ISPCS Portsmouth 2010

Using an IEEE 802.1AS Network as a Distributed IEEE 1588 Boundary, Ordinary, or Transparent Clock

Geoffrey M. Garner

Samsung Advanced Institute of Technology (SAIT), Consultant

Michel Ouellette

Huawei Technologies Co., Ltd.

Michael Johas Teener

Broadcom Corporation

Outline

- Introduction
- IEEE 802.1AS aspects needed here
- Equivalence of BC and Peer-to-Peer TC in transporting synchronization
- Distributed BC, OC, and Peer-to-Peer TC
- Examples
- Summary

Introduction – 1

ISPCS Portsmouth 2010

- IEEE 802.1AS is the Audio/Video Bridging (AVB) standard that specifies distribution of precise timing in an AVB network
  - One of a set of AVB standards to support the transport of time-sensitive applications over IEEE 802 bridged LANs
- IEEE 802.1AS includes a PTP profile that specifies timing transport over full-duplex, IEEE 802.3 links (Annex F of IEEE 1588)
  - 802.1AS time-aware systems (i.e., nodes) whose interfaces are full-duplex Ethernet are 1588 boundary clocks (BCs) or ordinary clocks (OCs)
    - Will be shown later that former is equivalent to 1588 peer-to-peer transparent clock (TC) in transporting synchronization

Introduction – 2

- IEEE 802.1AS specifies transport over other media
  - Not part of PTP profile, since these media are not described in IEEE 1588
- These other media include
  - IEEE 802.11 (WiFi)
  - IEEE 802.3 Ethernet Passive Optical Network (EPON)
  - Coordinate Shared Network (CSN)
    - E.g., Multimedia over Coax (MoCA), ITU-T G.9960/G.9961 (ex. G.hn)

Introduction – 3

- Synchronization transport over these other media uses timing facilities already defined (or being defined) in the standards for these media
  - Timing transport over 802.11 uses IEEE 802.11v (facilities being developed for location determination)
  - Timing transport over IEEE 802.3 EPON uses Multipoint Control (MPCP) of IEEE 802.3, clauses 64 and 77
  - Timing transport over CSN can use CSN-specific timing
- Since synchronization transport over the above media are not part of the 802.1AS PTP profile, a time-aware system that contains interfaces to those media does not, strictly speaking, act as a 1588 BC relative to those interfaces

Introduction – 4

ISPCS Portsmouth 2010

- However, it is possible to consider a network portion of an 802.1AS network where
  a) all links of the network portion use media that are not part of the PTP profile
  b) at least some node ports of this network portion are attached to links that are part of the PTP profile, and
  c) the network portion is *time-aware*, i.e., it has available a common source of time that is, in general, independent of the 802.1AS network clock and can be used for timestamping PTP messages at the network portion ingress and egress (i.e., at the ports of (b))
- This network portion can be considered a distributed IEEE 1588 clock

Introduction – 5

- The concept of a distributed IEEE 1588 BC and TC was introduced in [1] for the case where the non-PTP profile transport is GPON
- The concept is extended here to general time-aware subnetworks that are not part of a PTP profile
  - The functional equivalence of distributed and non-distributed clocks is shown
- The functional equivalence, with respect to synchronization, of a BC that uses the peer delay mechanism and a peer-to-peer TC, first described in [2], is also shown
  - The main difference between a BC and peer-to-peer TC is that the former invokes BMCA and the latter does not
- The distributed BC/TC concepts and the BC/peer-to-peer TC equivalence may be considered a new way of looking at synchronization transport in a network based on IEEE 1588

IEEE 802.1AS Aspects Needed Here – 1

- All network nodes are time-aware (i.e., meet the 802.1AS requirements)
- All clocks are two-step
- Each time-aware system syntonizes to the GM by measuring nearest neighbor rate ratio on every link and accumulating the GM rate ratio in a standards organization TLV attached to Follow_Up
  - For links that are part of the PTP profile (i.e., full-duplex IEEE 802.3), neighbor rate ratio is measured using Pdelay_Resp and Pdelay_Resp_Follow_Up messages
  - For other links, media-specific methods are used to measure neighbor rate ratio
  - Neighbor rate ratio is 1 on media for which the endpoints are syntonized via the physical layer, e.g., IEEE 802.3 EPON

IEEE 802.1AS Aspects Needed Here – 2

- Physical adjustment of the local oscillator frequency is not required (but is not prohibited)
  - Instead, synchronized (i.e., GM) time corresponding to a desired local time is computed using the measured GM rate ratio to convert a time interval value relative to the local oscillator to a value relative to the GM
  - This is described in more detail in the following slides (slides 11 – 15)
- This method of syntonization has two advantages
  - Fast convergence when GM changes, because neighbor rate ratios are measured persistently
  - Minimal gain peaking, because an error in one neighbor rate ratio measurement does not affect another neighbor rate ratio measurement

