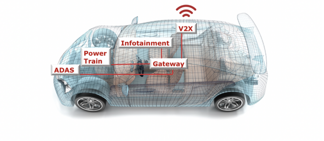
On precision and robustness of IEEE802.1AS synchronization in TSN networks

Jean-Luc Scharbarg
Université de Toulouse - IRIT/ENSEEIHT/INPT
Jean-Luc.Scharbarg@enseeiht.fr

DIVERSE Workshop - Tuesday 10th September 2024

Université de Toulouse

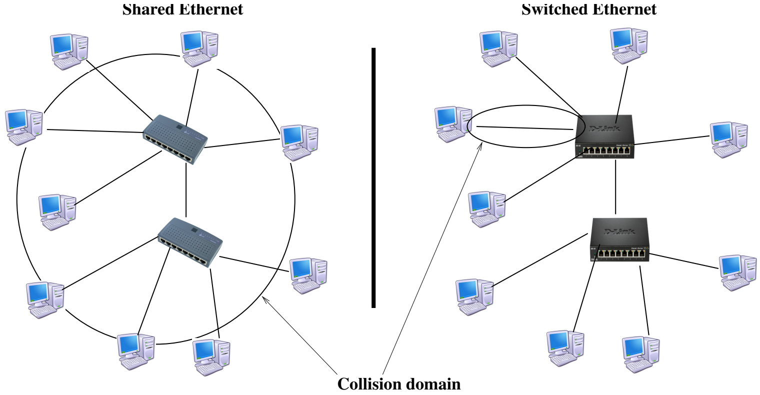
What is this talk about



- Embedded systems: heterogeneous communication needs
  - Hard real-time, soft real-time, best effort
  - High data rate, small data rate
- Flow differentiation is required
- EDEN project at IRT Saint-Exupery: Airbus Operation, Airbus Defense and Space, CNES, Continental Automotive, INPT/IRIT, ISAE/SUPAERO, ONERA, Safran Electronics and Defense, Thales Alenia Space, Thales Avionics
  - Can we define a subset of TSN for automotive, avionics and space?
  - PhD of Quentin Bailleul:
    can we have a precise and robust synchronization?

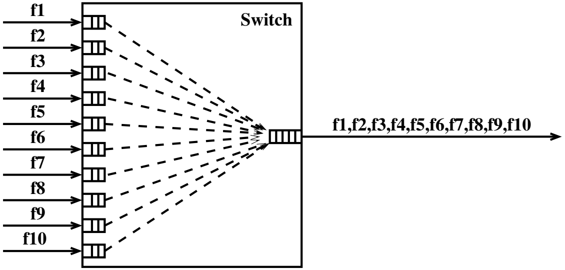
Real-time Ethernet solutions

- Ethernet provides high bandwidth
- Shared Ethernet is not predictable
  - Collision resolution is not deterministic
- Full duplex switched Ethernet: no more collision
  - Mono-emitter links
  - Frames are buffered in switch output ports



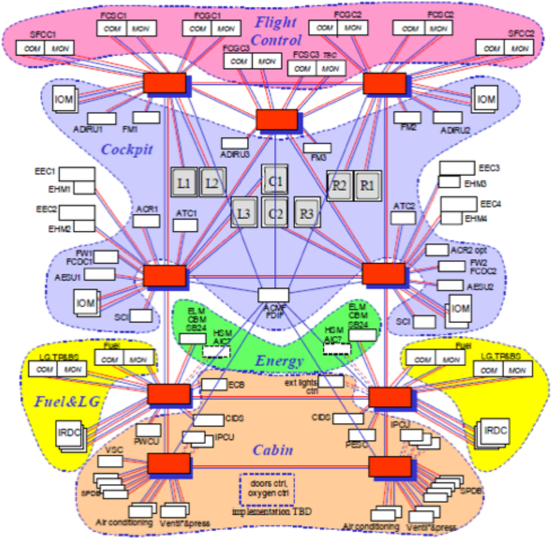
Full duplex switched Ethernet is not deterministic

- Temporary congestion on an output port
  - Increase of the waiting delay of frames in the Tx buffer
  - Frame loss by overflow of the Tx buffer
- An illustrative example



  - Five frames at the same time ⇒ one frame waits until the transmission of the four other ones
  - More than five frames at the same time ⇒ at least one frame is lost
- Addition of dedicated mechanisms to classical full duplex switched Ethernet in order to guarantee the determinism of transmissions

