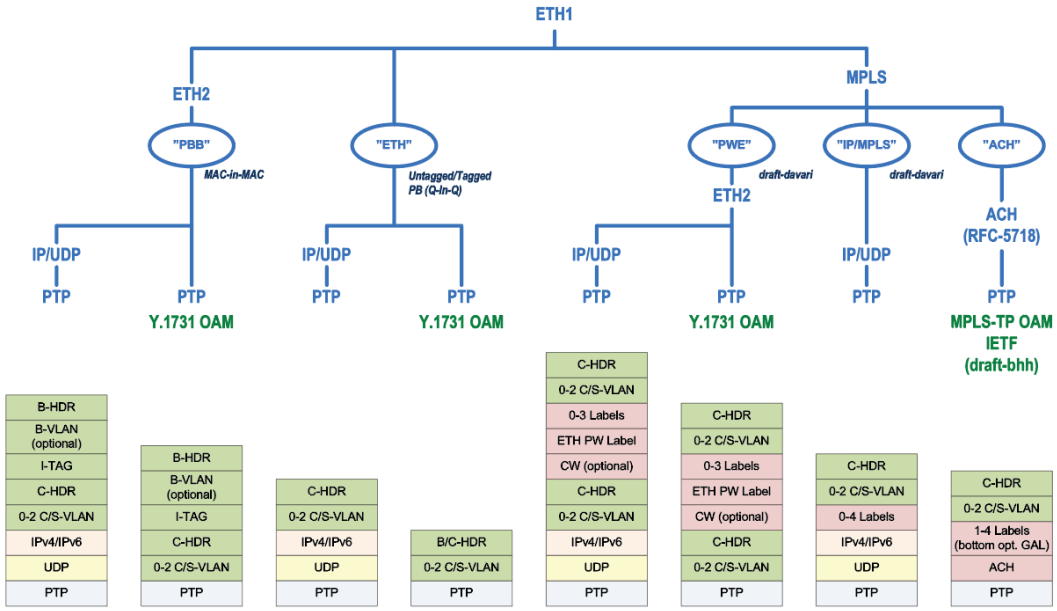
- A line card upgrade with Microsemi 1588 PHYs is shown on the left. The addition of a new line card to a legacy system with a 1588-aware Microsemi switch is shown on the right.
- The PTP application software is located on a central CPU. That CPU may be located on the sync module or it may be the general system CPU.
- Apart from the additional challenges of bridging multiple modules, the principle operation of this chassis solution does not differ from the pizza box system.

6.3 Encapsulations

Microsemi 1588-aware PHYs have sophisticated multi-protocol classifiers for timing and OAM. This includes IP/MPLS, Q-in-Q, MAC-in-MAC, and MPLS-TP encapsulations. Furthermore, 1G Cu/Optical, 10G LAN/WAN, and 10G OTN PHYs all have the same feature set, controlled via the same 1588-PHY API. Latest generation switches also support the full PHY features, whereas earlier generations focus on Ethernet encapsulations.

The PHYs can classify PTP frames encapsulated, as shown in the following illustration:

Figure 23 PTP Encapsulations Supported by VSC8488-15 (10 GbE PHY)



The corresponding frame formats are shown Figure 24 and Figure 25.