IEEE 802.1AS Aspects Needed Here – 3

- All nodes are required to participate in best master selection
  - However, a node is not required to be grandmaster capable
    - An OC that is not grandmaster-capable is a slave-only OC
    - A BC that is not grandmaster capable has exactly one slave port, unless no nodes are GM-capable
      - If no nodes are GM-capable, then all the nodes free-run
- See [3] and [4] for a more detailed description of IEEE 802.1AS

IEEE 802.1AS Aspects Needed Here – 4

ISPCS Portsmouth 2010

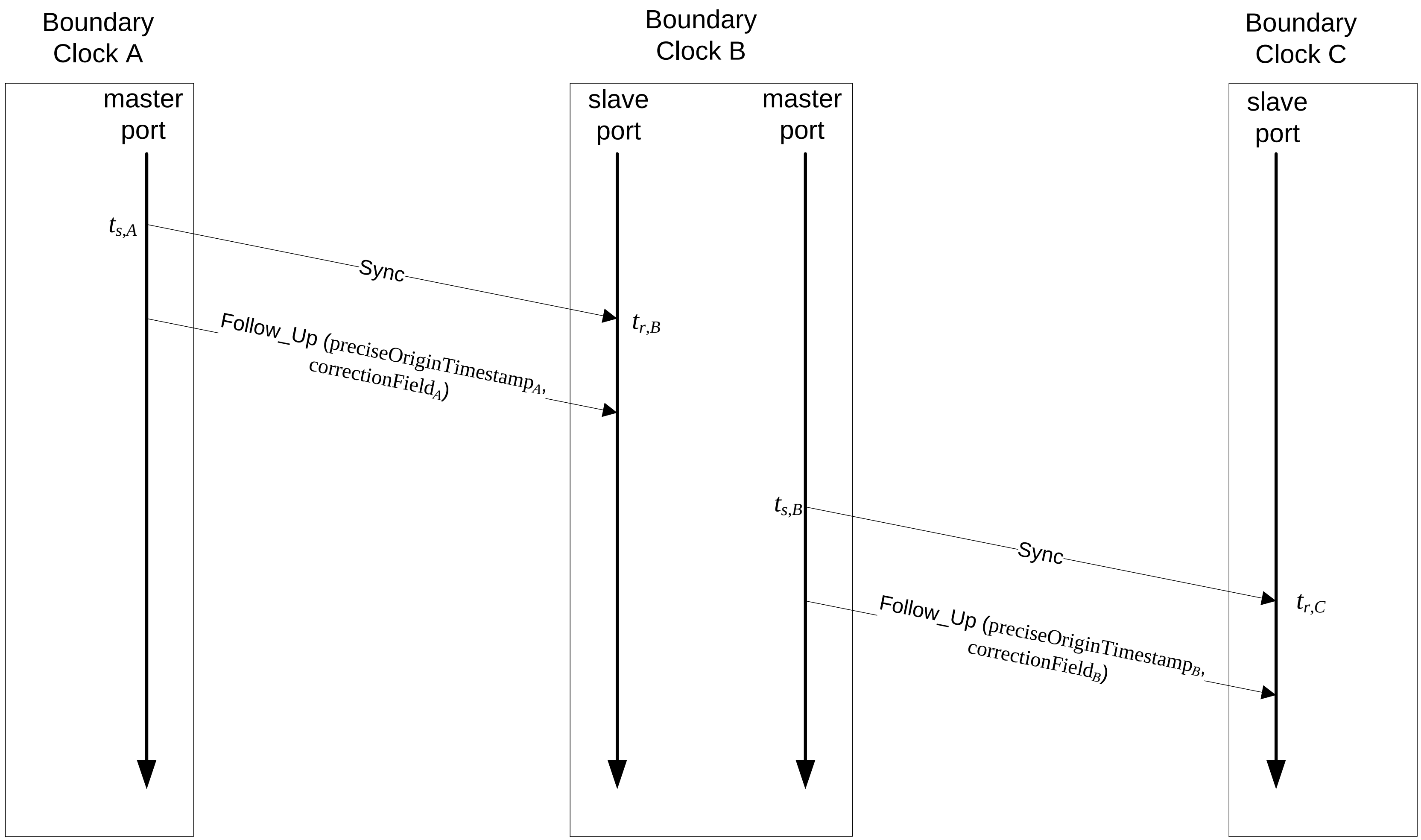
- All the examples here use two-step clocks, but they are easily adapted to the one-step case by replacing the preciseOriginTimestamp by the originTimestamp, and carrying any TLVs by Sync (or, if desired, Signaling messages) instead of Follow_Up
  - Note that it is assumed above that all nodes are time-aware, i.e., perform timestamping
  - It is also assumed that propagation delay is measured using the peer delay mechanism
  - This means that attaching a TLV to Sync will not cause delay asymmetry