AFDX: Real-time Ethernet success story

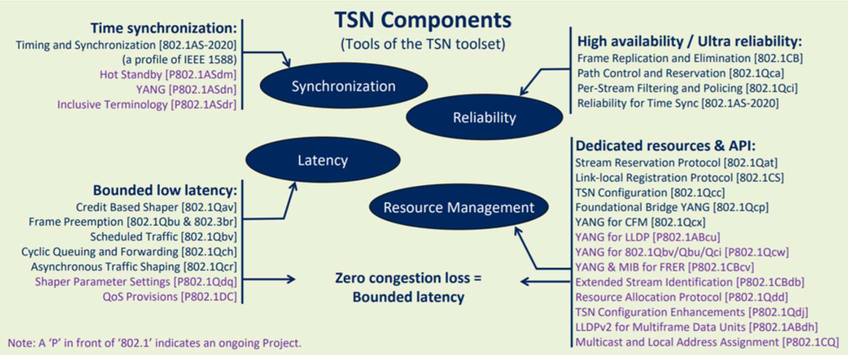


AFDX: Real-time Ethernet success story

- Static full duplex switched Ethernet
  - No spanning tree, static 802.1D tables
- Two priority levels ⇒ Two FIFO buffers per output port
- Sporadic flows statically defined
  - Mono emitter assumption
  - Multicast path with deterministic routing
  - Minimum inter frame gap
  - Bounded frame size
- Worst-case latency analysis
  - E.g. Network calculus
- Designed for homogeneous avionic flows
  - Unable to cope with control flows with very tight deadlines
  - Only two classes

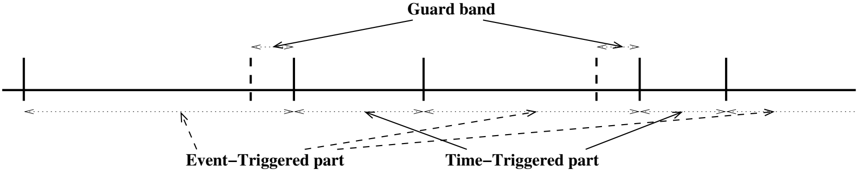
Time Sensitive Networking (IEEE TSN)

- A large set of standards



- Solution for time sensitive flows
  - Time synchronization is required
- Priority Queueing with shapers for the other flows

TSN: a possible frame scheduling

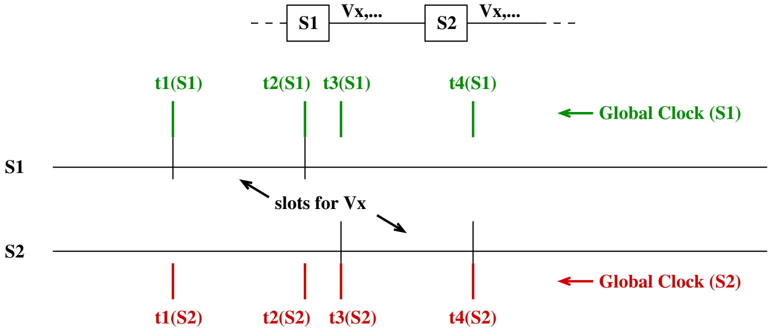


- Event-triggered (ET) traffic classes as well as time-triggered (TT) ones (e.g. for control data traffic)
- ET classes managed by priority queueing with Credit Based Shaper (CBS) for highest priority classes
- Each TT slot assigned to a TT class
- Guard bands prevent ET traffic to end transmission during a TT slot
- Global synchronization is required for TT

Additionnal solutions exist, e.g. Asynchronous Traffic Shaping (ATS)

Impact of an imprecise synchronization

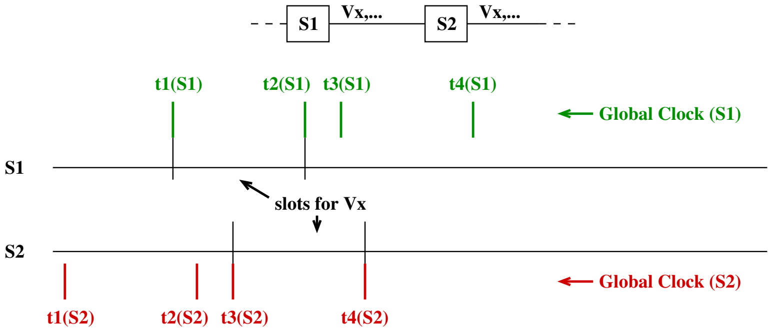
- Perfect synchronization
  - $S1$ and $S2$ have the same view of global time



