7. Time-synchronization model for a packet network

7.1 General

This clause provides a model for understanding the operation of the generalized precision time protocol (gPTP), which specifies the operation of time-aware systems on a packet network. Although this standard is based on the precision time protocol (PTP) described in IEEE Std 1588-2019 (and, indeed, is a proper profile of IEEE Std 1588-2019 in particular configurations), there are differences, which are summarized in 7.5.

Although this standard has been written as a stand-alone document, it is useful to understand the IEEE 1588 architecture as described in Clause 6 of IEEE Std 1588-2019.

7.2 Architecture of a time-aware network

7.2.1 General

A time-aware network consists of a number of interconnected time-aware systems that support the gPTP defined within this standard. These time-aware systems can be any networking device, including, for example, bridges, routers, and end stations. A set of time-aware systems that are interconnected by gPTP-capable network elements is called a *gPTP network*. Each instance of gPTP that the time-aware systems support is in one *gPTP domain*, and the instances of gPTP are said to be part of that gPTP domain. A time-aware system can support, and therefore be part of, more than one gPTP domain. The entity of a single time-aware system that executes gPTP in one gPTP domain is called a *PTP Instance*. A time-aware system can contain multiple PTP Instances, which are each associated with a different gPTP domain. There are two types of PTP Instances:

a) PTP End Instance, which, if not a Grandmaster PTP Instance, is a recipient of time information, and
b) PTP Relay Instance, which, if not a Grandmaster PTP Instance, receives time information from the Grandmaster PTP Instance (perhaps indirectly through other PTP Relay Instances), applies corrections to compensate for delays in the local area network (LAN) and the PTP Relay Instance itself, and retransmits the corrected information.

This standard defines mechanisms for delay measurements using standard-based procedures for the following:

c) IEEE 802.3 Ethernet using full-duplex point-to-point links (11)
d) IEEE 802.3 EPON links (Clause 13)
e) IEEE 802.11 wireless (Clause 12)
f) Generic coordinated shared networks (CSNs, e.g., MoCA and G.hn) (Clause 16)

7.2.2 Time-aware network consisting of a single gPTP domain

Figure 7-1 illustrates an example time-aware network consisting of a single gPTP domain, using all the above network technologies [i.e., item c) through item f) of 7.2.1], where end stations on several local networks are connected to a Grandmaster PTP Instance on a backbone network via an EPON access network.

