



Precision Time Protocol

- [About PTP, on page 1](#)
- [Cisco ACI and PTP, on page 24](#)

About PTP

The Precision Time Protocol (PTP) is a time synchronization protocol defined in IEEE 1588 for nodes distributed across a network. With PTP, you can synchronize distributed clocks with an accuracy of less than 1 microsecond using Ethernet networks. PTP's accuracy comes from the hardware support for PTP in the Cisco Application Centric Infrastructure (ACI) fabric spine and leaf switches. The hardware support allows the protocol to compensate accurately for message delays and variation across the network.



Note This document uses the term "client" for what the IEEE1588-2008 standard refers to as the "slave." The exception is instances in which the word "slave" is embedded in the Cisco Application Policy Infrastructure Controller (APIC) CLI commands or GUI.

PTP is a distributed protocol that specifies how real-time PTP clocks in the system synchronize with each other. These clocks are organized into a master-client synchronization hierarchy with the grandmaster clock, which is the clock at the top of the hierarchy, determining the reference time for the entire system. Synchronization is achieved by exchanging PTP timing messages, with the members using the timing information to adjust their clocks to the time of their master in the hierarchy. PTP operates within a logical scope called a PTP domain.

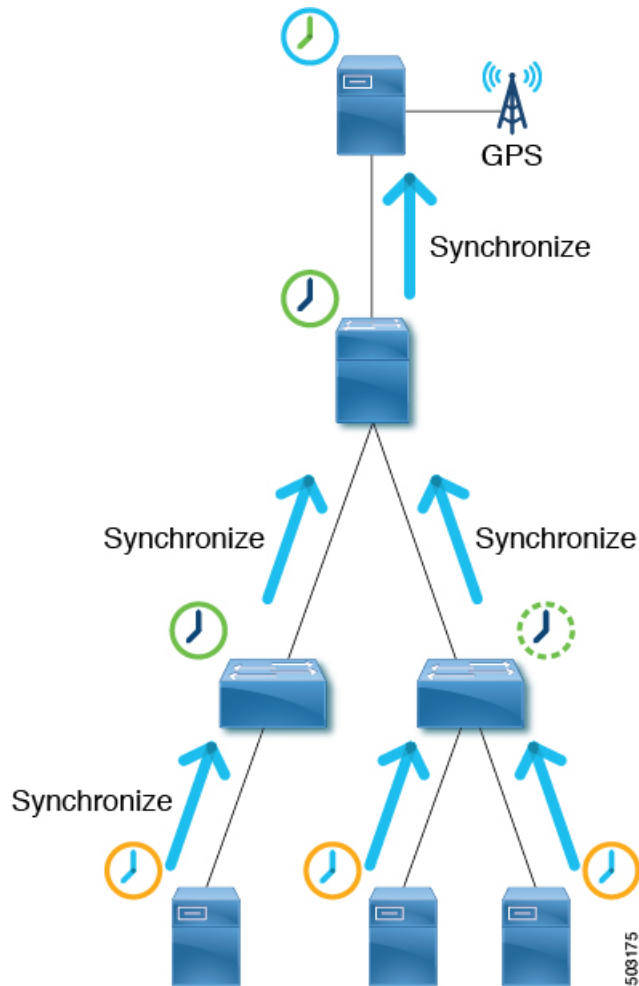
The PTP process consists of two phases: establishing the master-client hierarchy and synchronizing the clocks. Within a PTP domain, each port of an ordinary or boundary clock uses the following process to determine its state:

1. Establish the master-client hierarchy using the Best Master Clock Algorithm (BMCA):
 - Examine the contents of all received `Announce` messages (issued by ports in the master state).
 - Compare the data sets of the foreign master (in the `Announce` message) and the local clock for priority, clock class, accuracy, and so on.
 - Determine its own state as either master or client.
2. Synchronize the clocks:

- Use messages, such as `Sync` and `Delay_Req`, to synchronize the clock between the master and clients.

PTP Clock Types

The following illustration shows the hierarchy of the PTP clock types:



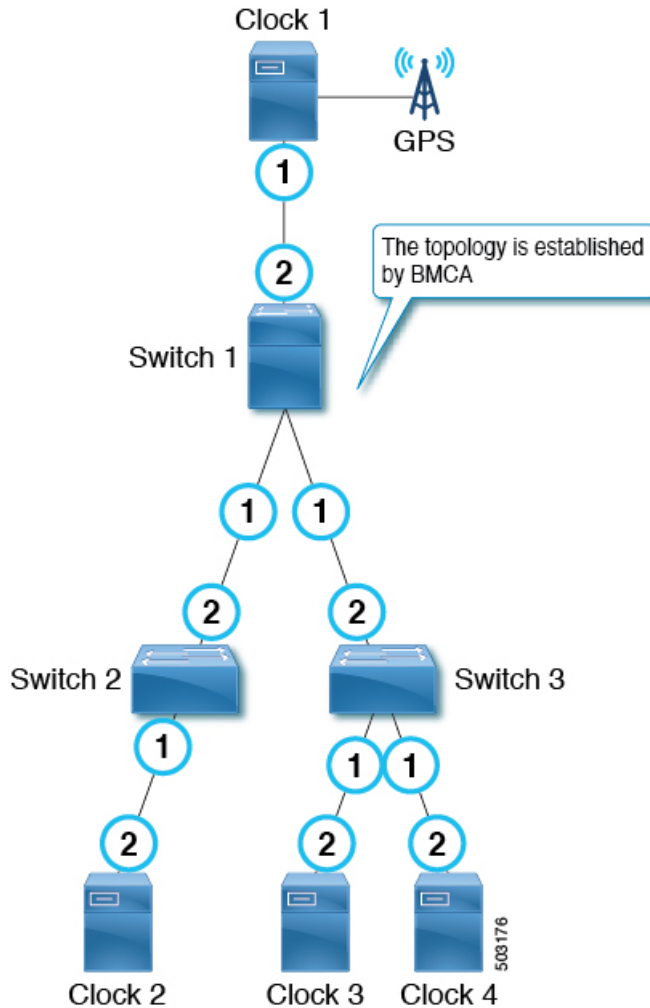
PTP has the following clock types:

| Type | Description |
|-----------------------------|--|
| Grandmaster Clock (GM, GMC) | The source of time for the entire PTP topology. The grandmaster clock is selected by the Best Master Clock Algorithm (BMCA). |

| Type | Description |
|------------------------|--|
| Boundary Clock (BC) | <p>A device with multiple PTP ports. A PTP boundary clock participates in the BMCA and each port has a status, such as master or client. A boundary clock synchronizes with its parent/master so that the client clocks behind itself synchronize to the PTP boundary clock itself. To ensure that, a boundary clock terminates PTP messages and replies by itself instead of forwarding the messages. This eliminates the delay caused by the node forwarding PTP messages from one port to another.</p> |
| Transparent Clock (TC) | <p>A device with multiple PTP ports. A PTP transparent clock does not participate in the BMCA. This clock type only transparently forwards PTP messages between the master clock and client clocks so that they can synchronize directly with one another. A transparent clock appends the residence time to the PTP messages passing by so that the clients can take the forwarding delay within the transparent clock device into account.</p> <p>In the case of a peer-to-peer delay mechanism, a PTP transparent clock terminates PTP Pdelay_XXX messages instead of forwarding the messages.</p> <p>Note Switches in the ACI mode cannot be a transparent clock.</p> |
| Ordinary Clock (OC) | <p>A device that may serve a source of time as a grandmaster clock or that may synchronize to another clock (such as a master) with the role as a client (a PTP client).</p> |

PTP Topology

Master and Client Ports



The master and client ports work as follows:

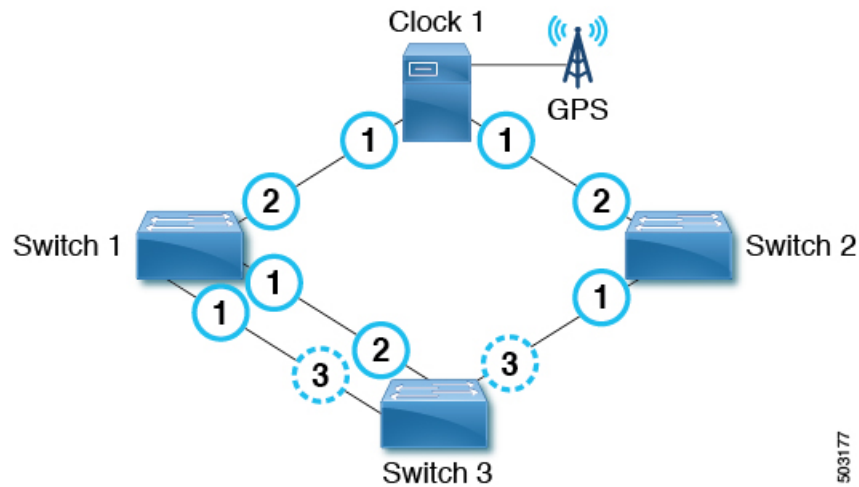
- Each PTP node directly or indirectly synchronizes its clock to the grandmaster clock that has the best source of time, such as GPS (Clock 1 in the figure).
- One grandmaster is selected for the entire PTP topology (domain) based on the Best Master Clock Algorithm (BMCA). The BMCA is calculated on each PTP node individually, but the algorithm makes sure that all nodes in the same domain select the same clock as the grandmaster.
- In each path between PTP nodes, based on the BMCA, there will be one master port and at least one client port. There will be multiple client ports if the path is point-to-multipoints, but each PTP node can have only one client port. Each PTP node uses its client port to synchronize to the master port on the other end. By repeating this, all PTP nodes eventually synchronize to the grandmaster directly or indirectly.
 - From Switch 1's point of view, Clock 1 is the master and the grandmaster.

- From Switch 2's point of view, Switch 1 is the master and Clock 1 is the grandmaster.
- Each PTP node should have only one client port, behind which exists the grandmaster. The grandmaster can be multiple hops away.
- The exception is a PTP transparent clock, which does not participate in BMCA. If Switch 3 was a PTP transparent clock, the clock would not have a port status, such as master and client. Clock 3, Clock 4, and Switch 1 would establish a master and client relationship directly.

Passive Ports

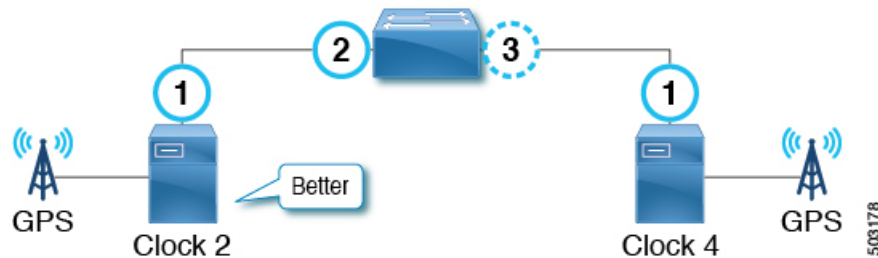
The BMCA can select another PTP port that is in the passive state on top of the master and client. A passive port does not generate any PTP messages, with a few exceptions such as PTP Management messages as a response to Management messages from other nodes.

Example 1



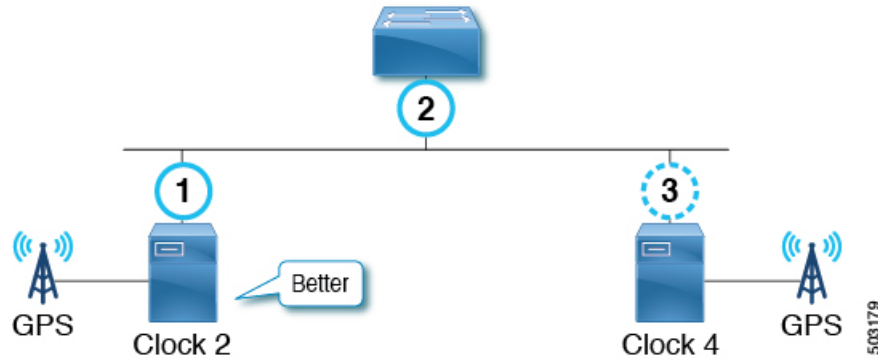
If a PTP node has multiple ports towards the grandmaster, only one of them will be the client port. The other ports toward the grandmaster will be passive ports.

Example 2



If a PTP node detects two master only clocks (grandmaster candidates), the port toward the candidate selected as the grandmaster becomes a client port and the other becomes a passive port. If the other clock can be a client, it forms a master and client relation instead of passive.

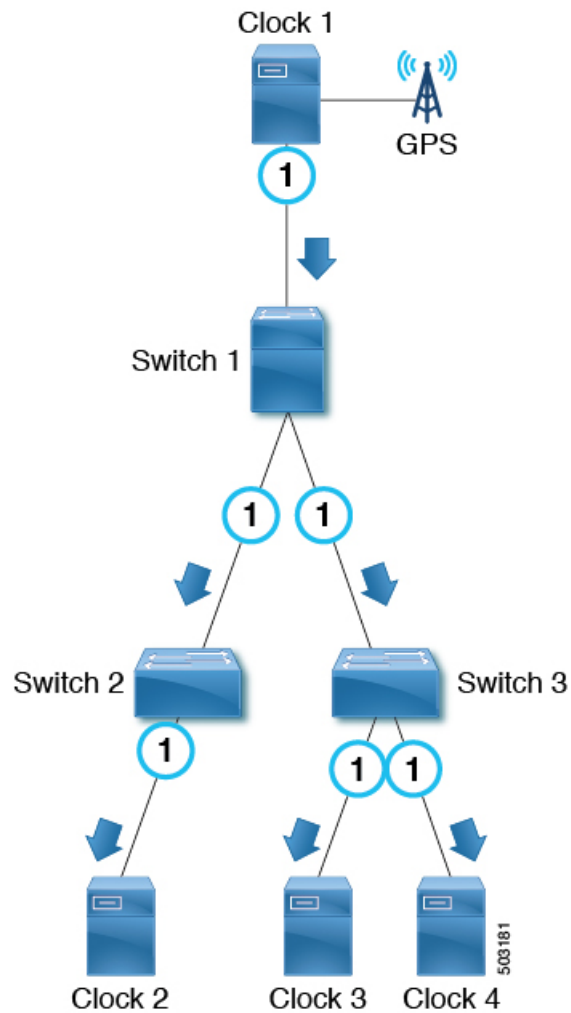
Example 3



If a master-only clock (grandmaster candidate) detects another master-only clock that is better than itself, the clock puts itself in a passive state. This happens when two grandmaster candidates are on the same communication path without a PTP boundary clock in between.

Announce Messages

The `Announce` message is used to calculate the Best Master Clock Algorithm (BMCA) and establish the PTP topology (master-client hierarchy).



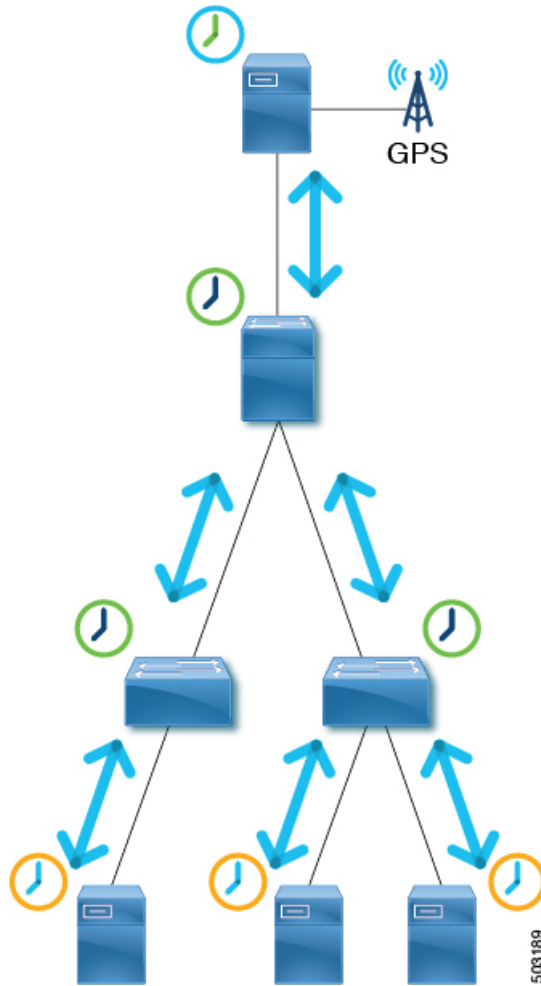
The message works as follows:

- PTP master ports send PTP `Announce` messages to IP address 224.0.1.129 in the case of PTP over IPv4 UDP.
- Each node uses information in the PTP `Announce` messages to automatically establish the synchronization hierarchy (master/client relations or passive) based on the BMCA.
- Some of the information that PTP `Announce` messages contain is as follows:
 - Grandmaster priority 1
 - Grandmaster clock quality (class, accuracy, variance)
 - Grandmaster priority 2
 - Grandmaster identity
 - Step removed
- PTP `Announce` messages are sent with an interval based on $2^{\log \text{AnnounceInterval}}$ seconds.

