IEEE 802.1ASbt Timing and Synchronization

Discussion of Proposals for Redundancy in 802.1ASbt

IEEE 802.1 Interim Meeting - Sept. 2014, Ottawa, Canada
Feng Chen, Franz-Josef Goetz - Siemens AG
Geoffrey M. Garner

SIEMENS

Overview

- This presentation focuses on two things:
  - explain why redundant time synchronization should be considered within one domain *in opposition to* the proposal in 1588 of handling redundancy using multiple domains
  - describe how redundancy could be specified in 802.1AS to meet industrial requirements with our proposals

- The purposes of this presentation
  - to ask for your opinions on the statements/proposals in this presentation
  - to let the group make decision on some fundamental issues so that we can really move forward with this topic

Why One Domain is Preferable to Multiple Domains for Redundant Time Synchronization (1)

SIEMENS

1. Domain is used to separate different time-scales, while redundancy is always considered within the same time-scale.

- One of major uses of “Domain” is to separate different time-scales, e.g. the universal time and working clocks. Each domain defines an independent area containing a set of time-aware systems running gPTP on the same time-scale. One time-aware system can support multiple domains by running one gPTP instance for each supported time-scale.

- Redundant time synchronization is realized through duplication of timing information including synchronized time and allows each involved time-aware system to receive multiple copies of the synchronized time that are distributed by one GM or redundant GMs and transmitted over separate paths.

- A general requirement for redundant time synchronization is that all involved time-aware systems must be on the same time-scale. For example, one time-aware system being synchronized to two GMs simultaneously but on different time-scales will not be considered to be a redundancy case.

Why One Domain is Preferable to Multiple Domains for Redundant Time Synchronization (2)

SIEMENS

2. Definition of "redundancy area"

- Redundancy as optional features of gPTP is typically applied to a set of time-aware systems within a specific area of the whole network that require high-available time synchronization.

- To specify the scope of redundancy, we need to define a new term "**redundancy area**". Since redundancy is always considered for the same scale (as stated in point 1), a redundancy area contains a set of time-aware systems running gPTP with redundancy on the same time-scale, including all involved redundant GMs and sync trees.

- Considering that one domain defines one time-scale, we can find that a redundancy area on a specific time-scale has the same scope of the domain for that time-scale. From this perspective, redundancy area can be identified by domain. In other words, redundancy on one time-scale should be handled within the same domain of that time-scale.

Why One Domain is Preferable to Multiple Domains for Redundant Time Synchronization (3)

SIEMENS

3. Identifying a redundancy area is required by the redundant path computation algorithm.

- As stated before, all time-aware systems of the same redundancy area can be identified using the same domain number. This is required by the redundant path computation algorithm, which can be explained as follows:
- In the computation of redundant sync trees, e.g. one GM with two maximally disjoint/redundant trees (MDT), the MDT are computed based on the knowledge of the topology of the redundancy area. The MDT algorithm is currently under development in the 802.1Qca for redundant data communications, which can be extended to support computation of MDT for redundant time synchronization at a later time.
- The topology information needed for computing MDT is collected by the IS-IS protocol, which requires that all involved stations forming that topology must be identified with the same ID. In other words, an instance of IS-IS runs at each of the stations of the targeted area that exchange information with each other using the same ID.
- Such a requirement also applies to redundant time synchronization, where all time-aware systems must be identified with the same ID. Using domain number as ID for a redundancy area is the best choice.

Why One Domain is Preferable to Multiple Domains for Redundant Time Synchronization (4)

SIEMENS

4. Domain is the primary way for independency/isolation, but redundancy contains dependency/coordination.

- First, in case of redundant GMs, a primary GM and a secondary GM need to selected, while the primary GM is typically the best clock and the secondary GM is the second best clock among all the clocks in the redundancy area. This requires that the selection of the redundant GMs must be done in the whole redundancy area, so that the best, the second best or even third best clocks can be determined through comparison.

  If using multiple domains for redundant GMs, since the BMCA (of the current version) is domain-specific, we will have to extend BMCA to support inter-domain operations.

- Furthermore, the secondary GM is generally required to be synchronized with the primary GM, regardless if the secondary GM is operating in either hot-stand-by or cold-standby mode. If they were separated into two independent domains, we would have to define an extra mechanism to support inter-domain synchronization, which to a certain extend breaks the independency of "domains".

Why One Domain is Preferable to Multiple Domains for Redundant Time Synchronization (4 cont.)

