Multicast (Group) Addresses for Layer 2 (Ethernet) Transport of IEEE 1588 PTP Messages, with Application to AVB

Geoffrey M. Garner
SAMSUNG Electronics (Consultant)

IEEE 802.1 AVB TG
IEEE 1588
2006.05.17
Revision 1

gmgarner@comcast.net

1

Outline

- Summary

- Introduction

- Terminology and Layering

- IEEE 1588 node types and desired forwarding behavior for each PTP message

- IEEE 802.1 reserved multicast address

- Generating and forwarding rules for PTP messages, for each PTP node type

Summary

❑ PTP messages in IEEE 802.1AS (AVB timing/synchronization) will be multicast

- All messages will be sent to a single reserved multicast address that is blocked in all 802.1 bridges
  - This address will be one of the addresses in Table 7-10 of 802.1D
- The AVB cloud will not contain any non-AVB bridges
- On arrival at a bridge, a PTP message will be handed to the PTP layer for processing
- An AVB node will send Pdelay_Req on all ports, including blocked ports
- An AVB node that receives Pdelay_Req will respond to it on the port that it arrived
- The Pdelay mechanism will be used to detect the presence of a non-AVB bridge (either by non-response or response with excessive delay)
- An AVB node will not forward (at either the PTP or Ethernet layer) any Pdelay messages
- Sync and Follow_Up messages will be processed at the PTP layer and forwarded as necessary
- Announce messages will be processed at the PTP layer via execution of the Best Master Clock (BMC) algorithm and PTP state machine

Summary (Cont.)

❑ IEEE 1588 should request an OUI for other applications of PTP over Layer 2 Ethernet that must work over non-1588 bridges

❑ IEEE 1588 applications of PTP over Layer 2 Ethernet can use the 802.1D Table 7-10 reserved multicast address for AVB if

- All bridges are 1588-enabled

- Only Pdelay messages are sent on blocked ports

- Note that such applications can use BCs and E2E TCs, along with Delay_Req and Delay_Resp messages, if desired

❑ Note that the material here does not preclude non-AVB applications from using unicast addressing; this presentation is simply limited to describing multicast addressing

Introduction

❑At present, all addresses in IEEE 1588 Version 2 will be multicast by default

- The only exception to this will be for communication within fault-tolerant master clusters, which will use unicast addresses
- Note that IEEE 1588 Version 1 uses only multicast addressing

❑Unicast addresses will be available for at least some messages as an option

❑The AVB synchronization network architecture is (see [1] for details)

- AVB networks will consist of a single Grandmaster (GM) that synchronizes slave clocks through Peer-to-Peer (P2P) Transparent Clocks (TCs)
  - The AVB cloud will be a single IEEE 1588 subdomain
- The GM and Slave Clocks will be Ordinary Clocks (OCs)
- Each OC (both GM and slaves) will be collocated with a P2P TC
  - I.e., an AVB node will consist of a P2P TC and collocated OC function, with an implied link between the two
- AVB networks will not use IEEE 1588 Boundary Clocks (BCs)

Introduction (Cont.)

❑ IEEE 1588 Precision Time Protocol (PTP) messages in AVB networks will be multicast

- With multicast addressing for Sync and Follow_Up, the GM only needs to send one Sync and one Follow_Up message at each sync interval (on the implied link between the GM and collocated P2P TC)

- If the addressing were unicast, the GM would have to send a separate Sync and Follow_Up to each slave clock, i.e., to each AVB network node
  - The GM would not scale well in this case with increasing network size

- In addition, each AVB network node will exchange Pdelay messages with the neighboring node on each link to measure propagation time on that link
  - Multicast addressing will simplify this process, as the same address can be used in every Pdelay message at every node
    - To make this work, we need to ensure that Pdelay messages are not forwarded (this is described more fully in the following slides; here we are only motivating why multicast

Terminology

❑ IEEE 1588 refers to the "forwarding" of various PTP messages (e.g., Sync, Follow_Up, etc.)

