# Choosing the Right TSN Tools to meet a Bounded Latency

Don Pannell

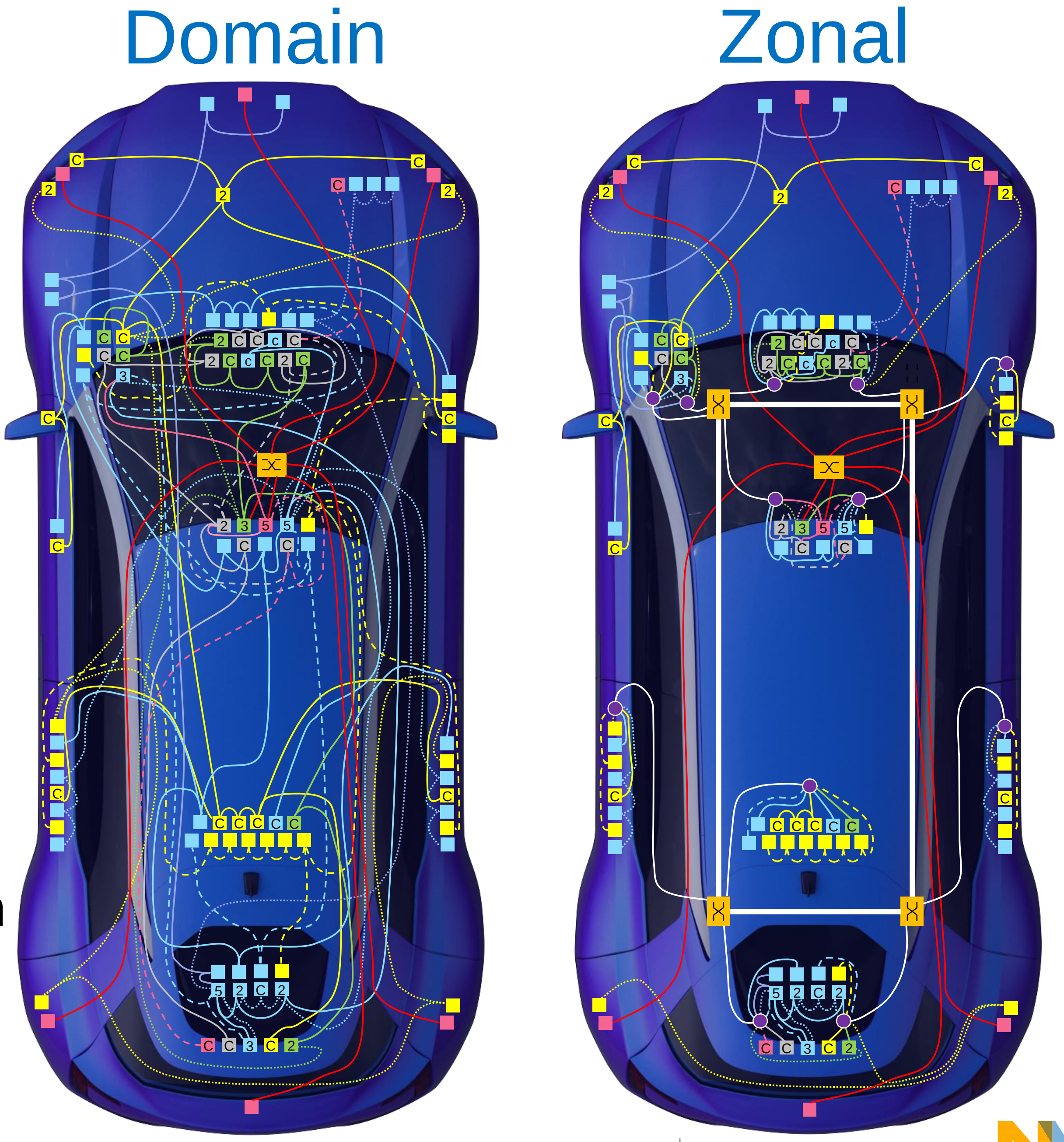
Fellow
Automotive Ethernet Networking

Sep 2019 | IP Tech Day - Detroit

SECURE CONNECTIONS FOR A SMARTER WORLD

# The Need

- Ethernet's high speeds saves wires in Zonal networks
- And Zonal networks bring new requirements that (TSN) solves
  - Multiple Domains using the same wire
  - Yet each Domain needs to know it's data will get delivered in the needed maximum time – as it no longer has its own dedicated wire!
  - How to guarantee & plan the maximum bounded latency for each flow is the focus of this presentation