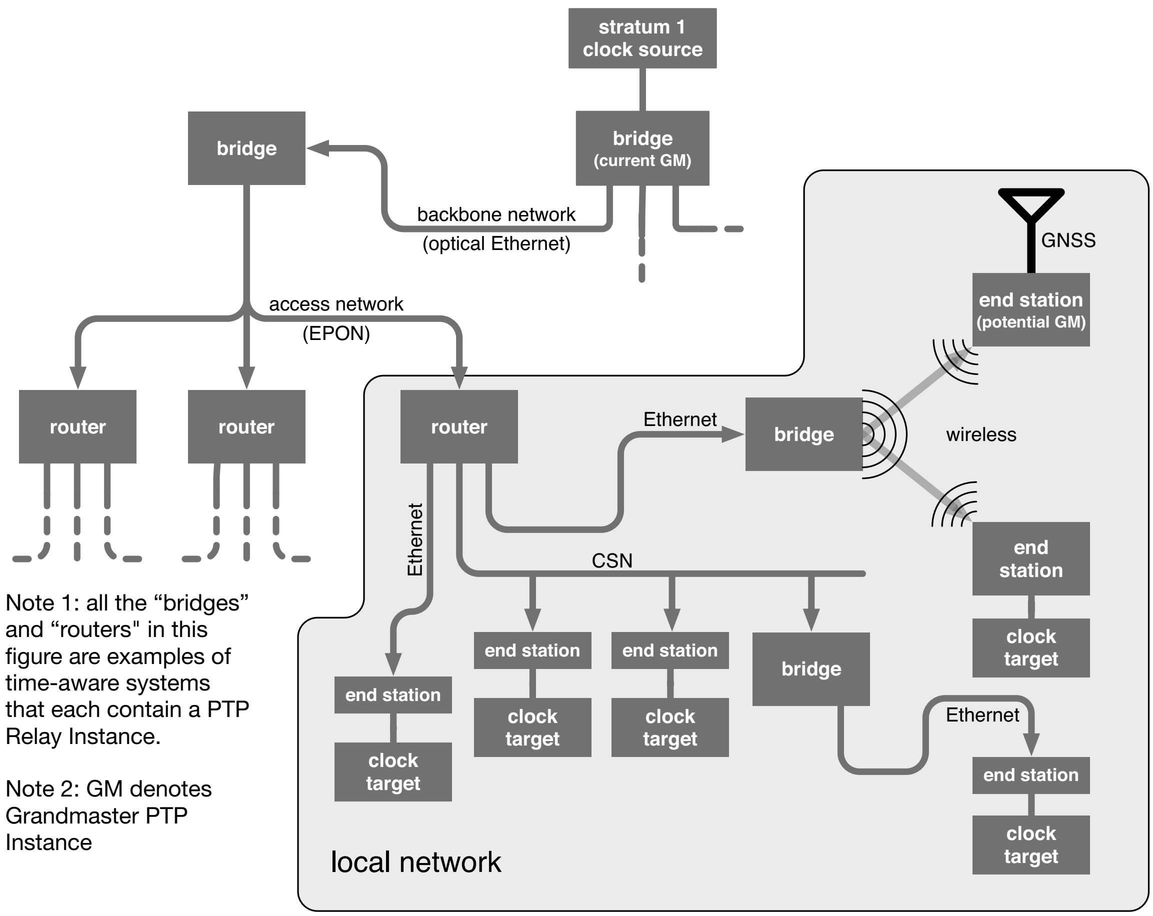
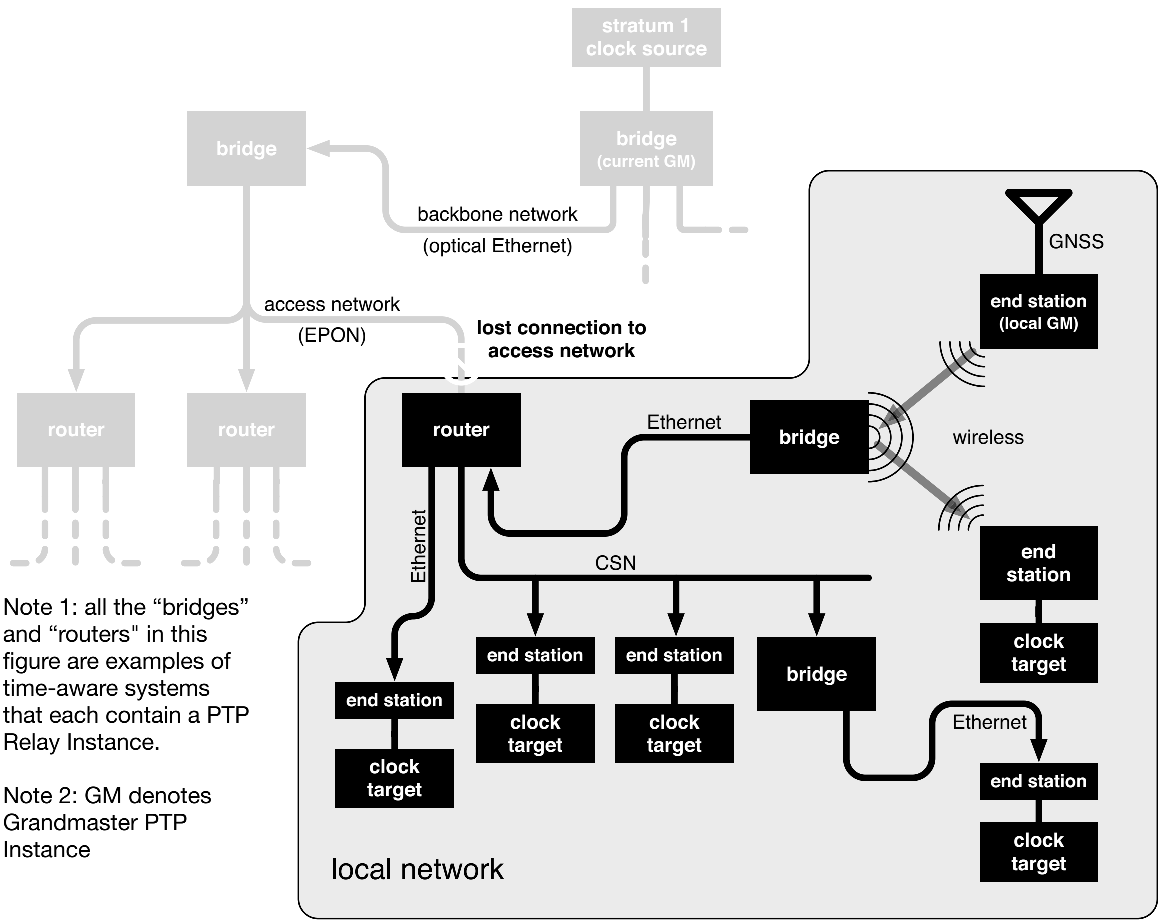


Figure 7-1—Time-aware network example

Any PTP Instance with clock sourcing capabilities can be a potential Grandmaster PTP Instance, and a selection method (the *best master clock algorithm*, or BMCA) ensures that all of the PTP Instances in a gPTP domain use the same Grandmaster PTP Instance.[12] The BMCA is largely identical to that used in IEEE Std 1588-2019, but somewhat simplified. In Figure 7-1 the BMCA process has resulted in the Grandmaster PTP Instance being on the network backbone. If, however, the access network fails, the systems on a local network automatically switch over to one of the potential Grandmaster PTP Instances on the local network that is as least as "good" as any other. For example, in Figure 7-2, the access network link has failed, and a potential Grandmaster PTP Instance that has a GNSS reference source has become the active Grandmaster PTP Instance. As a result, now two gPTP domains exist where there used to be one. Finally, note that when a time-aware system supports more than one domain, one of the domains supported must be domain 0 for backward compatibility with the 2011 edition of this standard, though domain 0 is not necessarily active in a time-aware system.



Note 1: all the "bridges" and "routers" in this figure are examples of time-aware systems that each contain a PTP Relay Instance.

Note 2: GM denotes Grandmaster PTP Instance

Figure 7-2—Time-aware network of Figure 7-1 after an access network link failure

[12] There are, however, short periods during network reconfiguration when more than one Grandmaster PTP Instance might be active while the BMCA process is taking place.

7.2.3 Time-aware network consisting of multiple gPTP domains

Figure 7-3 illustrates an example time-aware network consisting of multiple gPTP domains that could be used in an industrial application. Specifically, in this example the network has two timescales/domains, where domain 0 uses the PTP timescale and domain 1 uses the arbitrary (ARB) timescale (see 8.2). Notice that not all PTP Instances in domain 1 (within the blue shorter-dashed area) have domain 0 active in this example, even though every time-aware system supports domain 0 for backward compatibility with the 2011 edition of this standard. In addition, it is required that all PTP Instances belonging to the same domain have direct connections among them in their physical topology (e.g., time cannot be transported from one PTP Instance in domain 0 to another PTP Instance in domain 0 via a time-aware system that does not have domain 0 active). In addition, the time-aware systems for which both domains are active are depicted by slanted internal hatching, representing two independent, active PTP Instances.