- Forwarding occurs after possibly altering the PTP payload
  - By "PTP payload," we mean the Ethernet payload, i.e., everything between (and not including) the Ethertype and FCS fields

❑ At the Ethernet (IEEE 802.1) layer, these messages are not forwarded by participating (i.e., by 1588 enabled or AVB enabled) bridges

- Rather, the Ethernet bridge blocks the message and passes it to the PTP layer
- The PTP layer Processes the message and, based on the processing, does one of several things
  - Sends one or more related messages of the same type on one or more ports
    - Related messages may essentially be the original message with one or more fields altered
      » E.g., may send Sync messages on several ports, with the PTP payload of each Sync message sent identical to the PTP payload of the received Sync message except for the addition of the residence time (for each respective message sent) to the correction field
  - Responds to the received message with a different PTP message
  - Does not send any messages

❑ In this presentation, we try to be clear in using the term "forwarding" whether it is at the Ethernet layer or the PTP layer

❑ Next slide illustrates processing and forwarding of PTP message at PTP layer

Processing and Forwarding PTP Messages at PTP Layer
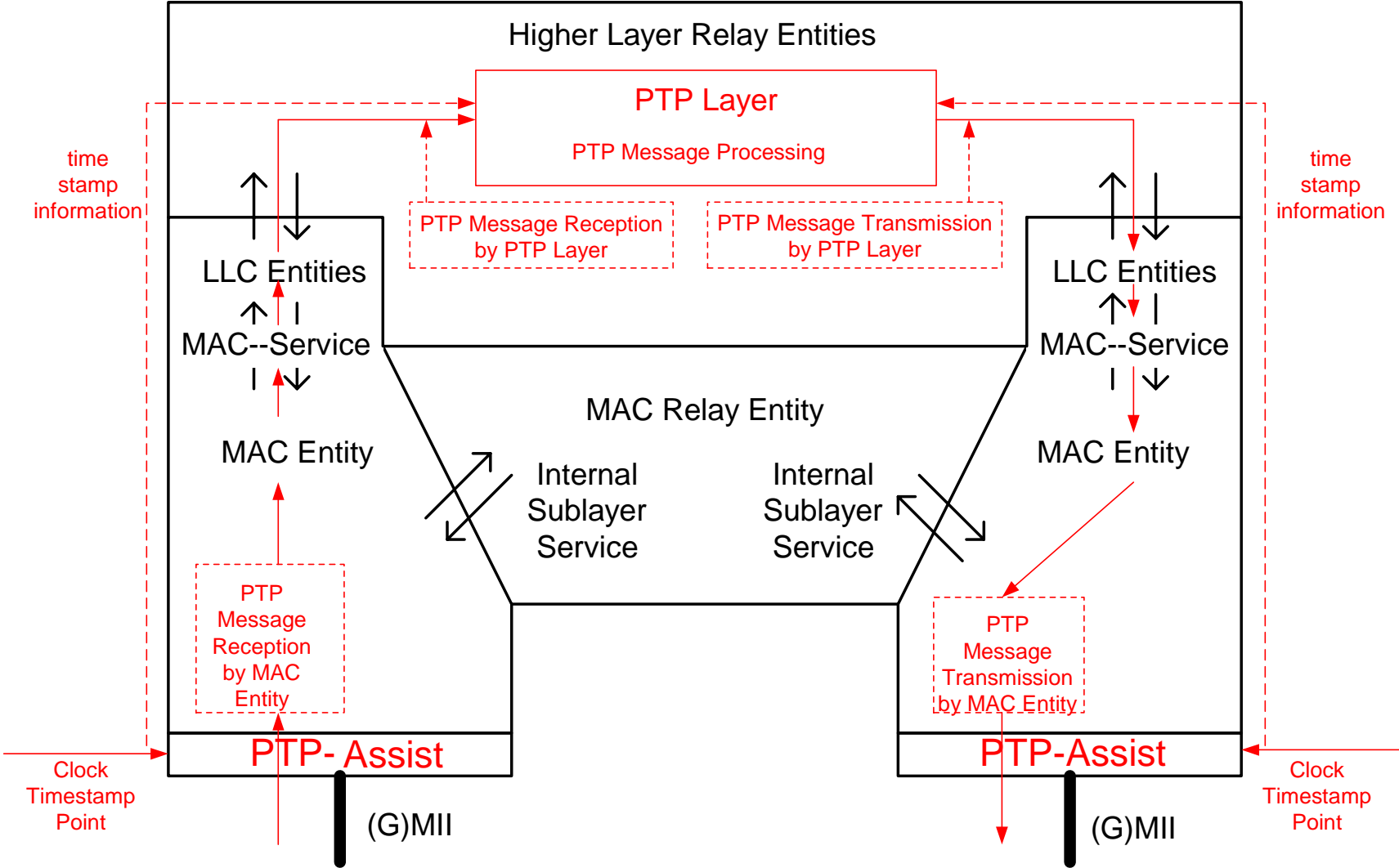
Node Types and Desired Forwarding Behavior

