IEEE 802 Time-Sensitive Networking: Extending Beyond AVB

Michael D. Johas Teener
Plumbing Architect and
Senior Technical Director

Agenda

- Requirements for Time-Sensitive Networking (TSN):
  - And where did those requirements come from?
- Audio Video Bridging (AVB): a good start:
  - What we have now.
- New standards in IEEE 802:
  - Further reductions in maximum delays.
  - Improved robustness and reliability.
  - Scalability to larger and smaller systems.
  - Status.
- Architectural implications for endpoints and switches:
  - Queuing and preemption.
  - Multipathing and stream redundancy.
  - Synchronization.

TSN Objectives

- Extend use cases from audio/video applications to control systems
- Reduced worst-case delays
  - 4 µs or less per hop @ 1 Gbps for short messages (plus cable delays)
- Improved robustness:
  - Alternative paths with “instant” switchover
  - Seamless redundancy using multiple simultaneous streams
  - Multiple clock sources with “instant” switchover
- Scalability
  - Reduced management traffic for reservations and configuration

Example: Industrial Networks

- Large physical extent:
  - Think refinery or automotive assembly line: 1000m or more
  - Work cell (robot) up to 5 hops, factory up to 64 hops
  - Coexistence of bulk traffic (video/quality reports)
- Precise timing:
  - Within the work cell ±500 ns
  - Within the factory ±100 µs
- Deterministic and very small delays
  - Within the work cell < 5 µs
  - Within the factory < 125 µs (≈ 4 µs per hop)
- Safety!
  - Redundant control/data paths with “instant” switchover
    - Seamless or at the very least < 1 µs

Example: Automotive Network

- Small physical size, many ports, many different data requirements:
  - 30m, 5 hops, perhaps 100 devices
  - Control, sensors, driver-assist video, radar, entertainment A/V
- Deterministic and very small delays:
  - < 100 µs through 5 hops using 100 Mbps PHY
- Safety!
  - Redundant paths/fault detection

802.1 AVB: A Good Start

What Works Now

- 802.1AS (the 802 profile of 1588) provides better than ±1 µs synchronization
  - Typical implementations are better than ±300 ns for 7 hops.
- 802.1Q stream reservation plus credit-based shaper provides delays about 130 µs per hop at 1 Gbps.
- 802.1Q Shortest Path Bridging is based on IS-IS:
  - Topology discovery, and separation of paths.
  - Reconfiguration after network fault at near optimal delays.
- 802.1Q VLANs can be configured to route redundant control data:
  - Requires manual configurations (which may be OK), but very hard to control if there is a fault.
- Most importantly, an architecture for managed traffic:
  - A standards-based way to manage different classes of time-sensitive traffic.
  - Control and monitoring via protocols running on endpoint devices (vs. requiring an SNMP-based centralized network manager).

Desired Improvements for Control

- 802.1AS can take up to 1 second to switch to a new grand master:
  - Prefer switchover in ¼ second or less
  - Allow for a second grand master to be active at the same time
- Switch delays need to be reduced by almost a factor of 100
- Standardized ways to specify multiple “minimally coincident” paths for critical data and control streams:
  - Very fast switchover and/or simultaneous transmission of streams
- Path selection and reservation needs to be simpler and faster:
  - Bigger networks (thousands of streams), lower-cost networks

Safety and Redundancy

IEEE 802.1Q provides basic tools

- “Shortest Path Bridging”, based on IS-IS:
  - Topology discovery and separation of paths
  - Reconfiguration after network fault at near optimal delays
- VLANs can be configured to route redundant control data:
  - Requires manual configurations (which may be OK), but very hard to control if there is a fault

Meeting the New Requirements: IEEE 802.1 Time-Sensitive Networking[1]

[1]Formerly known as AVB Gen 2.

Limits on Delay

The fundamental problem is interfering traffic!

If a packet is to be transmitted on a particular egress port, then all traffic, regardless of the priority, must wait until the egress port has completed transmitting that packet.

What is the best we could do?