As in the single-domain case, any of the network technologies of 7.2.1 can be used. The Grandmaster PTP Instance of each domain is selected by the BMCA; in this case, a separate, independent instance of the BMCA is invoked in each domain.

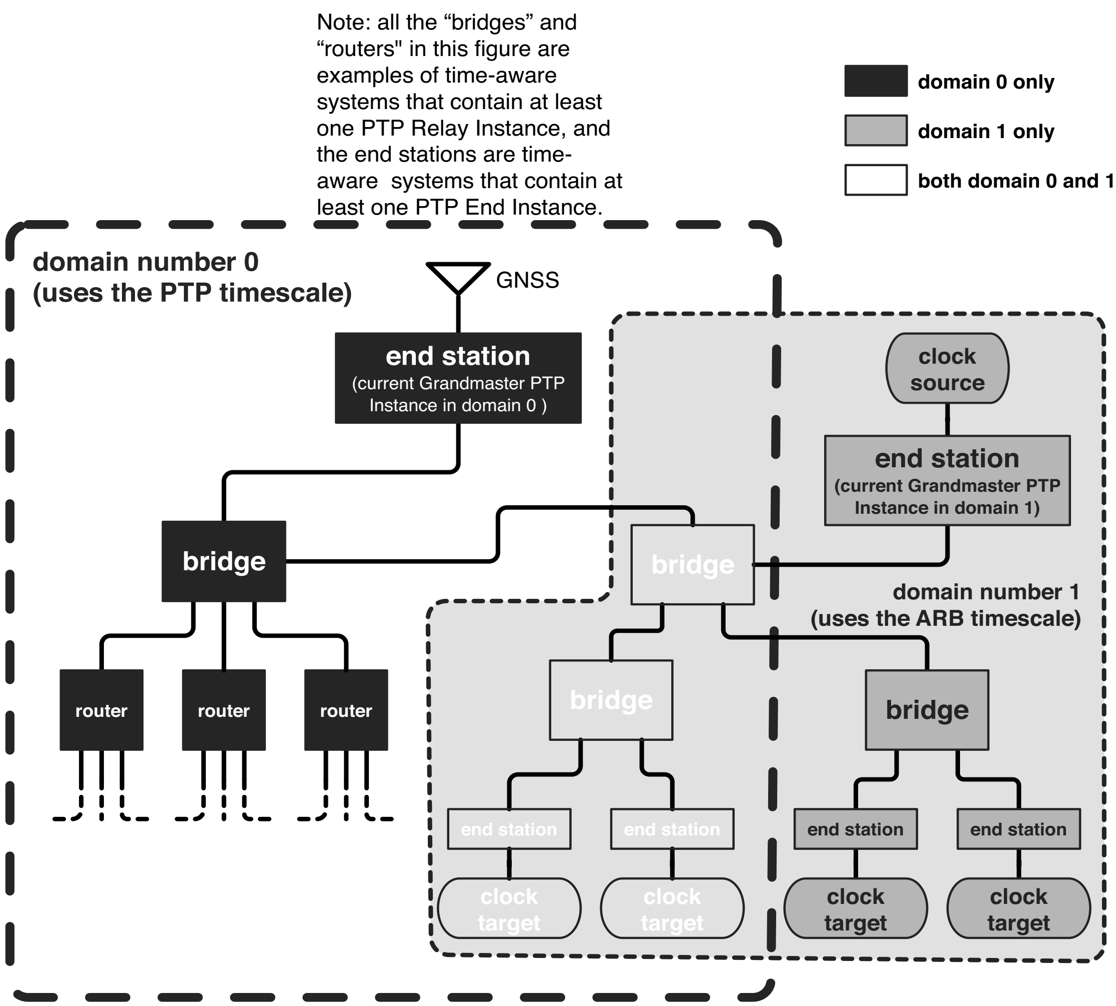


Figure 7-3—Time-aware network example for multiple gPTP domains

7.2.4 Time-aware networks with redundant Grandmaster PTP Instances and/or redundant paths

7.2.4.1 General

Redundancy has many levels of sophistication, performance, and cost. Therefore, the appropriate level and/or amount of redundancy required in a time-aware network can be very different for each application. Nonetheless, all solutions for redundancy consist of a detection component, a correction component, and an action component. The detection component detects that something is not working correctly. The correction component determines the appropriate corrective action. The action component performs the required action(s) to fix the detected problem.

7.2.4.2 Redundancy specified in this standard (BMCA)

This standard provides a basic level of redundancy as follows:

— A detection component that triggers when the current Grandmaster PTP Instance stops working (i.e., loss of Sync messages and Announce messages for a period of time) or if the link to the Grandmaster PTP Instance goes down (i.e., immediate loss of Sync messages and Announce messages).
— A correction component that triggers the Best Master Clock Algorithm (BMCA) and the sending of Announce messages so that a new Grandmaster PTP Instance can be elected.
— An action component, where the winning Grandmaster PTP Instance starts sending Announce messages and Sync messages and all the PTP Instances listen to this new Grandmaster PTP Instance.