❑ **Types of PTP nodes**

- Ordinary clocks
- Boundary clocks (not used in AVB)
- Peer-to-Peer (P2P) transparent clocks
- End-to-End (E2E) transparent clocks (not used in AVB)

❑ **More precisely, the above actually are functions**

- An AVB node will consist of an OC and P2P TC function (see [1])

❑ **PTP messages**

- Sync
- Follow_Up
- Delay_Req (not used in AVB)
- Delay_Resp (not used in AVB)
- Announce
- Pdelay_Req
- Pdelay_Resp
- Pdelay_Resp_Follow_Up
- Transport
- Management messages

Node Types and Desired Forwarding Behavior

❑ BCs and OCs do not forward any of the above PTP messages, except for Management messages

- A BC or OC port that is in the master state sends Sync and, if follow-up capable, Follow_Up

- A BC or OC port that is in the slave state processes received Sync and Follow_Up messages

- A BC or OC port that is in the slave state send Delay_Req to the master BC or OC on the communication path for that port, provided there are no P2P TCs on the communication path

- A BC or OC port that is in the master state responds, with Delay_Resp, to Delay_Req received from each slave BC or OC on the communication path for that port

- BCs and OCs send Announce messages and process received Announce messages, in the form of executing the Best Master Clock (BMC) algorithm and PTP state machine

- BCs and OCs send Pdelay_Req messages on links that lead to P2P TCs

- BCs and OCs respond to Pdelay_Req messages with Pdelay_Resp and possibly Pdelay_Resp_Follow_Up messages, on links that lead to P2P TCs

Node Types and Desired Forwarding Behavior

❑ **E2E TCs forward all of the above PTP messages (except that they will not see any Pdelay messages and therefore won't have any to forward)**

- Depending on the message type and whether the E2E TC is on-the-fly or follow-up, the E2E TC may or may not alter the message

  - The alteration would generally be the addition of a residence time to a correction field, if needed (Announce, Transport, and Management messages are not altered)

  - See [2] for details

❑ **P2P TCs forward Sync, Follow_Up, Announce, Transport, and Management messages**

- Depending on the message type and whether the P2P TC is on-the-fly or follow-up, the P2P TC may or may not alter the message

  - The alteration would generally be the addition of a residence time to a correction field, if needed (Announce, Transport, and Management messages are not altered)

  - See [2] for general details, and [1] for details specific to AVB networks

❑ **P2P TCs do not forward Pdelay_Req, Pdelay_Resp, and Pdelay_Resp_Follow_Up messages**

Node Types and Desired Forwarding Behavior

❑P2P TCs will not process Delay_Req and Delay_Resp messages

▪Therefore, BCs and OCs attached to P2P TCs must send and respond to Pdelay messages

•AVB network nodes will process Pdelay messages, and AVB networks will not use Delay_Req and Delay_Resp

IEEE 802.1 Reserved Multicast Addresses

❑ The OUI 00-80-C2 is assigned to IEEE 802.1

- This includes unicast addresses (that begin with 00-80-C2) and multicast addresses (that begin with 01-80-C2)

❑ IEEE 802.1D-2004 defines in Table 7-10 Reserved Multicast Addresses 01-80-C2-00-00-00 through 01-80-C2-00-00-0F

