IEEE 802.1 Time-Sensitive Networking (TSN)

János Farkas, Norman Finn, Patricia Thaler

Ericsson          Huawei           Broadcom

IETF 99 – Tutorial           July 16, 2017
Before We Start

This presentation should be considered as the personal view of the presenters not as a formal position, explanation, or interpretation of IEEE 802.1.
Outline

• Introduction
• Reliability
• Deterministic latency
• Resource management
• TSN Summary
• Related work: DetNet
INTRODUCTION
Potential Markets (not comprehensive)

- Industrial Automation
- Wind
- Nuclear
- Power Gen
- Healthcare
- Aviation
- Transportation
- 5G
- Oil & Gas
- Water

High Traffic Mix, Deterministic, Low Latency, Secure, Reliable, High Throughput
IEEE 802 and 802.1

• IEEE 802 LAN/MAN Standards Committee (aka IEEE 802 or LMSC)
  – Develop LAN and MAN standards
  – Mainly for link and physical layers of the network stack

• IEEE 802.1
  – 802 LAN/MAN architecture
  – Internetworking among 802 LANs, MANs, and other wide area networks
  – 802 Security
  – 802 overall network management, and protocol layers above the MAC & LLC layers.
From AVB to TSN

• IEEE 802.1 Audio Video Bridging (AVB) Task Group (TG)
  – Started in 2005
  – Address professional audio, video market
  – Consumer electronics
  – Automotive infotainment
  – Avnu Alliance: associated group for compliance and marketing

• IEEE 802.1 Time-Sensitive Networking (TSN) TG
  – AVB features become interesting for other use cases, e.g.
    • Industrial
    • Automotive
  – AVB was not an appropriate name to cover all use cases
  – AVB TG was renamed to TSN TG in 2012
  – Interworking TG and TSN TG were merged in 2015
Time-Sensitive Networking

TSN Components

**Time sync:**
- Timing and Synchronization

**Bounded low latency:**
- Credit Based Shaper
- Frame Preemption
- Scheduled Traffic
- Cyclic Queueing & Forwarding
- Asynchronous Traffic Shaping

**Ultra reliability:**
- Frame Replication & Elimination
- Path Control
- Per-Stream Filtering & Policing
- Time sync reliability

**Resource Mgmt**

**Reliability**

**Latency**

**Zero congestion loss**

Guaranteed data transport with bounded low latency, low delay variation, and extremely low loss
Bounded Latency

• TSN’s target applications, real-time networks, require a guaranteed not-to-exceed end-to-end latency for critical data

• Average/mean/best-case latencies are irrelevant

• Many ways to accomplish bounded latency:
  — Throw away late packets; grossly overprovision the network; intensive engineering and testing.
  — Provide zero congestion loss
0 Loss = Bounded Latency

• Given:
  – Constant input rate
  – Finite buffer capacity
  – 0 packets lost

• End-to-end latency is bounded
How to Get 0 Congestion Loss

• At every hop:
  – Packets/interval in == packets/interval out

• But:
  – Packetized data is not a constant-rate bit stream
  – Different flows’ optimal transmit times can conflict

• So, gaps and bursts are inevitable
Gaps and Bursts

1. Reserve buffer space and bandwidth resources before the critical flow starts

2. Use queuing/reservation disciplines that strictly limit inter-flow interference and provide predictable gap/burst behavior

3. Use extra buffers for known delay variations (e.g., forwarding delay)
Traditional Service

- Curve have long tails
- Average latency is good
- **Lowering the latency** means **losing packets** (or grossly overprovisioning)
TSN Service

- Packet loss is now due to equipment failure
- Average latency may be larger, but no tails
Bottom Line: Why TSN?