# Overview

- This presentation focuses on the TSN standards that affect bounded latency of flows through the Automotive Ethernet network
- It briefly lists the unique problems each of these Time Sensitive Networking (TSN) standard solves & the relative ‘costs’ of using each tool
- Based on these numbers, a per-hop metric is proposed, to help determine which TSN tool should be used and when
- This priority of tool usage, makes the job of “Engineering” the network easier via to the step-by-step process described

# List of Available TSN Tools for Controlling Latency

| Standard's Name: | Also Known As: |
| --- | --- |
| Strict Priority | 802.1p-1998 / QoS |
| Forwarding & Queueing for Time-Sensitive Streams | 802.1Qav-2009 / Credit Based Shaper or FQTSS |
| Enhancements for Scheduled Traffic | 802.1Qbv-2015 / Time Aware Shaper |
| Frame Preemption | 802.1Qbu-2016 & 802.3br-2016 |
| Cyclic Queueing & Forwarding | 802.1Qch-2017 / Peristaltic Shaper |

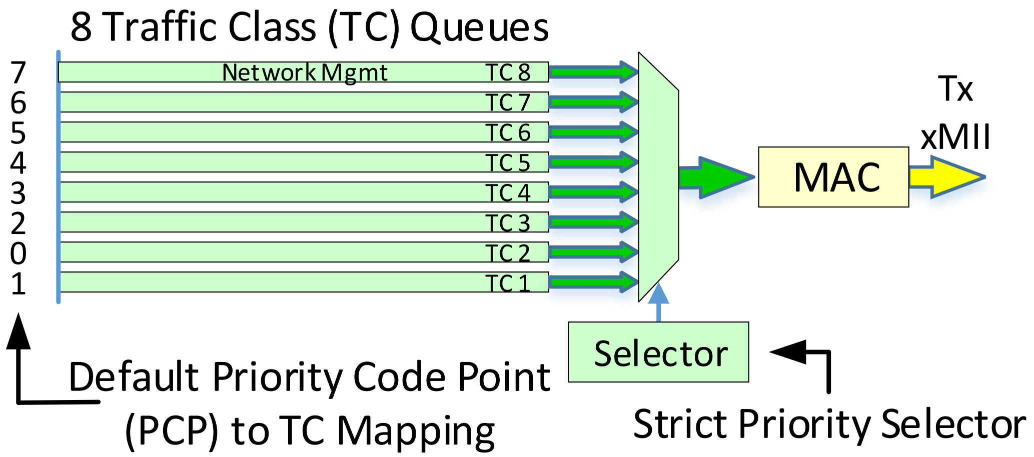
Note: New tools will become available, e.g., the Asynchronous Traffic Shaping, 802.1Qcr