Figure 24 PBB and ETH Frame Formats

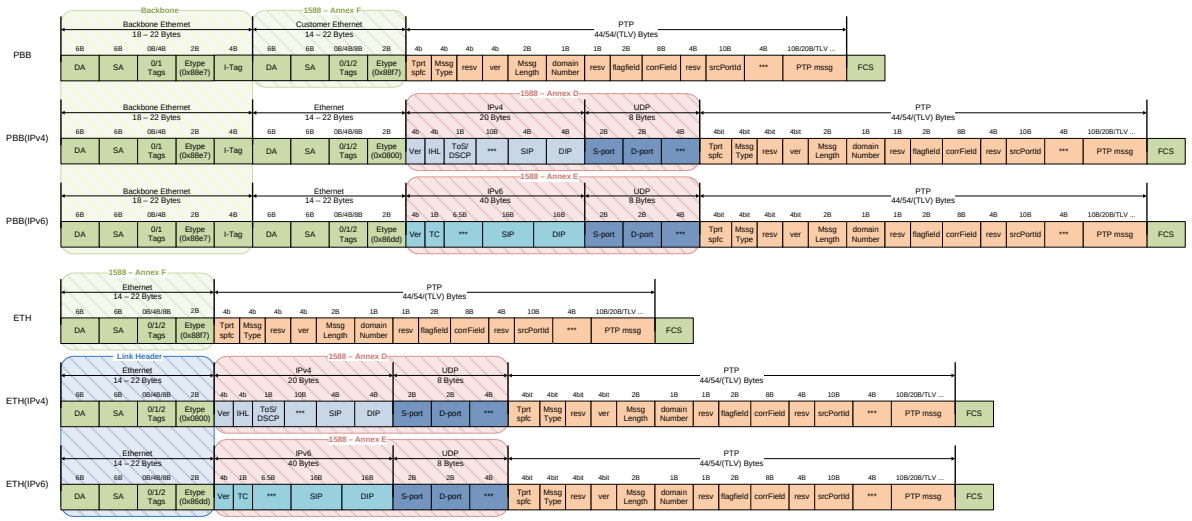
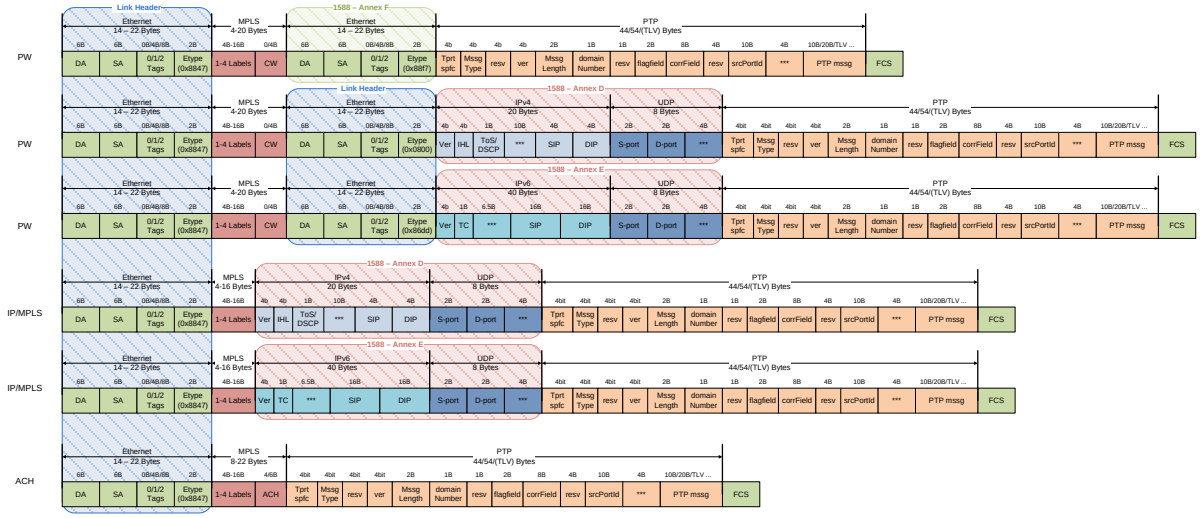


Figure 25 PW, IP/MPLS, and ACH Frame Formats



6.3.1 Carrying Time Stamps Internally

The time stamp can flexibly be inserted in the frame for transport between ingress and egress, including the CPU. The Microsemi default position is in the PTP reserved bytes. The time stamp is inserted in the reserved bytes at ingress and removed again at egress (overwritten with "0" per IEEE 1588v2 standard). This method has the following advantages:

- The same method can be used in support of TC and OC/BC operation.
- If for some reason the system leaks frames, the format is identifiable at end nodes and only results in minimal performance degradation.
- It is fully IEEE 1588v2 compliant externally.
- All calculation of residence time is done in one place. The egress ensures that the ToD counter-wrap does not affect the result.