• Without TSN
  – Network engineering
  – Bandwidth, over-provisioning
  – Testing

• With TSN
  – Way easier to engineer
  – Works even in hard-to-test corner cases
  – Way cheaper
RELIABILITY
Frame Replication and Elimination

• Avoid frame loss due to equipment failure (802.1CB)
• Per-packet 1+1 (or 1+n) redundancy
  – NO failure detection / switchover
• Send packets on two (or more) disjoint paths, then combine and delete extras
Illustration of QoS & Reliability Functions

Per-Stream Filtering and Policing

Packet Replication / Elimination

Per-Stream Shaping

Queuing

Transmission Selection

Frame reception

Per Stream

can be viewed as a hierarchical approach

Per Class

Frame transmission
Policing

• Every frame can be marked “green” or “yellow” using the Drop Eligible bit of VLAN tags
• “red” are dropped
• “yellow” frames have a higher probability of being discarded than “green” frames
• Policing is done per input port, but only after it is determined that a frame can be delivered to some port. Frames that are dropped by the forwarding mechanism are not policed.
• Policing algorithm is from MEF Forum spec 10.3 (see also RFC 2963)
Per-Stream Filtering and Policing

- Protection against bandwidth violation, malfunctioning, malicious attacks, etc. (802.1Qci)
- Decisions on per-stream, per-priority, etc.
- Stream Filter
  - Filters, Counters
- Stream Gate
  - Open or Closed
  - can be time-scheduled
- Meter
  - Bandwidth Profile of MEF 10.3
  - Red/Yellow/Green Marking
DETERMINISTIC LATENCY
Priority and Weighted Queuing

- **Strict Priority** (802.1Q-1998)
  
  ![Strict Priority Diagram]

  
  Highest priority: 7

- **Weighted queues** (802.1Qaz)
  
  - Standard management hooks for weighted priority queues without over-specifying the details
  
  ![Weighted Queues Diagram]

  Priority selection
Credit Based Shaper

- Credit Based **Shaper ▼** (CBS - 802.1Qat)
  - Shaped queues have higher priority than unshaped queues
  - Shaping still guarantees bandwidth to the highest unshaped priority (7)

```
  Weighted
  1 0 4 5 6 7 2 3
```

- Highest priority for shaped queues

- CBS is similar to the typical run rate/burst rate shaper, but with really useful mathematical properties
  - Only parameter = bandwidth
  - The impact on other queues of any number of adjacent shapers is the same as the impact of one shaper with the same total bandwidth.
Credit Based Shaper – Example

- CBS spaces out the frames in order to reduce bursting and bunching
Scheduled Traffic

- Reduces latency variation for Constant Bit Rate (CBR) streams, which are periodic with known timing
- Time-based control/programming of the 8 bridge queues (802.1Qbv)
- Time-gated queues
- Gate: Open or Closed
- Periodically repeated time-schedule
- Time synchronization is needed

Diagram:

- Weighted
- Priority selection

<table>
<thead>
<tr>
<th>1</th>
<th>0</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>
Cyclic Queuing and Forwarding

- **Double buffers** (802.1Qch) are served alternate using time-gated control.

- Two pairs: 2–3 and 4–5 in this example.

- Shapers ensure fair access for 0, 1, 6, 7 traffic.

- Alternately open green and purple.

- If the wire length and bridge transit time are negligible compared to the cycle time, double buffers are sufficient.

  - Frames being received.
  - Output in progress.
Asynchronous Traffic Shaping

- Zero congestion loss without time sync (P802.1Qcr)
- Similar to per-flow IntServ shaping, except that:
  - All flows from one input port to same output port share the same queue
  - One shaper state machine per flow, and the right shaper applied to the packet upfront of the queue

- Fewer queues, but same number of shapers
Frame Preemption

- **Express** frames suspend the transmission of preemptable frames (802.3br and 802.1Qbu)
  - It is link local per hop, i.e., it is not IP fragmentation

- **Scheduled rocks of critical packets** in each cycle:

- Conflict excessively with **non-guaranteed packet rocks**:

- Problem solved by preemptive **sand** between the rocks:
Without Hold and Release

- Preemption isn’t instantaneous.
- Packets with less than min packet size (64 octets) left to transmit or packets less than 123 octets can’t be preempted.
- In many use cases, this delay is short enough but not in all cases.
With Hold and Release

Hold primitive can preempt packets before the start of a scheduled rock.
Preemption with Scheduling

Whole packets cross the MAC service interface.