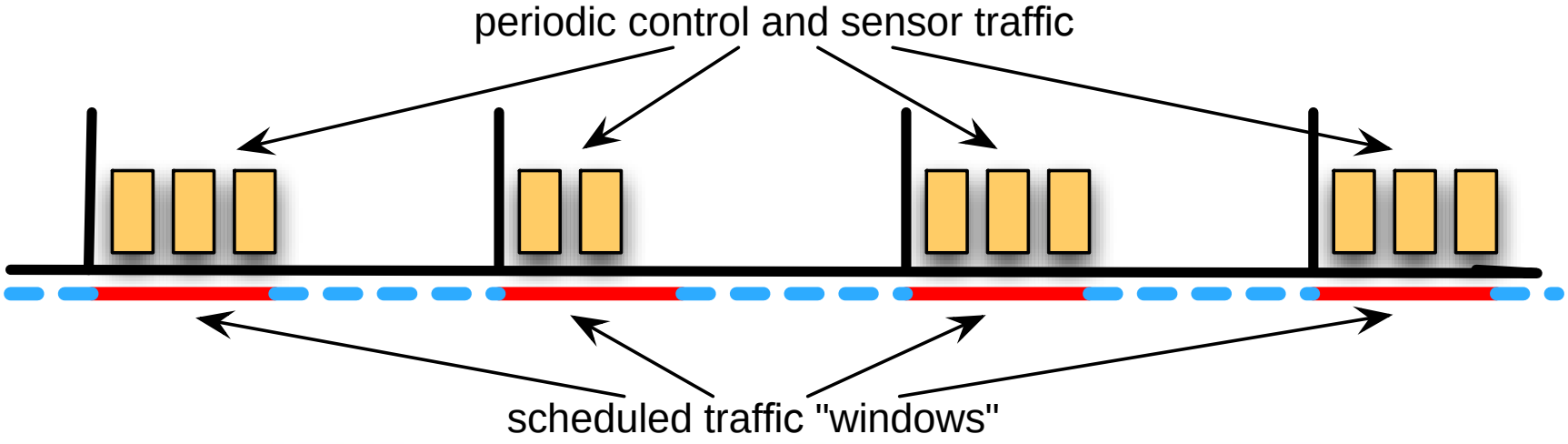
- Assume no interfering traffic:
  - Gigabit Ethernet switch could have delays as low as 4.122 µs.
- Add cut-through switching:
  - Gigabit Ethernet switch could have delays as low as 2.074 µs!
  - Note that cut-through normally does not help if there is interfering traffic, but in this case, we assume no interference.
- So, for a 32-hop network, we would have 66.368 µs delays.

Not bad!

Avoiding Interfering Traffic

Time-Aware Shaper

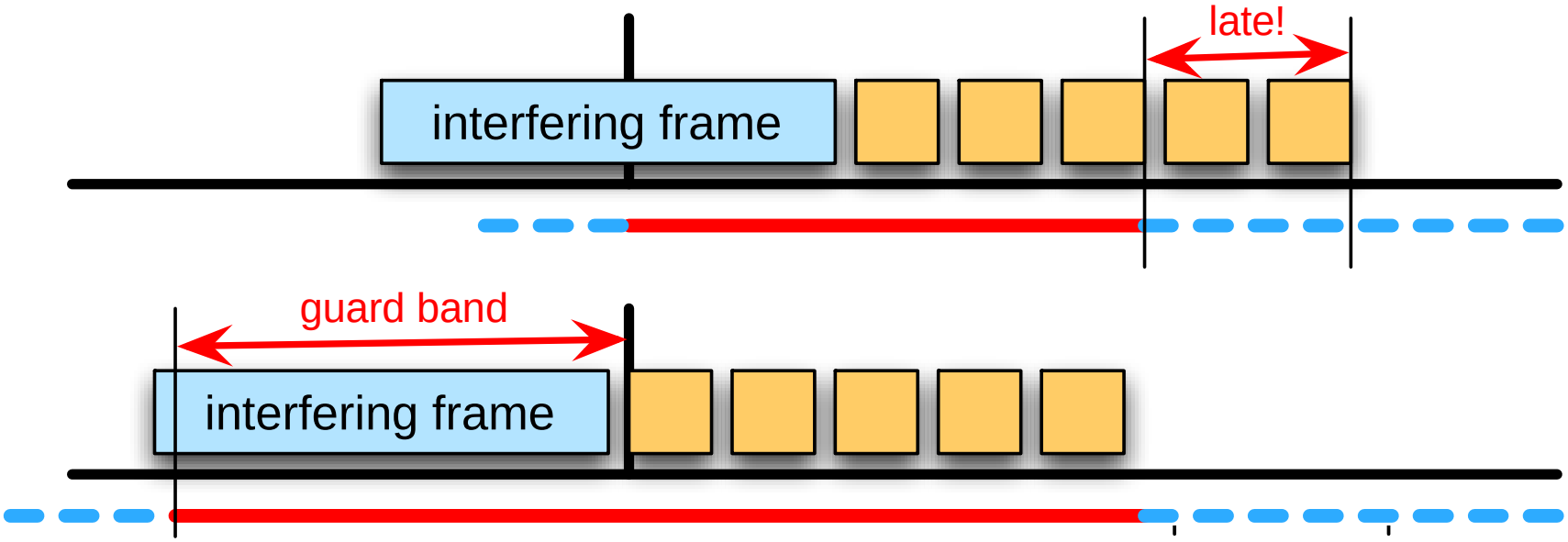
- Make switches aware of the cycle time for control traffic:
  - Block non-control traffic during particular windows of time to ensure that the egress port for a control stream is idle when the control traffic is expected.
  - Each egress port would have a separate schedule.
- Nontrivial calculation in nontrivial networks:
  - Requires a fully managed network.
  - This is a well-understood but difficult problem currently implemented in proprietary networks such as Siemens' "Profinet."