SIEMENS

4. Domain is the primary way for independency/isolation, but redundancy contains dependency/coordination. (cont.)

- It is true that forwarding redundant synchronized time is independent for each redundant GM and each redundant tree. However, each time-aware system using redundant synchronized time to synchronize its local clock needs coordination of redundant information received from all redundant paths, e.g. in terms of comparison or combination.

In summary, if a redundancy area is divided into multiple independent domains for each redundant GM and each redundant path, we will have to spend a lot of effort to define a set of inter-domain mechanisms to collect information from the whole redundancy area, in order to perform the operations based on information global to that redundancy area. Putting a redundancy area in one domain can save much effort.

Proposals for Redundancy in 802.1ASbt
1. One Domain for One Redundancy Area

In general, we propose to use domain to identify time-scale and to identify a redundancy area as well if redundancy is applied to that time-scale. In this way, all redundancy related operations belonging to that redundancy area can be handled within one domain. This allows also isolation of multiple redundancy areas existing on a global network, making them operate independently from each other and also from other domains without redundancy.

Use domain as separator of different time-scales, while treating redundancy as optional features of one specific domain, e.g. a working clock domain where high-available time synchronization is required. In this sense, a domain implementing redundancy becomes a redundancy area.

One typical use-case for industrial control applications is to use .1AS (without redundancy features) to synchronize the global network within a universal time domain, while using .1ASbt (with redundancy features) for a sub-network to implement a redundant working clock domain. In this use case, domain is also used as a separator for different profiles, i.e. .1AS and .1ASbt.

Since domain number is used to identify a redundancy area if redundancy is supported and enabled in that domain. Thus, we need a flag to indicate whether redundancy is applied in a domain. A possible way to set this flag may be to use one of the flag bits defined in Table 10-6 of IEEE P802.1ASbt/D0.5.

Proposals for Redundancy in 802.1ASbt
2. Redundancy Types

High available time synchronization in a domain with redundancy can be implemented in either of or a combination of the following two forms:

- **Redundant GMs**: i.e., a primary GM and a secondary GM, while the secondary GM can operate in either hot-stand-by or cold-stand-by mode. The secondary GM must be synchronized with the primary GM.
- **Redundant sync trees:** defined to be two maximally disjoint sync trees rooted from the same GM)

(Note: some applications with higher redundancy requirements might need more than two GMs or more than two redundant trees from one GM.)

Thus, depending on the underlying network topology, node capability and application requirement, possible configurations for redundancy can be:

a. One GM with redundant sync trees

b. Redundant GMs with one sync tree for each GM

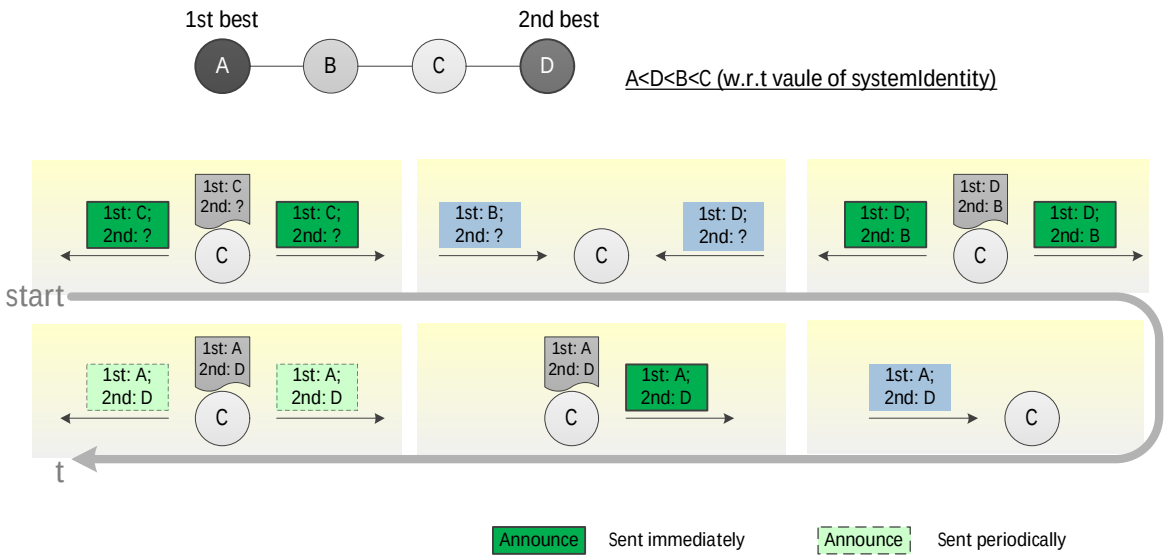
c. Redundant GMs with redundant sync trees for each GM