Fragments exist only below the MAC.
mPacket Format

Legend:
Start mPacket delimiter (SMD)
SMD-E: Express mPacket
SMD-Sx: Start Fragment
SMD-Cx: Continuation Fragment

MCRC is the CRC of a non-final fragment.
Value is the same as the FCS of the frame bytes transmitted XOR FFFF0000
MCRC indicates that the frame has been preempted

Payload of each fragment (DATA plus CRC) ≥ min packet size
DEDICATED RESOURCES
TSN Configuration

- TSN configuration (P802.1Qcc)
- Information model & YANG
- Configuration Models
  - Fully Distributed Model
  - Fully Centralized Model
  - Centralized Network / Distributed User Model
Reservation Protocol

• Stream Reservation Protocol (SRP - 802.1Qat)
  – Advertises streams
  – Registers the path of streams
  – Calculates the worst-case latency
  – Establishes an AVB domain
  – Reserves the bandwidth for streams

• SRP enhancements (P802.1Qcc)

• Link-local Registration Protocol (LRP - P802.1CS)
  – Replicate a registration including changes
  – Optimized for databases on the order of 1 Mbyte
  – Not tied to bridges
NO TIME TO TALK ABOUT
Timing & Synchronization

- A profile of IEEE 1588v2 for Layer 2 Ethernet (P802.1AS-Rev)
- Redundancy
  - Redundant paths
  - Redundant GMs
- Improved scalability
- Improved support for long chains, rings
- More responsive
- Faster Grand Master change over
- Reduce Best Master Clock Algorithm (BMCA) convergence time
- Multiple domains with synchronization information
Security

• Port-based Network Access Control (802.1X)
  – Defines encapsulation of Extensible Authentication Protocol (EAP) over IEEE 802
  – Widely deployed on both wired and Wi-Fi networks

• MAC Security (MACsec) (802.1AE)
  – MACsec secures a link not a conversation
  – MACsec counters 802.1X man-in-the-middle attacks

• Secure Device Identity (802.1AR)
  – Supports trail of trust from manufacturer to user
  – Defines how a Secure Device Identifier may be cryptographically bound to a device to support device identity authentication
SUMMARY
Guaranteed data transport with bounded low latency, low delay variation, and extremely low loss.
Related Work: DetNet

- IETF Deterministic Networking WG provides Layer 3 aspects in support of applications requiring deterministic networking
- Collaboration between DetNet and TSN to define a common architecture
- DetNet covers
  - characterization of flows
  - data plane, including encapsulation
  - required forwarding behaviors
Q & A

Survey:
https://www.surveymonkey.com/r/99ieee
Thank You!
FURTHER READING
Further Reading

• Introduction to IEEE 802.1 TSN
• Tutorial on IEEE 802 Ethernet Networks for Automotive
• IEEE 802.1 TSN for Automotive Networks – flyer
• IEEE 802.1 TSN for Industrial Networks – flyer
• “A Time-Sensitive Networking Primer: Putting It All Together”
  https://drive.google.com/file/d/0B6Xurc4m_PVsZ1lzWWoxS0pTNVE/view?usp=sharing
• “Heterogeneous Networks for Audio and Video: Using IEEE 802.1 Audio Video Bridging”
  http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6595589
• Tutorial on IEEE 802 Ethernet Networks for Automotive
  http://www.ieee802.org/Tutorials.shtml
• Tutorial on IEEE 802.3br Interspersing Express Traffic (IET) and IEEE 802.1 Time-Sensitive Networking
  http://www.ieee802.org/802_tutorials/2015-03/8023-IET-TF-1501-Winkel-Tutorial-20150115_r06.pptx
• Tutorial on Deterministic Ethernet
• Tutorial on IEEE 802.1Q at IETF 86
Further Reading

- IEEE Std 802.1AE-2006 MAC Security
- IEEE Std 802.1AEbn-2011 Amendment: GCM-AES-256 Cipher Suite
- IEEE Std 802.1AEbw-2013 Amendment: Extended Packet Numbering
- IEEE Std 802.1X-2010 Port-Based Network Access Control
- IEEE Std 802.1Xbx-2014 Amendment: MAC Security Key Agreement Protocol (MKA) Extensions
- P802.1AR-Rev/D2.2 Secure Device Identity
- P802.1Xck Amendment: YANG Data Model
- RFC 7030 Enrollment over Secure Transport
IEEE 802 is here: a standards committee formed by the Computer Society

aka NesCom

aka RevCom

IEEE Standards Organization
All Those Dots ..... 