7 Example of a Chassis System Upgrade

This section shows an example of upgrading a legacy system, assumed to be a CESR, with line cards containing a Microsemi PHY and switch. Although this shows the VSC8488-15 (10G PHY) and the VSC7460 (Jaguar CE switch) as examples, many of these considerations are also relevant for new designs.

7.1 Centralized and Distributed Approaches

Some networks are based on a BC-hierarchy in which each BC connects to another BC with no intermediate switches (1588-aware or -unaware). The root BC connects with the GM and the leaf BCs connect with multiple slaves (OC-S). In this case, the PTP traffic and session load is not at all large for a BC, apart from the leaf BCs. This scenario points to an implementation using centralized PTP software.

Other networks rely on fewer BCs, selecting TC-switches instead. This means more slaves are connected to the BC, which results in higher PTP traffic and session load. In this case, it might be necessary to distribute the relevant BC functions to the line cards in order to offload the central CPU and provide better scalability.

The example upgrade considered here is a distributed approach. A centralized approach is almost the same, only the PTP software functions are located on one CPU. Therefore, such a system can easily be derived from the distributed.

7.2 System Requirements

It is assumed that the chassis system is upgraded to support one boundary clock and to be a transparent clock for multiple clock domains. Also, the CESR still acts as a PTP-unaware element for frames that are intentionally hidden by the end user (for instance, in an MPLS-TP tunnel).

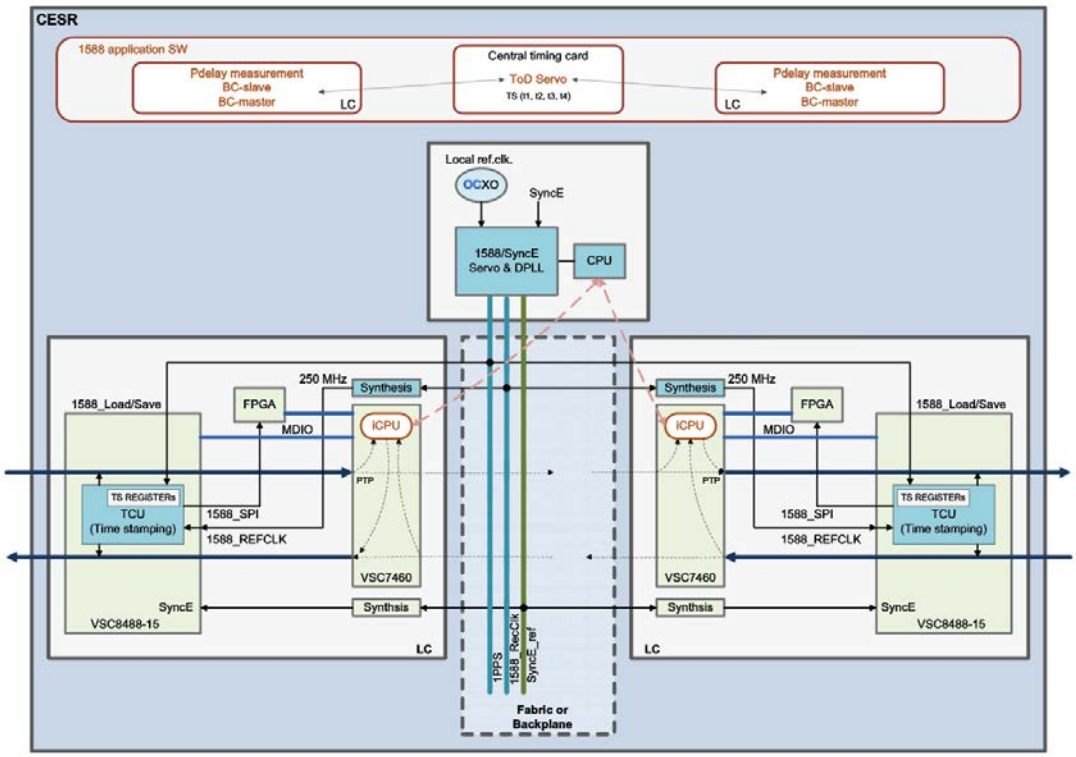
The following requirements are in overview:

- BC for one 1588 distribution network
- P2P/E2E TC for one or more 1588 distribution networks
- 1-step and 2-step operation
- SyncE, which may be used for better stability
- Equipment redundancy for BC-slave

7.3 The CESR Architecture

The parts of the system architecture relevant for 1588-operation are shown in the following illustration.

Figure 26 The CESR 1588 Architecture



It is here assumed that some form of 1 PPS distribution from a central synchronization module is available. The following lists the main elements of the architecture:

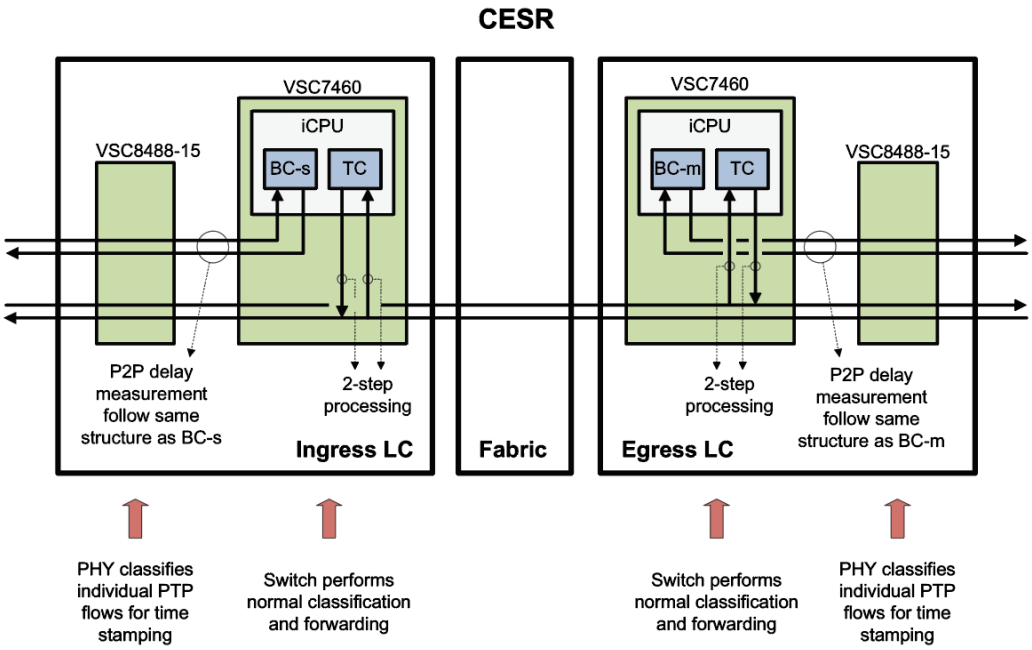
- Central BC-slave servo. All PTP applications needing software assistance can be handled centrally, but the solution is more scalable if certain functions are distributed to the LCs. It is assumed that the BC-S, BC-M, and Pdelay measurement functions are run on the VSC7460 internal CPUs on each LC.
- The control plane is used to transfer PTP and time stamp information between servo and line card iCPUs.
- There is a SyncE clock distribution and synchronization source selection system available, but the SyncE synchronization signal collection is not shown.
- There is a separate 1588 timing distribution system (for instance, a 155 MHz clock across the backplane). This provides a timing domain for PHY time stamping that is separate from SyncE. The backplane clock may be up-converted on the LC (for example, to 250 MHz) for better accuracy.
- 1 PPS signal distribution to all 1588 LCs. A fixed delay can be compensated for in the VSC8488-15. The 1588_load/save input is used for simultaneous update of ToD on the LCs. For BC and OC-S operation, the time stamping domains are the same as the upstream OC-M. For TC operation, the time stamping is in the local ToD domain that need not be the same domain as the OC-M initiating the TC flows.
- Egress time stamp harvesting is required for Delay_Req time stamp (t3) in 1-step solutions and when implementing 2-step solutions, in general. The PHY holds a time stamp FIFO, which can be accessed using MDIO, SPI, or parallel bus (PHY-dependent).