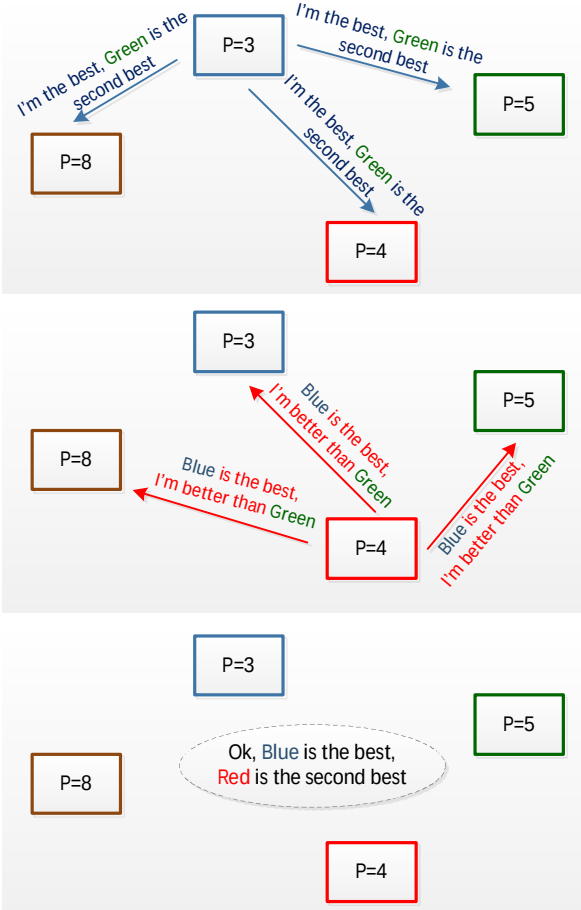
Proposals for Redundancy in 802.1ASbt
3. Selection of Redundant GMs

To support redundant GMs, we need an algorithm/mechanism to select the best, the second best (or even 3rd best) clocks to be the primary and secondary (or 3rd) GMs, for which the existing BMCA does not support.

There are two options for selecting redundant GMs

Option 1: extended BMCA with announce msg





Execution w/ Announce msgs

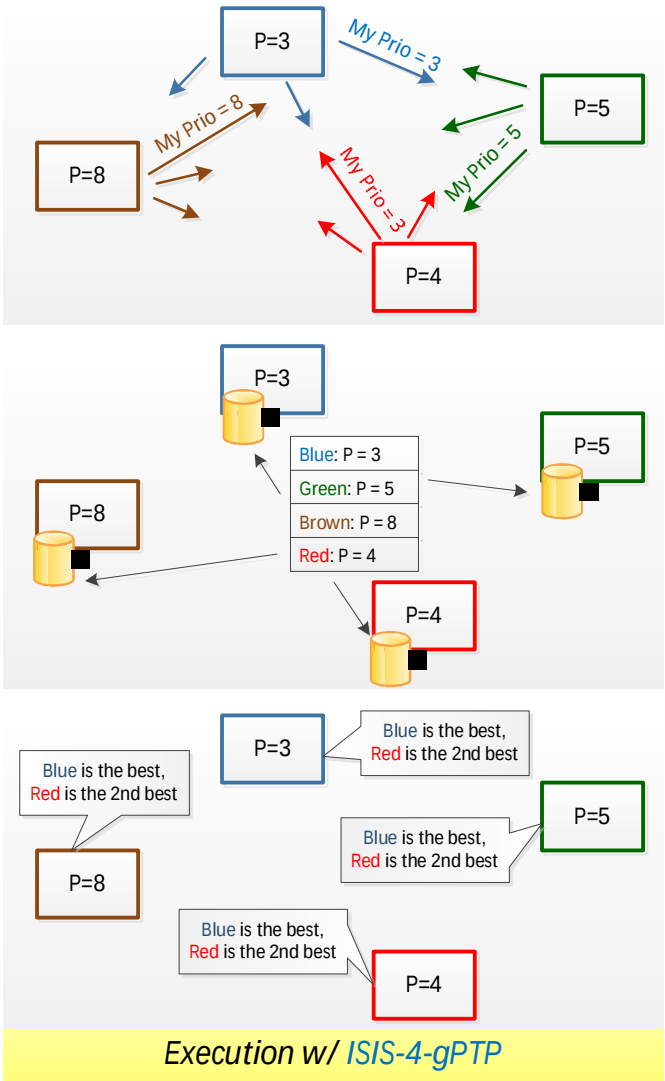
We might only need to compare the *rootSystemIdentity* values (Extended for 1st, 2nd or more) conveyed in a received Announce Message (i.e. *messagePriorityVector*), with the ones maintained locally (i.e. *gmPriorityVector*).