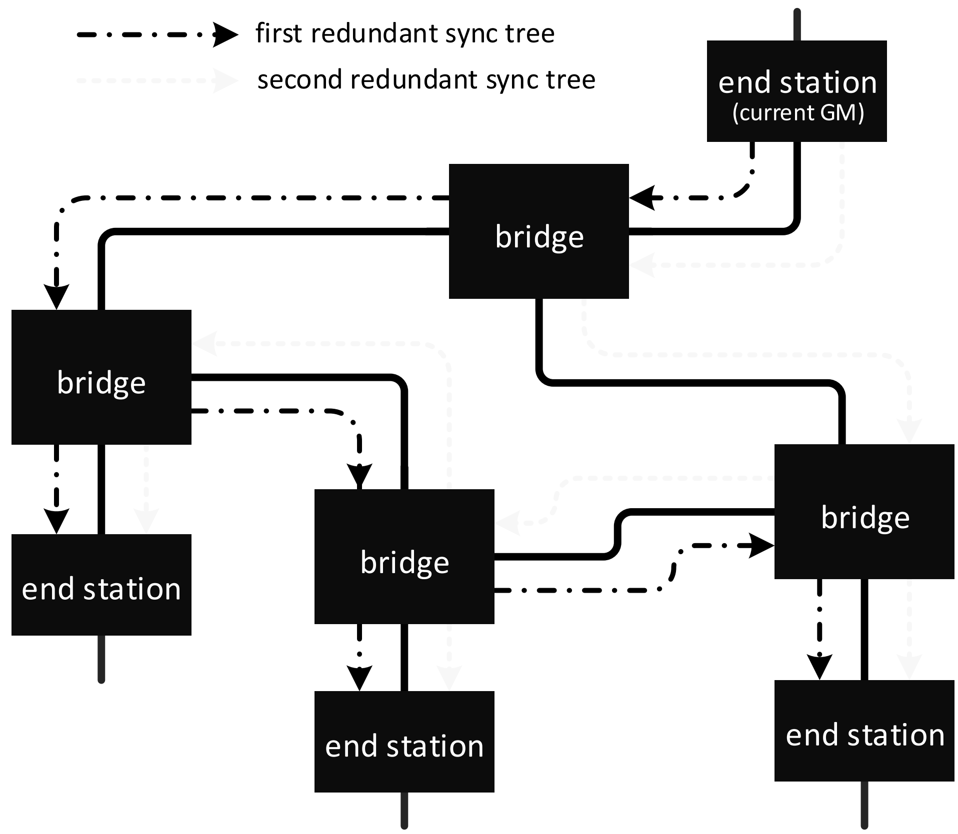
7.2.4.3 Redundacy not fully specified in this standard

In addition to providing the basic level of redundancy, this standard provides the ability to support more sophisticated network configurations that provide additional levels of Grandmaster PTP Instance and clock path redundancy. Figure 7-4 through Figure 7-6 are examples of such networks that provide these additional levels of redundancy as a way to deal with these failures. The information necessary to implement and configure these network configurations is contained in this standard.

In order to take advantage of these failure correction configurations, new types of fault detection are required. The category of fault detection where a Grandmaster PTP Instance completely fails and stops sending clock information is supported as mentioned above.

Other types of faults involve instability of the Grandmaster Clock, such as time glitches, excess jitter, or wander, or various other impairments that could occur in the Grandmaster Clock. Techniques for identifying these types of failures, and the appropriate correction necessary, are not specified in this standard. However, if other techniques or standards are used for detection and correction of these types of failures, this standard provides the means to recover from these errors.

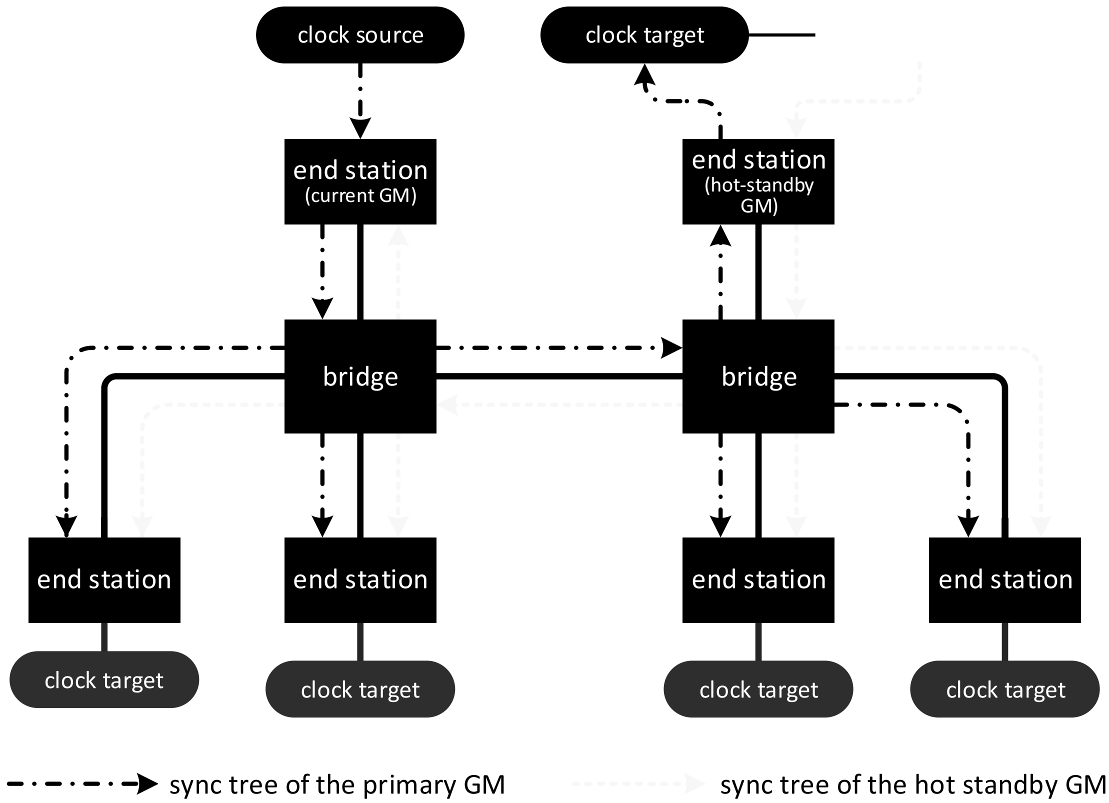
Figure 7-4 shows an example network realizing two redundant synchronization trees from a single GM, with each synchronization tree in a different gPTP domain (there are a total of two gPTP domains).