- Vx frame at the output port of $S2$ no later than $t3(S2)$

Impact of an imprecise synchronization

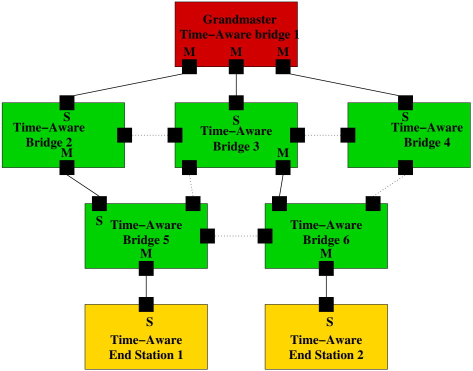
- Poor synchronization
  - $S1$ and $S2$ have a different view of global time



- Vx frame at the output port of $S2$ might miss the slot

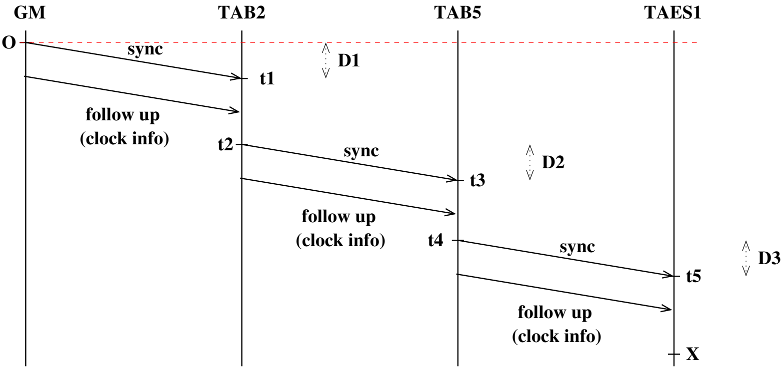
TSN: clock synchronization

- IEEE802.1AS: Precision Time Protocol (IEEE1588) profile for TSN
- A grandmaster node is selected
  - Statically: at design time
  - Dynamically: Best Master Clock Algorithm
- The grandmaster periodically broadcasts its clock: spanning tree
  - Correction of clock drift



- Periodic measurement of the delay of each link

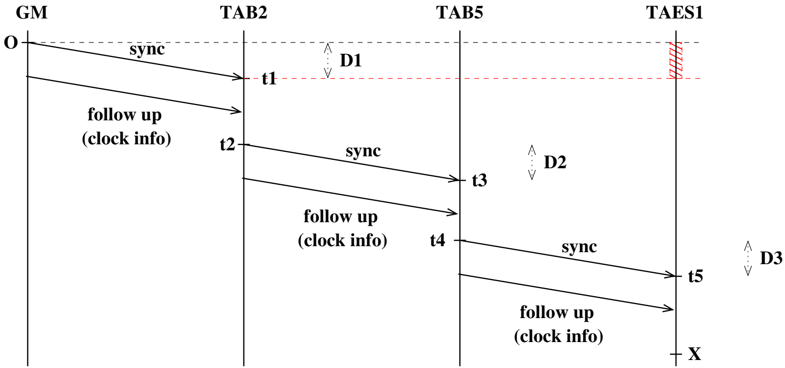
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O$$

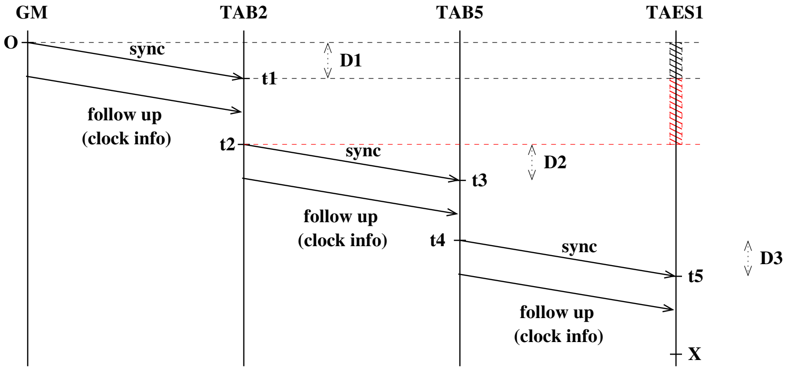
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1$$

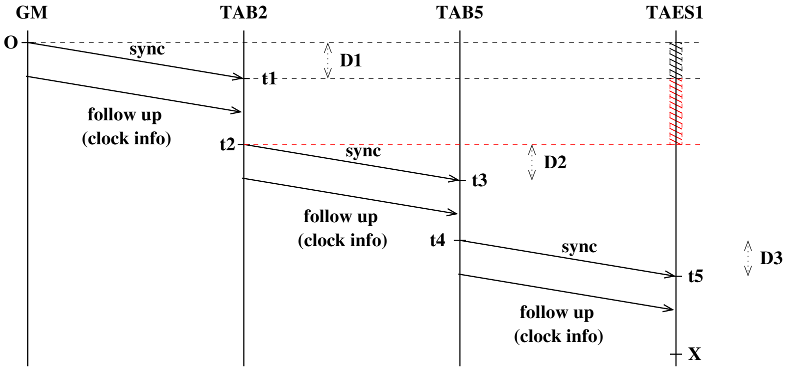
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1 + (t2 - t1)$$