BC/Peer-to-Peer TC Functional Equivalence – 1

- Consider 3 IEEE 1588 BCs (A, B, and C on next slide)
- Assume the BCs measure propagation delay using the peer-delay mechanism
- Assume the BCs syntonize to the GM by measuring GM rate ratio (but how they do it is up to the implementation)
  - e.g., can use successive Sync/Follow_Up messages
  - e.g., can measure neighbor rate ratio and accumulate GM rate ratio in TLV
- B uses the Sync/Follow_Up information received from A, and the measured GM rate ratio and propagation delay, to synchronize to A
  - It does this by computing the time relative to the GM that corresponds to any desired time relative to the local oscillator

BC/Peer-to-Peer TC Functional Equivalence – 2

ISPCS Portsmouth 2010



BC/Peer-to-Peer TC Functional Equivalence – 3

ISPCS Portsmouth 2010

- The grandmaster time when the local time at B is t is the sum of
  - The GM time when the most recent Sync message was received
    - This is the sum of the preciseOriginTimestamp, correctionField, and propagation delay on the upstream link
  - The elapsed local time since the most recent Sync message was received, multiplied by the GM rate ratio

$$T_{GM,B}(t) = \text{preciseOriginTimestamp}_A + \text{correctionField}_A + D_{A,B} + R \cdot (t - t_{r,B}) \quad \text{(Eq. 1)}$$

- $D_{A,B}$ = measured propagation delay
- R = measured GM rate ratio
- $T_{GM,B}(t)$ = GM time corresponding to local time t at B

BC/Peer-to-Peer TC Functional Equivalence – 4

- Next, consider the case where B is a peer-to-peer TC
- When B sends Sync and Follow_Up to C, the correctionField is set equal to the sum of
  - The correctionField of the Follow_Up message received from A
  - The measured Propagation delay on the link between A and B
  - The residence time, which is equal to the elapsed local time between the receipt of Sync from A and the sending of Sync to C multiplied by the GM rate ratio
- The sum of the preciseOriginTimestamp and correctionField sent to C is

$$T_{GM,B}(t_{s,B}) = \text{preciseOriginTimestamp}_A + \text{correctionField}_A + D_{A,B} + R \cdot (t_{s,B} - t_{r,B}) \quad \text{(Eq. 2)}$$

- This is the same as the previous equation, evaluated at the time the Sync message is sent from B to C