- 802.1 Bridging and Architecture  
  – generally the top of the link layer  
- 802.3 Ethernet  
- 802.11 Wireless LAN (WLAN)  
- 802.15 Wireless Specialty Networks (WSN)  
- 802.16 Broadband Wireless Access (BWA)  
- 802.18 Radio Regulatory TAG  
- 802.19 Coexistence  
- 802.21 Media Independent Handover  
- 802.22 Wireless Regional Area Networks (WRAN)  
- 802.24 Vertical Applications TAG  

TAG = Technical Advisory Group
IEEE 802.1 Working Group

- 802 LAN/MAN architecture, internetworking among 802 LANs, MANs and other wide area networks, 802 Security, 802 overall network management, and protocol layers above the MAC & LLC layers.
- Chair: Glenn Parsons
- Vice-chair: John Messenger
- Addressing and Data Center Bridging (DCB) TG
  - Chair: Patricia Thaler
- Maintenance TG
  - Chair: John Messenger
- OmniRAN TG (Model of IEEE 802 Access Networks)
  - Chair: Maximilian Riegel
- Security TG
  - Chair: Michael Seaman
- Time-Sensitive Networking (TSN) TG
  - Chair: János Farkas
IEEE 802.1 Standards

- The ones with capital letters, e.g. 802.1Q or 802.1AX are independent standards
- Amendments to these standards are identified by lower case letters e.g., 802.1Qbv or 802.1AEcg
- Periodically the amendments get merged into a revision of the main standard, e.g., 802.1Qav is now part of 802.1Q-2014
- 802.1Q can be considered as many individual standards (RFCs) integrated into a single document
  - Clauses 6 through 9 give a general overview of the 802.1Q bridge architecture
  - To get oriented on an additional area, it’s best to read the Clause titled the “Principles of <area>”
  - Once oriented, references in the subclause of Clause 5 Conformance for the relevant device can be helpful
Basic Principles

• MAC addresses are “identifier” addresses, not “location” addresses
  – This is a major Layer 2 value, not a defect!

• Bridge forwarding is based on
  – Destination MAC
  – VLAN ID (VID)

• Frame filtering for only forwarding to proper outbound ports(s)
  – Frame is forwarded to every port (except for reception port) within the frame's VLAN if it is not known where to send it
  – Filter (unnecessary) ports if it is known where to send the frame (e.g. frame is only forwarded towards the destination)

• Quality of Service (QoS) is implemented after the forwarding decision based on
  – Priority
  – Drop Eligibility
  – Time
AVB Standards

- IEEE Std. 802.1AS-2011 – generalized Precision Time Protocol (gPTP)
  - A Layer 2 profile of the IEEE 1588 Precision Time Protocol (PTP)
- IEEE Std. 802.1Qav – Forwarding and Queuing of Time-Sensitive Streams (FQTSS):
  - Specifies Credit-Based Shaper (CBS)
- IEEE Std. 802.1Qat – Stream Reservation Protocol (SRP)
  - Registration and reservation of time-sensitive streams
- IEEE Std. 802.1BA – AVB Systems
  - Provides an overall AVB architecture and AVB profiles
- CBS + SRP to provide delays under 250 μs per bridge
TSN Standards and Projects

- 802.1Qbu – Frame Preemption
- 802.1Qbv – Enhancements for Scheduled Traffic
- 802.1Qca – IS-IS Path Control and Reservation (PCR)
- 802.1Qch – Cyclic Queuing and Forwarding
- 802.1Qci – Per-Stream Filtering and Policing
- P802.1Qcc – Stream Reservation Protocol (SRP) Enhancements & Performance Improvements and TSN configuration
- P802.1Qcj – Auto-attach to PBB services
- P802.1Qcp – YANG Data Model
- P802.1Qcr – Asynchronous Traffic Shaping (ATS)
- P802.1AS-Rev – Timing and Synchronization - Revision
- 802.1CB – Frame Replication and Elimination for Reliability
- P802.1CM – Time-Sensitive Networking for Fronthaul
- P802.1CS – Link-local Registration Protocol (LRP)
Forwarding Process in 802.1Q

- Reception Port State
- Transmission Port State
- Active topology enforcement (8.6.1)
- Ingress (8.6.2)
- Frame filtering (8.6.3)
- Egress (8.6.4)
- Flow metering (8.6.5)
- Queueing frames (8.6.6)
- Queue management (8.6.7)
- Transmission selection (8.6.8)
- Transmission Port

Filtering Database