- With Microsemi hardware and software, this can be processed with less processor load, as the software can write the time stamp and the hardware updates this by adding the correction to the correction field. The software need not extract the time stamp from the time stamp FIFO.
- The BC implementation is distributed.
    - A distributed implementation divides the BC functionality into suitable blocks and allocates these to modules. One important division is into slave and master functionality. The BC-S communicates with the OC-M or BC-M and locks to its frequency and ToD. The BC-M communicates with the OC-Ss and provides them phase and ToD.
    - BC-S functionality is allocated to the LCs. This includes communication with the master (selection of the best master among those available on the LC); P2P (P2P delay measurement, Sync message processing, and transmit of ToD offset to the central servo); and E2E (t_1-t_4 harvesting and transmit to the central servo).
    - The selection of the best master amongst the LCs using the V-BMC algorithm must be handled centrally on the timing module for reasons of equipment redundancy. The V-BMC basically selects the GM to use based on connection availability, quality, and priority (as does the LC). The timing module works in a redundant configuration, meaning that the working servo and other central functions impose state on the standby functions for minimum switching time and disturbance.
    - The BC-master function is assigned to line cards facing the access network (or at least other slaves). One BC-M on a LC can support down-stream slaves over the available ports. If this does not suffice, more independent BC-M LCs can be activated.
    - BC-S and BC-M belong to the same PTP domain.
- TC occurs between TC-enabled ports. TC flows may belong to the same PTP domain (for example, OC-M or BC-M connecting to multiple BC-S or OC-S), or they may belong to different PTP domains, such as different operators.
- BC and TC traffic may run over the same ports.