Note 1: The methods used for merging the redundant sync msgs received at each end station are not specified in this standard.
Note 2: All the “bridges” in this figure are examples of time-aware systems that contain PTP Relay Instances, and the end stations are examples of time-aware systems that contain PTP End Instances.
Note 3: GM denotes Grandmaster PTP Instance

Figure 7-4—Time-aware network example for synchronization path redundancy, with one clock source providing time to two domains

Figure 7-5 shows an example network with two redundant GMs, one as primary GM and the other as secondary GM, where each GM has one of the two redundant synchronization trees originating from it. This example supports hot-standby operating mode. In this mode, the secondary GM has to be synchronized to the primary GM, because it is part of the synchronization tree of the primary GM as shown in the figure.

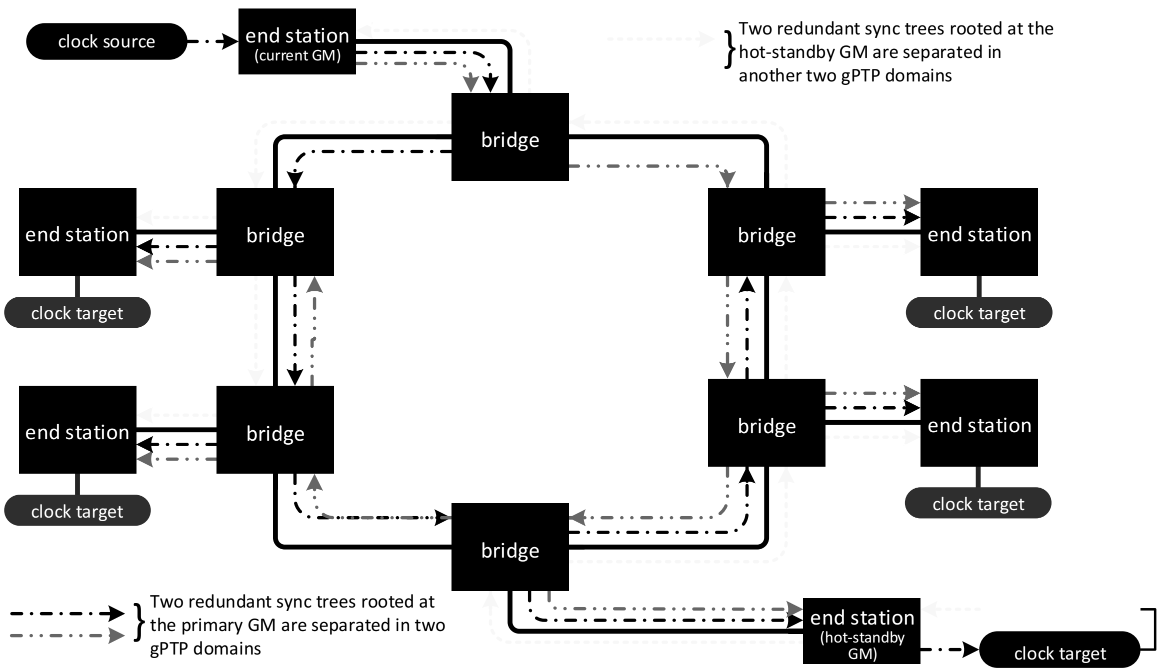


Note 1: The methods used for merging the redundant sync msgs received at each end station are not specified in this standard.
Note 2: All the “bridges” in this figure are examples of time-aware systems that contain PTP Relay Instances, and the end stations are examples of time-aware systems that contain PTP End Instances.
Note 3: GM denotes Grandmaster PTP Instance

Figure 7-5—Time-aware network example for GM redundancy with one primary GM and one hot-standby GM, which are separated in two gPTP domains

Figure 7-6 shows another example network, which involves ring topology, using the redundancy features of both Figure 7-4 and Figure 7-5.

For the techniques shown in the examples of Figure 7-4, Figure 7-5, and Figure 7-6, the detection component, correction component, and action component are not fully specified in this standard.



Note 1: The methods used for merging the redundant sync msgs received at each end station are not specified in this standard.
Note 2: All the "bridges" in this figure are examples of time-aware systems that contain PTP Relay Instances, and the end stations are examples of time-aware systems that contain PTP End Instances.
Note 3: GM denotes Grandmaster PTP Instance

Figure 7-6—Time-aware network example for GM+synchronziation path redundancy, with one primary and one hot-standby GM. Each GM establishes two sync trees, resulting in a total of four Sync trees that are separated in four gPTP domains

7.3 Time synchronization

7.3.1 General

Time synchronization in gPTP is done the same way (in the abstract) as is done in IEEE Std 1588-2019: a Grandmaster PTP Instance sends information including the current synchronized time to all directly attached PTP Instances. Each of these PTP Instances must correct the received synchronized time by adding the propagation time needed for the information to transit the gPTP communication path from the Grandmaster PTP Instance. If the PTP Instance is a PTP Relay Instance, then it must forward the corrected time information (including additional corrections for delays in the forwarding process) to all the other attached PTP Instances.

To make this all work, there are two time intervals that must be precisely known: the forwarding delay (called the *residence time*), and the time taken for the synchronized time information to transit the gPTP communication path between two PTP Instances. The residence time measurement is local to a PTP Relay Instance and easy to compute, while the gPTP communication path delay is dependent on many things including media- dependent properties and the length of the path.