BC/Peer-to-Peer TC Functional Equivalence – 5

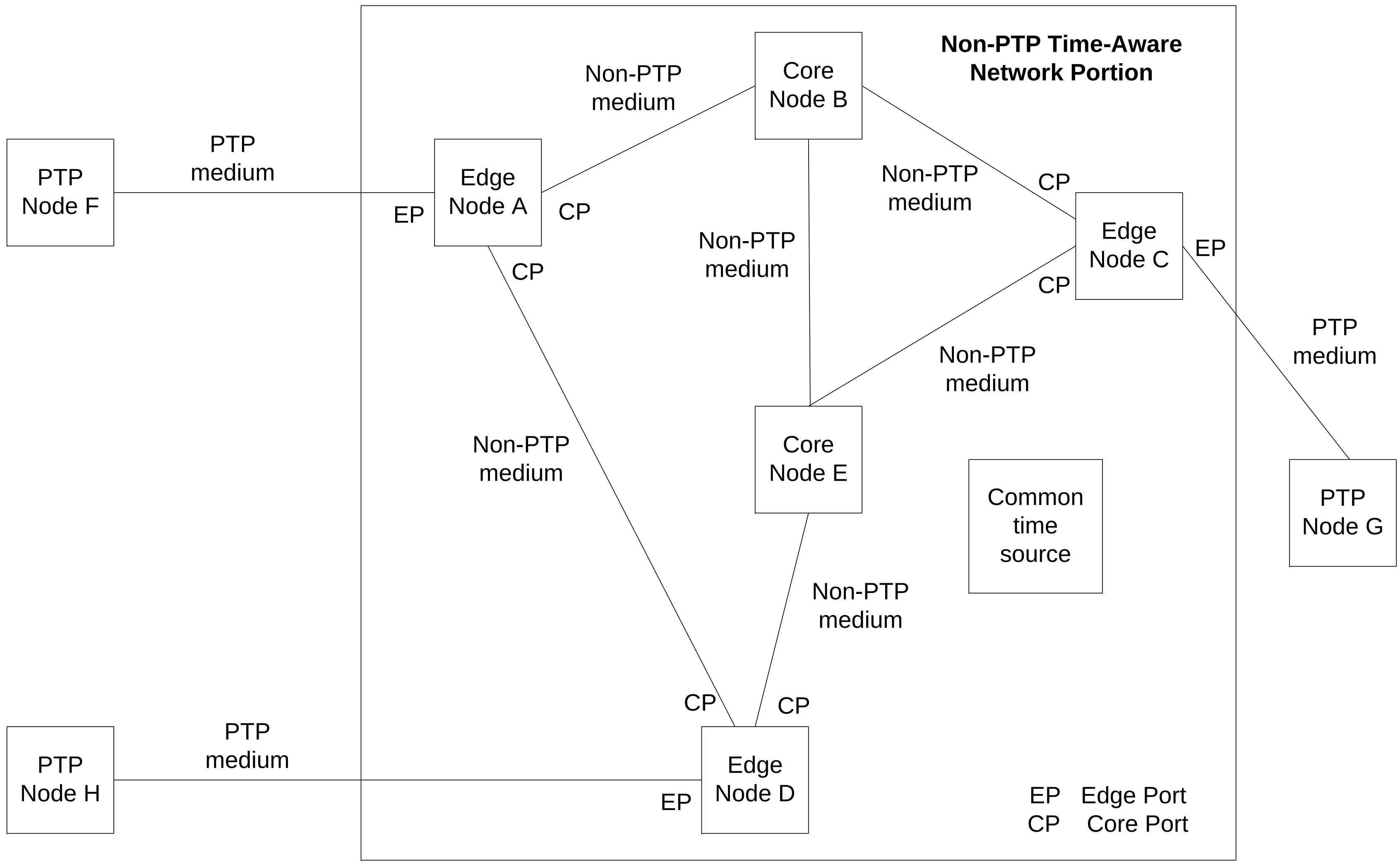
ISPCS Portsmouth 2010

- The above shows that a BC and peer-to-peer TC functionally perform the same computations
  - The only difference is in how the synchronized time (time relative to the GM) is divided between the preciseOriginTimestamp and correctionField
- Note that the above computations are still equivalent if the BC and TC do not syntonize
  - In this case, rate ratio is not measured, and $R = 1$ in the equations
- Ordinarily, a TC does not provide synchronized time at an arbitrary time t, but only computes the correctionField value at the time Sync is sent
  - A TC with the added function of computing synchronized time at an arbitrary time t may be considered to be a TC plus OC synchronization function

BC/Peer-to-Peer TC Functional Equivalence – 6

- Note that a BC and TC implementation or simulation will give the same phase error accumulation performance if the various design aspects are the same
  - Use/non-use of PLL filtering and, in former case, same PLL parameters (e.g., bandwidth, gain peaking)
  - Local oscillator noise generation
  - Sync interval
  - Residence time
  - Pdelay turnaround time
  - Measurement of rate ratio and algorithm used
  - One-step/two-step clock
- The primary functional difference between a BC and peer-to-peer TC is that the former invokes BMCA, while the latter does not

Distributed Clocks – 1



Distributed Clocks – 2

- A non-PTP time-aware network portion may be considered a distributed clock
  - *Time-aware* means that the network portion technology provides a common source of time that can be used to timestamp incoming and outgoing PTP event messages
    - The common source of time is, in general, independent of the PTP GM
  - Nodes completely within the non-PTP network portion are *core nodes*
  - Nodes that have both links that are not part of the PTP profile and links that are part of the profile are *edge nodes*
    - Ports attached to PTP profile links are *edge ports*
    - Ports attached to non-PTP profile links are *core ports*
  - Note that edge/core terminology did not originate here

Distributed Clocks – 3

ISPCS Portsmouth 2010

- The non-PTP network portion must transport the following information from an ingress edge port to each egress edge port
  - preciseOriginTimestamp
  - correctionField
  - GM rate ratio at ingress
  - Propagation delay on ingress link
  - Sync ingress timestamp, relative to the non-PTP network portion common time source, of the arriving Sync message at the ingress edge port
- With the above information, Eq. (1) (slide 13) may be used to compute the synchronized time (GM time) corresponding to any local time, at any edge or core node