7.3.1 PTP Packet Injection/Extraction

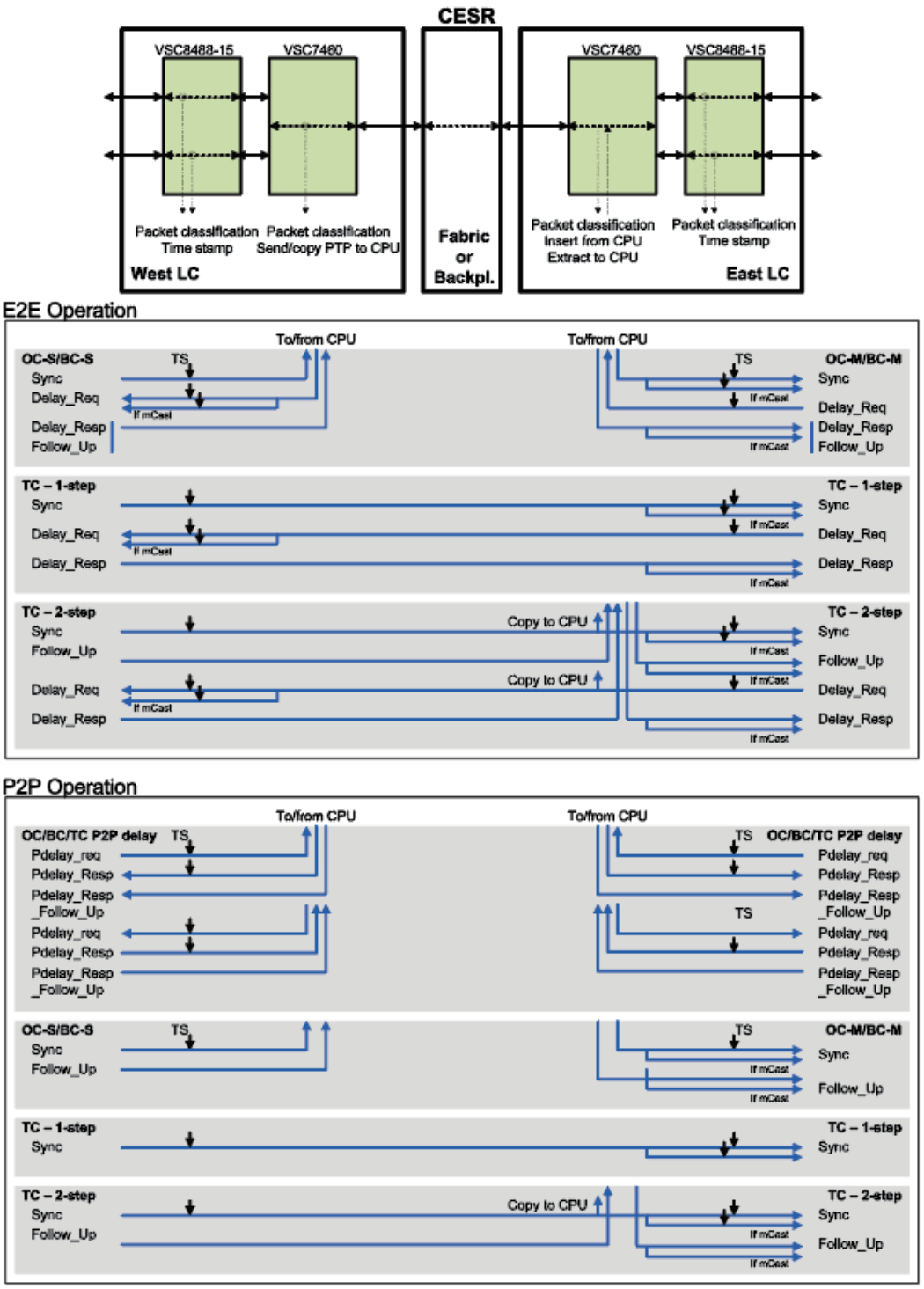
Figure 27 and Figure 28 show the PTP traffic initiated by, passing through, and terminating in the CESR. Message packets in these three cases are all time stamped and depend on clock function. Other types of messages (Announce, Follow_Up, Delay_Resp, Pdelay_Resp_Follow_Up, Signaling, and Management) are not time stamped.

Figure 27 BC/TC PTP Frame Transport and Points of Classification



In the previous illustration, initiation and termination require receipt upon sending to and receiving from the iCPU. Extraction to iCPU (forward or copy to iCPU) and insertion from iCPU is done in the VSC7460.

Figure 28 PTP Traffic Through the CESR



Notes:
Multicast traffic is sent to multiple ports and is indicated with "if mCast"
**_Follow_Up messages are only transmitted from 2-step nodes*

The previous illustration does not show frame types.

7.4 Classification Hierarchy Considerations

In general, arriving frames are analyzed twice: first in the PHY (or switch) to detect if the frame must be time stamped, and then in the switch to determine the egress port. This section proposes methods to reduce the amount of classifier configuration necessary and save classification resources.

The approach taken here is also illustrated in Figure 27. The PHY must classify specifically for all PTP flows, whereas by proper network design, the switch can classify for groups of flows.

In a PTP network consisting of multiple domains, each domain is structured in a hierarchy. A slave (BC-S, OC-S) has a few masters and a master (OC-M, BC-M) can have many slaves. A particular CESR participates in a single PTP domain. This expands to multiple domains—for example, a BC and may act as TC for multiple domains. Therefore, we arrive at the following assumptions:

- BC-S: a few encapsulations may apply, such as one for the primary connection with the GM and one for the backup GM. If unicast (UC), use one address per GM.
- BC-M: a few encapsulations may apply because the downstream network is most likely is only one or a few types (for instance, Ethernet, IP, MPLS-TP, and so on). If UC has one address per OC-S, then there are many active flows.
- TC: more encapsulations may apply. If UC has many addresses, then there are many active flows.