Time-Aware Shaper Issues

A “guard band” is necessary

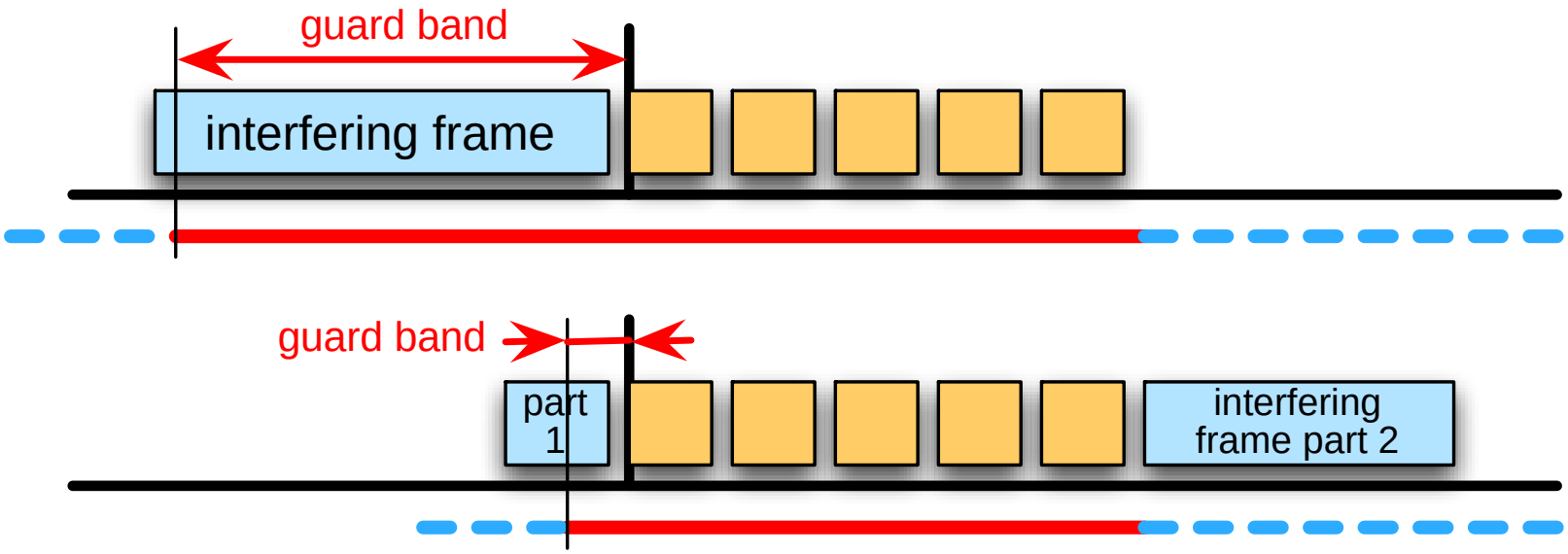
- If an interfering frame begins transmission just before the start of a reserved time period, it can extend critical transmissions outside the window.
- Therefore, a guard band is required before the window starts, equal in size to the largest possible interfering frame.