- Frames sent to any of these addresses are never forwarded by bridges

❑ It was indicated in the May, 2006 interim 802.1 AVB TG meeting that an address from this table could be used by 802.1AS for PTP messages

❑ Any frames sent to this address will be blocked by a bridge and passed to the PTP layer for processing

- Only need to use one reserved address, because Ethernet layer simply recognizes the address and passes the message to the PTP layer; PTP layer does all the processing

❑ Since 802.1AS will require all bridges to be participating, i.e., synchronization will not be transported over non-AVB bridges, this scheme will work

- If a PTP message is sent to a non-participating bridge, there will be no PTP layer present in that bridge and the message will be dropped

IEEE 802.1 Reserved Multicast Addresses

❑ If non-participating (i.e., non-1588) bridges are present, then the PTP messages should use a reserved address with forwarding behavior analogous to those of the GARP Application Addresses in 802.1D, Table 12-1

❑ A single unused address can be taken from one of the addresses owned by IEEE 802.1 (taken from the 00-80-C2 OUI, i.e., beginning with 01-80-C2) and assigned to be used as a "PTP Application Address"

❑ If a bridge is participating (i.e., 1588 enabled, which includes AVB-enabled), it should not forward (at the Ethernet layer) frames sent to this address

  ▪ Instead, the frames will be passed to the PTP layer

❑ If a bridge is not participating (i.e., not 1588 enabled, which means not AVB enabled), it should forward frames sent to this address on all non-blocked ports excluding the port the frame arrived on

  ▪ If there are no non-blocked ports other than the port the frame arrived on, the frame is dropped

  ▪ This ensures that the PTP frames reach 1588 enabled bridges when non-1588 enabled bridges are present

❑ Only need one reserved address, because Ethernet layer simply recognizes the address and either forwards the frame or passes the message to the PTP layer; PTP layer does all the processing

Forwarding Rules for PTP Messages

❑ BCs and OCs will operate as indicated on Slide 10; they do not forward (at the PTP layer) any of the above PTP messages except for Management Messages

- A BC or OC sends (i.e., originates) Sync and, if follow-up capable, Follow_Up on each port that is in the master state
  - This means that Sync and Follow_Up are not sent on ports that are:
    - Blocked at the Ethernet layer by STP (or any equivalent protocol used at the Ethernet layer to ensure there are no loops)
    - Not blocked at the Ethernet layer but are in the slave state (note that an OC has only one port, and therefore the OC is either a master or a slave)
    - Not blocked at the Ethernet layer but in the PTP-passive state; this means the port is blocked at the PTP layer
  - The PTP Best Master Clock (BMC) algorithm is able to produce a spanning tree for the BC and OC (but not TC) nodes at the PTP layer
    - For layer 2 Ethernet operation, a spanning tree may already be produced by STP or an equivalent protocol at layer 2
    - This issue is not relevant for AVB, because AVB will not use BCs
- A BC or OC processes Sync and Follow_Up messages received on the single port that is in the slave state (and uses the information to synchronize to its master)
  - Sync and Follow_Up messages will not be received on ports blocked at the Ethernet layer, nor ports that are in the PTP-passive state, i.e., blocked at the PTP layer

Forwarding Rules for PTP Messages

❑BCs and OCs (Cont.)

- A BC or OC sends Delay_Req to the master BC or OC on its single port that is in the slave state, provided there are no P2P TCs on the communication path
  - Delay_Req is not sent on ports blocked at the Ethernet layer, nor ports in the PTP-passive state or master state

- A BC or OC responds, with Delay_Resp, to Delay_Req received from each slave BC or OC on each port that is in the master state
  - Delay_Resp is not sent on ports blocked at the Ethernet layer, nor ports in the PTP-passive state nor slave state

- BCs and OCs send Announce messages and process received Announce messages, in the form of executing the Best Master Clock (BMC) algorithm and the PTP state machine
  - Announce messages are not sent and not received on ports blocked at the Ethernet layer

Forwarding Rules for PTP Messages

❑BCs and OCs (Cont.)