# The Shaper Standards:

# What Problems the Standards Solve & How They were Envisioned to be Used

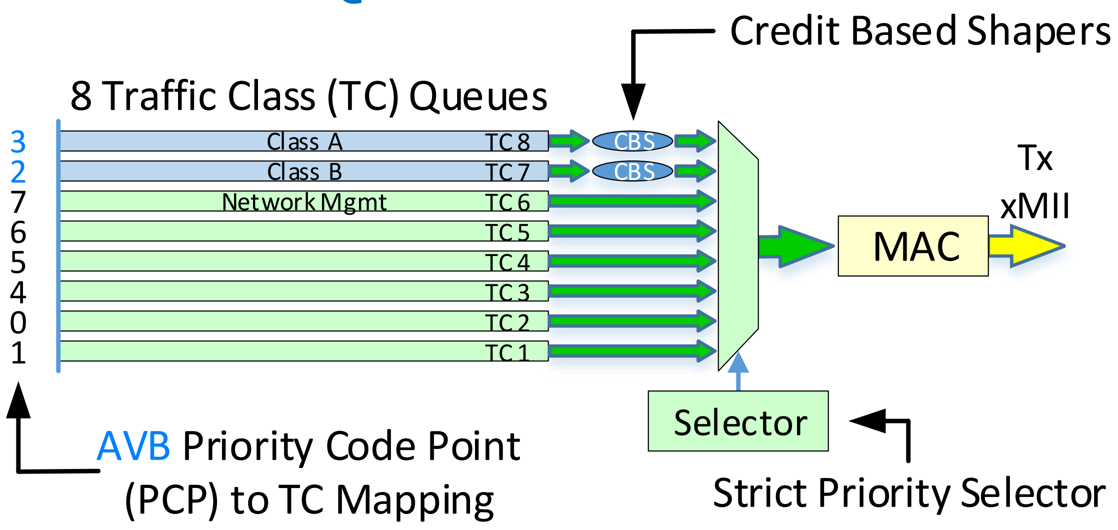
# Strict Priority Shaper (Strict) – 802.1p-1998

- Priority solves the problem that some frames are more important than others
- It was needed so Network Management could work




- Management frames had to get through the Network in order to fix any Network problem
- The Strict hardware selector is defined as: “Frames are selected from the corresponding queue for transmission only if all queues corresponding to numerically higher values of traffic class … are empty at the time of selection.”

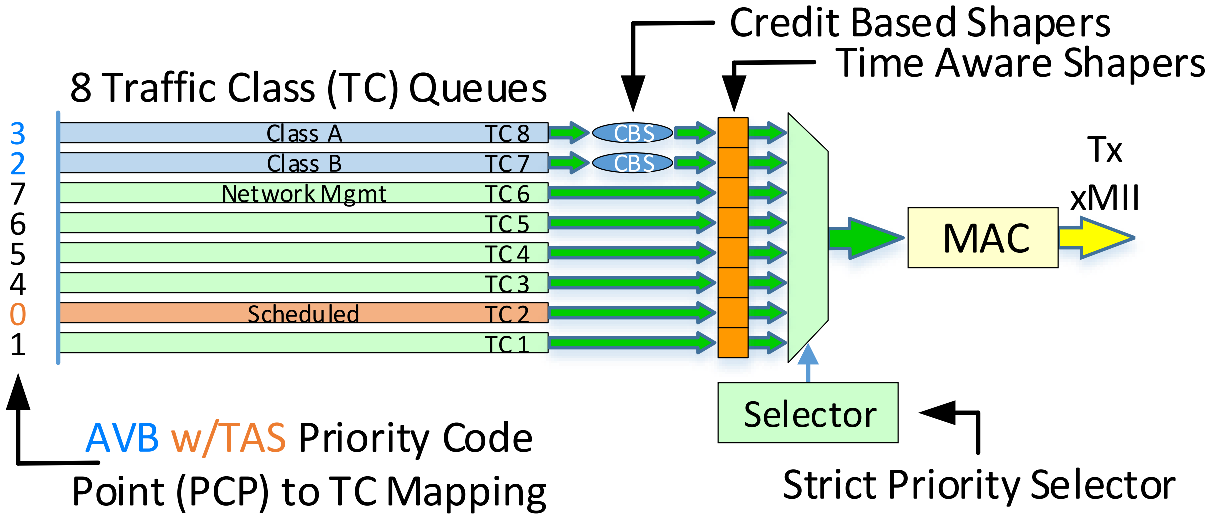
# Credit Based Shaper (CBS) – 802.1Qav-2009




- CBS solves the problem that long bursts of data are really bad for the Bridges
- It was needed so Reserved frames are not dropped
- It caps the bandwidth a queue can transmit with hardware
- It de-bursts flows in hardware so that optimized software stacks that try to burst can be used (for streams that are not self-shaping)
  - I.e., audio from a USB drive vs. audio from a microphone or radio
  - It allows very small bursts of data to ‘catch-up’ due to momentary interference so the Reserved data rate can be maintained
  - In AVB, PCP 2,3 are re-mapped above Mgmt since they don’t use 100% of the wire