Reducing the Guard Band

Preemption is a solution

- If preemption is used, the guard band needs to be only as large as the largest possible interfering fragment instead of the largest possible interfering frame.
- It is easy to see that the smaller the size of the time-reserved windows, the larger the impact of preemption.



Efficiency of Time-Aware Shaper

Assume the control traffic consists of a burst of four 128-byte packets, and half of the window is needed to compensate for delivery jitter.

Control traffic (as a % of link bandwidth)	Overhead (margin, preamble, interframe gap)	With preemption	Without preemption
0.1%	0.2%	0.3%	0.5%
1.0%	2.3%	2.6%	5.3%
10%	23%	26%	53%
30%	69%	78%	160%

802.1Qca: Path Control for Redundancy

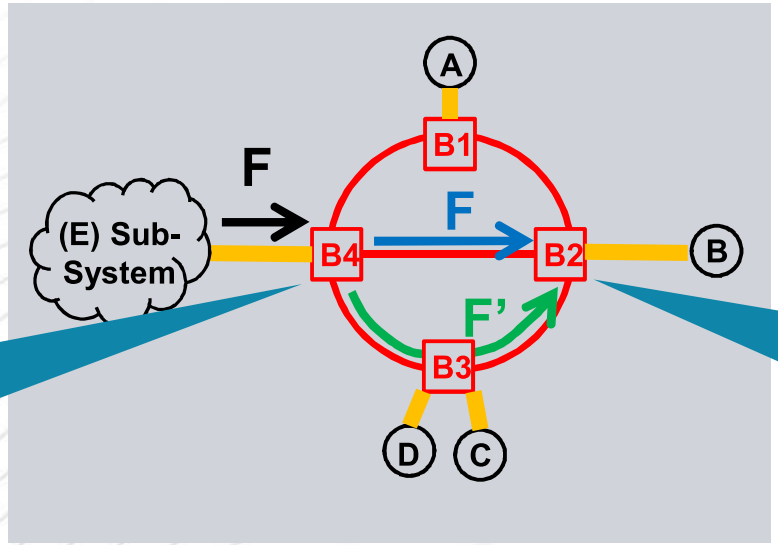
- Defining new ways to use shortest path bridging concepts to configure preferred routes for redundant paths:
  - And to ensure the paths remain “redundant.”
- Adding rules for path selection that are not just based on link speed:
  - But also for reliability and worst-case delay

802.1Qcc: Enhanced Stream Reservation

- Reduce the size and frequency of reservation messages:
  - Compatible with existing 802.1Qat Stream Reservation Protocol
  - Relaxed timers, updates only on link state or reservation changes
  - Reservation status updates done incrementally:
    - Only changes are explicitly transmitted, along with a “digest” of existing reservations.
- Explicit interoperation with “God box” preconfigured systems
- Explicit interoperation with higher-level reservations
  - e.g., RSVP/IGMP

802.1CA: Seamless Redundancy

- Objective–Introduce the seamless redundancy mechanism that is compatible with the TSN traffic classes:
  - Supports redundant rate-constrained and redundant scheduled traffic.
  - Redundancy concept is simple and lightweight.
  - Suitable for resource-constrained embedded systems.



Replication of critical frames

Elimination of redundant frames

New Tools for Time Synchronization

- IEEE 802.1AS-2011 is the IEEE 802-preferred profile of IEEE 1588:
  - No layer violations, scalable to thousands of ordinary clocks.
  - Meets the synch requirements of the most stringent current requirements.
- However …
  - Only two-step clocks were defined and one-step clocks are much less processor-intensive, and so are much preferred for very large port count switches.
  - Reestablishment of a grand master in very large networks can take one second:
    - Industrial networks want much faster recovery, seamless, if possible.
  - Multiple levels of synchronization needed in industrial networks:
    - Much tighter within a work cell than within a whole factory.
- New 802.1ASbt amendment to address all of these issues

When?