Distributed Clocks – 4

ISPCS Portsmouth 2010

- The local time, *t*, is the time relative to the non-PTP network portion common time source
- The synchronized time may be provided to any higher-layer application at any edge or core node, as needed
- Eq. (2) may be used to compute the correctionField and/or preciseOriginTimestamp at an egress edge port
  - As indicated previously, the distributed clock may be considered functionally a distributed BC or TC with respect to synchronization
    - The only difference is in how the synchronized time is distributed between the preciseOriginTimestamp and correctionField

Distributed Clocks – 5

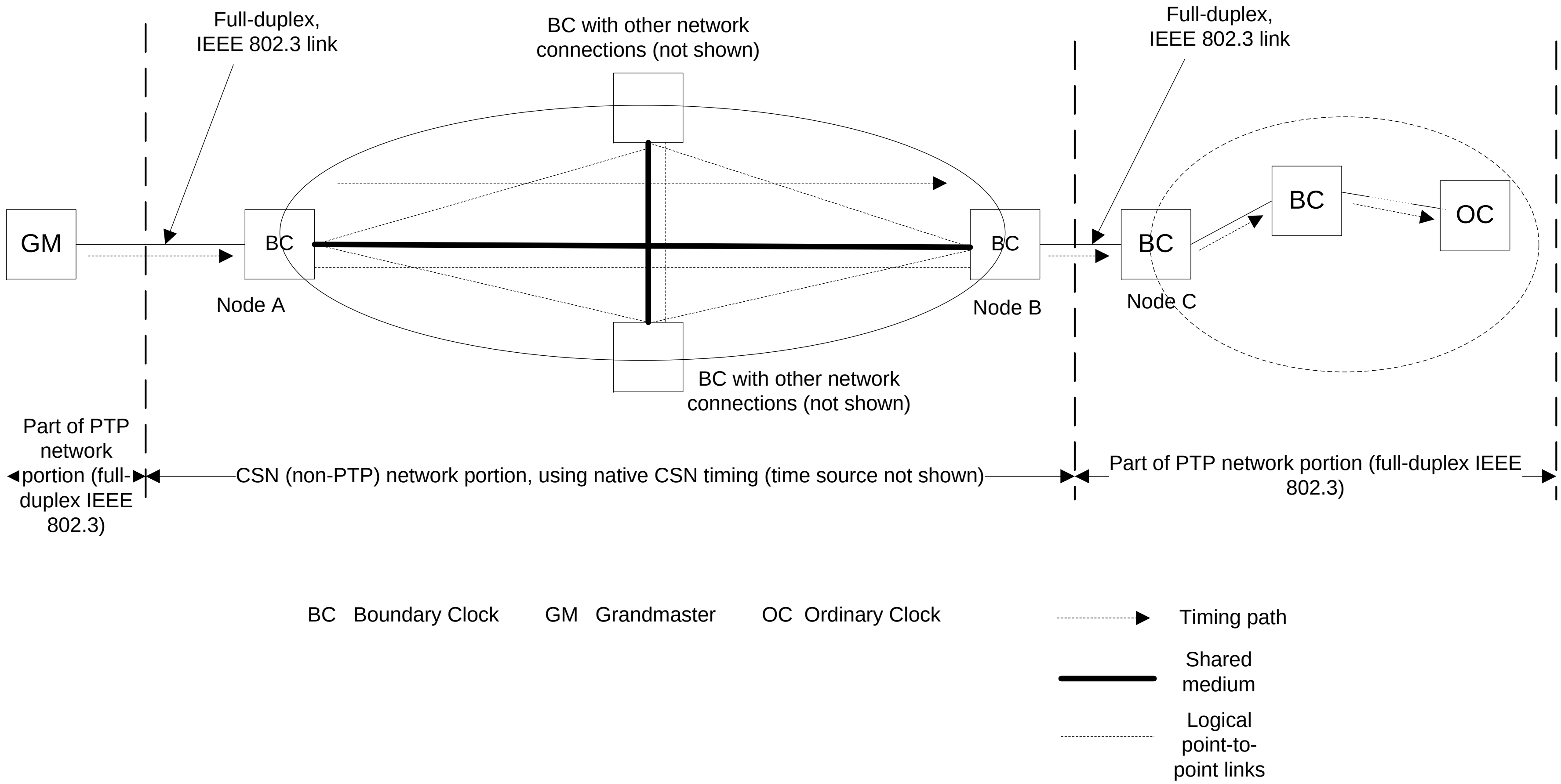
- The non-PTP network portion is a distributed BC if it participates in best master selection and sets the states of the edge ports
  - The exact implementation of BMCA in the non-PTP network portion is not specified by PTP
  - Note that the best master selection function refers to the PTP grandmaster selection, and not the selection of the master to serve as the common time source of the non-PTP portion
    - The selection of this common time source is outside PTP
- If the non-PTP network portion does not participate in BMCA, it is a distributed peer-to-peer TC
- If the non-PTP network portion has only one edge port, it is a distributed OC