# Time Aware Shaper (TAS) – 802.1Qbv-2015

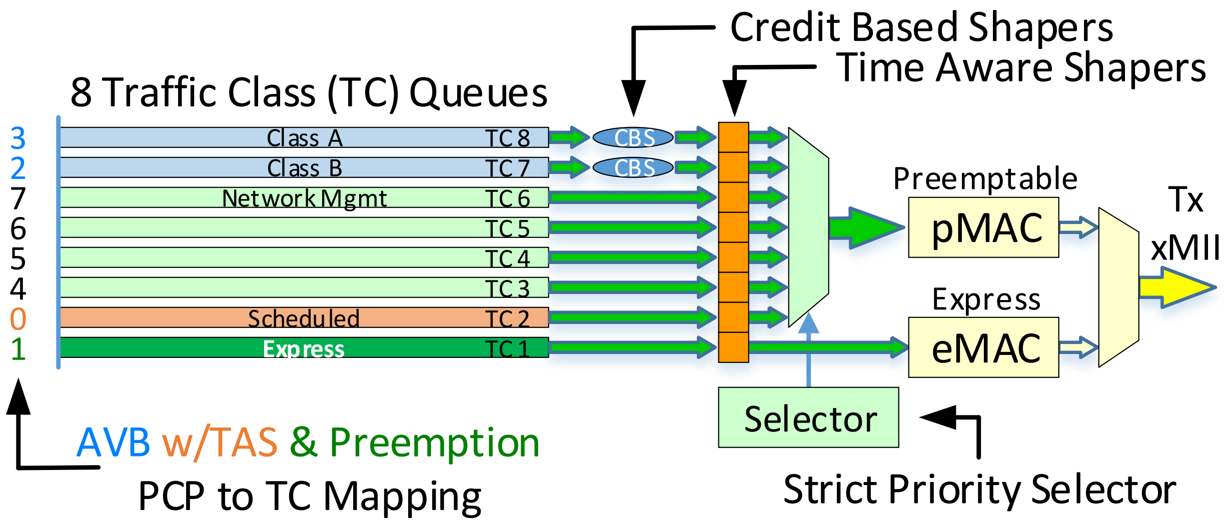
- TAS delivers the theoretically lowest possible latency for scheduled periodic data
- It uses significant bandwidth, so is best used as a last resort




- Transmission Gates are added for ALL queues just before the Strict Priority Selector
  - Following a defined periodic schedule, the gates on the queues are opened or closed for a period of time – allowing critical traffic to pass without interference
  - ALL queues are time-gated, but really only 1 or 2 queues are actually “Scheduled” and the “non-Scheduled” queues are left open during the remainder of the time
  - Any TC can be used for “critical” scheduled traffic (TC 2 in the figure)

# Preemption – 802.1Qbu-2016 & 802.3br-2016

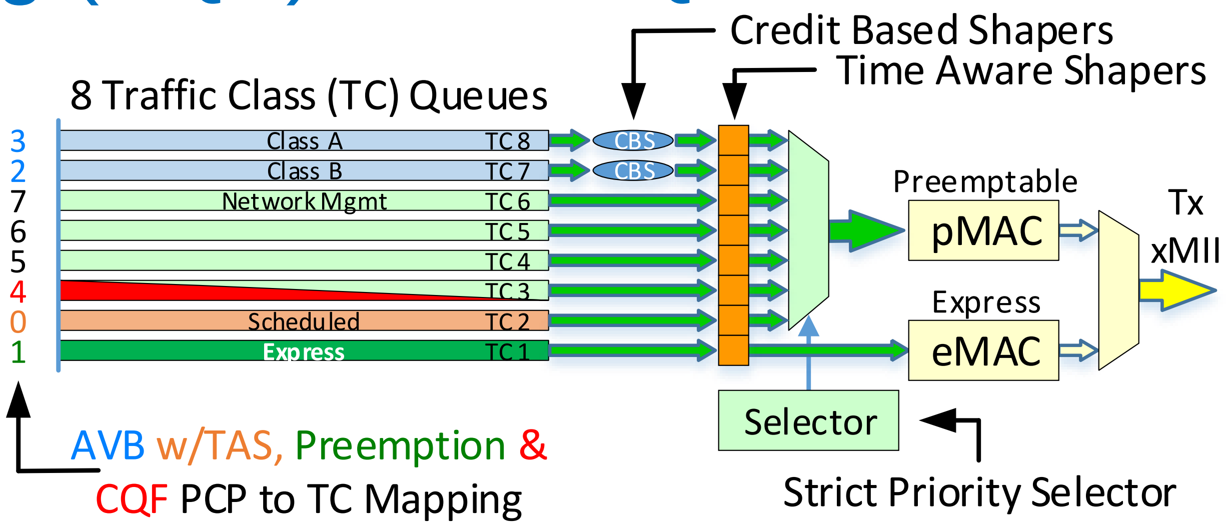
- Preemption delivers very low latency for a limited set of non-scheduled data
- Preemption gains the most on slow data links (100 Mb/s)