PTP Topology With Various PTP Node Types

PTP Topology With Only End-to-End Boundary Clocks

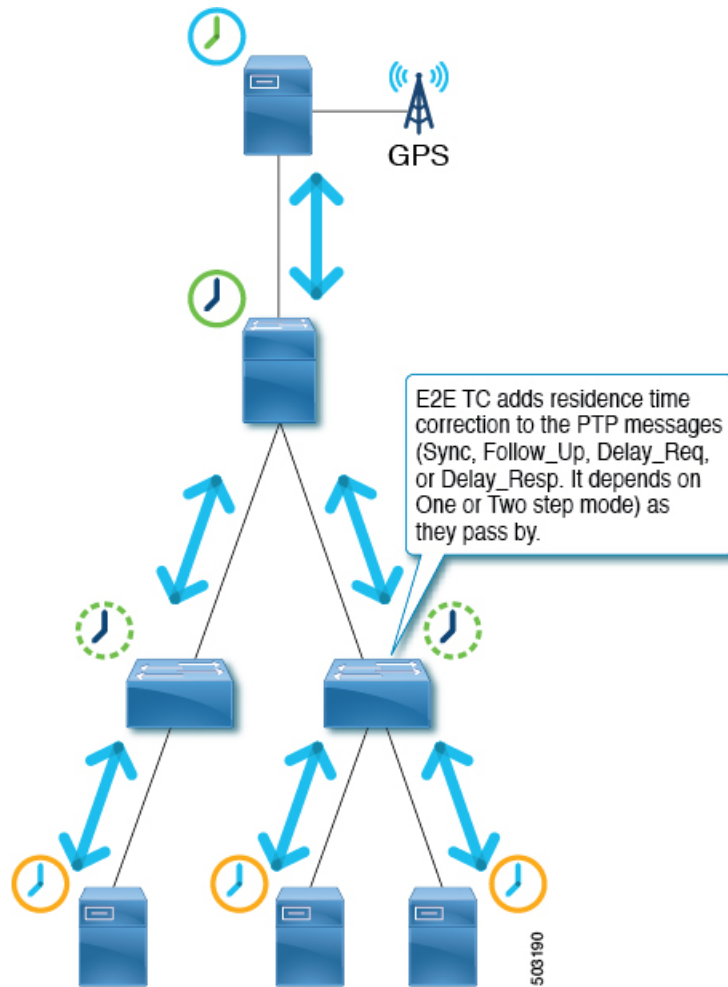
In this topology, the boundary clock nodes terminate all multicast PTP messages, except for `Management` messages.



This ensures that each node processes the `sync` messages from the closest parent master clock, which helps the nodes to achieve high accuracy.

PTP Topology With a Boundary Clock and End-to-End Transparent Clocks

In this topology, the boundary clock nodes terminate all multicast PTP messages, except for `Management` messages.



End-to-end (E2E) transparent clock nodes do not terminate PTP messages, but simply add a residence time (the time the packet took to go through the node) in the PTP message correction field as the packets pass by so that clients can use them to achieve better accuracy. But, this has lower scalability as the number of PTP messages that need to be handled by one boundary clock node increases.

PTP BMCA

PTP BMCA Parameters

Each clock has the following parameters defined in IEEE 1588-2008 that are used in the Best Master Clock Algorithm (BMCA):

| Parameter | Possible Values | Description |
|------------|-----------------|---|
| Priority 1 | 0 to 255 | A user configurable number. The value is normally 128 or lower for grandmaster-candidate clocks (master-capable devices) and 255 for client-only devices. |

| Parameter | Possible Values | Description |
|--------------------------|--|--|
| Clock Quality - Class | 0 to 255 | Represents the status of the clock devices. For example, 6 is for devices with a primary reference time source, such as GPS. 7 is for devices that used to have a primary reference time source. 127 or lower are for master-only clocks (grandmaster candidates). 255 is for client-only devices. |
| Clock Quality - Accuracy | 0 to 255 | The accuracy of the clock. For example, 33 (0x21) means < 100 ns, while 35 (0x23) means < 1 us. |
| Clock Quality - Variance | 0 to 65535 | The precision of the timestamps encapsulated in the PTP messages. |
| Priority 2 | 0 to 255 | Another user-configurable number. This parameter is typically used when the setup has two grandmaster candidates with identical clock quality and one is a standby. |
| Clock Identity | This is an 8-byte value that is typically formed using a MAC address | This serves as the final tie breaker, and is typically a MAC address. |

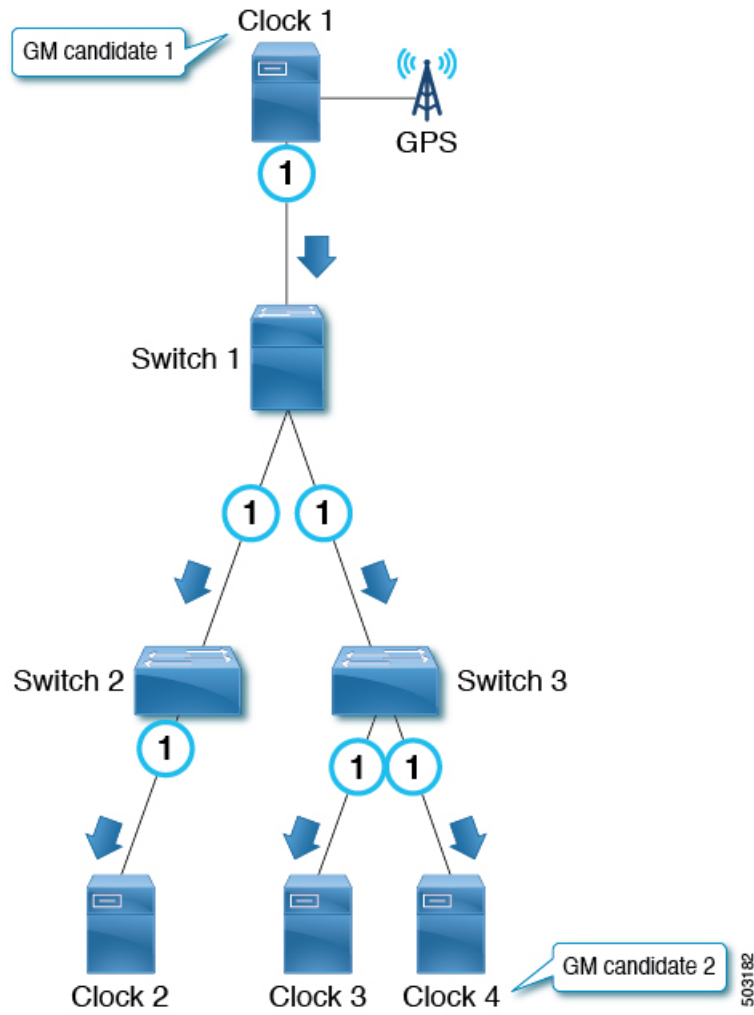
These parameters of the grandmaster clock are carried by the PTP `Announce` messages. Each PTP node compares these values in the order as listed in the table from all `Announce` messages that the node receives and also the node's own values. For all parameters, the lower number wins. Each PTP node will then create `Announce` messages using the parameters of the best clock among the ones the node is aware of, and the node will send the messages from its own master ports to the next client devices.



Note For more information about each parameter, see clause 7.6 in IEEE 1588-2008.

PTP BMCA Examples

In the following example, Clock 1 and Clock 4 are the grandmaster candidates for this PTP domain:



Clock 1 has the following parameter values:

| Parameter | Value |
|--------------------------|----------------|
| Priority 1 | 127 |
| Clock Quality (Class) | 6 |
| Clock Quality (Accuracy) | 0x21 (< 100ns) |
| Clock Quality (Variance) | 15652 |
| Priority 2 | 128 |
| Clock Identity | 0000.1111.1111 |
| Step Removed | * |

Clock 4 has the following parameter values:

| Parameter | Value |
|--------------------------|----------------|
| Priority 1 | 127 |
| Clock Quality (Class) | 6 |
| Clock Quality (Accuracy) | 0x21 (< 100ns) |
| Clock Quality (Variance) | 15652 |
| Priority 2 | 129 |
| Clock Identity | 0000.1111.2222 |
| Step Removed | * |

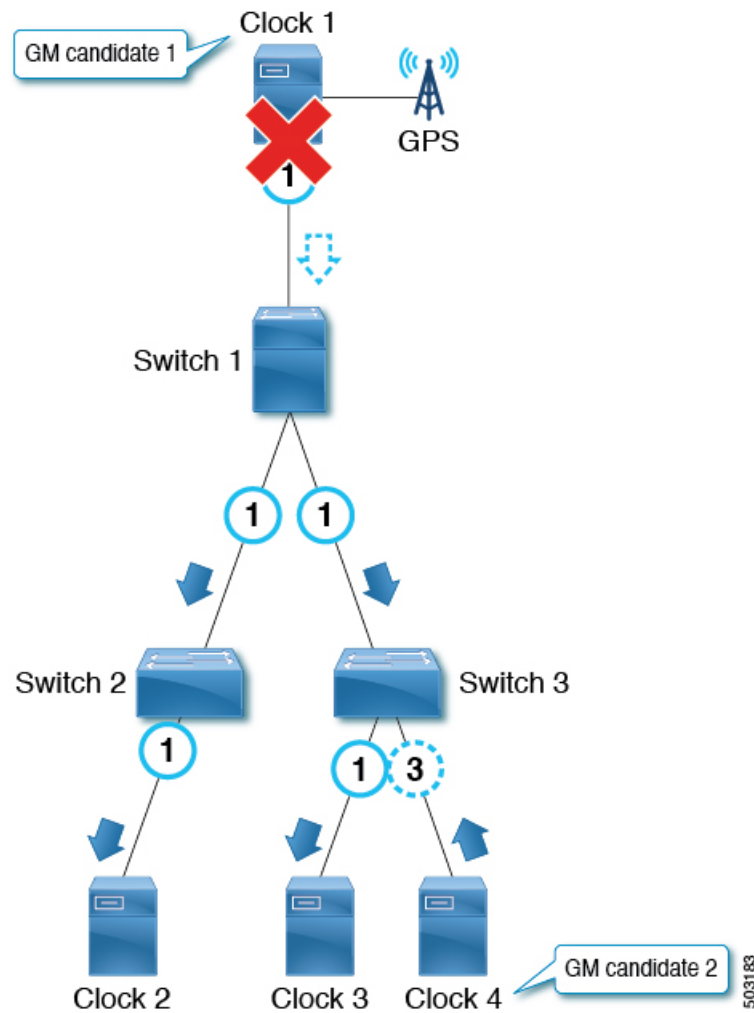
Both clocks send PTP *Announce* messages, then each PTP node compares the values in the messages. In this example, because the first four parameters have the same value, *Priority 2* decides the active grandmaster, which is Clock 1.

After all switches (1, 2, and 3) have recognized that Clock 1 is the best master clock (that is, Clock 1 is the grandmaster), those switches send PTP *Announce* messages with the parameters of Clock 1 from their master ports. On Switch 3, the port connected to Clock 4 (a grandmaster candidate) becomes a passive port because the port is receiving PTP *Announce* messages from a master-only clock (class 6) with parameters that are not better than the current grandmaster that is received by another port.

The *Step Removed* parameter indicates the number of hops (PTP boundary clock nodes) from the grandmaster. When a PTP boundary clock node sends PTP *Announce* messages, it increments the *Step Removed* value by 1 in the message. In this example, Switch 2 receives the PTP *Announce* message from Switch 1 with the parameters of Clock 1 and a *Step Removed* value of 1. Clock 2 receives the PTP *Announce* message with a *Step Removed* value of 2. This value is used only when all the other parameters in the PTP *Announce* messages are the same, which happens when the messages are from the same grandmaster candidate clock.

PTP BMCA Failover

If the current active grandmaster (Clock 1) becomes unavailable, each PTP port recalculates the Best Master Clock Algorithm (BMCA).



The availability is checked using the `Announce` messages. Each PTP port declares the timeout of the `Announce` messages after the `Announce` messages were consecutively missing for `Announce Receipt Timeout` times. In other words, for $\text{Announce Receipt Timeout} \times 2^{\log \text{AnnounceInterval}}$ seconds. This timeout period should be uniform throughout a PTP domain as mentioned in Clause 7.7.3 in IEEE 1588-2008. When the timeout is detected, each switch starts recalculating the BMCA on all PTP ports by sending `Announce` messages with the new best master clock data. The recalculation can result in a switch initially determining that the switch itself is the best master clock, because most of the switches are aware of only the previous grandmaster.

When the client port connected toward the grandmaster goes down, the node (or the ports) does not need to wait for the announce timeout and can immediately start re-calculating the BMCA by sending `Announce` messages with the new best master clock data.

The convergence can take several seconds or more depending on the size of the topology, because each PTP port recalculates the BMCA from the beginning individually to find the new best clock. Prior to the failure of the active grandmaster, only Switch 3 knows about Clock 4, which should take over the active grandmaster role.

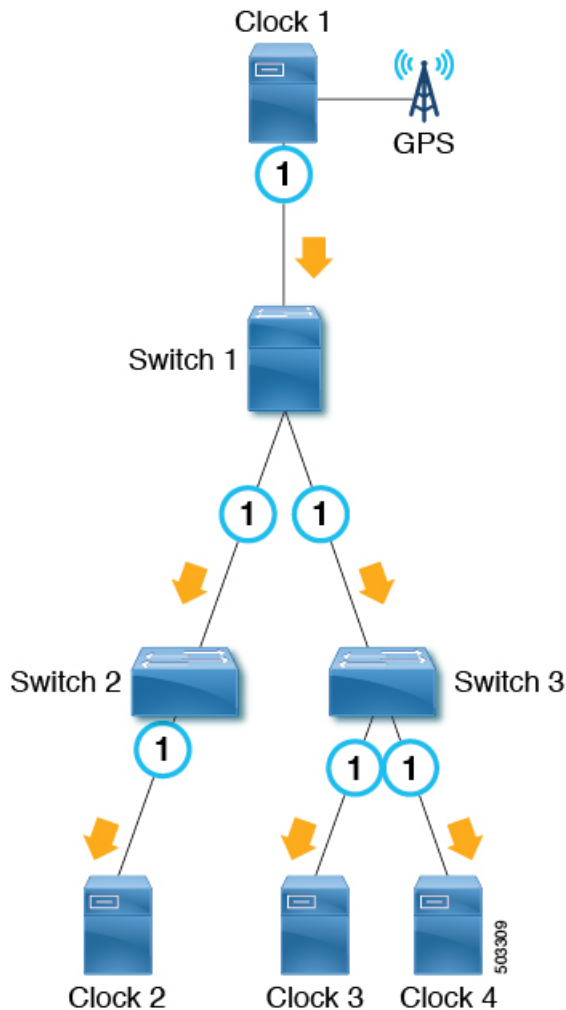
Also, when the port status changes to master from non-master, the port changes to the `PRE_MASTER` status first. The port takes `Qualification Timeout` seconds for the port to become the actual master, which is typically equal to:

(Step Removed + 1) x the announce interval

This means that if the other grandmaster candidate is connected to the same switch as (or close to) the active grandmaster, the port status changes will be minimum and the convergence time will be shorter. See Clause 9.2 in IEEE 1588-2008 for details.

PTP Clock Synchronization

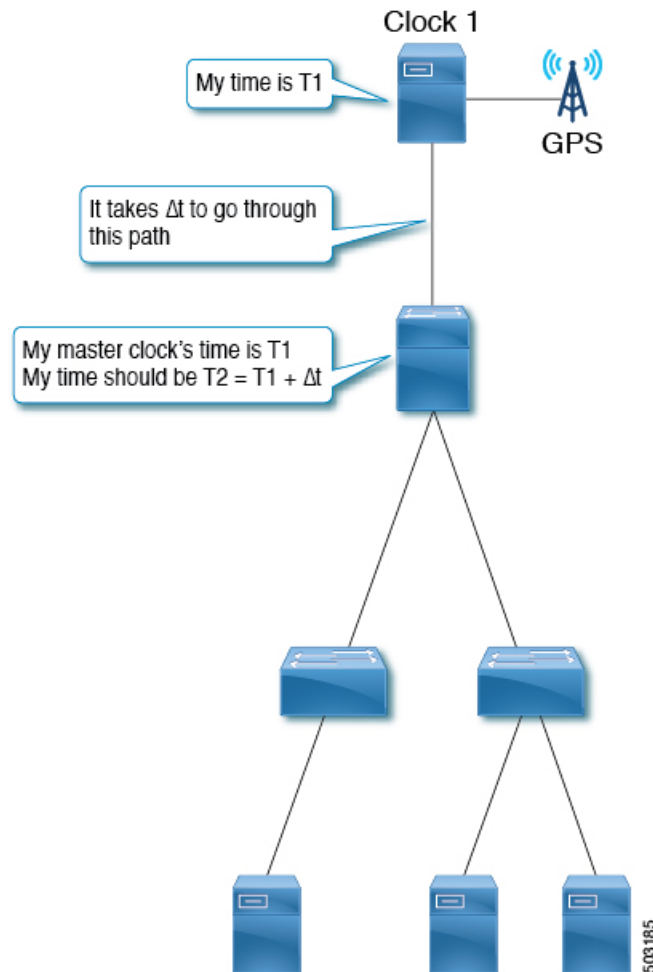
The PTP master ports send PTP `Sync` and `Follow_Up` messages to IP address 224.0.1.129 in the case of PTP over IPv4 UDP.