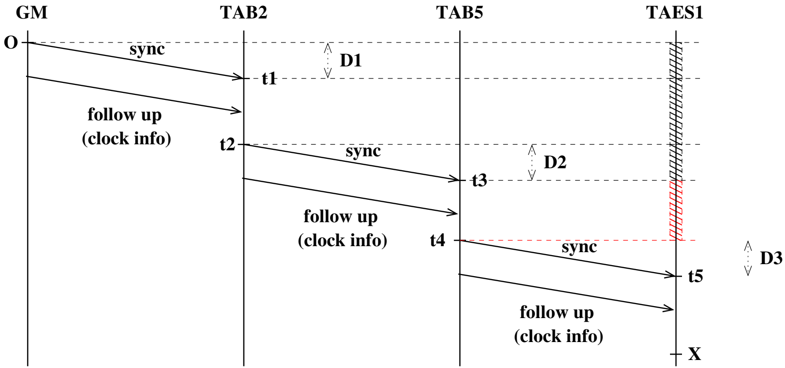
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1 + (t2 - t1) + \textcolor{red}{D2}$$

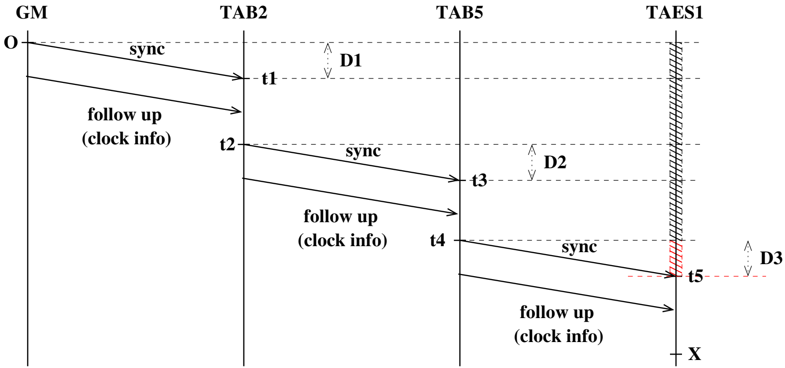
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1 + (t2 - t1) + D2 + {\color{red}(t4 - t3)}$$

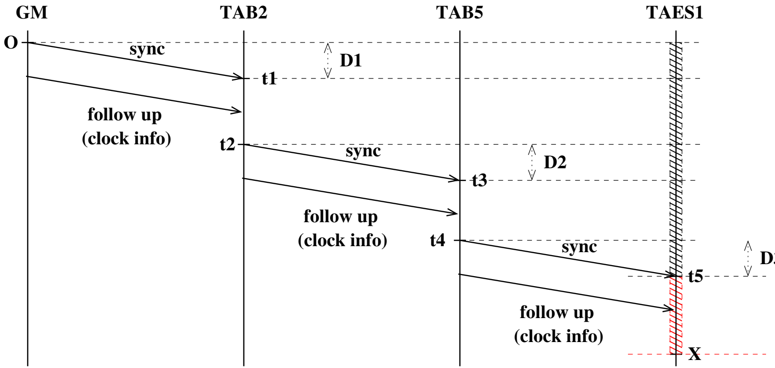
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1 + (t2 - t1) + D2 + (t4 - t3) + D3$$

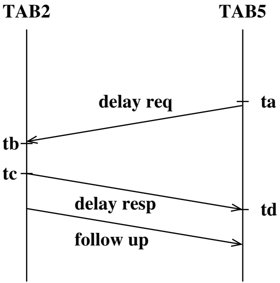
TSN: clock synchronization - Clock broadcasting



- Global clock in Time Aware End Station 1 at X:

$$O + D1 + (t2 - t1) + D2 + (t4 - t3) + D3 + (X - t5)$$

TSN: clock synchronization - Link delay

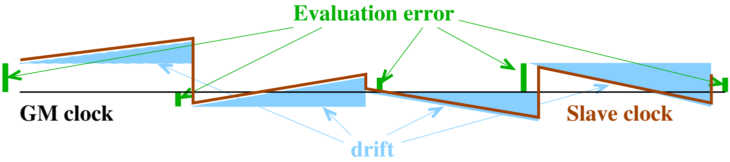


- The link is assumed to be symetric
- We have:

$$D2 = \frac{(td - ta) - (tc - tb)}{2}$$

TSN: clock synchronization - Precision

- Precision can be impacted by
  - Drift between two synchronization points
  - Evaluation error during clock broadcasting and/or link delay measurement