- Two 802.3 MACs are used, one for “normal” preemptable traffic (pMAC) and one for “express” preempting traffic (eMAC)
  - Only 1 level of preemption is supported & frames < 126 bytes can’t be preempted
  - 802.1 allows connection of each TC queues to either MAC – if more than one queue connects to a MAC, the Strict selector algorithm is assumed
  - In the figure, TC 1 is effectively above all the other TC’s since it can preempt them!

# Cyclic Queuing & Forwarding (CQF) – 802.1Qch-2017

- A fully accurate worst case equation for the Credit Based Shaper (CBS) does not exist
- CQF is a replacement for the Credit Based Shaper for very large, higher speed Networks




- CQF receives data in an Even time window and transmits it in an Odd time window and visa versa using parallel buffers/queues
  - Can be built using two Traffic Class queues with Even/Odd alternate Time Aware Shaper gating to select only one queue at a time, as long as the input mapping to the two queues is synchronously time switched as well

# The Shaper Standards: Their Metrics

# Latency TSN Tool Comparison

| TSN Tool | Silicon Complexity | Engineering Complexity | Wire Efficiency | Comments |
| --- | --- | --- | --- | --- |
| Strict Priority | Low | Easy | 100% | Needed component, but it is not deterministic by itself |
| Credit Based Shaper | Medium | Easy | 100% | All CBS queues are deterministic + next highest TC (for Mgmt) |
| Time Aware Shaper | Medium | Hard (>1 TC)<br>Medium (1 TC) | - Guard band<br>- Idle Opens | All TAS queues are deterministic |
| Frame Preemption | High | Medium but only 1 level deep | - Fragment overhead | Fragmentation may affect determinism on the other flows |
| Cyclic Queueing | Medium | Medium | 50% - FE<br>90% - GE | Effects are not well know<br>% numbers due to Guard band |

Engineering Complexity is the expected user difficulty or effort, needed to get proper results

Wire Efficiency is how much data can go down the wire – this includes critical data and background data

Note: Preemption is the only standard that requires support on both sides of the wire

# Per Hop Latency – Credit Based Shaper

- Class A ≈ $_t$Interval + $_t$MaxFrameSize
  - $_t$Interval = observation interval of the Class (125 uSec for AVB – but can be changed)
  - $_t$MaxFrameSize = the maximum size of an interfering frame + gaps, etc.
  - This is a good rule-of-thumb equation that results in slightly larger numbers than the equation in 802.1BA-2011 subclause 6.5
- Class B ≈ $_t$Interval + $_t$MaxFrameSize + $_t$TimeForAllHigherFrames
  - $_t$TimeForAllHigherFrames = the time to transmit all Class A frames (+ gaps, etc.) for the duration of Class B's tInterval (which is typically multiple Class A tIntervals)
- Class C ≈ $_t$Interval + $_t$MaxFrameSize + $_t$TimeForAllHigherFrames
  - Where tTimeForAllHigherFrames includes Class A & Class B frames
- Etc.