In One-Step mode, the `Sync` messages carry the timestamp of when the message was sent out. `Follow_Up` messages are not required. In Two-Step mode, `Sync` messages are sent out without a timestamp. `Follow_Up` messages are sent out immediately after each `Sync` message with the timestamp of when the `Sync` message was sent out. Client nodes use the timestamp in the `Sync` or `Follow_Up` messages to synchronize their clock along with an offset calculated by `meanPathDelay`. `Sync` messages are sent with the interval based on $2^{\log_{\text{SyncInterval}}}$ seconds.

PTP and meanPathDelay

`meanPathDelay` is the mean time that PTP packets take to reach from one end of the PTP path to the other end. In the case of the E2E delay mechanism, this is the time taken to travel between a PTP master port and a client port. PTP needs to calculate `meanPathDelay` (Δt in the following illustration) to keep the synchronized time on each of the distributed devices accurate.



There are two mechanisms to calculate `meanPathDelay`:

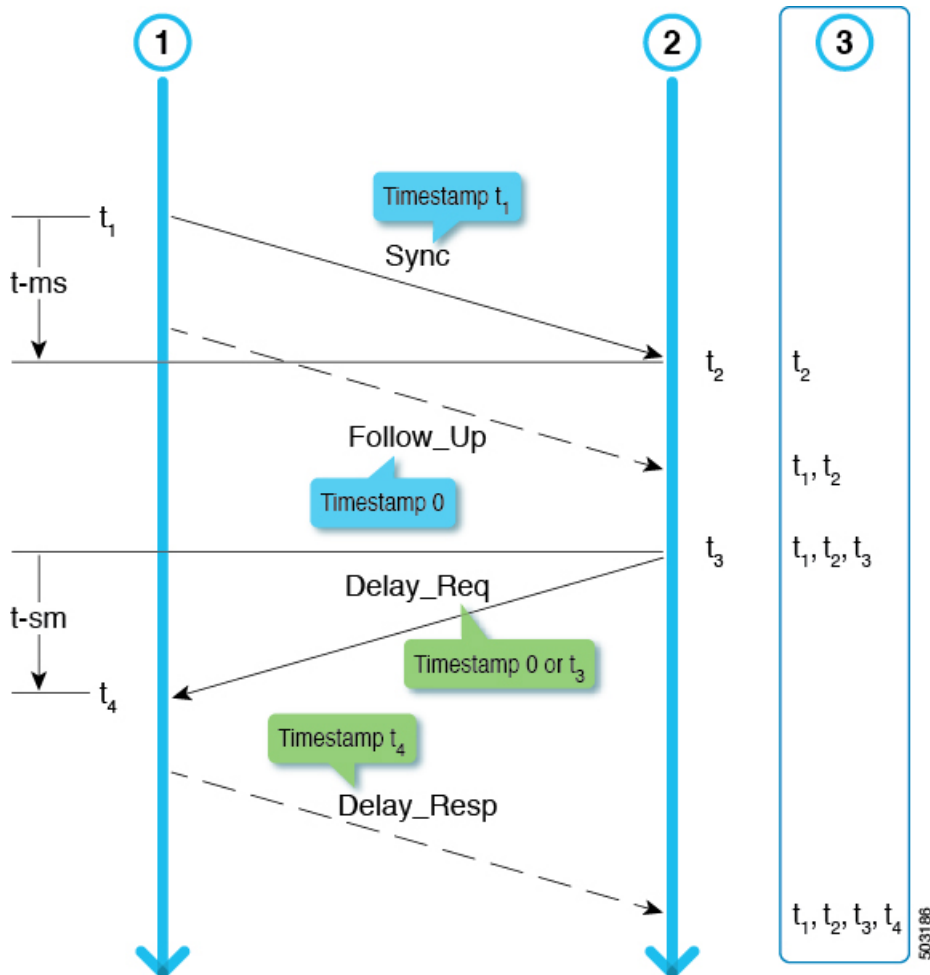
- Delay Request-Response (E2E): End-to-end transparent clock nodes can only support this.
- Peer Delay Request-Response (P2P): Peer-to-peer transparent clock nodes can only support this.

Boundary clock nodes can support both mechanisms by definition. In IEEE 1588-2008, the delay mechanisms are called "Delay" or "Peer Delay." However, the Delay Request-Response mechanism is more commonly referred to as the "E2E delay mechanism," and the Peer Delay mechanism is more commonly referred to as the "P2P delay mechanism."

meanPathDelay Measurement

Delay Request-Response

The delay request-response (E2E) mechanism is initiated by a client port and the `meanPathDelay` is measured on the client node side. The mechanism uses `Sync` and `Follow_Up` messages, which are sent from a master port regardless of the E2E delay mechanism. The `meanPathDelay` value is calculated based on 4 timestamps from 4 messages.



$t - ms$ ($t_2 - t_1$) is the delay for master to client direction. $t - sm$ ($t_4 - t_3$) is the delay for client to master direction. `meanPathDelay` is calculated as follows:

$$(t - ms + t - sm) / 2$$

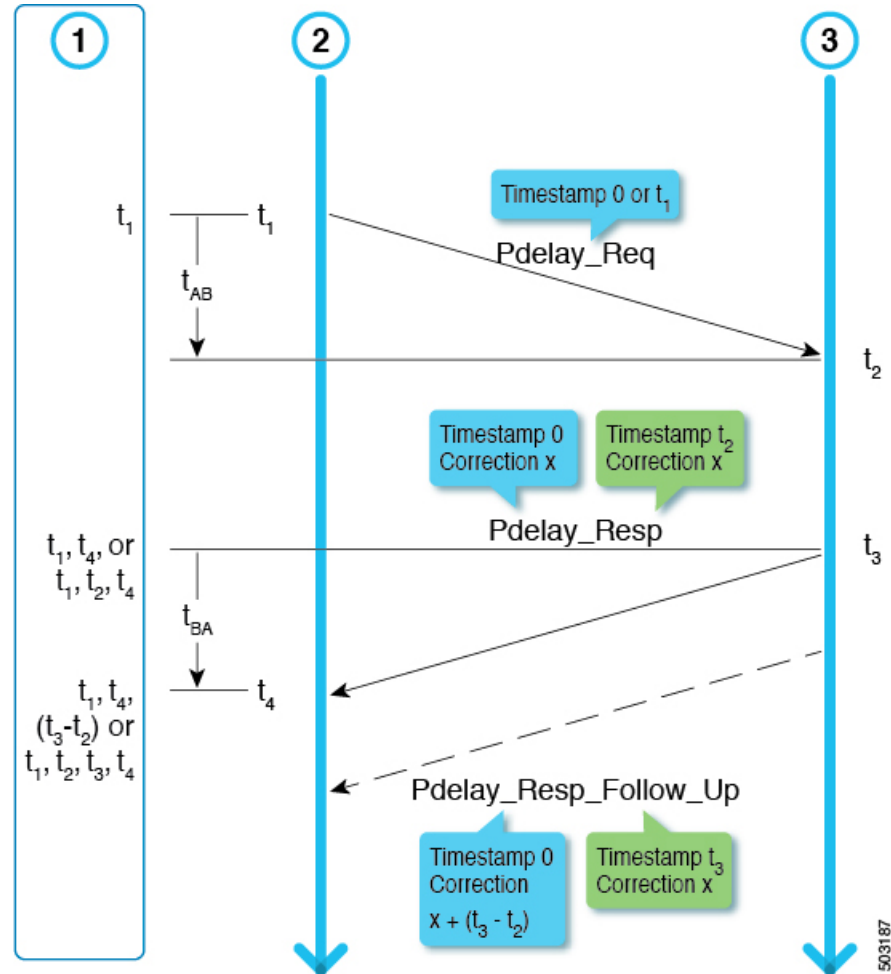
`Sync` is sent with the interval based on $2^{\log_{\text{SyncInterval}}}$ sec. `Delay_Req` is sent with the interval based on $2^{\log_{\text{MinDelayReqInterval}}}$ sec.



Note This example focuses on Two-Step mode. See Clause 9.5 from IEEE 1588-2008 for details on the transmission timing.

Peer Delay Request-Response

The peer delay request-response (P2P) mechanism is initiated by both master and client port and the `meanPathDelay` is measured on the requester node side. `meanPathDelay` is calculated based on 4 timestamps from 3 messages dedicated for this delay mechanism.



In the two-step mode, t_2 and t_3 are delivered to the requester in one of the following ways:

- As $(t_3 - t_2)$ using `Pdelay_Resp_Follow_Up`
- As t_2 using `Pdelay_Resp` and as t_3 using `Pdelay_Resp_Follow_Up`

`meanPathDelay` is calculated as follows:

$$(t_4 - t_1) - (t_3 - t_1) / 2$$

`Pdelay_Req` is sent with the interval based on $2^{\log \text{MinPDelayReqInterval}}$ seconds.



Note Cisco Application Centric Infrastructure (ACI) switches do not support the peer delay request-response (P2P) mechanism.

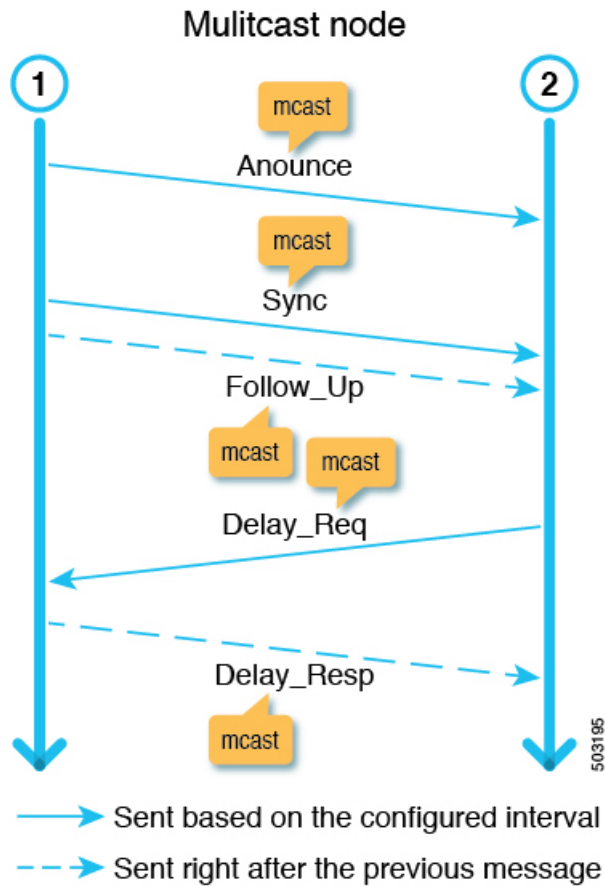
See clause 9.5 from IEEE 1588-2008 for details on the transmission timing.

PTP Multicast, Unicast, and Mixed Mode

The following sections describe the different PTP modes using the delay request-response (E2E delay) mechanism.

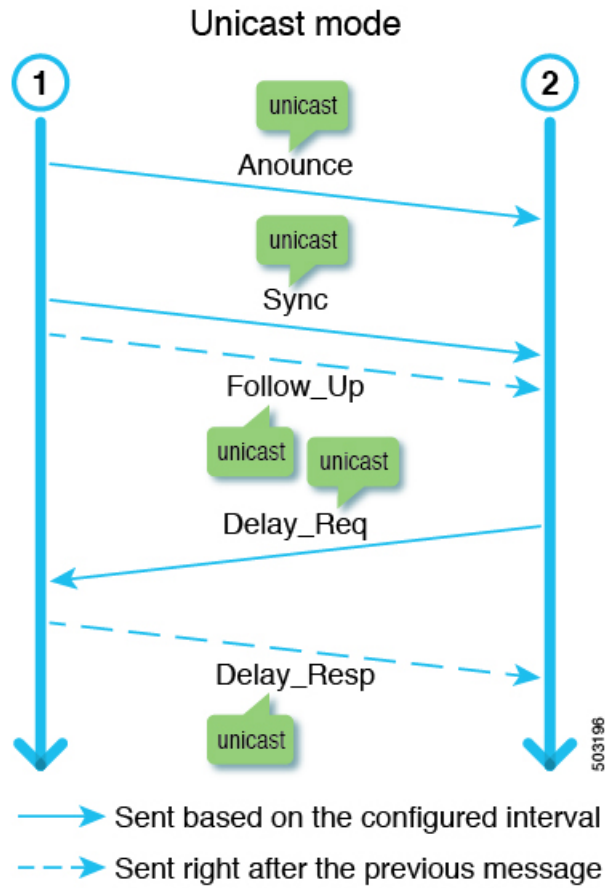
Multicast Mode

All PTP messages are multicast. Transparent clock or PTP unaware nodes between the master and clients result in inefficient flooding of the `Delay` messages. However, the flooding is efficient for `Announce`, `Sync`, and `Follow_Up` messages because these messages should be sent toward all client nodes.



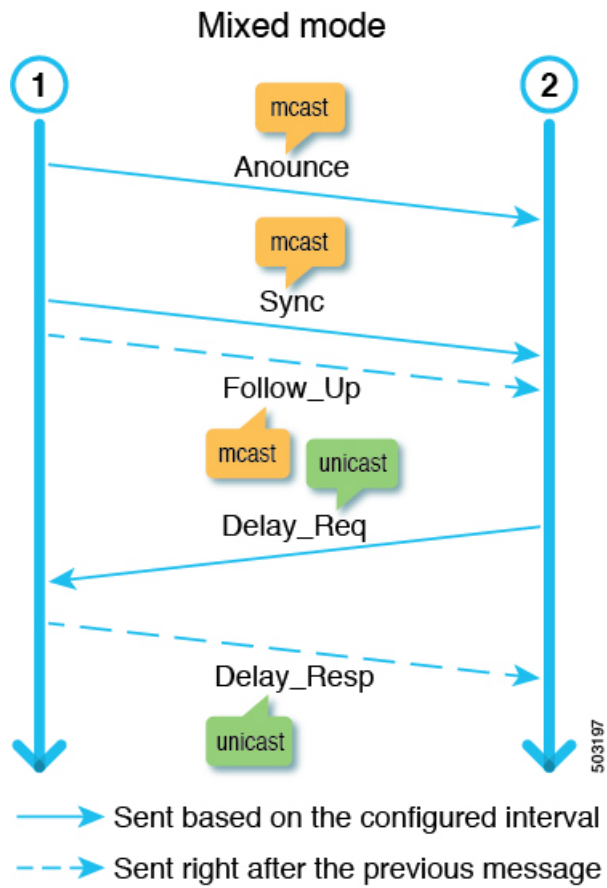
Unicast Mode

All PTP messages are unicast, which increases the number of messages that the master must generate. Hence, the scale, such as the number of client nodes behind one master port, is impacted.



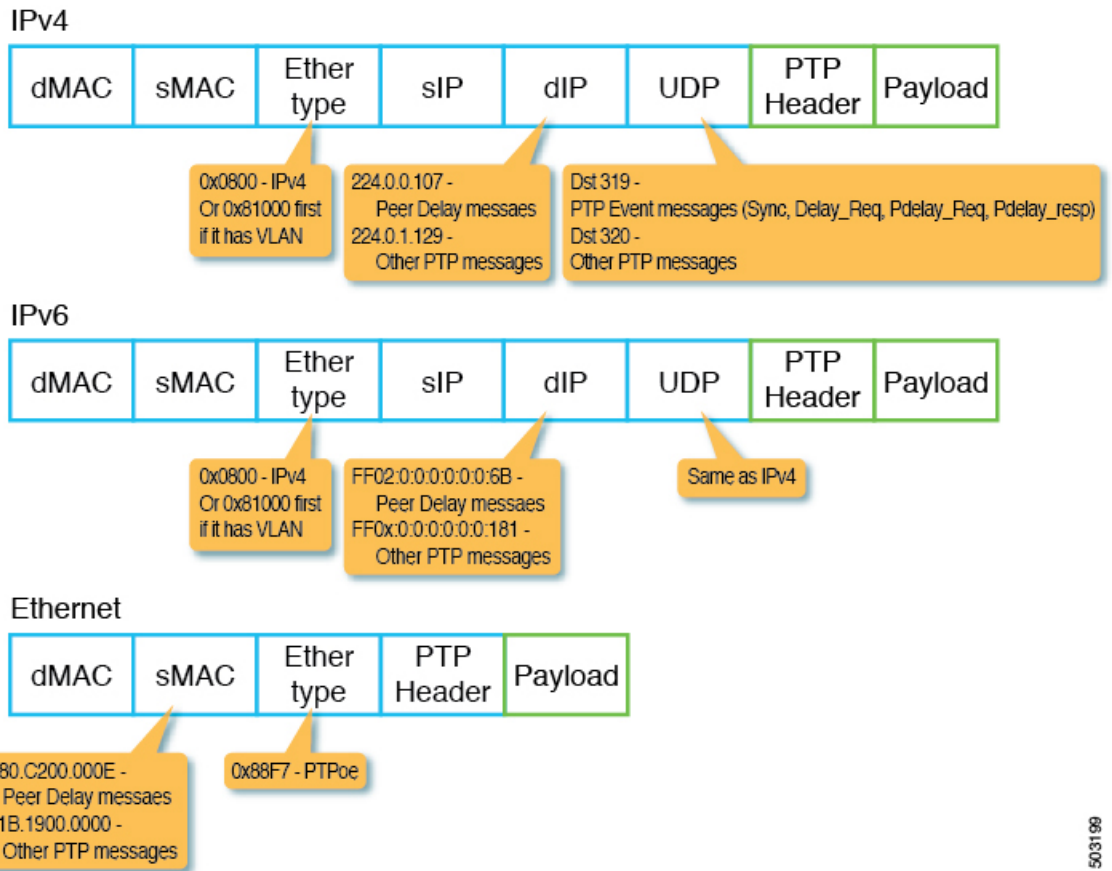
Mixed Mode

Only `Delay` messages are unicast, which resolves the problems that exist in multicast mode and unicast mode.