Distributed Clocks – 6

- The above ideas hold if the non-PTP network portion is replaced by a PTP network portion whose profile is different from that of the existing PTP network portion
  - The profile must include provision to transport the information on slide 19, e.g., using a TLV
  - The GM of the PTP network portion of the different profile is independent of the GM of the rest of the network (i.e., the network portion also forms its own domain)

Example 1

IEEE 802.1AS network with a CSN portion acting as a distributed BC



Example 1

- CSN acts as a distributed BC
- CSN physical layer is a shared bus
  - Therefore, all CSN nodes are edge nodes, with a single core port (and act as BCs with respect to rest of IEEE 802.1AS network)
  - CSN native time source, and its distribution, is outside PTP and not shown
- 802.1AS GM sends Sync, Follow_Up, and Announce to Node A
  - Node A invokes BMCA on receipt of Announce, and on completion places each other port, including the core port, in appropriate PTP state
  - Node A transmits Announce to all other CSN edge nodes
    - On receipt of Announce message on a core port, each edge node invokes BMCA, places other edge nodes in master or passive state, and transmits Announce on ports in master state

Example 1

ISPCS Portsmouth 2010

- Node A transmits Sync and Follow_Up messages to Node B (and to other CSN edge nodes)
- The Sync message is timestamped on transmission relative to the CSN time source
- The timestamp value, propagation delay on the link between node A and the GM, and grandmaster rate ratio are transported to node B in a TLV attached to Follow_Up
  - This is a simplified description; full details are given in IEEE 802.1AS
- After receipt of Sync and Follow_Up, node B transmits Sync to node C and timestamps the Sync transmission
- TLV is removed at node B