7.3.2 Delay measurement

Each type of LAN or gPTP communication path has different methods for measuring propagation time, but they are all based on the same principle: measuring the time that a well-known part of a message is transmitted from one device and the time that the same part of the same message is received by the other device, then sending another message in the opposite direction and doing the same measurement as shown in Figure 7-7.

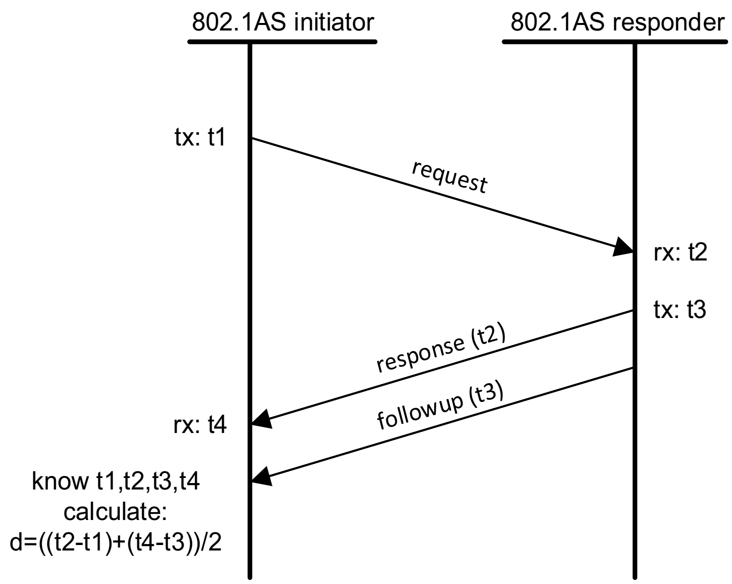


Figure 7-7—Conceptual medium delay measurement

This basic mechanism is used in the various LANs in the following ways:

a) Full-duplex Ethernet LANs use the two-step peer-to-peer (P2P) path delay algorithm as defined in IEEE Std 1588-2019, where the messages are called Pdelay_Req, Pdelay_Resp, and Pdelay_Resp_Follow_Up (see Figure 11-1).

b) IEEE 802.11 wireless LANs use the Timing Measurement (TM) procedure or the Fine Timing Measurement (FTM) procedure defined in IEEE Std 802.11-2016. The Timing measurement messages are the “Timing Measurement frame” and its corresponding “Ack” (see Figure 12-1). The Fine Timing Measurement messages are the “initial FTM request frame” and the “Fine Timing Measurement frame” and its “Ack” (see Figure 12-2).

c) EPON LANs use the discovery process, where the messages are “GATE” and “REGISTER_REQ” (see Clause 64 and Clause 77 of IEEE Std 802.3-2018).

d) CSNs either use the same mechanism as full-duplex Ethernet or use a method native to the particular CSN (similar to the way native methods are used by IEEE 802.11 networks and EPON) (see Figure 16-5).

7.3.3 Logical syntonization

The time-synchronization correction previously described is dependent on the accuracy of the delay and residence time measurements. If the clock used for this purpose is frequency locked (syntonized) to the Grandmaster Clock, then all the time interval measurements use the same time base. Since actually adjusting the frequency of an oscillator (e.g., using a phase-lock loop) is slow and prone to gain peaking effects, PTP Relay Instances can correct time interval measurements using the Grandmaster Clock frequency ratio.

Each PTP Instance measures, at each PTP Port, the ratio of the frequency of the PTP Instance at the other end of the link attached to that PTP Port to the frequency of its own clock. The cumulative ratio of the Grandmaster Clock frequency to the local clock frequency is accumulated in a standard organizational type, length, value (TLV) attached to the Follow_Up message (or the Sync message if the optional one-step processing is enabled). The frequency ratio of the Grandmaster Clock relative to the local clock is used in computing synchronized time, and the frequency ratio of the neighbor relative to the local clock is used in correcting the propagation time measurement.

The Grandmaster Clock frequency ratio is measured by accumulating neighbor frequency ratios for two main reasons. First, if there is a network reconfiguration and a new Grandmaster PTP Instance is elected, the nearest neighbor frequency ratios do not have to be newly measured as they are constantly measured using the Pdelay messages. This results in the frequency offset relative to the new Grandmaster Clock being known when the first Follow_Up message (or first Sync message if the optional one-step processing is enabled) is received, which reduces the duration of any transient error in synchronized time during the reconfiguration. This is beneficial to many high-end audio applications. Second, there are no gain peaking effects because an error in frequency offset at one PTP Relay Instance, and resulting residence time error, does not directly affect the frequency offset at a downstream PTP Relay Instance.

7.3.4 Grandmaster PTP Instance (best master) selection and network establishment

All PTP Instances participate in best master selection so that the IEEE 802.1AS protocol can determine the synchronization spanning tree. This synchronization spanning tree can be different from the forwarding spanning tree determined by IEEE 802.1Q™ Rapid Spanning Tree Protocol (RSTP) since the spanning tree determined by RSTP can be suboptimal or even inadequate for synchronization or can be for a different topology of nodes from the synchronization spanning tree.

gPTP requires that all systems in the gPTP domain be time-aware systems, i.e., the protocol does not transfer timing over systems that are not time-aware (e.g., those that meet the requirements of IEEE Std 802.1Q-2018, but do NOT meet the requirements of the present standard). A time-aware system uses the peer-to-peer delay mechanism on each PTP Port to determine if a non-time-aware system is at the other end of the link or between itself and the Pdelay responder. If, on sending Pdelay_Req,

a) No response is received,
b) Multiple responses are received, or
c) The measured propagation delay exceeds a specified threshold, then

the protocol concludes that a non-time-aware system or end-to-end Transparent Clock (TC) (see IEEE Std 1588-2019) is present. In this case, the link attached to the PTP Port is deemed not capable of running gPTP, and the BMCA ignores it. However, the PTP Port continues to attempt the measurement of propagation delay using the peer-to-peer delay mechanism (for full-duplex IEEE 802.3 links), multipoint control protocol (MPCP) messages (for EPON), or IEEE 802.11 messages (for IEEE 802.11 links), and periodically checks whether the link is or is not capable of running the IEEE 802.1AS protocol.