PTP Transport Protocol

The following illustration provides information about the major transport protocols that PTP supports:



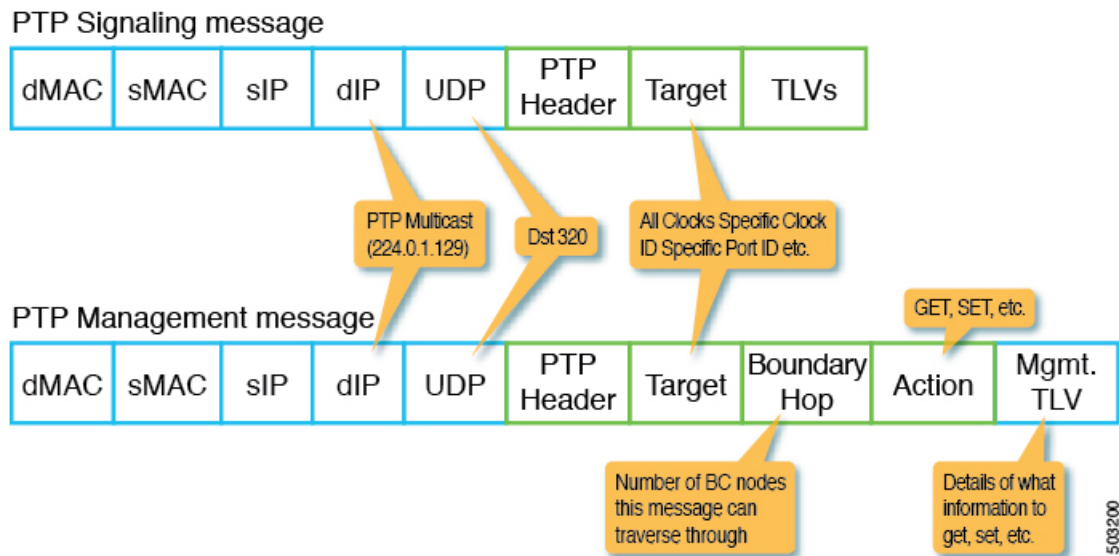
503199



Note Cisco Application Centric Infrastructure (ACI) switches only support IPv4 as a PTP transport protocol.

PTP Signaling and Management Messages

The following illustration shows the Signaling and Management message parameters in the header packet for PTP over IPv4 UDP:



A *Management* message is used to configure or collect PTP parameters, such as the current clock and offset from its master. With the message, a single PTP management node can manage and monitor PTP-related parameters without relying on an out-of-band monitoring system.

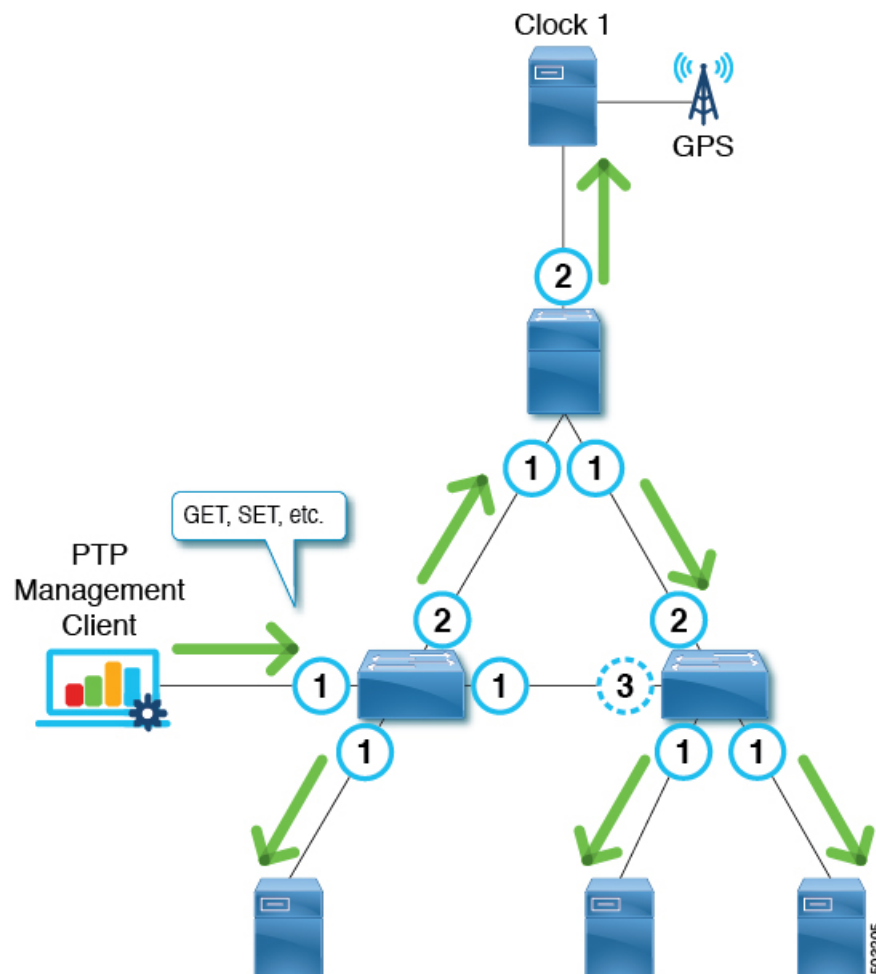
A *Signaling* message also provides various types of type, length, and value (TLVs) to do additional operations. There are other TLVs that are used by being appended to other messages. For example, the `PATH_TRACE` TLV as defined in clause 16.2 of IEEE 1588-2008 is appended to *Announce* messages to trace the path of each boundary clock node in the PTP topology.



Note Cisco Application Centric Infrastructure (ACI) switches do not support management, signal, or other optional TLVs.

PTP Management Messages

PTP Management messages are used to transport management types, lengths, and values (TLVs) toward multiple PTP nodes at once or to a specific node.



The targets are specified with the targetPortIdentity (clockID and portNumber) parameter. PTP Management messages have the actionField that specify actions such as GET, SET, and COMMAND to inform the targets of what to do with the delivered management TLV.

PTP Management messages are forwarded by the PTP boundary clock, and only to the Master, Client, Uncalibrated, or Pre_Master ports. A message is forwarded to those ports only when the message is received on a port in the Master, Client, Uncalibrated, or Pre_Master port. BoundaryHops in the message is decremented by 1 when the message is forwarded.

The SMTPE ST2059-2 profile defines that the grandmaster should send PTP Management messages using the action COMMAND with the synchronization metadata TLV that is required for the synchronization of audio/video signals.



Note Cisco Application Centric Infrastructure (ACI) switches do not process Management messages, but forward them to support the SMTPE ST2059-2 PTP profile.

PTP Profiles

PTP has a concept called PTP profile. A PTP profile is used to define various parameters that are optimized for different use cases of PTP. Some of those parameters include, but not limited to, the appropriate range of PTP message intervals and the PTP transport protocols. A PTP profile is defined by many organizations/standards in different industries. For example:

- IEEE 1588-2008: This standard defines a default PTP profile called the `Default Profile`.
- AES67-2015: This standard defines a PTP profile for audio requirements. This profile is also called the `Media Profile`.
- SMPTE ST2059-2: This standard defines a PTP profile for video requirements.

The following table shows some of the parameters defined in each standard for each PTP profile:

| Profiles | logAnnounce Interval | logSync Interval | logMinDelayReq Interval | announceReceipt Timeout | Domain Number | Mode | Transport Protocol |
|----------------------------|-----------------------------------|-------------------------------------|---|--------------------------------|----------------|---------------------|--------------------|
| Default Profile | 0 to 4 (1) [= 1 to 16 sec] | -1 to +1 (0) [= 0.5 to 2 sec] | 0 to 5 (0) [= 1 to 32 sec] | 2 to 10 announce intervals (3) | 0 to 255 (0) | Multicast / Unicast | Any/IPv4 |
| AES67-2015 (Media Profile) | 0 to 4 (1) [= 1 to 16 sec] | -4 to +1 (-3) [= 1/16 to 16 sec] | -3 to +5 (0) [= 1/8 to 32 sec] Or <i>logSyncInterval</i> to <i>logSyncInterval</i> + 5 seconds | 2 to 10 announce intervals (3) | 0 to 255 (0) | Multicast / Unicast | UDP/IPv4 |
| SMTPE ST2059-2-2015 | -3 to +1 (-2) [= 1/8 to 2 sec] | -7 to -1 (-3) [= 1/16 to 16 sec] | <i>logSyncInterval</i> to <i>logSyncInterval</i> + 5 seconds | 2 to 10 announce intervals (3) | 0 to 127 (127) | Multicast / Unicast | UDP/IPv4 |

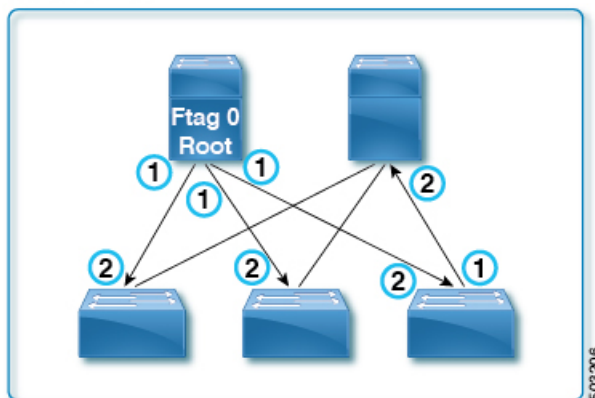
Cisco ACI and PTP

In the Cisco Application Centric Infrastructure (ACI) fabric, when the PTP feature is globally enabled in the Cisco Application Policy Infrastructure Controller (APIC), the software automatically enables PTP on specific interfaces of all the supported spine and leaf switches to establish the PTP master-client topology within the fabric. Starting in Cisco APIC release 4.2(5), you can enable PTP on leaf switch front panel ports and extend PTP topology to outside of the fabric. In the absence of an external grandmaster clock, one of the spine switch is chosen as the grandmaster. The master spine switch is given a different PTP priority that is lower by 1 than the other spines and leaf switches.

Implementation in Cisco APIC Release 3.0(1)

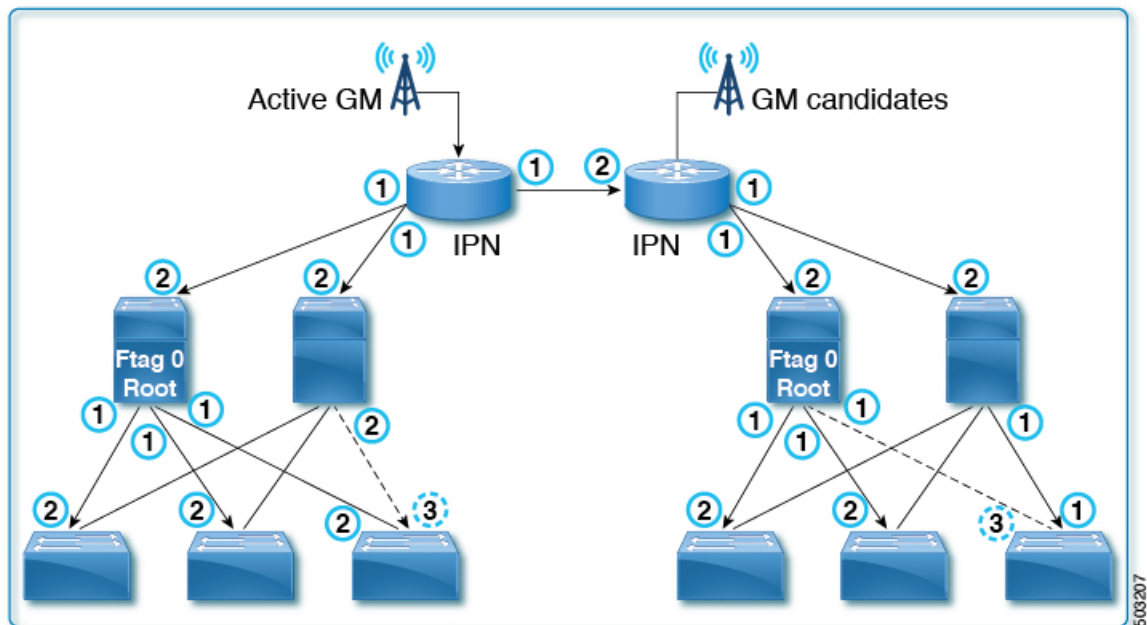
Starting in Cisco Application Policy Infrastructure Controller (APIC) release 3.0(1), PTP was partially introduced to synchronize time only within Cisco Application Centric Infrastructure (ACI) fabric switches. PTP was required to provide the latency measurement feature that was also introduced in Cisco APIC release 3.0(1). For this purpose, a single option was introduced to enable or disable PTP globally. When PTP is enabled globally, all leaf and spine switches are configured as PTP boundary clocks. PTP is automatically enabled on all fabric ports that are used by the ftag tree with ID 0 (ftag0 tree), which is one of the internal tree topologies that is automatically built based on Cisco ACI infra ISIS for loop-free multicast connectivity between all leaf and spine switches in each pod. The root spine switch of the ftag 0 tree is automatically configured with PTP priority1 254 to be the grandmaster when there are no external grandmasters in the inter-pod network (IPN). Other spine and leaf switches are configured with PTP priority1 255.

Single Pod



In a Cisco ACI Multi-Pod setup, when PTP is enabled globally, PTP is automatically enabled on the spine sub-interfaces configured for IPN connectivity in the tn-infra Multi-Pod L3Out. Until Cisco APIC release 4.2(5) or 5.1(1), this was the only way to enable PTP on external-facing interfaces. With this, you can have the Cisco ACI fabric look to an external grandmaster through IPN. When a high accuracy is required, we recommend that you have an external grandmaster with a primary reference time source, such as GPS/GNSS. When enabling PTP in a Cisco ACI Multi-Pod setup without an external grandmaster, one of the spine switches can become a grandmaster for all pods assuming PTP is enabled on IPN and IPN's PTP BMCA parameters, such as PTP priorities, are not better than the spine switch's parameters. When using a spine switch as the grandmaster, adding a new pod may unintentionally result in the new grandmaster being selected from the new pod, which can temporarily cause churn in PTP synchronization throughout the fabric. Regardless of external grandmasters, for a better PTP topology that is fewer hops from the grandmaster, we recommend that you connect all spine switches in the fabric to IPN because users do not have control over how ftag0 tree is formed, which decides the PTP topology inside each pod.

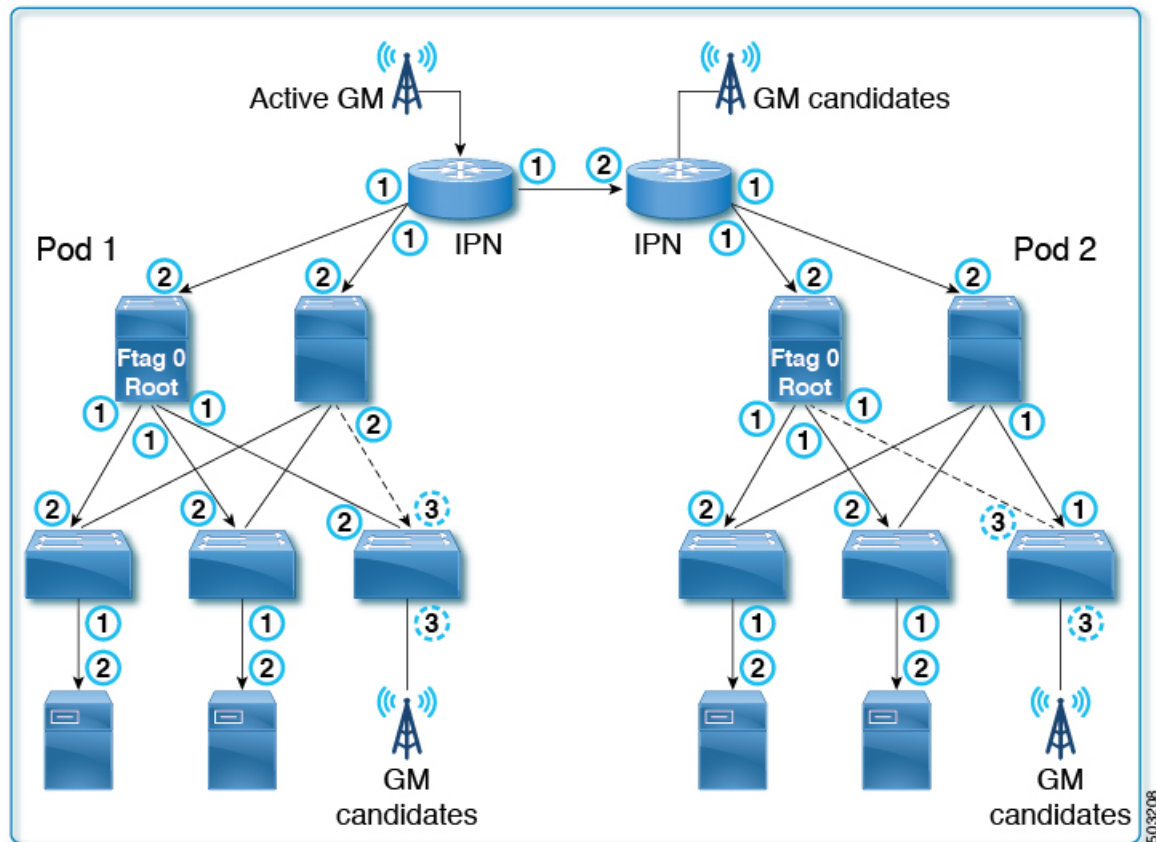
Multi-Pod (External GMs are mandatory)



In Cisco APIC release 3.0(1), PTP cannot be enabled on any other interfaces on demand, such as the down link (front panel ports) on leaf switches.