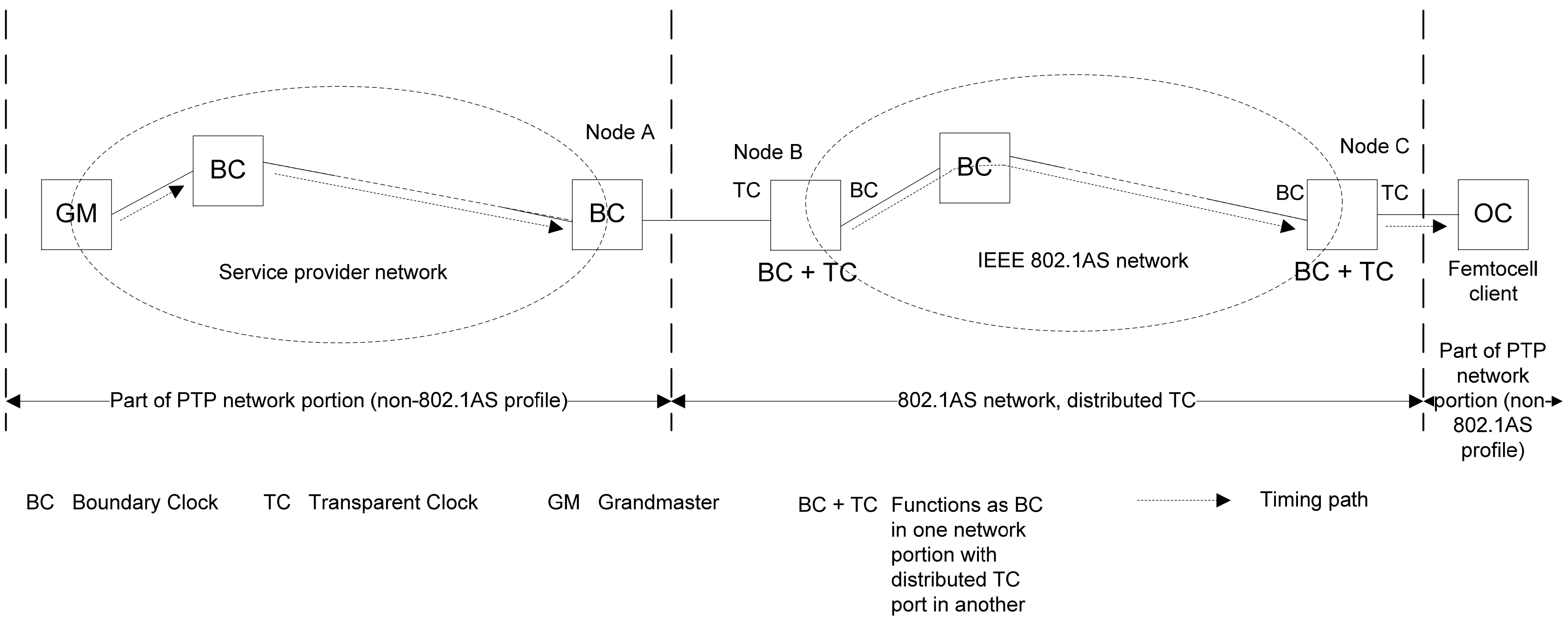
Example 1

ISPCS Portsmouth 2010

- Node B transmits a Follow_Up message whose preciseOriginTimestamp is equal to that in the received Follow_Up message, and correctionField set equal to the sum of
  - correctionField of received Follow_Up message
  - Received propagation delay on link between GM and node A
  - Timestamp for sending of Sync by node B minus timestamp for receipt of Sync by node A, multiplied by received GM rate ratio
    - Note that the CSN is syntonized at the physical layer, so the neighbor rate ratio of node B relative to node A is 1
- The sum of the preciseOriginTimestamp and correctionField are equal to the GM time when node B transmits the Sync, in accordance with Eqs. 1 and 2

Example 2

IEEE 802.1AS network acting as a distributed TC for synchronization transport using non-802.1AS PTP profile



Example 2

- Illustration of synchronization transport to Femtocell client from service provider network, over IEEE 802.1AS network that acts as distributed TC
- Service provider network uses non-802.1AS PTP profile
  - In addition, service provider and 802.1AS networks are different domains, and each network has its own GM
  - Within the 802.1AS network, the nodes act as BCs (and invoke BMCA with respect to their domain
- Sync messages received at node B are timestamped relative to the 802.1AS network time (traceable to the 802.1AS GM)

Example 2

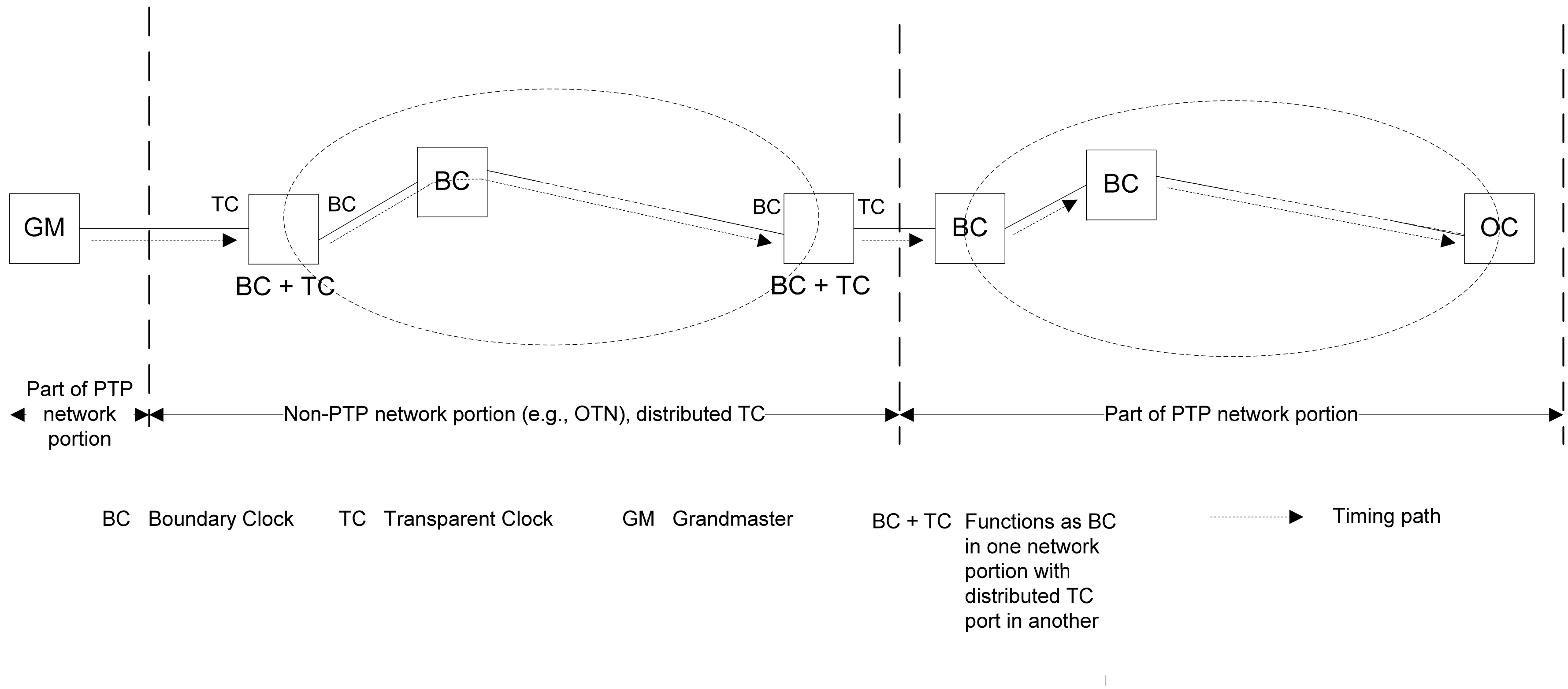
- 802.1AS network must transport the time synchronization information to node C
  - Sync ingress timestamp at node B relative to 802.1AS GM
  - Measured propagation delay on link between nodes A and B
  - IEEE 1588 GM rate ratio measured at node A
  - preciseOriginTimestamp and correctionField of Follow_Up message received at node B
- The method of transporting this information across the 802.1AS network is not specified in either 802.1AS or IEEE 1588