7.3.5 Energy efficiency

Sending PTP messages at relatively high rates when there is otherwise little or no traffic conflicts with the goal of reducing energy consumption. This standard specifies a way to request that a neighbor PTP Port reduce the rate of sending Sync (and Follow_Up if optional one-step processing is not enabled), peer delay, and Announce messages and also to inform the neighbor not to compute neighbor rate ratio and/or propagation delay on this link. A time-aware system could do this when it enters low-power mode, but this standard does not specify the conditions under which this is done; it specifies only the actions a time-aware system takes.

7.4 PTP Instance architecture

The model of a PTP Instance is shown in Figure 7-8.

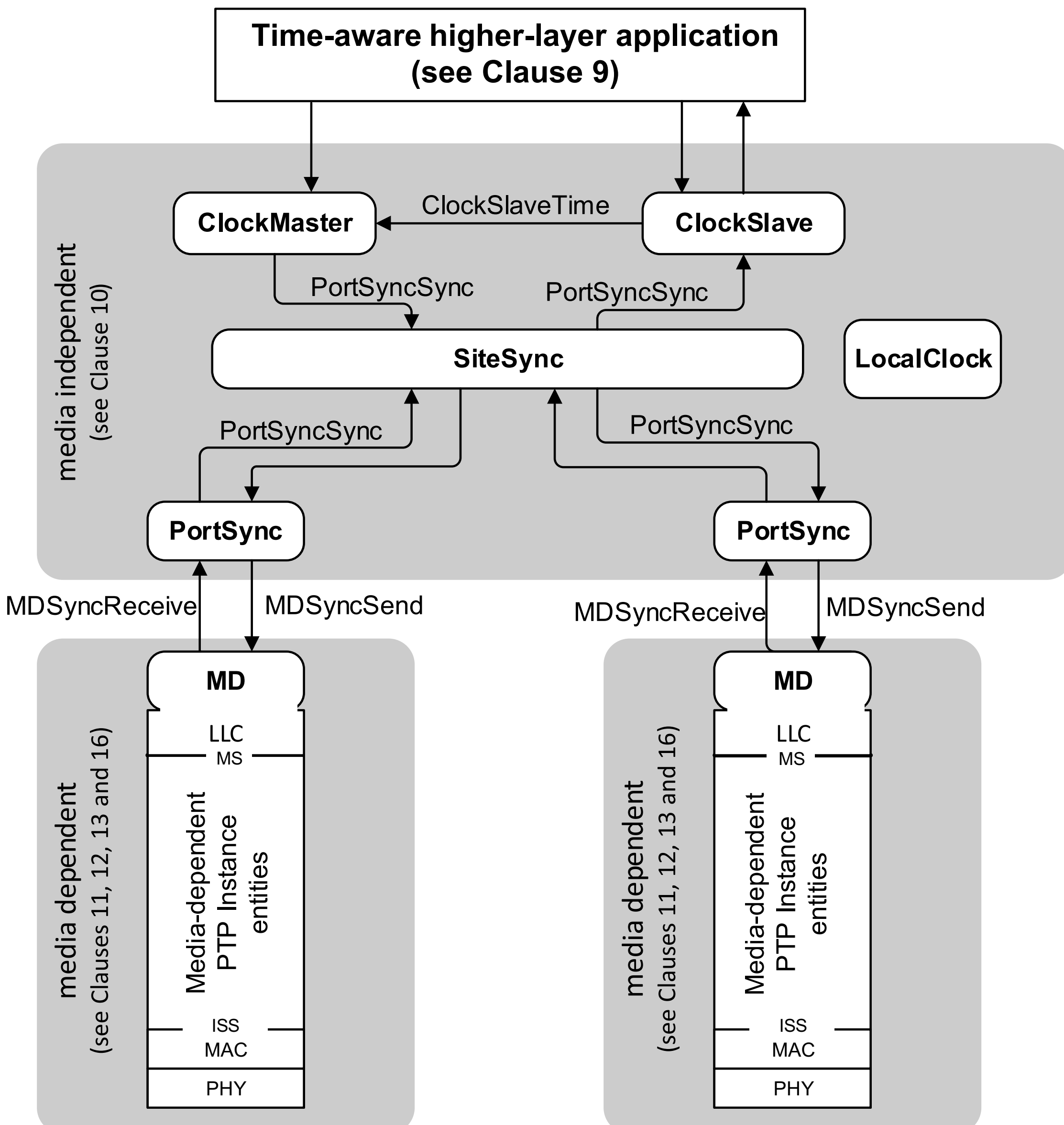


Figure 7-8—PTP Instance model

A PTP Instance consists of the following major parts:

a) If the PTP Instance includes application(s) that either use or source time information, then they interface with the gPTP information using the application interfaces specified in Clause 9.

b) A single media-independent part that consists of ClockMaster, ClockSlave, and SiteSync logical entities, one or more PortSync entities, and a LocalClock entity. The BMCA and forwarding of time information between logical ports and the ClockSlave and ClockMaster is done by the SiteSync entity, while the computation of PTP Port-specific delays needed for time-synchronization correction is done by the PortSync entities.

c) Media-dependent ports, which translate the abstract “MDSyncSend” and “MDSyncReceive” structures received from or sent to the media-independent layer and corresponding methods used for the particular LAN attached to the port.