The following table suggests methods to confine the variation in encapsulation, depending on the application.

Table 2 Encapsulation and Network Operation

Clock Type	Encapsulation Branch	MC/UC	Network Operation
BC	PBB	ETH MC/UC	Confine PTP traffic to one ISID, with or without B-tag. Ethernet S-VLANs should only be used if part of the network is non-PBB. In that case, confine to a single S-VLAN (single/double tagged).
	PBB(IP)	IP MC/UC	Same as PBB.
	ETH	ETH MC/UC	Confine PTP traffic to one VLAN, single or double tagged.
	ETH(IP)	IP MC/UC	Same as ETH.
	PW-ETH	ETH MC/UC	Confine PTP traffic to one PW or to one label range, with or without tunnel. Ethernet VLANs should only be used if part of the network is non-MPLS. In that case, confine to a single VLAN (single/double tagged).
	PW-ETH(IP)	IP MC/UC	Same as PW-ETH.
	IP/MPLS	IP MC/UC	Confine PTP traffic to one VC LSP, with or without tunnel.
	ACH	p-2-p	Same as IP/MPLS.
TC	PBB	ETH MC/UC	Confine PTP traffic to a set of B-VLANs. S-VLANs should only be used if part of the network is non-PBB. In that case, confine to a single VLAN (single/double tagged).

Clock Type	Encapsulation Branch	MC/UC	Network Operation
	PBB(IP)	IP MC/UC	Same as PBB.
	ETH	ETH MC/UC	Confine PTP traffic to a set of VLANs (single tagged), or one outer tag (indicating PTP traffic) and N inner tags (separating customers).
	ETH(IP)	IP MC/UC	Same as ETH.
	PW-ETH	ETH MC/UC	Confine PTP traffic to a set of PWs (single label), or one outer tunnel (indicating PTP traffic) and N PWs (separating customers). Ethernet VLANs should only be used if part of the network is non-MPLS. In that case, confine to a single VLAN (single/double tagged).
	PW-ETH(IP)	IP MC/UC	Same as PW-ETH.
	IP/MPLS	IP MC/UC	Confine PTP traffic to a set of VC LSPs (single label), or one outer tunnel (indicating PTP traffic) and N PWs (separating customers).
	ACH	p-2-p	Same as IP/MPLS.

7.4.1 PHY and Switch Classification Details

This section describes the combined classification operation of the VSC8488-15 PHY and VSC7460 switch as an example.

The following table shows classifications in the ingress and egress PHYs and switches. ETH and IP encapsulations are considered for the cases of ingress BC and ingress-to-egress TC/1-step.

Table 3 Classifications (clock type)

	Addressing	ETH MC (BC)	ETH UC (BC)	IP MC (TC/1-Step)	IP UC (TC/1-Step)
iPHY	M-DA	d/c	d/c	d/c	d/c
	O-VID	Match	Match	Match	Match
	I-VID	Match	Match	Match	Match
	Etype	Match	Match	= v4/v6	= v4/v6
	DIP	N/A	N/A	d/c	d/c
	UDP-Port	N/A	N/A	= 319	= 319
	Message Type	MSB=0	MSB=0	MSB=0	MSB=0
	Domain No.	Match	Match	Match	Match
	Action	Place TS in Rsvr Bytes	Place TS in Rsvr Bytes	Place TS in Rsvr Bytes	Place TS in Rsvr Bytes
iSwitch	Port	Line	Line	Line	Line
	M-DA	d/c	d/c	Match	Match
	O-VID	Match	Match	Match	Match
	I-VID	Match	Match	Range	Range
	Etype	Match	Match	d/c	d/c
	DIP	N/A	N/A	d/c	d/c
	UDP-Port	N/A	N/A	d/c	d/c
	Message Type	d/c	d/c	d/c	d/c