# Per Hop Latency – Credit Based Shaper – part 2

- Class A ≈ $_t$Interval + $_t$MaxFrameSize
  - For a 64 byte frame in a 125uSec observation interval the worst case # is:
  - On a 100BASE link ≈ 125 uSec + 124 uSec = 249 uSec per hop
  - On a 1000BASE link ≈ 125 uSec + 13 uSec = 138 uSec per hop
- The observation interval is a significant portion of these latencies
  - Lower worst case latency numbers are possible on 1000BASE links by using shorter observation intervals, but don't go below the time of $_t$MaxFrameSize
  - But lowering this number reduces latency at the cost of capacity
    - 1000 vs 100 is either 10x lower latency or 10x the capacity or somewhere in between

Note: The simplified equation on the previous page is useful for calculating the worst case latency range for a fully loaded (i.e., 75% bandwidth allocation) on a Class A link. A scheduling tool needs to use the equation that is in IEEE 802.1BA. Also see: [http://www.ieee802.org/1/files/public/docs2011/ba-boiger-per-hop-class-a-wc-latency-0311.pdf](http://www.ieee802.org/1/files/public/docs2011/ba-boiger-per-hop-class-a-wc-latency-0311.pdf)

# Per Hop Latency – Time Aware Shaper

- Store & Forward with Gate Open ≈ $_t$Device + $_t$FrameSize
  - $_t$Device = the delay through a Store & Forward bridge
    - Good Rule-of-Thumb is 2 x 512 bit times + Cable delay
      - or 10.5 uSec for 100BASE & 1.5 uSec for 1GBASE
  - $_t$FrameSize = the size of the frame passing through the bridge
      - For more information see http://www.ieee802.org/1/files/public/docs2011/new-pannell-latency-options-0311-v1.pdf
- For a 64 byte frame the worst case # is:
  - On a 100BASE link ≈ 10.5 uSec + 5.2 uSec = 15.7 uSec per hop
  - On a 1000BASE link ≈ 1.5 uSec + 0.5 uSec = 2.0 uSec per hop

# Per Hop Latency – Frame Preemption

- Store & Forward w/ Preemption ≈ $_t$Device + $_t$FrameSize + $_t$Framelet
  - $_t$Device = the delay through a Store & Forward bridge
    - Good Rule-of-Thumb is 2 x 512 bit times + Cable delay
      - or 10.5 uSec for 100BASE & 1.5 uSec for 1GBASE
  - $_t$FrameSize = the size of the frame passing through the bridge
  - $_t$Framelet = 126 bytes + overhead, as this size interfering frame can't be preempted
      - For more information see http://www.ieee802.org/1/files/public/docs2011/new-avb-pannell-latency-options-1111-v2.pdf
- For a 64 byte frame the worst case # is:
  - On a 100BASE link ≈ 10.5 uSec + 5.2 uSec + 11.7 uSec = 27.4 uSec per hop
  - On a 1000BASE link ≈ 1.5 uSec + 0.5 uSec + 1.2 uSec = 3.2 uSec per hop

Note: Preemption requires support on both sides of the wire

# Per Hop Latency – Cyclic Queueing

- Even/Odd Buffering ≈ $_{t}$Interval * 2
  - $_{t}$Interval = observation interval of the Class (must be at least 2x $_{t}$MaxFrameSize)
    - $_{t}$MaxFrameSize = the maximum size of an interfering frame + gaps, etc. which is needed because, when using Qbv for the Even/Odd queue selection, a guard band is needed
    - The 1$^{st}$ $_{t}$Interval is the reception of the data
    - The 2$^{nd}$ $_{t}$Interval is the transmission of the data where the worst case frame is at the end
- For a 64 byte frame the worst case # is:
  - On a 100BASE link ≈ 250 uSec * 2 = 500 uSec per hop
  - On a 1000BASE link ≈ 125 uSec * 2 = 250 uSec per hop

Note: The $_{t}$Interval for 1000BASE could be smaller, but the limit needs to be studied

# Latency TSN Tool Comparison in Lowest Latency Order

| TSN Tool | Engineering Complexity | Wire Efficiency | Worst Case Latency for the Examples | Ranking |
|---|---|---|---|---|
| Time Aware Shaper | Hard (>1 TC)<br>Medium (1 TC) | - Guard band<br>- Idle Opens | 15.7 uSec FE Hop<br>2.0 uSec GE Hop | 2 |
| Frame Preemption | Medium but only 1 level deep | - Fragment overhead | 27.4 uSec FE Hop<br>3.2 uSec GE Hop | 3 |
| Credit Based Shaper | Easy | 100% | 249 uSec FE Hop<br>138 uSec GE Hop | 1 |
| Cyclic Queueing | Medium | 50% - FE<br>90% - GE | 500 uSec FE Hop<br>250 uSec GE Hop | 4 |
| Strict Priority | Easy | 100% | Can’t determine | N/A |