- BCs and OCs send Pdelay_Req messages on links that lead to P2P TCs

  - Pdelay_Req messages are sent on all ports that lead to P2P TCs, regardless of whether they are in the master, slave, or PTP-passive states, or whether they are blocked at the Ethernet layer

  - This is because propagation delay must be measured on all links and made available in the event of reconfiguration

- BCs and OCs respond to Pdelay_Req messages with Pdelay_Resp and possibly Pdelay_Resp_Follow_Up messages, on links that lead to P2P TCs

  - As above, Pdelay_Resp messages are received on all ports that lead to P2P TCs, regardless of whether they are in the master, slave, or PTP-passive states, or whether they are blocked at the Ethernet layer

  - This is because propagation delay must be measured on all links and made available in the event of reconfiguration

- As indicated earlier, there will be no E2E TCs, nor non-1588 bridges, between a BC or OC and a P2P TC

Forwarding Rules for PTP Messages

❑ P2P and E2E TCs will process received Sync and Follow_Up messages in accordance with the PTP protocol, i.e., they will measure residence time and update the respective correction field

❑ They will then send the Sync and Follow_Up messages with the correction fields updated as needed on all respective ports that are not blocked by STP (or any equivalent protocol used at the Ethernet layer to ensure there are no loops), except the port the message arrived on

❑ E2E TCs will also process Delay_Req and Delay_Resp by measuring residence time and updating the respective correction field, and then send the possibly altered Delay_Req and Delay_Resp messages on non-blocked ports except the port the message arrived on

❑ P2P TCs will not receive and therefore not process Delay_Req and Delay_Resp

Forwarding Rules for PTP Messages

❑ P2P TCs will generate and send Pdelay_Req on all ports, including blocked ports

- ▪ Pdelay messages are sent on blocked ports so that propagation delays can be measured on these links and made available in the event of reconfiguration
- ▪ As indicated earlier, there will be no E2E TCs, nor non-1588 bridges, between two P2P TCs

❑ P2P TCs will respond to Pdelay_Req only on the port that the Pdelay_Req arrived on; the response will be with Pdelay_Resp and possibly Pdelay_Resp_Follow_Up

❑ E2E TCs will not receive, and therefore not process, Pdelay_Req, Pdelay_Resp, and Pdelay_Resp_Follow_Up messages

Multicast Addresses for Non-AVB Applications of PTP over Layer 2 Ethernet

❑ Non-AVB applications of IEEE 1588 over Layer 2 Ethernet can use the same Table 7-10 Reserved Multicast Address, provided

- All bridges are 1588-enabled

- Only Pdelay messages are sent on blocked ports

- Note that such applications can use BCs and E2E TCs, along with Delay_Req and Delay_Resp messages, if desired

❑ IEEE 1588 should request an OUI for other applications of PTP over Layer 2 Ethernet that must work over non-1588 bridges

- One or more Application Multicast Addresses can be defined within this OUI for use by non-AVB applications of PTP over Layer 2 Ethernet

- Care should be taken to avoid looping in such applications if Pdelay messages are sent on blocked ports to non-1588 bridges

❑ Note that the material here does not preclude non-AVB applications from using unicast addressing; this presentation is simply limited to describing multicast addressing

Detection of a non-AVB bridge by an AVB Bridge

❑At startup, an AVB bridge will send Pdelay_Req on all ports

❑If Pdelay_Resp is not received, the AVB bridge will conclude that the bridge at the other end of the link is not AVB-enabled, and will not send PTP messages on this link

❑If Pdelay_Resp is received after a time interval that exceeds a threshold to be specified in 802.1AS, the bridge will conclude that there are one or more non-AVB bridges between it and the bridge that sends Pdelay_Resp

  ▪The bridge will not send PTP messages on this link

References

1. Geoffrey M. Garner, *Description of Use of IEEE 1588 Peer-to-Peer Transparent Clocks in A/V Bridging Networks*, Revision 1.1, April 24, 2006.

2. *Transparent Clock – Working Technical Description, Revision 13*, IEEE 1588 Transparent Clock Subcommittee, February 17, 2006.