Implementation in Cisco APIC Releases 4.2(5) and 5.1(1)

Starting in Cisco APIC releases 4.2(5) and 5.1(1), you can enable PTP on a leaf switch's front panel ports to connect the PTP nodes, clients, or grandmaster. The PTP implementation on fabric ports are still the same as the previous releases, except that the PTP parameters for fabric ports can now be adjusted. With this change, you can use the Cisco ACI fabric to propagate time synchronization using PTP with Cisco ACI switches as PTP boundary clock nodes. Prior to that, the only approach Cisco ACI had was to forward PTP multicast or unicast messages transparently as a PTP unaware switch from one leaf switch to another as a tunnel.



Note The 5.0(x) releases do not support the PTP functionality that was introduced in the 4.2(5) and 5.1(1) releases.

Cisco ACI Software and Hardware Requirements

Supported Software for PTP

The following feature is supported from Cisco Application Policy Infrastructure Controller (APIC) release 3.0(1):

- PTP only within the fabric for the latency measurement feature

The following features are supported from Cisco APIC release 4.2(5):

- PTP with external devices by means of the leaf switches
- PTP on leaf switch front panel ports
- Configurable PTP message intervals
- Configurable PTP domain number
- Configurable PTP priorities

- PTP multicast port
- PTP unicast master port on leaf switch front panel ports
- PTP over IPv4/UDP
- PTP profile (Default, AES67, and SMTPE ST2059-2)

Supported Hardware for PTP

Leaf switches, spine switches, and line cards with -EX or later in the product ID are supported. For example, N9K-X9732C-EX or N9K-C93180YC-FX.

The following leaf switches are not supported:

- N9K-C9332PQ
- N9K-C9372PX
- N9K-C9372PX-E
- N9K-C9372TX
- N9K-C9372TX-E
- N9K-C9396PX
- N9K-C9396TX
- N9K-C93120TX
- N9K-C93128TX

The following spine box switch is not supported:

- N9K-C9336PQ

The following spine switch line card is not supported:

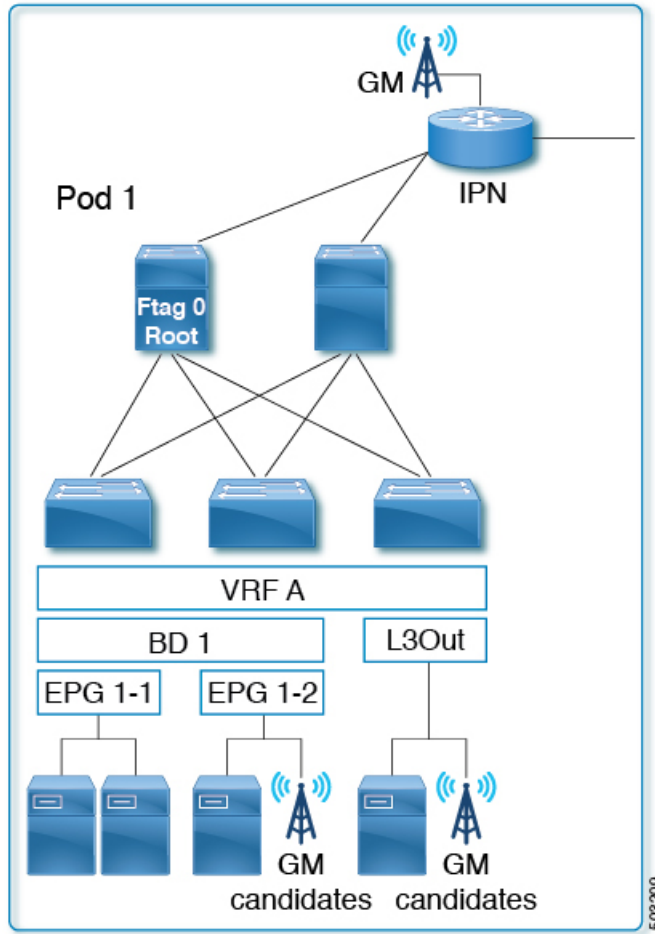
- N9K-X9736PQ

PTP Connectivity

Supported PTP Node Connectivity

External PTP nodes can be connected to the Cisco Application Centric Infrastructure (ACI) fabric using the following methods:

- Inter-pod network
- EPG (on a leaf switch)
- L3Out (on a leaf switch)



PTP is VRF-agnostic, the same as with a standalone NX-OS switch. All PTP messages are terminated, processed, and generated at the interface level on each Cisco ACI switch node as a PTP boundary clock. Regardless of the VRF, bridge domain, EPG, or VLAN, the Best Master Clock Algorithm (BMCA) is calculated across all of the interfaces on each Cisco ACI switch. There is only one PTP domain for the entire fabric.

Any PTP nodes with the E2E delay mechanism (delay req-resp) can be connected to the Cisco ACI switches that are running as a PTP boundary clock.



Note Cisco ACI switches do not support the Peer Delay (P2P) mechanism. Therefore, a P2P transparent clock node cannot be connected to Cisco ACI switches.

Supported PTP Interface Connectivity

| Connection Type | Interface Type | Leaf Switch Type | Supported / Not Supported |
|---|------------------------|------------------|---------------------------|
| Fabric Link (between a leaf and spine switch) | Sub-interface (non-PC) | - | Supported |

| Connection Type | Interface Type | Leaf Switch Type | Supported / Not Supported |
|---|-----------------------------|-----------------------|---------------------------|
| Fabric Link (between a tier-1 and tier-2 leaf switch) | Sub-interface (non-PC) | - | Supported |
| Spine (toward an IPN) | Sub-interface (non-PC) | - | Supported |
| Remote leaf (toward an IPN) | Sub-interface (non-PC) | - | Supported |
| Normal EPG (trunk, access, 802.1P) | Physical, port channel, vPC | On leaf switch | Supported |
| | | On remote leaf switch | Supported |
| | | On tier-2 leaf switch | Supported |
| L3Out (routed, routed-sub) | Physical, port channel | On leaf switch | Supported |
| | | On remote leaf switch | Supported |
| | | On tier-2 leaf switch | Supported |
| L3Out (SVI – trunk, access, 802.1P) | Physical, port channel, vPC | On leaf switch | Not supported |
| | | On remote leaf switch | Not supported |
| | | On tier-2 leaf switch | Not supported |
| L2Out (trunk) | Physical, port channel, vPC | On leaf switch | Not supported |
| | | On remote leaf switch | Not supported |
| | | On tier-2 leaf switch | Not supported |
| EPG/L3Out in tn-mgmt | Physical, port channel, vPC | On leaf switch | Not supported |
| | | On remote leaf switch | Not supported |
| | | On tier-2 leaf switch | Not supported |
| Service EPG (trunk) ¹ | Physical, port channel, vPC | On leaf switch | Not supported |
| | | On remote leaf switch | Not supported |
| | | On tier-2 leaf switch | Not supported |
| Any type of FEX interface | Any | Any | Not supported |
| Out-of-band management interface | Physical | - | Not supported |

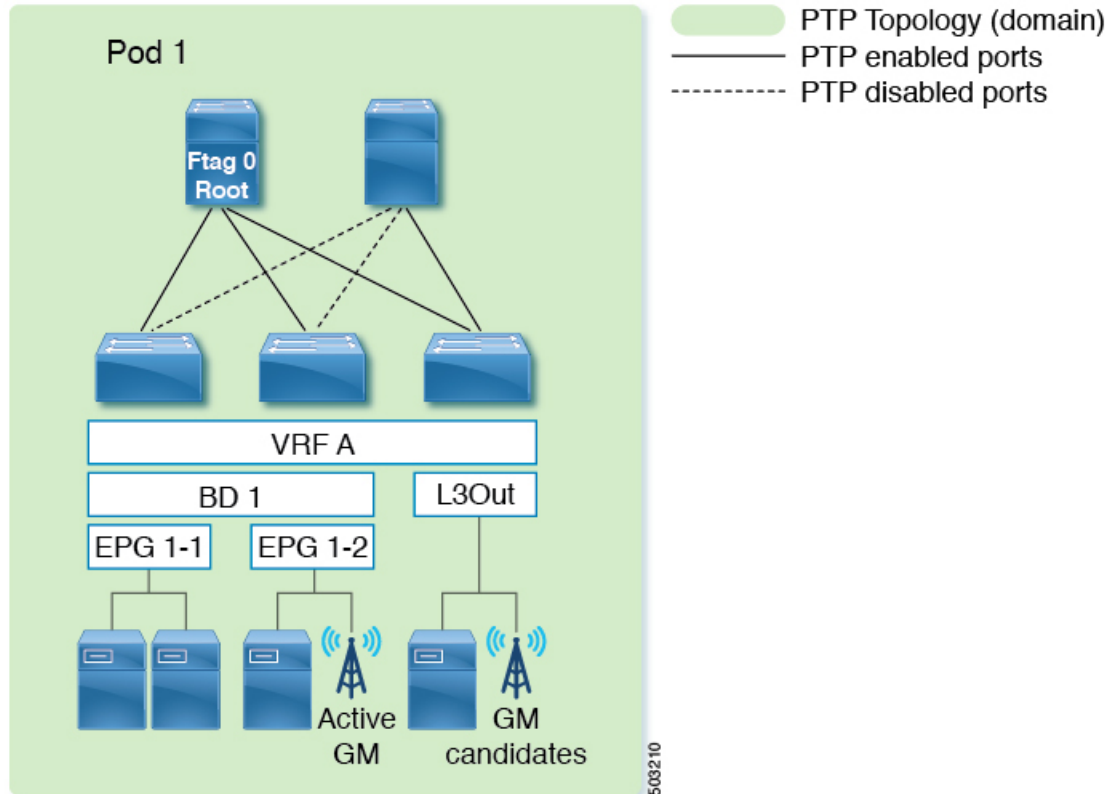
¹ The service EPG is an internal EPG created for a Layer 4 to Layer 7 service graph.

Grandmaster Deployments

You can deploy the grandmaster candidates using one of the following methods:

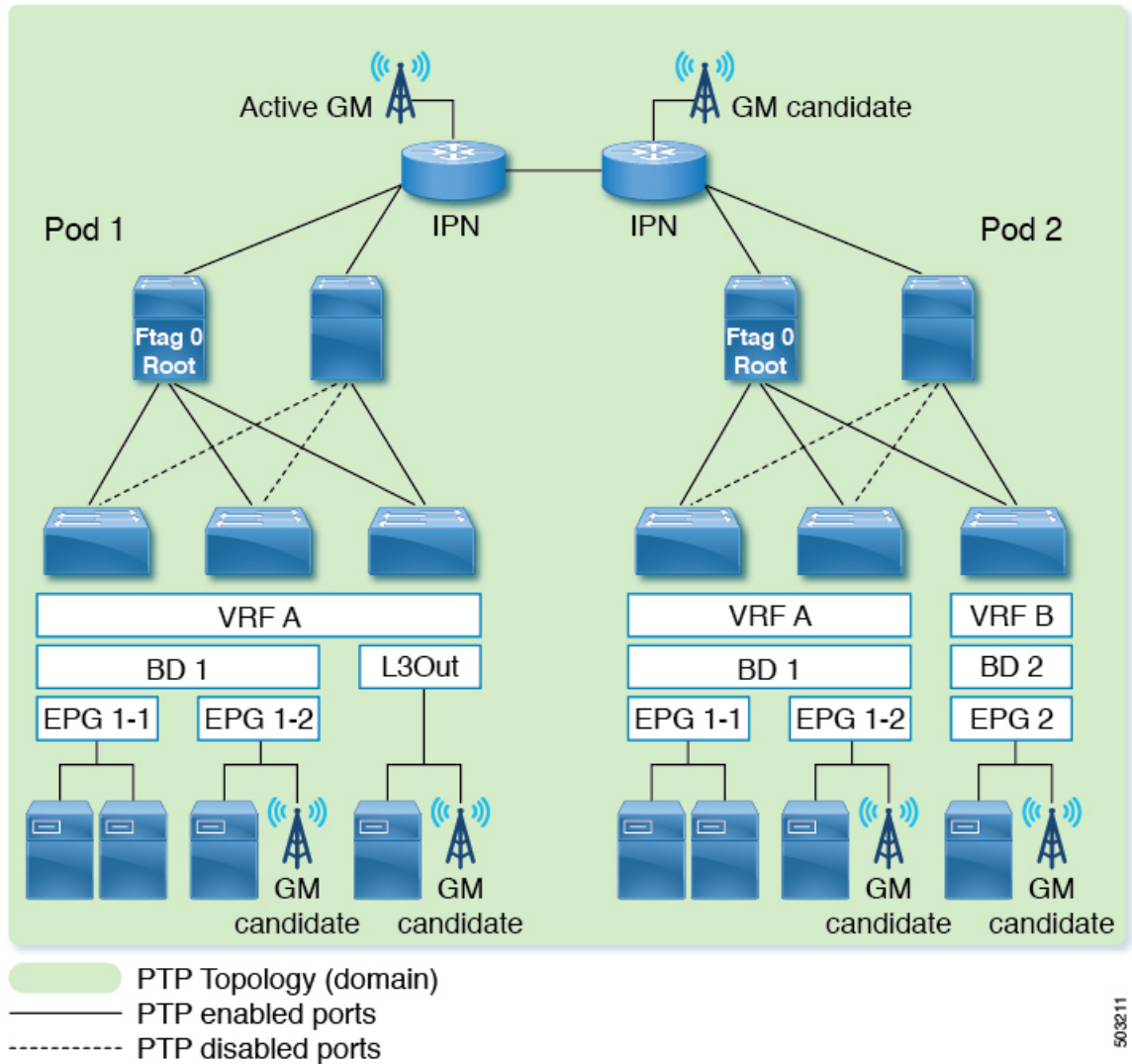
Single Pod

In a single pod deployment, grandmaster candidates can be deployed anywhere in the fabric (L3Out, EPG, or both). The Best Master Clock Algorithm (BMCA) picks one active grandmaster from among all of them.



Multipod With BMCA Across Pods

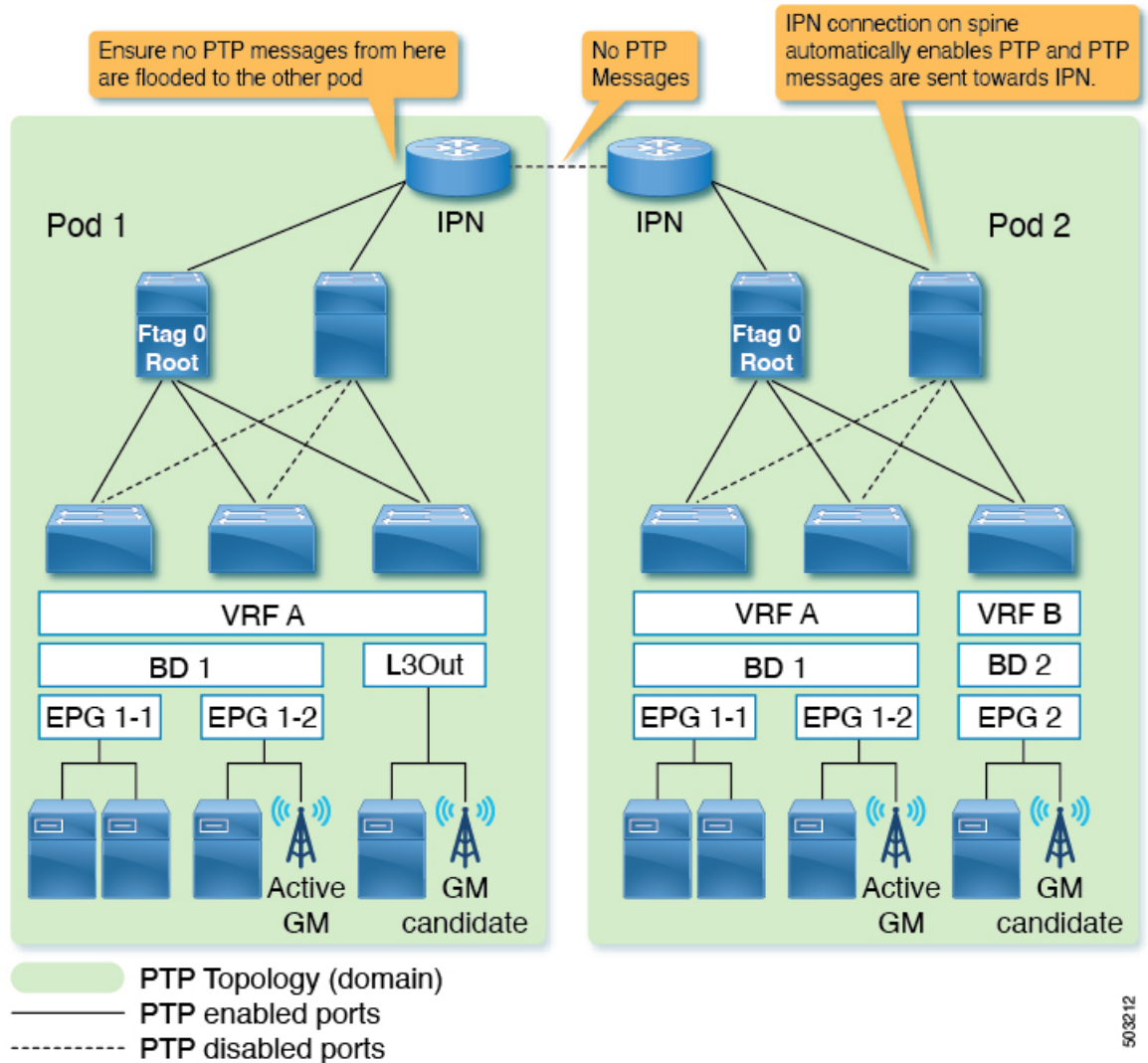
Grandmaster candidates can be deployed anywhere in the fabric (an inter-pod network, L3Out, EPG, or all of them). The BMCA picks one active grandmaster from among all of them across pods. We recommend that you place your grandmasters on inter-pod networks (IPNs) so that the PTP clients in any pod have a similar number of hops to the active grandmaster. In addition, the master/client tree topology will not change drastically when the active grandmaster becomes unavailable.