Proposals for Redundancy in 802.1ASbt
3. Selection of Redundant GMs (cont.)

SIEMENS

Option 2: ISIS-like mechanism with Link-State-PDU

- very similar to the way of using ISIS to collect and compute topology information
- exchanging only local clock property information (such as clockIdentity, GM capability, priority …) in LSPDU, instead of the best clock information like in Announce msg of BMCA
- selection will be done locally at each station, because all of them have the same database with information of all clocks.
- **Pros**: ISIS becomes a unified mechanism if also using ISIS for redundant tree computation
- **Cons:** ISIS might carry too much information



Proposals for Redundancy in 802.1ASbt
4. Computation of Redundant Sync Trees

SIEMENS

Redundant sync trees are defined to be maximally disjoint trees (MRT). As stated before, the MRT algorithm based on the topology information collected by ISIS is being standardized in 802.1Qca, which can be adopted to compute redundant sync trees for gPTP (we temporally name it ISIS-4-gPTP).

This requires 802.1ASbt to define an interface to ISIS-4-gPTP, because

- **ISIS needs access to gPTP world**

MRT for gPTP must be computed based on the gPTP topology that takes some gPTP-specific information into account, e.g. who are GMs (as MRT roots), which gPTP ports are disabled, etc.

- **ISIS must install the computation results into gPTP**

Computed MRT for gPTP will be installed (in which form to be discussed later) on each involved time-aware station. Note that gPTP does not utilize MAC relay to forward sync msgs, thus, MRT must be installed into gPTP.

Proposals for Redundancy in 802.1ASbt
5. Identification of Redundant GMs/Sync Trees in a Redundancy Domain

SIEMENS

Redundant GMs selection algorithm is responsible for notifying who are selected

- clockIdentity is the ID of gPTP clocks, thus also for redundant GMs
- Question: does each of redundant GMs need to put its ID into sync msgs?

Redundant Sync trees from the same GM must be distinguished

- ISIS-4-gPTP is responsible for allocating such IDs (e.g. treeID)
  - Question: need to associate treeID to GM ID?
- The GM distributing sync msgs over redundant sync trees must tag sync msgs with corresponding treeID (for sync follow-up msgs as well)
- Each time-aware system receiving sync msgs will use their embedded treeID to perform the following operations:
  - forwarding sync msgs and their follow-up msgs via corresponding sync trees
  - calculating rateRatio values on different (redundant) trees
  - using received redundant sync time to synchronize its local clock
  - performing monitoring on different redundant paths for maintenance

Proposals for Redundancy in 802.1ASbt
6. Encoding of treeID

SIEMENS

A possible way of encoding treeID is to use the "**domainnumber**" field in the header of all gPTP msgs, where some of the bits (e.g. higher 3 bits of the total 8 bits) can be deployed to store treeID and the rest bits are for the actual domain number.

- Such an encoding scheme requires a special interpretation of the "domainnumber" field, which is implemented only for time-aware systems supporting sync redundancy, i.e. within a domain implementing .1ASbt where redundancy features are enabled.
- For the time-aware systems, which reside outside that redundancy area/domain and neither support nor enable sync redundancy, they will treat the "domainnumber" field in any received gPTP msg normally, i.e. interpreting all bits as a whole.

However, the following operations do not rely on redundant trees, thus do not need to embed treeID in their msgs.

- PDelay measurement and calculation of neighborRateRatio (domain independent)
- Using extended BMCA and Announce msgs to select redundant GMs
- all others not related to redundancy

Thank you for your attention!

Feng Chen

I IA ATS TM5 1

Gleiwitzer Str. 555

90475 Nürnberg

Phone: +49 (911) 895-4955

Fax: +49 (911) 895-3762

E-Mail: chen.feng@siemens.com

siemens.com/answers