For full-duplex Ethernet ports, IEEE 1588 Sync and Follow_Up (or just Sync if the optional one-step processing is enabled) messages are used, with an additional TLV in the Follow_Up (or the Sync if the optional one-step processing is enabled) used for communication of rate ratio and information on phase and frequency change when there is a change in Grandmaster PTP Instance. The path delay is measured using the two-step IEEE 1588 peer-to-peer delay mechanism. This is defined in Clause 11.

For IEEE 802.11 ports, timing information is communicated using the MAC Layer Management Entity to request a “Timing Measurement” or “Fine Timing Measurement” (as defined in IEEE Std 802.11-2016), which also sends everything that would be included in the Follow_up message for full-duplex Ethernet. The Timing Measurement or Fine Timing Measurement result includes all the information to determine the path delay. This is defined in Clause 12.

For EPON, timing information is communicated using a “slow protocol” as defined in Clause 13. CSNs use the same communication system used by full-duplex Ethernet, as defined in Clause 16.

7.5 Differences between gPTP (IEEE Std 802.1AS) and PTP (IEEE Std 1588-2019)

a) gPTP assumes all communication between PTP Instances is done only using IEEE 802 MAC PDUs and addressing, while IEEE Std 1588-2019 supports various layer 2 and layer 3-4 communication methods.

b) gPTP specifies a media-independent sublayer that simplifies the integration within a single timing domain of multiple different networking technologies with radically different media access protocols. gPTP specifies a media-dependent sublayer for each medium. The information exchanged between PTP Instances has been generalized to support different packet formats and management schemes appropriate to the particular networking technology. IEEE Std 1588-2019, on the other hand, has introduced a new architecture based on media-independent and media-dependent sublayers (see 6.5.2, Figure 5, and Figure 6 of IEEE Std 1588-2019); however, this architecture is optional. The architecture of IEEE Std 1588-2008 [B10], which is not based on media-independent and media-dependent layers, has been retained for Internet Protocol (IP) version 4, IP version 6, Ethernet LANs, and several industrial automation control protocols. The intent in IEEE Std 1588-2019 is that the new architecture, based on media-independent and media-dependent layers, will be used for IEEE 802.11 networks, IEEE 802.3 EPON, and CSN using the specifications of gPTP, and that the architecture must be used for transports that define native timing mechanisms if those native timing mechanisms are used.

c) In gPTP there are only two types of PTP Instances: PTP End Instances and PTP Relay Instances, while IEEE Std 1588-2019 has Ordinary Clocks, Boundary Clocks, end-to-end Transparent Clocks, and P2P Transparent Clocks. A PTP End Instance corresponds to an IEEE 1588 Ordinary Clock, and a PTP Relay Instance is a type of IEEE 1588 Boundary Clock where its operation is very tightly defined, so much so that a PTP Relay Instance with Ethernet ports can be shown to be mathematically equivalent to a P2P Transparent Clock in terms of how synchronization is performed, as shown in 11.1.3. In addition, a PTP Relay Instance can operate in a mode (i.e., the mode where the variable syncLocked is TRUE; see 10.2.5.15) where the PTP Relay Instance is equivalent to a P2P Transparent Clock in terms of when time-synchronization messages are sent. A time-aware system measures link delay and residence time and communicates these in a correction field. In summary, a PTP Relay Instance conforms to the specifications for a Boundary Clock in

IEEE Std 1588-2019, but a PTP Relay Instance does not conform to the complete specifications for a P2P Transparent Clock in IEEE Std 1588-2019 because:

1) When syncLocked is FALSE, the PTP Relay Instance sends Sync according to the specifications for a Boundary Clock, and
2) The PTP Relay Instance invokes the BMCA and has PTP Port states.

d) PTP Instances communicate gPTP information only directly with other PTP Instances. That is, a gPTP domain consists ONLY of PTP Instances. Non-PTP Relay Instances cannot be used to relay gPTP information. In IEEE Std 1588-2019, it is possible to use non-IEEE-1588-aware relays in an IEEE 1588 domain, although this slows timing convergence and introduce extra jitter and wander that must be filtered by any IEEE 1588 clock.

e) For full-duplex Ethernet links, gPTP requires the use of the peer-to-peer delay mechanism, while IEEE Std 1588-2019 also allows the use of end-to-end delay measurement.

f) For full-duplex Ethernet links, gPTP requires the use of two-step processing (use of Follow_Up and Pdelay_Resp_Follow_Up messages to communicate timestamps), with an optional one-step processing mode that embeds timestamps in the Sync "on the fly" as they are being transmitted (gPTP does not specify one-step processing for peer delay messages). IEEE Std 1588-2019 allows either two-step or one-step processing to be required (for both Sync and peer delay messages) depending on a specific profile.

g) All PTP Instances in a gPTP domain are logically syntonized; in other words, they all measure time intervals using the same frequency. This is done by the process described in 7.3.3 and is mandatory. Syntonization in IEEE Std 1588-2019 is optional. The syntonization method used by gPTP is supported as an option in IEEE Std 1588-2019, but uses a TLV standardized as part of IEEE Std 1588-2019 (this feature is new for IEEE Std 1588-2019), while gPTP uses the ORGANIZATION_EXTENSION TLV specified in 11.4.4.3.

h) Finally, this standard includes formal interface definitions, including primitives, for the time-aware applications (see Clause 9). IEEE Std 1588-2019 describes external interfaces without describing specific interface primitives.