Multipod With BMCA in Each Pod

If you must have active grandmasters in each pod because PTP accuracy suffers too much degradation through an IPN domain, PTP messages must not traverse through an IPN across pods. You can accomplish this configuration in one of the following ways:

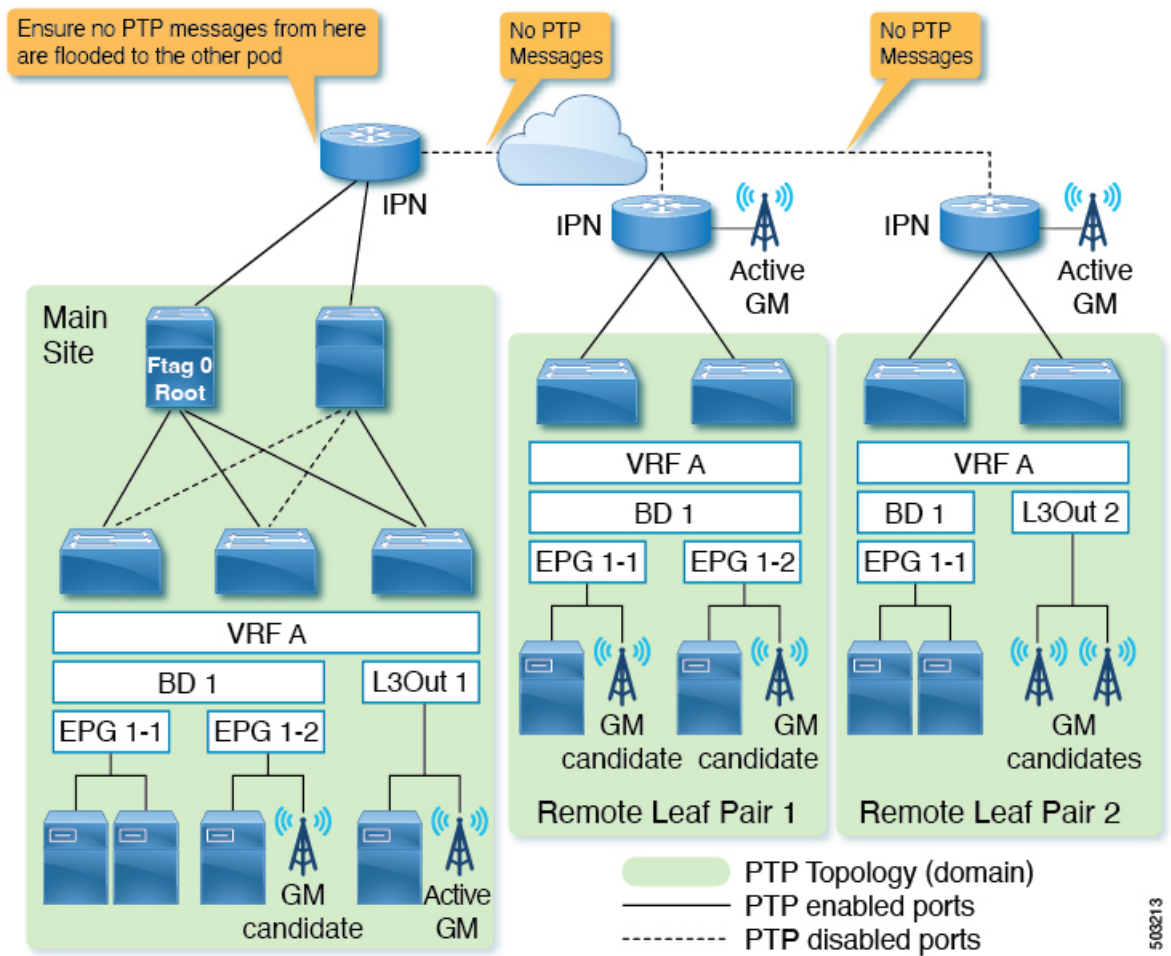
- Option 1: Ensure sub-interfaces are used between the IPN and spine switches, and disable PTP on the IPN.
- Option 2: If the PTP grandmaster is connected to the IPN in each pod, but the PTP topologies still must be separated, disable PTP on the IPN interfaces that are between the pods.



503212

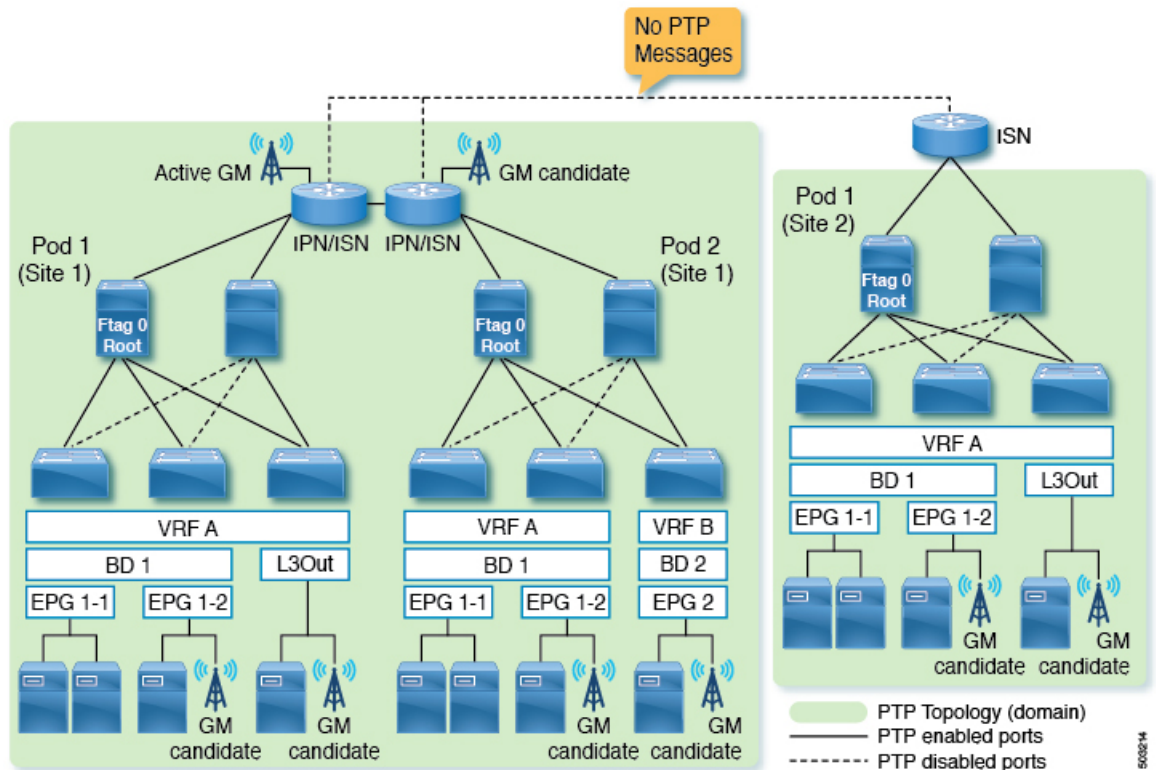
Remote Leaf Switch

Remote leaf switch sites are typically not close to the main data center or to each other, and it is difficult to propagate PTP messages across each location with accurate measurements of delay and correction. Hence, we recommend that you prevent PTP messages from traversing each site (location) so that the PTP topology is established within each site (location). Some remote locations may be close to each other. In such a case, you can enable PTP between those IPNs to form one PTP topology across those locations. You can use the same options mentioned in *Multipod With BMCA in Each Pod* to prevent PTP message propagation.



Cisco ACI Multi-Site

Each site is typically not close to each other, and it is difficult to propagate PTP messages across each site with accurate measurements of delay and correction. Hence, we recommend that you prevent PTP messages from traversing each site so that the PTP topology is established within each site. You can use the same options mentioned in *Multipod With BMCA in Each Pod* to prevent PTP message propagation. Also, Cisco ACI Multi-Site has no visibility or capability of configuring PTP.



PTP Limitations

The following limitations apply to PTP:

- Cisco Application Centric Infrastructure (ACI) leaf and spine switches can work as PTP boundary clocks. The switches cannot work as PTP transparent clocks.
- Only the E2E delay mechanism (delay request/response mechanism) is supported. The P2P delay mechanism is not supported.
- Only PTP over IPv4/UDP is supported. PTP over IPv6 or Ethernet is not supported.
- Only PTPv2 is supported.
 - Although PTPv1 packets are still redirected to the CPU when PTP is enabled on any of the front panel ports on the leaf switch, the packets will be discarded on the CPU.
- PTP Management TLVs are not recognized by Cisco ACI switches, but they are still forwarded as defined in IEEE1588-2008 to support the SMTPE PTP profile.
- PTP cannot be used as the system clock of the Cisco ACI switches.
- PTP is not supported on Cisco Application Policy Infrastructure Controller (APIC).
- NTP is required for all switches in the fabric.
- PTP offload is not supported. This functionality is the offloading of the PTP packet processing to each line card CPU on a modular spine switch for higher scalability.

- Due to a hardware limitation, interfaces with 1G/100M speed have lower accuracy than 10G interfaces when there are traffic loads.
- `Sync` and `Delay_Request` messages can support up to a -4 interval (1/16 seconds). Interval values of -5 to -7 are not supported.
- For leaf switch front panel ports, PTP can be enabled per interface and VLAN, but PTP is automatically enabled on all appropriate fabric links (interfaces between leaf and spine switches, tier-1 and tier-2 leaf switches, and interfaces toward the IPN/ISN) after PTP is enabled globally. The appropriate fabric links are the interfaces that belong to `ftag0` tree.
- PTP on Cisco ACI interfaces toward the IPN/ISN is enabled with the native VLAN 1 and sent out without a VLAN tag. The interfaces on the ISN/IPN node can send PTP packets toward Cisco ACI spine switches without a VLAN tag or with VLAN ID 4, which is enabled automatically for IPN/ISN connectivity regardless of PTP.
- PTP must be enabled globally for leaf switch front panel interfaces to use PTP. This means that you cannot enable PTP on leaf switch front panel ports without enabling PTP on the fabric links.
- PTP configuration using `tn-mgmt` and `tn-infra` is not supported.
- PTP can be enabled only on one VLAN per interface.
- PTP cannot be enabled on the interface and the VLAN for an L3Out SVI. PTP can be enabled on another VLAN on the same interface using an EPG.
- Only the leaf switch front panel interfaces can be configured as unicast master ports. The interfaces cannot be configured as unicast client ports. Unicast ports are not supported on a spine switch.
- Unicast negotiation is not supported.
- Unicast mode does not work with a PC or vPC when the PC or vPC is connected to a device such as NX-OS, which configures PTP on individual member ports.
- PTP and MACsec should not be configured on the same interface.
- When PTP is globally enabled, to measure the latency of traffic traversing through the fabric, Cisco ACI adds Cisco timestamp tagging (TTag) to traffic going from one ACI switch node to another ACI switch node. This results in an additional 8 bytes for such traffic. Typically, users do not need to take any actions regarding this implementation because the TTag is removed when the packets are sent out to the outside of the ACI fabric. However, when the setup consists of Cisco ACI Multi-Pod, user traffic traversing across pods will keep the TTag in its inner header of the VXLAN. In such a case, increase the MTU size by 8 bytes on the ACI spine switch interfaces facing toward the Inter-Pod Network (IPN) along with all non-ACI devices in the IPN. IPN devices do not need to support nor be aware of the TTag, as the TTag is embedded inside of the VXLAN payload.
- When PTP is globally enabled, ERSPAN traffic traversing through spine nodes to reach to the ERSPAN destination will have Cisco timestamp tagging (TTag) with ethertype 0x8988. There is no impact to the original user traffic.
- In the presence of leaf switches that do not support PTP, you must connect an external grandmaster to all of the spine switches using IPN or using leaf switches that support PTP. If a grandmaster is connected to one or a subset of spine switches, PTP messages from the spine may be blocked by the unsupported leaf switch before they reach other switches depending on the `ftag0` tree status. PTP within leaf and spine switches are enabled based on `ftag0` tree, which is automatically built based on Cisco ACI infra ISIS for loop free multicast connectivity between all leaf and spine switches in each pod.

- You cannot enable PTP on a VLAN that is deployed on a leaf node interface using VMM domain integration.

Configuring PTP

PTP Configuration Basic Flow

The following steps provide an overview of the PTP configuration process:

-
- Step 1** Enable PTP globally and set PTP parameters for all fabric interfaces.
 - Step 2** Create PTP user profile for leaf front panel interfaces under **Fabric > Access Policies > Policies > Global**.
 - Step 3** Enable PTP under **EPG > Static Ports** with the PTP user profile.
 - Step 4** Enable PTP under **L3Out > Logical Interface Profile > Routed or Sub-Interface** with the PTP user profile.
-

Configuring the PTP Policy Globally and For the Fabric Interfaces Using the GUI

This procedure enables PTP globally and for the fabric interfaces using the Cisco Application Policy Infrastructure Controller (APIC) GUI. When PTP is enabled globally, ongoing TEP to TEP latency measurements get enabled automatically.

-
- Step 1** On the menu bar, choose **System > System Settings**.
 - Step 2** In the Navigation pane, choose **Precision Time Protocol**.
 - Step 3** In the Work pane, set the interface properties as appropriate for your desired configuration. At the least, you must set the **Admin State** to **Enabled**.

See the online help page for information about the fields. If any interval value that you specify is outside of the chosen PTP profile standard range, the configuration is rejected.

The PTP profile, intervals, and timeout fields apply to fabric links. The other fields apply to all of the leaf and spine switches.
 - Step 4** Click **Submit**.
-

Creating the PTP User Profile for Leaf Switch Front Panel Ports Using the GUI

This procedure creates the PTP user profile for leaf switch front panel ports using the Cisco Application Policy Infrastructure Controller (APIC) GUI. A PTP user profile is applied to the leaf switch front panel interfaces using an EPG or L3Out.

Before you begin

You must enable PTP globally to use PTP on leaf switch front panel ports that face external devices.

-
- Step 1** On the menu bar, choose **Fabric > Access Policies**.

- Step 2** In the Navigation pane, choose **Policies > Global > PTP User Profile**.
- Step 3** Right-click **PTP User Profile** and choose **Create PTP User Profile**.
- Step 4** In the **Create PTP User Profile** dialog, set the values as desired for your configuration.
- See the online help page for information about the fields. If any interval value that you specify is outside of the chosen PTP profile standard range, the configuration is rejected.
- Step 5** Click **Submit**.
-

Enabling PTP on EPG Static Ports Using the GUI

This procedure enables PTP on EPG static ports using the Cisco Application Policy Infrastructure Controller (APIC) GUI. You can enable PTP with multicast or unicast master mode.

Before you begin

You must first create a PTP user profile for the leaf switch front panel ports and enable PTP globally.

- Step 1** On the menu bar, choose **Tenants > All Tenants**.
- Step 2** In the Work pane, double-click the tenant's name.
- Step 3** In the Navigation pane, choose **Tenant *tenant_name* > Application Profiles > *app_profile_name* > Application EPGs > *app_epg_name* > Static Ports > *static_port_name***.
- Step 4** In the Work pane, for the **PTP State** toggle, choose **Enable**. You might need to scroll down to see **PTP State**. PTP-related fields appear.
- Step 5** Configure the PTP fields as required for your configuration.
- **PTP Mode:** Choose **multicast** or **unicast master**, as appropriate.
 - **PTP Source Address:** PTP packets from this interface and VLAN are sent with the specified IP address as the source. The leaf switch TEP address is used by default or when you enter "0.0.0.0" as the value. This value is optional for multicast mode. Use the bridge domain SVI or EPG SVI for unicast mode. The source IP address must be reachable by the connected PTP node for unicast mode.
 - **PTP User Profile:** Choose the PTP user profile that you created for the leaf switch front panel ports to specify the message intervals.
- See the online help page for additional information about the fields.
- Step 6** Click **Submit**.
-

Enabling PTP on L3Out Interfaces Using the GUI

This procedure enables PTP on L3Out interfaces using the Cisco Application Policy Infrastructure Controller (APIC) GUI. You can enable PTP with multicast or unicast master mode.