	Addressing	ETH MC (BC)	ETH UC (BC)	IP MC (TC/1-Step)	IP UC (TC/1-Step)
	Domain No.	Match	Match	d/c	d/c
	Action	Send to CPU	Send to CPU	Forward to X port	Forward to X port
eSwitch	Port			Host	Host
	M-DA			Match	Match
	O-VID			Match	Match
	I-VID			Range	Range
	Etype			d/c	d/c
	DIP			d/c	d/c
	UDP-Port			d/c	d/c
	Message Type			d/c	d/c
	Domain No.			d/c	d/c
	Action			Forward to X port	Forward to X port
ePHY	M-DA			d/c	d/c
	O-VID			Match	Match
	I-VID			Match	Match
	Etype			= v4/v6	= v4/v6
	DIP			d/c	d/c
	UDP-Port			= 319	= 319
	Message Type			MSB=0	MSB=0
	Domain No.			Match	Match
	Action			Update CF	Update CF

The basic principles are:

- The PHY classifies individual PTP flows for time stamping and delay/asymmetry compensation.
- If PTP traffic is confined to a VLAN or a DMAC and all traffic within them is destined to the iCPU, the classification in the switch can be done in the basic classification without use of TCAM resources.
- Similarly, for TC operation, if PTP traffic is confined to a VLAN or a DMAC and all traffic within them is destined to the same egress port(s), the classification in the switch can be done in the basic classification without use of TCAM resources.
- Using TCAMs, the VSC7460 classifier configuration for BC and P2P link delay measurement is detailed to the required level (but this will in practice only be relevant for a few concurrent encapsulation types).
- VSC7460 classifier configuration for TC basically does normal forwarding.

The PTP Classification spreadsheet provides more detail and shows classifications in ingress and egress PHYs and switches. ETH, IP, and IP/MPLS encapsulations are considered for the cases of BC, TC 1- and 2-step, and P2P delay measurement.

7.5 Analogue System Considerations

This section summarizes resulting ToD accuracy as well as the influence of analogue parameters and quantization noise.

Four basic elements determine slave performance:

- Time stamp accuracy
- Oscillator stability
- Tolerance to PDV
- Servo filter algorithm

Time stamp uncertainty directly affects the accuracy of the recovered ToD. High-precision time stamping cannot simply be "bolted on"—only careful design results in the performance needed. Improvement can be made by minimizing the number of clock domain crossings, introducing logic on framed interfaces (for example, WAN and OTN interfaces) that predict when the PTP packet entered the device, and using a high sampling frequency.

The oscillator, from which the 1 PPS signal transmits, used for synchronizing the devices must be stable because its quality determines the noise floor of the recovered ToD. Depending on the required accuracy, ovenized crystal oscillators (OCXO) may be needed. If the clock is part of a SyncE network, an even more stable 1 PPS can be generated and an OCXO may be unnecessary.

The PTP packets' delay through the network depends on network load. ToD variations may seem to exhibit master clock wander patterns. The use of transparent clocks, as supported by Microsemi PHYs and CE switches between master and slave, eliminates this effect because residence time is compensated for.

Alternatively, if the network consists of legacy equipment that does not support 1588, an advanced filter algorithm must be used in the slave's loop filter, which discards PTP packets exhibiting extraordinary large delays. These filters are included in the Microsemi solution.

8 Summary

The number of mobile users and their traffic is constantly growing. To transport this traffic, a cost-efficient packet technology is required. Concurrently, the advanced multipath antenna technologies of LTE and LTE-Advanced put stringent requirements on frequency and phase relationships.

Microsemi's carrier Ethernet switches and PHYs are designed to meet these challenges and provide advanced QoS, OAM, and synchronization features.

The Microsemi 1588 solution is applicable to both pizza boxes and chassis-based systems for upgrading existing systems and new designs.

As evidenced by the measurement results shown in the following illustration, excellent performance is possible with the Microsemi 1588 solution.

Figure 29 Chain of Transparent Clocks (achieving 80 ns MTIE over 20 ks)