- Standards development in process:
  - Project 802.1ASbt amendment to time-sync
  - Project 802.1Qbu amendment to switch standard for preemption
    - Call for interest within 802.3 at November IEEE 802 Plenary for Ethernet-specifics
  - Project 802.1Qbv amendment to switch standard for time-aware queuing
  - Project 802.1Qca amendment to support path control and registration for redundant networks
  - Project 802.1CB to support seamless redundancy in IEEE 802 networks
  - Project 802.1Qcc amendment to update existing Stream Reservation Protocol to support Qbu/Qbv/Qca/CB technology
- Most projects will be “technically stable” in a year:
  - Final standardization is two to three years away.
- Exception: 802.1Qca and 802.1CB are difficult projects:
  - Very careful analysis required to make sure they **work.**
  - Technical stabilization is 2 or 3 years away, final standards 3 to 5 years.

Implications for Designs of Future Products

Architectural Implications

Time-Aware Shaper

- Talkers (transmitting devices) need a “scheduled queue”:
  - If the device is very simple, only has one stream, and is not running complex software with varying real-time requirements, then this can be done in firmware.
  - Best if there is a separate queue, with a “launch me as soon as possible after time x” attribute per packet.
    - This is the method used in the existing BCM58701 controller used by Apple in existing Macs.
- Switches:
  - Every egress port needs a scheduled blocking and enabling service per class:
    - Within a 30 µs to 10 ms cycle, there should be “n” (an unknown value at this time, but not too many) enable and disable events per class.
    - Note that for some periods, all queues are blocked (the guard band).

Architectural Implications

Preemption

- Talkers ***may*** need to implement preemption:
  - Required for complex talkers that support multiple traffic classes and have significant software delays *(anything running Linux/Unix/Windows/etc.*).
  - Separate queue/DMA channel for “preempting” traffic.
- Switches ***must*** provide one level of preemption:
  - Each egress port needs an indication of which traffic classes (priority code points) to preempt.
- Ethernet MAC transmitter in switches (and perhaps talkers) ***must*** be able to stop transmitting preemptable traffic, insert preempting packet, then resume preempted traffic when no preempting traffic is in a queue.
- Ethernet MAC receiver in ***all*** devices ***must*** detect boundary between preempted packet and preempting packet and reassemble the preempted packet.

Architectural Implications

Cut through

- Switches ***may*** implement cut-through, particularly for 100 Mbps and 1 Gbps
- Will reduce switch delays for preempting traffic by at least 2 µs @ 1 Gbps:
  - Much more for longer frames
  - Dependent on link rate of ingress and egress ports

Architectural Implications

Redundant paths

- All switches ***may*** need to implement 802.1Qca IS-IS topology discovery and/or 802.1Qcc reservation in their firmware:
  - New control protocol enhancements to enable multiple simultaneous paths, not just the “best” path.
  - Alternative is to preconfigure switch, or to use “God box” centralized control.
- All switches ***must*** implement flexible forwarding capabilities:
  - Forwarding/routing/queuing rules need to make decisions based on destination address (including the possibility of MANY different “multicast” addresses in use), VLAN addresses, traffic class, etc.
- Endpoint firmware ***should*** support updated 802.1Qcc reservation protocol:
  - Currently defined SRP does not include all the needed information.

Seamless Redundancy

- Selective packet replication based on address/traffic class and path information:
  - Possible encapsulation/additional tagging to aid duplicate frame elimination
  - May use Stream Reservation Protocol to set up paths and labeling rules
- Duplicate frame elimination based on address/traffic class and timing:
  - Timing information needed to limit memory needed for duplicate frame detection
- NOT TRIVIAL!
- May also be L3/IP version in the near future, but using the same techniques

Architectural Implications

Enhanced time synchronization

- All switches ***should*** implement one-step timestamping:
  - Fortunately, almost all current designs already do this.
    - Plus, we have PHYs with built-in timestamping.
  - Firmware needs to be updated.
- Endpoints ***should*** also implement one-step timestamping:
  - Straight-forward update, but needs to be done carefully:
    - Support for “all” of the various 1588 modes.
    - Also, some non-1588 modes used in some ITU diagnostic services.
  - One-step services could be used at the application layer as well:
    - Financial services, data center management, high-speed-computing complexes.

Thank you!

Michael Johas Teener

mikejt@broadcom.com

References

Hyperlinks to the public documents areas of the IEEE 802.1 website:

- AVB - Generation 2 Latency Improvement Options
- QoS requirements for Automotive Ethernet backbone systems
- AVB + Extensions for Industrial Communication
- Preemptive Transmission advantages
- Support for Seamless Redundancy in AVB 2