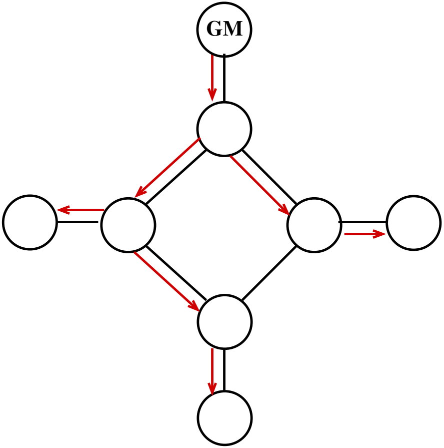


- Evaluation error can be bounded
  - Error comes from clock granularity, link asymmetry, ...
  - Can be bounded (Bailleul et al, ETFA'2023)
- Drift depends on
  - Slave clock quality: bound on the drift per second (ppm)
  - Time elapsed since the last synchronization point

⇒ Bound the duration between two synchronization points for any node

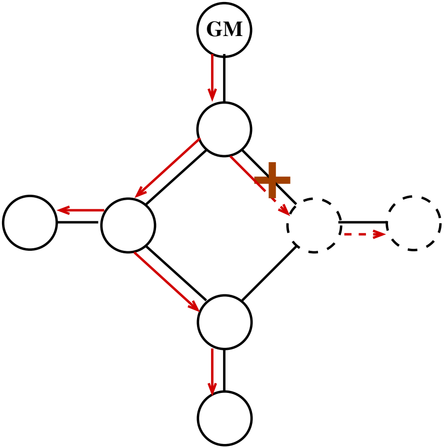
IEEE802.1AS and robustness

- A spanning tree for the broadcasting of its clock by the grandmaster



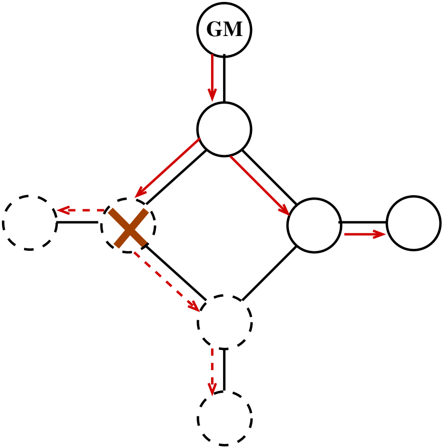
IEEE802.1AS and robustness

- A spanning tree for the broadcasting of its clock by the grandmaster
- Link failure ⇒ no more synchronization messages for a subset of the nodes



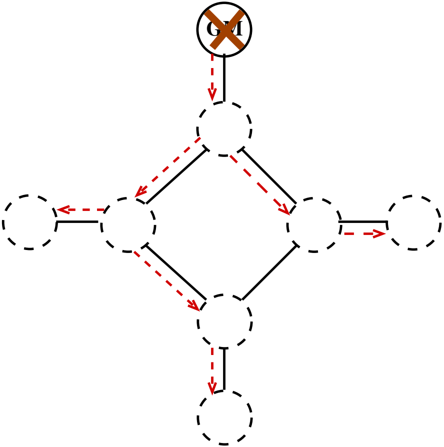
IEEE802.1AS and robustness

- A spanning tree for the broadcasting of its clock by the grandmaster
- Node failure ⇒ no more synchronization messages for a subset of the nodes



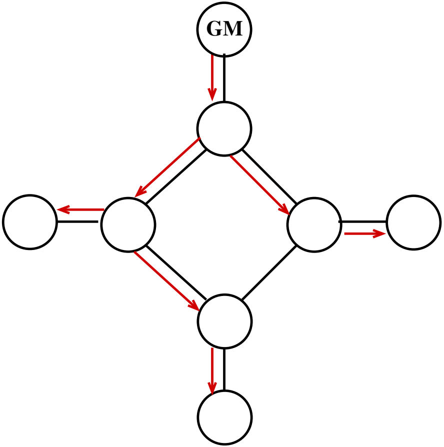
IEEE802.1AS and robustness

- A spanning tree for the broadcasting of its clock by the grandmaster
- GM failure ⇒ no more synchronization message