Before you begin

You must first create a PTP user profile for the leaf switch front panel ports and enable PTP globally.

-
- Step 1** On the menu bar, choose **Tenants > All Tenants**.
- Step 2** In the Work pane, double-click the tenant's name.
- Step 3** In the Navigation pane, choose **Tenant *tenant_name* > Networking > L3Outs > *l3out_name* > Logical Node Profiles > *node_profile_name* > Logical Interface Profiles > *interface_profile_name***.
- Step 4** In the Work pane, choose **Policy > Routed Sub-Interfaces** or **Policy > Routed Interfaces**, as appropriate.
- Step 5** If you want to enable PTP on an existing L3Out, perform the following sub-steps:
- Double-click the desired interface to view its properties.
 - Scroll down if necessary to find the PTP properties, set the **PTP State** to **Enable**, and enter the same values that you used for the EPG static ports.

See the online help page for information about the fields.
 - Click **Submit**.
- Step 6** If you want to enable PTP on a new L3Out, perform the following sub-steps:
- Click + at the upper right of the table.
 - In **Step 1 > Identity**, enter the appropriate values.
 - In **Step 2 > Configure PTP**, set the **PTP State** to **Enable**, and enter the same values that you used for the EPG static ports.

See the online help page for information about the fields.
 - Click **Finish**.
-

Configuring the PTP Policy Globally and For the Fabric Interfaces Using the REST API

This procedure enables PTP globally and for the fabric interfaces using the REST API. When PTP is enabled globally, ongoing TEP to TEP latency measurements get enabled automatically.

To configure the PTP policy globally and for the fabric interfaces, send a REST API POST similar to the following example:

POST: /api/mo/uni/fabric/ptpmode.xml

```
<latencyPtpMode
  state="enabled"                # PTP admin state
  systemResolution="11"         # Latency Resolution (can be skipped for PTP)
  prio1="255"                   # Global Priority1
  prio2="255"                   # Global Priority2
  globalDomain="0"              # Global Domain
  fabProfileTemplate="aes67"    # PTP Profile
  fabAnnounceIntvl="1"         # Announce Interval (2^x sec)
  fabSyncIntvl="-3"             # Sync Interval (2^x sec)
  fabDelayIntvl="-2"           # Delay Request Interval (2^x sec)
  fabAnnounceTimeout="3"       # Announce Timeout
/>
```

Creating the PTP User Profile for Leaf Switch Front Panel Ports Using the REST API

A PTP user profile is applied to the leaf switch front panel interfaces using an EPG or L3Out. You also must enable PTP globally to use PTP on leaf switch front panel ports that face external devices.

To create the PTP user profile, send a REST API POST similar to the following example:

```

POST: /api/mo/uni/infra/ptpprofile-PTP_AES.xml

<ptpProfile
  name="PTP_AES" # PTP User Profile name
  profileTemplate="aes67" # PTP Profile
  announceIntvl="1" # Announce Interval (2^x sec)
  syncIntvl="-3" # Sync Interval (2^x sec)
  delayIntvl="-2" # Delay Request Interval (2^x sec)
  announceTimeout="3" # Announce Timeout
/>

```

Enabling PTP on EPG Static Ports Using the REST API

Before you can enable PTP on EPG static ports, you must first create a PTP user profile for the leaf switch front panel ports and enable PTP globally.

To enable PTP on EPG static ports, send a REST API POST similar to the following example:

```

POST: /api/mo/uni/tn-TK/ap-AP1/epg-EPG1-1.xml

```

Multicast Mode

```

<fvRsPathAtt
  tDn="topology/pod-1/paths-101/pathep-[eth1/1]"
  encap="vlan-2011">
  <ptpEpgCfg
    ptpMode="multicast" # PTP mode
    <ptpRsProfile # PTP user profile
      tDn="uni/infra/ptpprofile-PTP_AES"/>
    </ptpEpgCfg>
  </fvRsPathAtt>

```

Unicast Mode

```

<fvRsPathAtt
  tDn="topology/pod-1/paths-101/pathep-[eth1/1]"
  encap="vlan-2011">
  <ptpEpgCfg
    srcIp="192.168.1.254" # PTP source IP address
    ptpMode="unicast-master" # PTP mode
    <ptpRsProfile # PTP user profile
      tDn="uni/infra/ptpprofile-PTP_AES"/> # PTP unicast destination
    <ptpUcastIp dstIp="192.168.1.11"/> IP address
    </ptpEpgCfg>
  </fvRsPathAtt>

```

If `ptpEpgCfg` exists, that means that PTP is enabled. If PTP must be disabled on that interface, delete `ptpEpgCfg`.

Enabling PTP on L3Out Interfaces Using the REST API

This procedure enables PTP on L3Out interfaces using the REST API. Before you can enable PTP on the L3Out interfaces, you must first create a PTP user profile for the leaf switch front panel ports and enable PTP globally.

To enable PTP on L3Out interfaces, send a REST API POST similar to the following example:

```

POST: /api/node/mo/uni/tn-TK/out-BGP/lnodep-BGP_nodeProfile/lifp-BGP_IfProfile.xml

```


Multicast Mode

```
<l3extRsPathL3OutAtt
  tDn="topology/pod-1/paths-103/pathep-[eth1/11]"
  addr="11.0.0.1/30" ifInstT="l3-port">
  <ptpRtdEpgCfg
    ptpMode="multicast"> # PTP mode
    <ptpRsProfile # PTP user profile
      tDn="uni/infra/ptpprofile-PTP_AES"/>
    </ptpRtdEpgCfg>
  </l3extRsPathL3OutAtt>
```

Unicast Mode

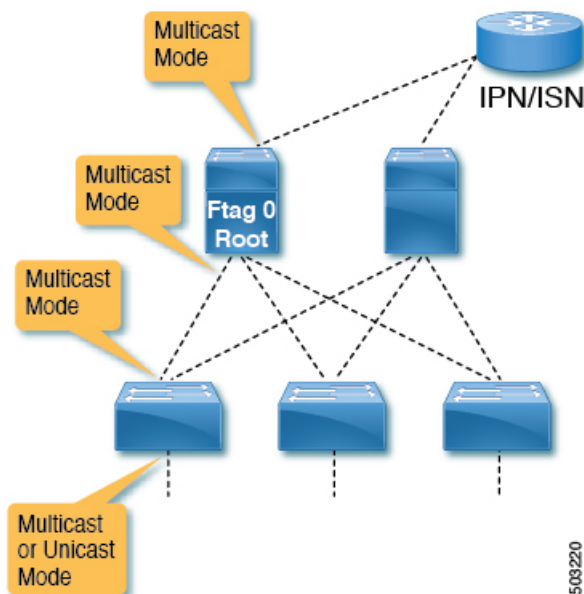
```
<l3extRsPathL3OutAtt
  tDn="topology/pod-1/paths-103/pathep-[eth1/11]"
  addr="11.0.0.1/30" ifInstT="l3-port">
  <ptpRtdEpgCfg
    srcIp="11.0.0.1" # PTP source IP address
    ptpMode="unicast-master"> # PTP mode
    <ptpRsProfile # PTP user profile
      tDn="uni/infra/ptpprofile-PTP_AES"/> # PTP unicast destination
    <ptpUcastIp dstIp="11.0.0.4"/> IP address
    </ptpRtdEpgCfg>
  </l3extRsPathL3OutAtt>
```

If `ptpRtdEpgCfg` exists, that means that PTP is enabled. If PTP needs to be disabled on that interface, delete `ptpRtdEpgCfg`.

PTP Unicast, Multicast, and Mixed Mode on Cisco ACI

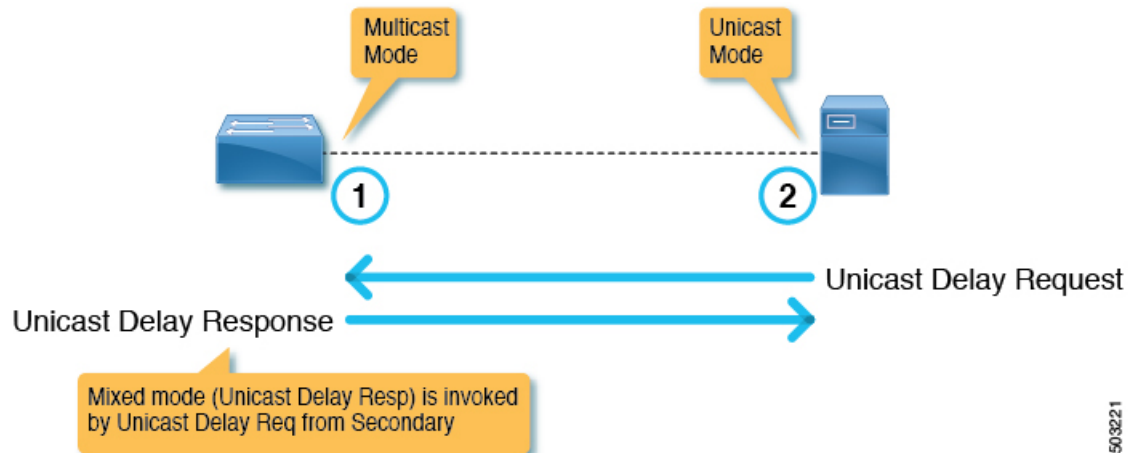
By default, all PTP interfaces run in multicast mode. Only the leaf switch front panel interfaces can be configured in unicast mode. Only unicast master ports are supported; unicast client ports are not supported.

Figure 1: Multicast or Unicast Mode



Mixed mode (a PTP multicast port replying with a unicast delay response) will be automatically activated on a PTP master port in multicast mode when the port receives a unicast delay request. Mixed mode is essentially a multicast master and unicast client.

Figure 2: Mixed Mode



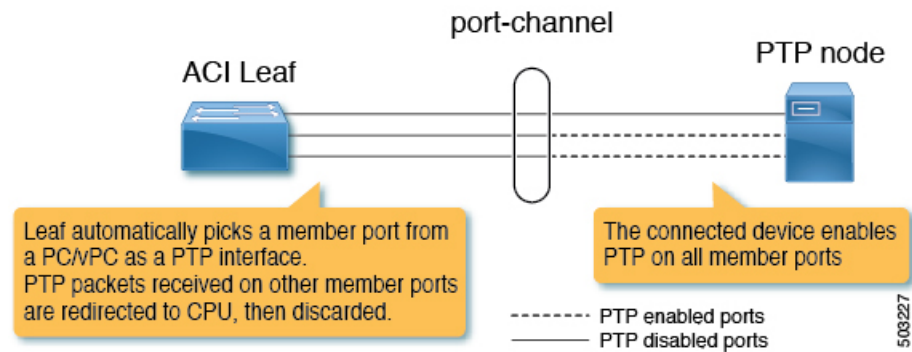
One leaf switch can have multiple PTP unicast master ports. The supported number of client switch IP addresses on each unicast master port is 2. More IP addresses can be configured, but not qualified. The PTP unicast master ports and PTP multicast ports can be configured on the same switch.

PTP Unicast Mode Limitations on Cisco ACI

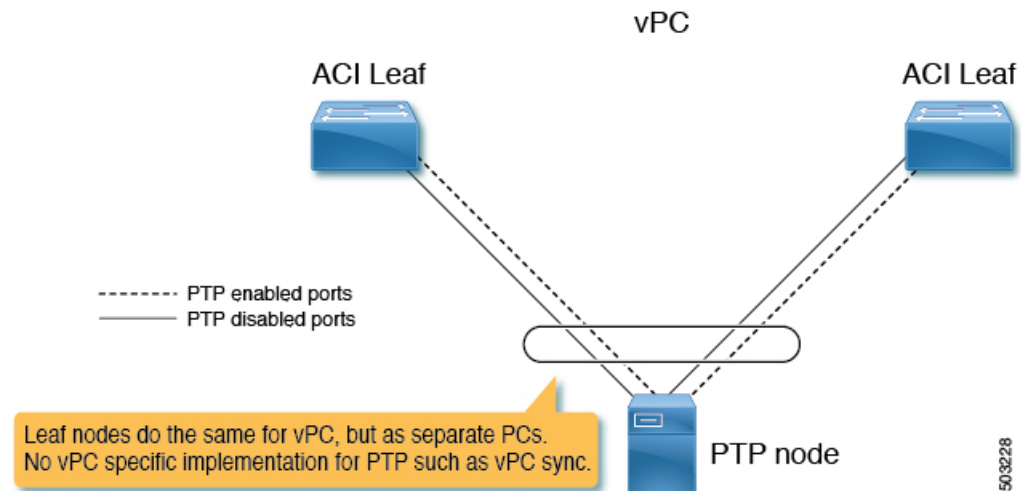
PTP unicast negotiation is not supported. Because Cisco Application Centric Infrastructure (ACI) does not have unicast negotiation to request the messages that Cisco ACI wants or to grant those requests from other nodes, Cisco ACI PTP unicast master ports will send `Announce`, `Sync`, and `Follow_Up` messages with intervals configured using the Cisco Application Policy Infrastructure Controller (APIC) without receiving any requests from its client nodes. `Unicast_Delay_Response` messages are sent out as a response to `Delay_Request` messages from the unicast client nodes. Because a unicast master port sends PTP messages such as `Sync` without listening to unicast requests, the Best Master Clock Algorithm (BMCA) is not calculated on the Cisco ACI PTP unicast ports.

PTP PC and vPC Implementation on Cisco ACI

For port channels (PCs) and virtual port channels (vPCs), PTP is enabled per PC or vPC instead of per member port. Cisco Application Centric Infrastructure (ACI) does not allow PTP to be enabled on each member port of the parent PC or vPC individually.



When PTP is enabled on a Cisco ACI PC or vPC, the leaf switch automatically picks a member port from the PC on which PTP is enabled. When the PTP-enabled member port fails, the leaf switch picks another member port that is still up. The PTP port status is inherited from the previous PTP-enabled member port.



When PTP is enabled on a Cisco ACI vPC port, even though vPC is a logical bundle of two port channels on two leaf switches, the behavior is the same as PTP being enabled on a normal port channel. There is no specific implementation for the vPC, such as the synchronization of PTP information between vPC peer leaf switches.



Note Unicast mode does not work with a PC or vPC when the PC or vPC is connected to a device such as NX-OS, which configures PTP on individual member ports.

PTP Packet Filtering and Tunneling

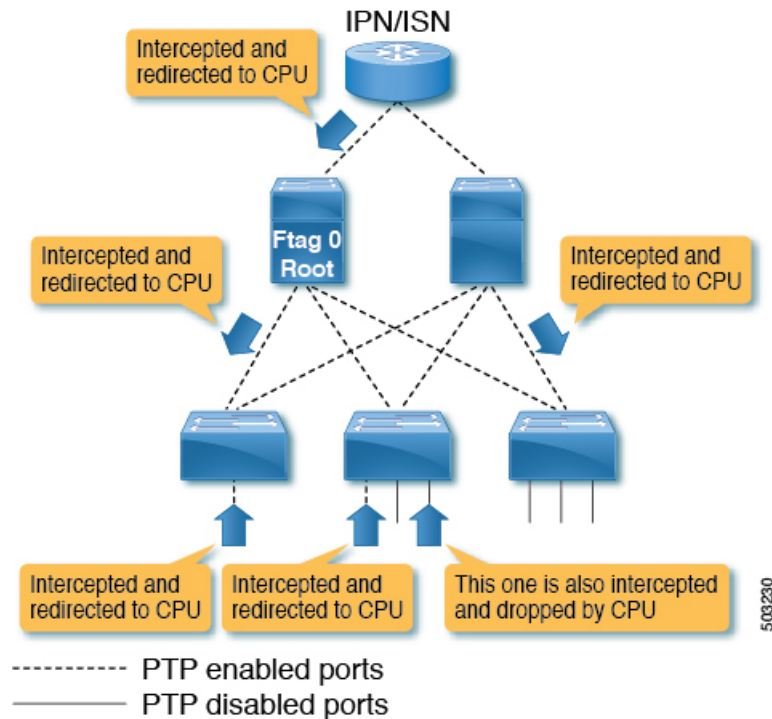
PTP Packet Filtering

When PTP handles packets on the fabric ports and PTP is enabled globally, all spine and leaf switches have internal filters to redirect all incoming PTP packets from any fabric ports to the CPU.