Example 2

- One method would be to follow the approach used for CSN
  - Transport the 1588 Sync and Follow_Up, and carry the received preciseOriginTimestamp and correctionField at node A
  - Transport the Sync ingress timestamp, propagation delay on the link attached to node A, and 1588 grandmaster rate ratio in a TLV attached to Follow_Up
  - Node C has all the necessary information to compute the preciseOriginTimestamp and correctionField according to Eqs. 1 and 2 for Sync and Follow_Up transmitted to Femtocell client
  - Node C removes the TLV
- The additions to the 802.1AS PTP profile would also have to specify the addressing for the transported Sync and Follow_Up, and ensure that they are not processed within the 802.1AS network

Example 3

PTP network with arbitrary profile, and an arbitrary non-PTP network portion acting as a distributed TC



Example 3

- PTP network with arbitrary profile, and arbitrary non-PTP network acting as a distributed TC
  - For example, the non-PTP network could be an Optical Transport Network (OTN), with time transported using the PTP protocol and the time synchronization and best master selection information carried in OTN overhead
  - The OTN network is referred to as “non-PTP” because there is no transport-specific annex in IEEE 1588 for OTN
- As in Example 2, the received Sync from the PTP GM is timestamped at the non-PTP network ingress relative to the OTN time source
- The received Sync, and associated Follow_Up, are transported over the non-PTP network

Example 3

- The Sync ingress time (relative to the non-PTP network time source), propagation delay on the upstream link to the GM, and GM rate ratio are transported over the non-PTP network in a manner not specified by PTP
  - For example, the information could be transported in a TLV as in the previous examples
- The non-PTP network egress node transmits and timestamps Sync, and has all the information needed to compute the preciseOriginTimestamp and correctionField of the associated Follow_Up message

Example 3

- The non-PTP network transmits ingress Announce messages to the egress edge nodes
- Since the non-PTP network acts as a distributed TC, the edge nodes do not process the PTP network Announce messages and do not invoke BMCA
- The non-PTP network does perform its own best master selection outside of PTP

Summary – 1

ISPCS Portsmouth 2010

- This paper has described how a network with a common source of time can be described as a distributed BC or TC
- When such a network is observed at its edge ports, its behavior is functionally the same as that of a non-distributed BC or TC
- The paper also has described how a BC and peer-to-peer TC are functionally equivalent in how they transport synchronization
  - From the standpoint of synchronization, the main difference is in the distribution of the network time between the preciseOriginTimestamp (or originTimestamp if the clock is one-step) and correctionField

Summary – 2

- The main difference between a BC and peer-to-peer TC is that the former invokes BMCA and the latter does not
  - A network of BCs will create the synchronization hierarchy, i.e., synchronization spanning tree
  - A network of TCs requires that the forwarding/routing of synchronization information be determined separately

References – 1

1. Richard Goodson, Hal Roberts, and Tim Frost, *Distributed Boundary and Transparent Clocks*, contribution to ITU-T SG 15, Q13, COM 15-C811-E, May, 2010
2. Geoffrey M. Garner, Michel Ouellette, and Wei Jianying, *Equivalence of the IEEE 1588 Boundary Clock and Peer-to-Peer Transparent Clock for Synchronization*, contribution to ITU-T SG 15, Q13, COM 15-C1001-E, May, 2010
3. Geoffrey M. Garner, Aaron Gelter, and Michael Johas Teener, *New Simulation and Test Results for IEEE 802.1AS Timing Performance,* ISPCS '09, Brescia, Italy, October 12 – 16, 2009.

References – 2

4. Geoffrey M. Garner and Hyunsurk (Eric) Ryu, *Synchronization of Audio/Video Bridging Networks using IEEE 802.1AS, submitted to IEEE* Communications Magazine (being reviewed for acceptance).