IEEE802.1AS and robustness

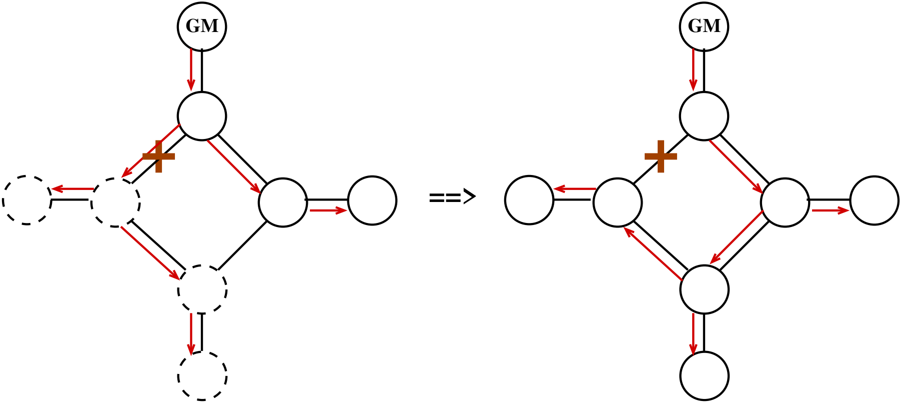
- A spanning tree for the broadcasting of its clock by the grandmaster



- Robustness: limit the impact of failures on synchronization

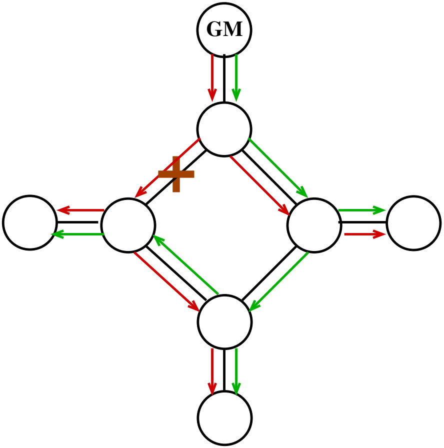
IEEE802.1AS and robustness: dynamic solution

- BMCA (Best Master Clock Algorithm): dynamic generation of a new spanning tree



IEEE802.1AS and robustness: static solution

- Several (e.g. 2) spanning trees are statically configured

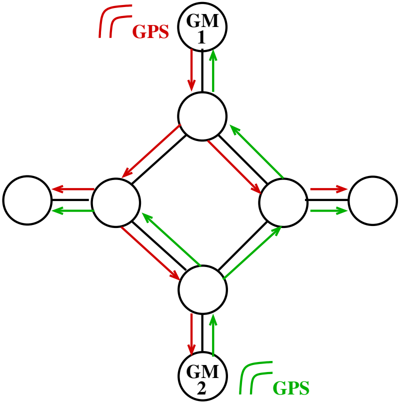


IEEE802.1AS and robustness

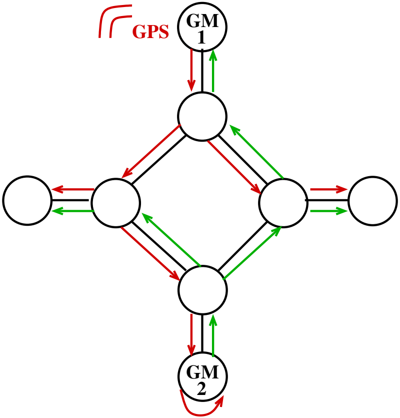
- Both solutions consume very few bandwidth
  - Dynamic solution consumes less than static one
- Configuration much more difficult for static solution
  - But done offline
- Very different reconfiguration delay for the solutions
  - Hard to predict for dynamic solution
  - Static solution induces no reconfiguration delay

$\Rightarrow$ Static solution is better for critical contexts (more predictable)

Static solution and GM failure

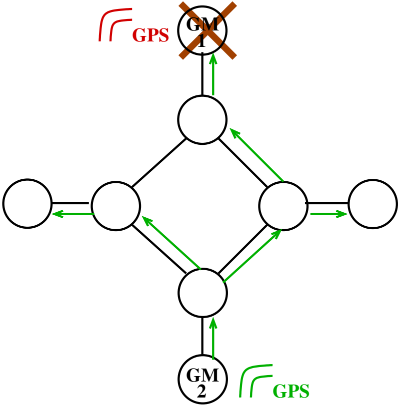


Backup Grandmaster with external time source (Classic GM)

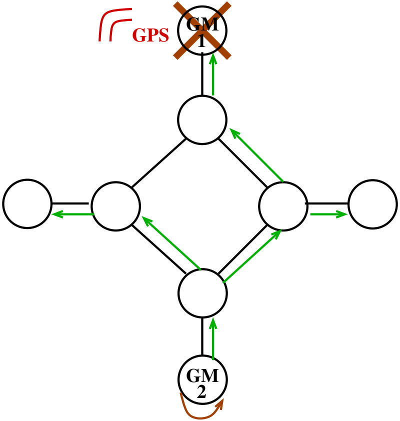


Backup Grandmaster without external time source (hot standby GM)