1 = Multiple queues can be used with different observation intervals/latencies

2 & 3 = Assuming only 1 TC is used for very limited, very critical traffic, 2 is more available with lower latencies

4 = Effectively requires GE or faster, it’s not generally available, and its effects over different link speeds is not well understood

# The Shaper Standards: Which Tool to Use First

# Proposed Tool Usage Order

- Process the critical flows in the lowest to highest latency order
- First insure the total bandwidth through any link is not more that 75% loaded with these flows
  - This # could go a bit higher, but 60% to 75% is a good place to start
- Start with the Credit Based Shaper
  - Select an Observation Interval that is as large as possible that delivers the required latency over the path(s) the flow uses
  - If the default 125 uSec Observation Interval is too long, reduce it, but don't go < 125 uSec on 100BASE links
  - If that doesn't work, use Time Aware Shaping &/or Preemption as last resorts
    - As these are limited resource that are less wire efficient
    - Subtract any wire efficiency loss as used bandwidth toward the 75% critical flow limit

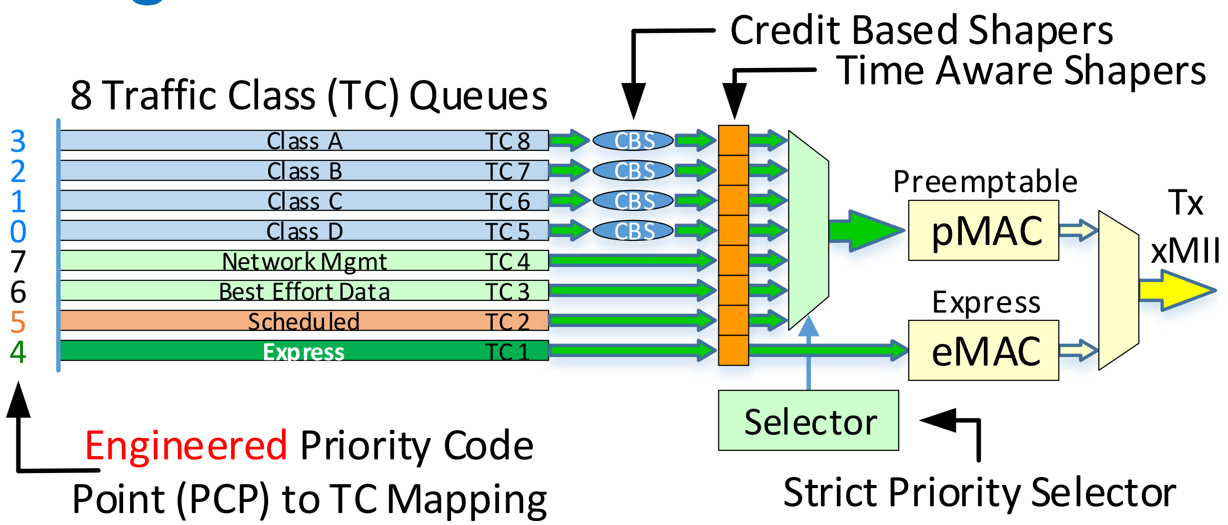
# Proposed Tool Usage Order

- Multiple Credit Based Shaper's w/increasing Observation Intervals can be used – More than two Classes can be used if needed!
  - Start by loading each Class no more than 20% of the link's bandwidth
  - Keep in mind that the sum total of ALL Reserved flows, & their frame (IFG, etc.) & scheduler overhead (Qbv & Qbu), must not exceed 75% of any one link's bandwidth
    - If this happens, try an alternate path for the flow
    - 60% may be a better starting number so that new flows can fit in easier
      - CAN network loading is typically started at 50% so new messages can be added
        - To fix bugs & oversites
        - And to add new features
- Network Mgmt must be the highest non-CBS Traffic Class
- The remaining "non-Reserved" flows will use the remainder of the unused bandwidth in a Best Effort fashion