When PTP handles packets on the front panel ports and PTP is enabled on at least one leaf switch front panel port on a given leaf switch, the leaf switch has internal filters to redirect all incoming PTP packets from any

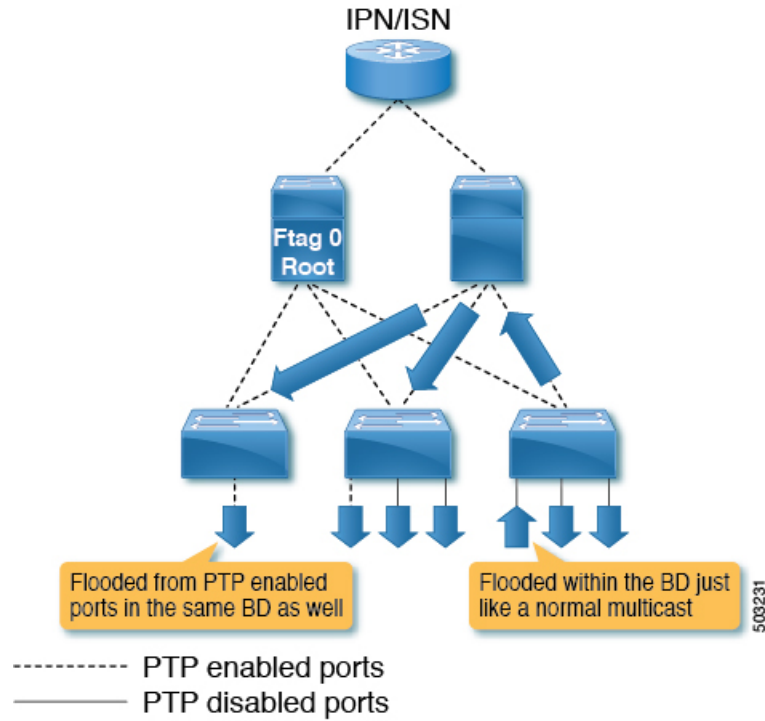
front panel ports. Even if a PTP packet is received from a front panel port on which PTP is not enabled, the packet is still intercepted and redirected to the CPU, then discarded.

Figure 3: Packet Filtering On the Front Panel on a Leaf switch With PTP-Enabled Front Panel Ports



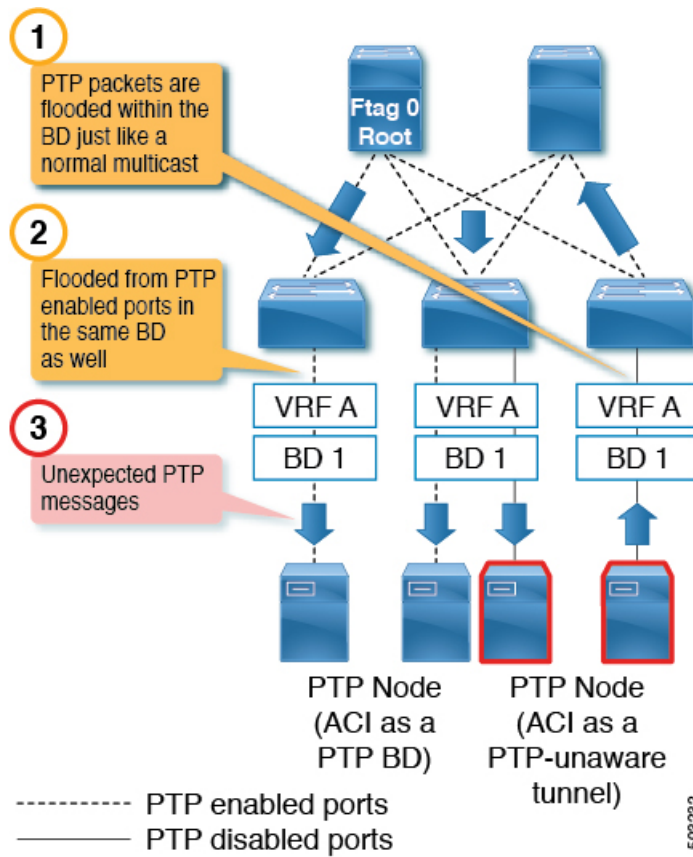
When PTP handles packets on the front panel ports and PTP is not enabled on any leaf switch front panel ports on a given leaf switch, the leaf switch does not have internal filters to redirect PTP packets from any front panel ports. If a PTP packet is received on a front panel port on such a leaf switch, the packet is handled as a normal multicast packet and is forwarded or flooded to other switches using VXLAN. The other switches will also handle this as a normal multicast packet because the PTP packets that are supposed to be intercepted by the Cisco Application Centric Infrastructure (ACI) switches are not encapsulated in VXLAN even between leaf and spine switches. This may cause unexpected PTP behavior on other leaf switches with PTP enabled on the front panel ports. For more information, see [Cisco ACI As a PTP Boundary Clock or PTP-Unaware Tunnel](#), on page 45.

Figure 4: Packet Filtering On the Front Panel on a Leaf switch Without PTP-Enabled Front Panel Ports

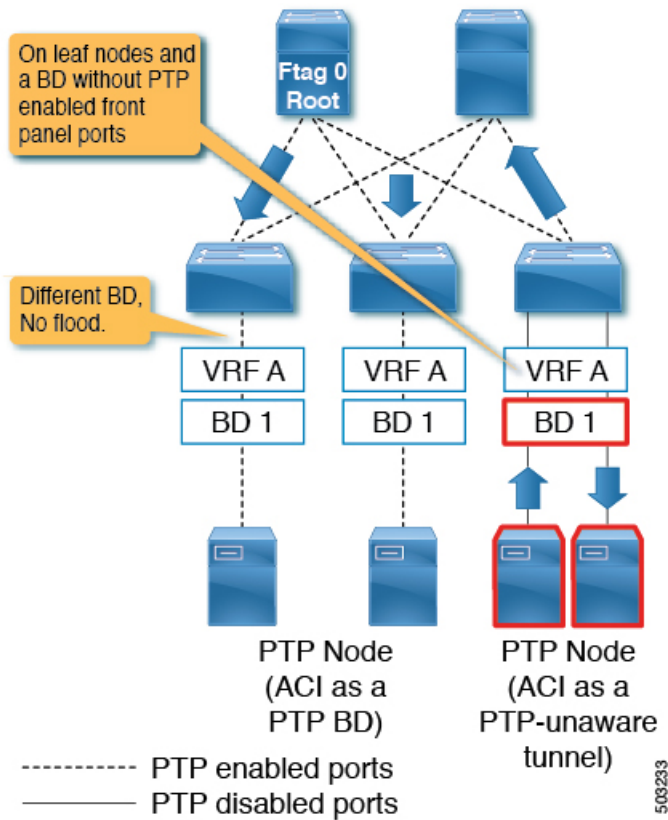


Cisco ACI As a PTP Boundary Clock or PTP-Unaware Tunnel

PTP packets from a leaf switch with no PTP front panel ports are flooded in the bridge domain. The packets are flooded even toward PTP nodes in the same bridge domain that expect Cisco Application Centric Infrastructure (ACI) to regenerate PTP messages as a PTP boundary clock, as shown in the following illustration:

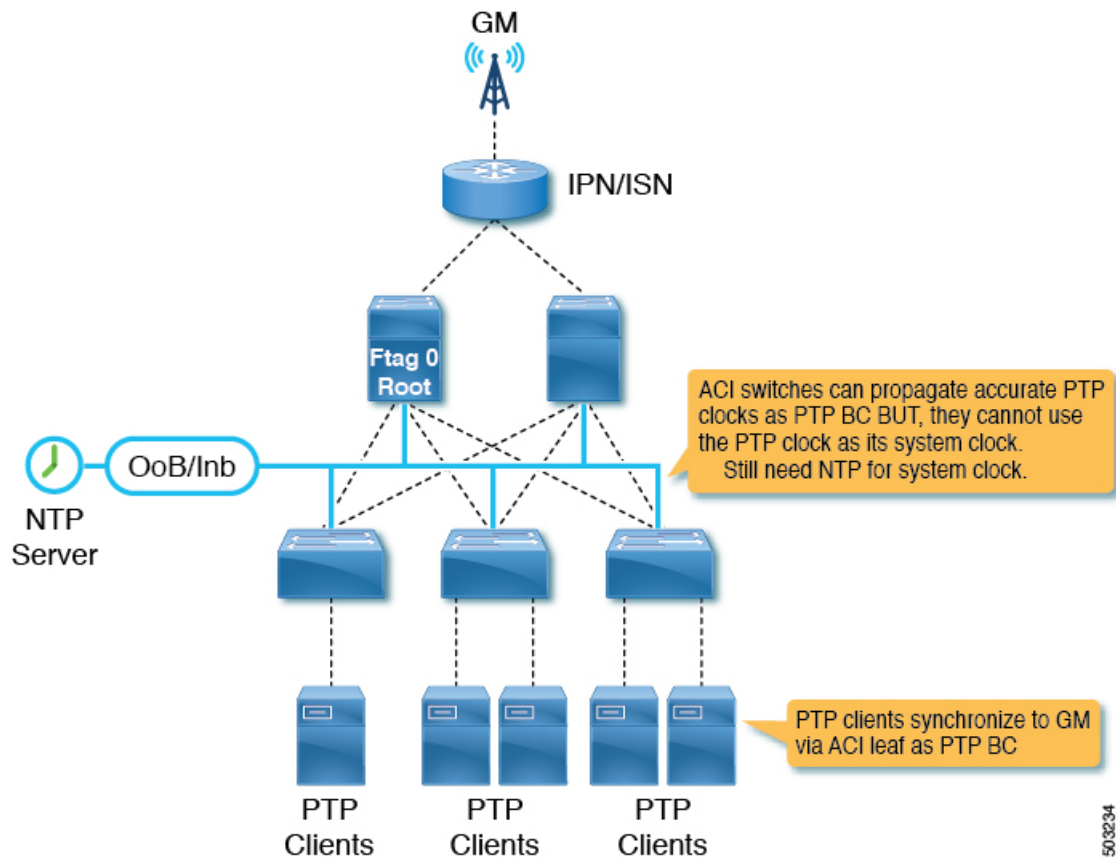


This will result in confusing PTP nodes and its time calculation due to unexpected PTP packets. On the other hand, PTP packets from a leaf switch with PTP front panel ports are always intercepted and never tunneled even if the packets are received on a port on which PTP is not enabled. Therefore, do not mix PTP nodes that need Cisco ACI to be a PTP boundary clock and that need Cisco ACI to be a PTP-unaware tunnel in the same bridge domain and on the same leaf switch. The configuration shown in the following illustration (different bridge domain, different leaf switch) is supported:



PTP and NTP

Cisco Application Centric Infrastructure (ACI) switches run as PTP boundary clocks to provide an accurate clock from the grandmaster to the PTP clients. However, the Cisco ACI switches and Cisco Application Policy Infrastructure Controllers (APICs) cannot use those PTP clocks as their own system clock. The Cisco ACI switches and Cisco APICs still need an NTP server to update their own system clock.



503234



Note For PTP to work accurately and constantly on Cisco ACI, NTP must be configured for all of the switches to keep their system clock as accurate as the PTP grandmaster in the 100 ms order. In other words, the system clock must have less than a 100 ms difference compared to the PTP grandmaster.

PTP Verification

Summary of PTP Verification CLI Commands

You can log into one of the leaf switches and the following commands to verify the PTP configuration:

| Command | Purpose |
|---|---|
| <code>show ptp port interface <i>slot/port</i></code> | Displays the PTP parameters of a specific interface. |
| <code>show ptp brief</code> | Displays the PTP status. |
| <code>show ptp clock</code> | Displays the properties of the local clock, including clock identity. |
| <code>show ptp parent</code> | Displays the properties of the PTP parent. |

| Command | Purpose |
|--|--|
| show ptp clock foreign-masters record | Displays the state of foreign masters known to the PTP process. For each foreign master, the output displays the clock identity, basic clock properties, and whether the clock is being used as a grandmaster. |
| show ptp counters [all interface Ethernet slot/port] | Displays the PTP packet counters for all interfaces or for a specified interface. |
| show ptp corrections | Displays the last few PTP corrections. |

Showing the PTP Port Information

The following example shows the port interface information:

```
f2-leaf1# vsh -c 'show ptp port int e1/1'
PTP Port Dataset: Eth1/1
Port identity: clock identity: 00:3a:9c:ff:fe:6f:a4:df
Port identity: port number: 0
PTP version: 2
Port state: Master
VLAN info: 20 <--- PTP messages are sent on this PI-VLAN
Delay request interval(log mean): -2
Announce receipt time out: 3
Peer mean path delay: 0
Announce interval(log mean): 1
Sync interval(log mean): -3
Delay Mechanism: End to End
Cost: 255
Domain: 0
```

The following example shows the information for the specified VLAN:

```
f2-leaf1# show vlan id 20 extended
```

| VLAN Name | Encap | Ports |
|------------------|-----------|------------------------|
| 20 TK:AP1:EPG1-1 | vlan-2011 | Eth1/1, Eth1/2, Eth1/3 |

Showing the PTP Port Status

The following example shows the brief version of the port status:

```
f2-leaf1# show ptp brief
```

```
PTP port status
-----
Port                State
-----
Eth1/1              Master
Eth1/51             Passive
Eth1/52             Slave
```

Showing the PTP Switch Information

The following example shows the brief version of the switch status:

```
f2-leaf1# show ptp clock
PTP Device Type : boundary-clock
PTP Device Encapsulation : layer-3
```

```

PTP Source IP Address : 20.0.32.64 <--- Switch TEP. Like a router-id.
                                     This is not PTP Source Address you
                                     configure per port.
Clock Identity : 00:3a:9c:ff:fe:6f:a4:df <--- PTP clock ID. If this node is
                                     the grandmaster, this ID is the
                                     grandmaster's ID.

Clock Domain: 0
Slave Clock Operation : Two-step
Master Clock Operation : Two-step
Slave-Only Clock Mode : Disabled
Number of PTP ports: 3
Configured Priority1 : 255
Priority1 : 255
Priority2 : 255
Clock Quality:
    Class : 248
    Accuracy : 254
    Offset (log variance) : 65535
Offset From Master : -8 <--- -8 ns. the clock difference from the
                             closest parent (master)
Mean Path Delay : 344 <--- 344 ns. Mean path delay measured by
                             E2E mechanism.
Steps removed : 2 <--- 2 steps. 2 PTP BC nodes between the
                             grandmaster.

Correction range : 100000
MPD range : 1000000000
Local clock time : Thu Jul 30 01:26:14 2020
Hardware frequency correction : NA

```

Showing the Grandmaster and Parent (Master) Information

The following example shows the PTP grandmaster and parent (master) information:

```

f2-leaf1# show ptp parent

PTP PARENT PROPERTIES

Parent Clock:
Parent Clock Identity: 2c:4f:52:ff:fe:e1:7c:1a <--- closest parent (master)
Parent Port Number: 30
Observed Parent Offset (log variance): N/A
Observed Parent Clock Phase Change Rate: N/A

Parent IP: 20.0.32.65 <--- closest parent's PTP
                             source IP address

Grandmaster Clock:
Grandmaster Clock Identity: 00:78:88:ff:fe:f9:2b:13 <--- GM
Grandmaster Clock Quality: <--- GM's quality
    Class: 248
    Accuracy: 254
    Offset (log variance): 65535
    Priority1: 128
    Priority2: 255

```

The following example shows the PTP foreign master clock records:

```

f2-leaf1# show ptp clock foreign-masters record

P1=Priority1, P2=Priority2, C=Class, A=Accuracy,
OSLV=Offset-Scaled-Log-Variance, SR=Steps-Removed
GM=Is grandmaster

-----
Interface          Clock-ID          P1  P2  C  A  OSLV  SR
-----

```

```

-----
Eth1/51  c4:f7:d5:ff:fe:2b:eb:8b  128 255 248 254 65535 1
Eth1/52  2c:4f:52:ff:fe:e1:7c:1a  128 255 248 254 65535 1

```

The output shows the master clocks that send grandmaster information to the switch and the switch's connected interface. The clock ID here is the closest master's ID. The ID is not the grandmaster's ID. Because this switch is receiving the grandmaster's data from two different ports, one of the ports became passive.

Showing the Counters

The following example shows the counters of a master port:

```
f2-leaf1# show ptp counters int e1/1
```

```
PTP Packet Counters of Interface Eth1/1:
```

```

-----
Packet Type          TX          RX
-----
Announce              4            0
Sync                  59           0
FollowUp              59           0
Delay Request         0            30
Delay Response        30           0
PDelay Request        0            0
PDelay Response       0            0
PDelay Followup       0            0
Management            0            0

```

A master port should send the following messages:

- Announce
- Sync
- FollowUp
- Delay Response

A master port should receive the following message:

- Delay Request

The following example shows the counters of a client port:

```
f2-leaf1# show ptp counters int e1/52
```

```
PTP Packet Counters of Interface Eth1/52:
```

```

-----
Packet Type          TX          RX
-----
Announce              0            4
Sync                  0            59
FollowUp              0            59
Delay Request         30           0
Delay Response        0            30
PDelay Request        0            0
PDelay Response       0            0
PDelay Followup       0            0
Management            0            0

```

The sent and received messages are the opposite of a master port. For example, if the Rx of `Delay Request` and the Tx of `Delay Response` are zero on a master port, the other side is not configured or not working as a client correctly since the client should initiate a `Delay Request` for the E2E delay mechanism.

In a real world, the counter information may not be as clean as the example, because the port state may have changed in the past. In such a case, clear the counters with the following command:

```
f2-leaf1# clear ptp counters all
```



Note The `PDelay_xxx` counter is for the P2P mechanism, which is not supported on Cisco Application Centric Infrastructure (ACI).