Static solution and GM failure



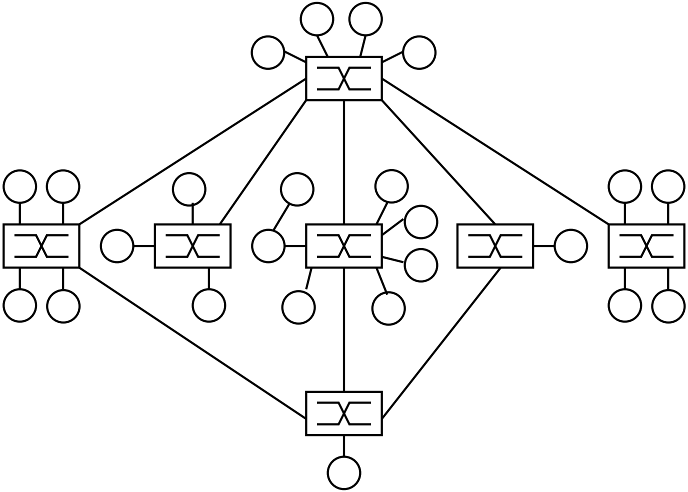
Backup Grandmaster
with external
time source
(Classic GM)
GM2 clock is used



Backup Grandmaster
without external
time source
(hot standby GM)
GM1 clock in freerunning
is used

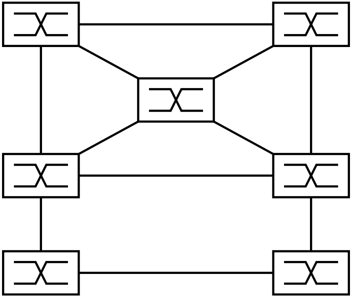
Large number of candidate static configurations

- Automotive example
  - 4356 candidates with 2 GMs and 2 trees per GM



Large number of candidate static configurations

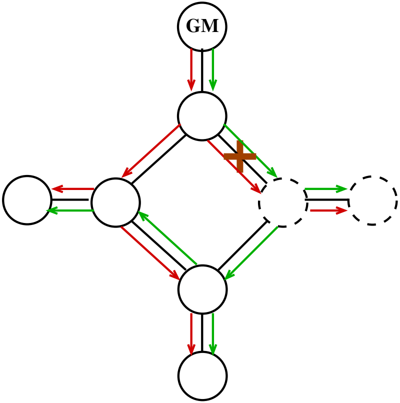
- Avionics (A350) example
  - 157 778 721 candidates with 2 GMs and 2 trees per GM (higher connectivity)



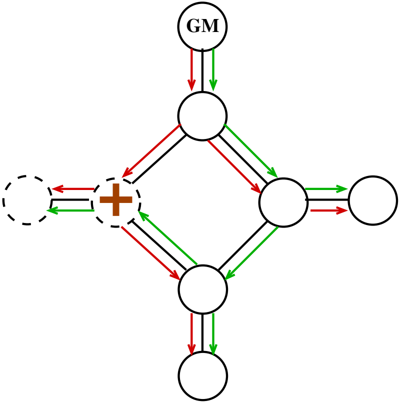
What is a robust configuration?

- Different possible metrics
  1. Disjointedness
     - ★ Number of nodes reachable from GM using disjoint pathes (no common links)
  2. Independence
     - ★ Number of nodes reachable from GM using independent pathes (no common nodes)
  3. Resistance
     - ★ Minimum number of link or node failures required to disconnect each node from GM
  4. Failure impact
     - ★ Minimum overall number of disconnected nodes per failure
- Failure impact metric gives the best results
  - Better evaluate the impact of failures
  - Can adapt to any topology

Failure impact metric



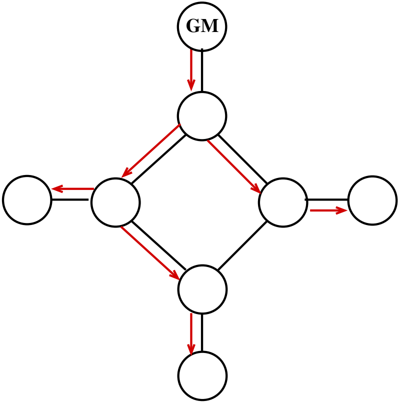
2 disconnected nodes



2 disconnected nodes

What is a precise configuration?