# Summary

# Summary & Proposed Queueing Model

- These standards are designed to work together
- Multiple different data delivery requirements/latencies can be supported on the same wire




- The Credit Based Shaper is not limited to just Audio & Video data & it is not limited to the AVP Profile's plug-&-play parameters
- There is a current limit of 8 Priority Code Points (PCP) that are effectively used to indicate the "type of service" a flow needs
- Automotive networks are Engineered, but let the hardware enforce the needed guarantees to make the job much simpler

SECURE CONNECTIONS
FOR A SMARTER WORLD

www.nxp.com

# Disclaimers

- This is a really hard concept that has been simplified so that an easy starting point on which shaper to use for a target flow can be made
- The listed latency numbers are in the correct range but they are still estimates.  For example:
  - A generic bridge delay is used vs. the actual delay in the specific bridge you are using
  - All latency numbers use 64 byte data frames.  In most cases, larger data frames will impact the latency numbers.
  - Cable delay is mostly ignored – which is approximately 80ns for 15 meters
  - Look at the referenced presentations & others on the same subject in the same areas
- As a rule-of-thumb for link speed conversion in a bridge:
  - For $_{t}$MaxFrameSize & $_{t}$Framelet use the egress link speed, for $_{t}$FrameSize use the ingress link speed and for $_{t}$Device use the faster link speed of the two

# IEEE 802.1 Automotive AVB and TSN Standards Handout

| | Transport | Synchronization | Stream Reservation | Quality of Service | Redundancy | Security |
|---|---|---|---|---|---|---|
| **AVB**<br>802.1BA-2009<br>The AVB Profile | 1722-2011<br>Media Transport Protocol | 802.1AS-2011<br>gPTP | 802.1Qat-2010<br>SRP (now Q clause 35) | 802.1Qav-2009<br>Credit Based Shaper (now Q clause 34) | - | 802.1X-2010<br>802.1Xbx-2014<br>802.1Xck-2018<br>Network Access |
| **TSN** | 1722-2016<br>Adds CAN, FlexRay, LIN, + more Audio/Video Transports | 802.1AS-Rev<br>Redundant gPTP | 802.1Qcc-2018<br>Enhanced SRP<br>802.1Qca-2015<br>Path Control & Reservation | 802.1Qbv-2015<br>Time Aware Shaper<br>802.1Qbu-2016 & 802.3br-2016<br>Preemption<br>802.1Qch-2017<br>Cyclic Queue Forwarding<br>802.1Qcr<br>Asynchronous Shaping | 802.1CB-2017<br>Frame Replication & Elimination<br>802.1AS-Rev<br>Redundant gPTP | 802.1Qci-2017<br>Policing<br>802.1AEcg-2017<br>(end-to-end) MACSec |

Standards without an appended year are not completed yet.

Updated 9-2019

# IEEE 802.3 Automotive Ethernet PHY Standards Handout

| | 10 Mbit/s | 100 Mbit/s | 1000 Mbit/s | 2500 Mbit/s | 10 Gbit/s | Next |
|---|---|---|---|---|---|---|
| **MAC Interface**<br>Digital/<br>SERDES | SNI, xMII/<br>SGMII<br>OC-SGMII | xMII/<br>SGMII<br>OC-SGMII | xGMII/<br>SGMII<br>OC-SGMII<br>1000BASE-X | OC-SGMII<br>2500BASE-X | USXGMII<br>XFI | ? |
| **Media Interface**<br>Single Twisted Pair | 802.3cg<br>10BASE-T1S<br>15m Point to Point<br>25m Multi-Drop<br>10BASE-T1L<br>1000m Point to Point | 802.3bw-2015<br>100BASE-T1<br>15m Point to Point | 802.3bp-2016<br>1000BASE-T1<br>15m Point to Point | 802.3ch<br>2500BASE-T1<br>15m Point to Point | 802.3ch<br>10GBASE-T1<br>15m Point to Point | ? |

Media Interface (PHY) Standards without an appended year are not completed yet. Updated 9-2019