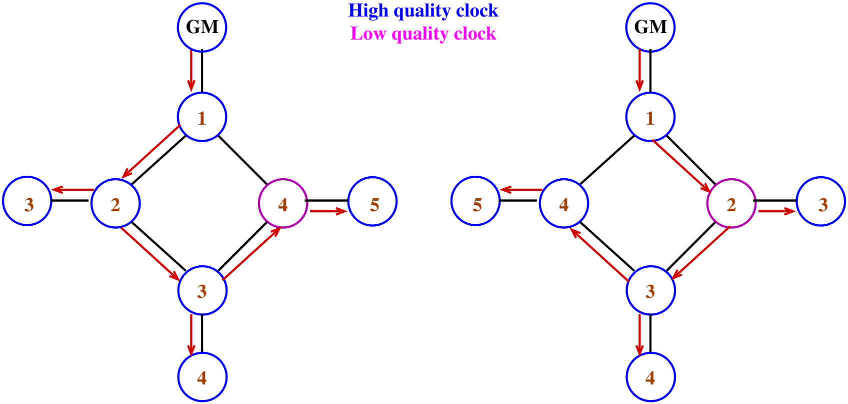
- Two candidate metrics
  - Average number of hops between GM and the nodes



- The synchronization error clearly increases with the number of hops between GM and node

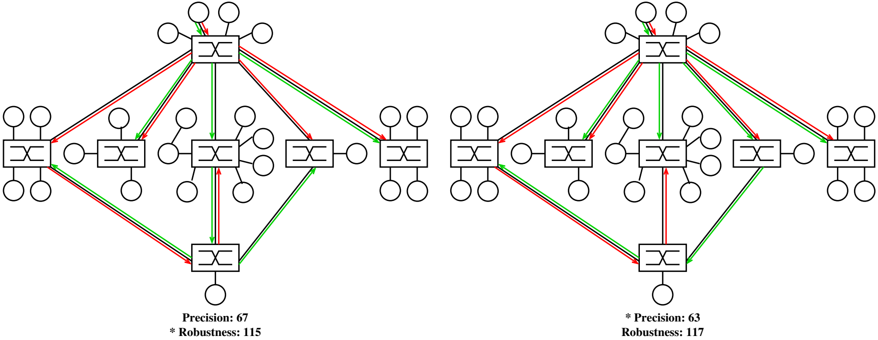
What is a precise configuration?

- Two candidate metrics
  - Average clock quality with weights depending on the number of hops from GM



- Synchronization in a given node impacted by the quality of its clock
- But no impact on the following nodes on the path

Precision vs robustness



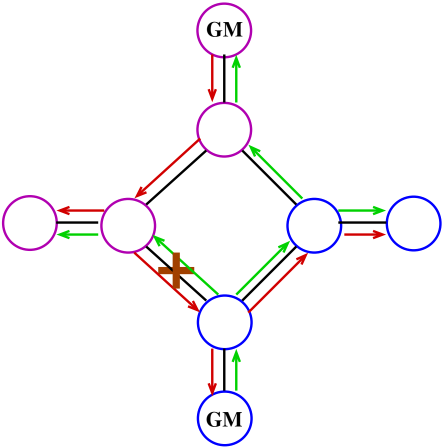
- Precision and robustness often not compatible
  - Precision requires short pachs
  - Robustness requires more or less disjoint paths which are not the shortest ones
- Precision more easily tunable, e.g. increase resynchronization frequency

⇒ Select the most precise solution among the most robust ones

The case with (a) backup GM(s)

- x GMs, y spanning trees for each of them
- GMs are selected, based on the quality of their clock, their position, ...
- Searching for the best solution in one step doesn't scale
  - too many groups of $x \times y$ spanning trees
- Process in two steps
  - Select the best solution(s) for each GM independently
  - Find the best combination between the selected solutions
- No significant impact observed on the results
- Multiple time base issue: due to failures
  - GM on hot standby does not receive synchronization message from primary GM
  - One node does not receive synchronization message from primary GM, but receives synchronization message from GM on hot standby

Time base issue



- One solution: cover all possible paths from primary GM to GM on hot standby

Conclusion

- Synchronization is required in real-time switched Ethernet networks if
  - scheduled traffic, typically based on TAS,
  - synchronization constraint between applications on different nodes
- Synchronization has to be reliable
  - Insure that all the nodes share a common clock with a bounded error
- $\Rightarrow$ build a robust and precise synchronization
  - Robustness measured as the number of disconnected nodes per failure
  - Precision measured as the average number of hops between GM and nodes
  - Precision more easily tunable $\Rightarrow$ select the most precise solution among the most robust ones

Thank you for your attention
Any question?