Time Aware Shaper Options

Review of latency and variability, a proposal plus other topics

Daniel Sexton
GE Global Research
Niskayuna, NY

imagination at work

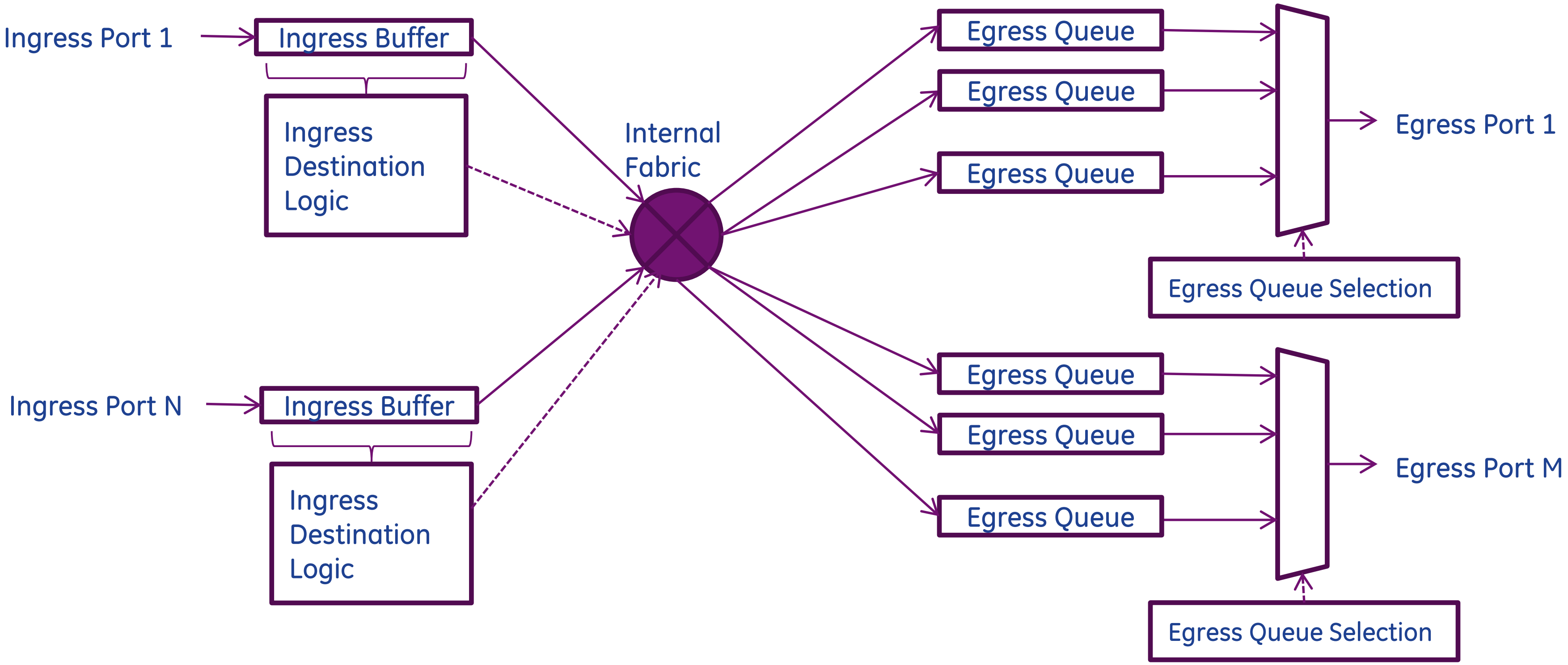
Introduction

This presentation will:

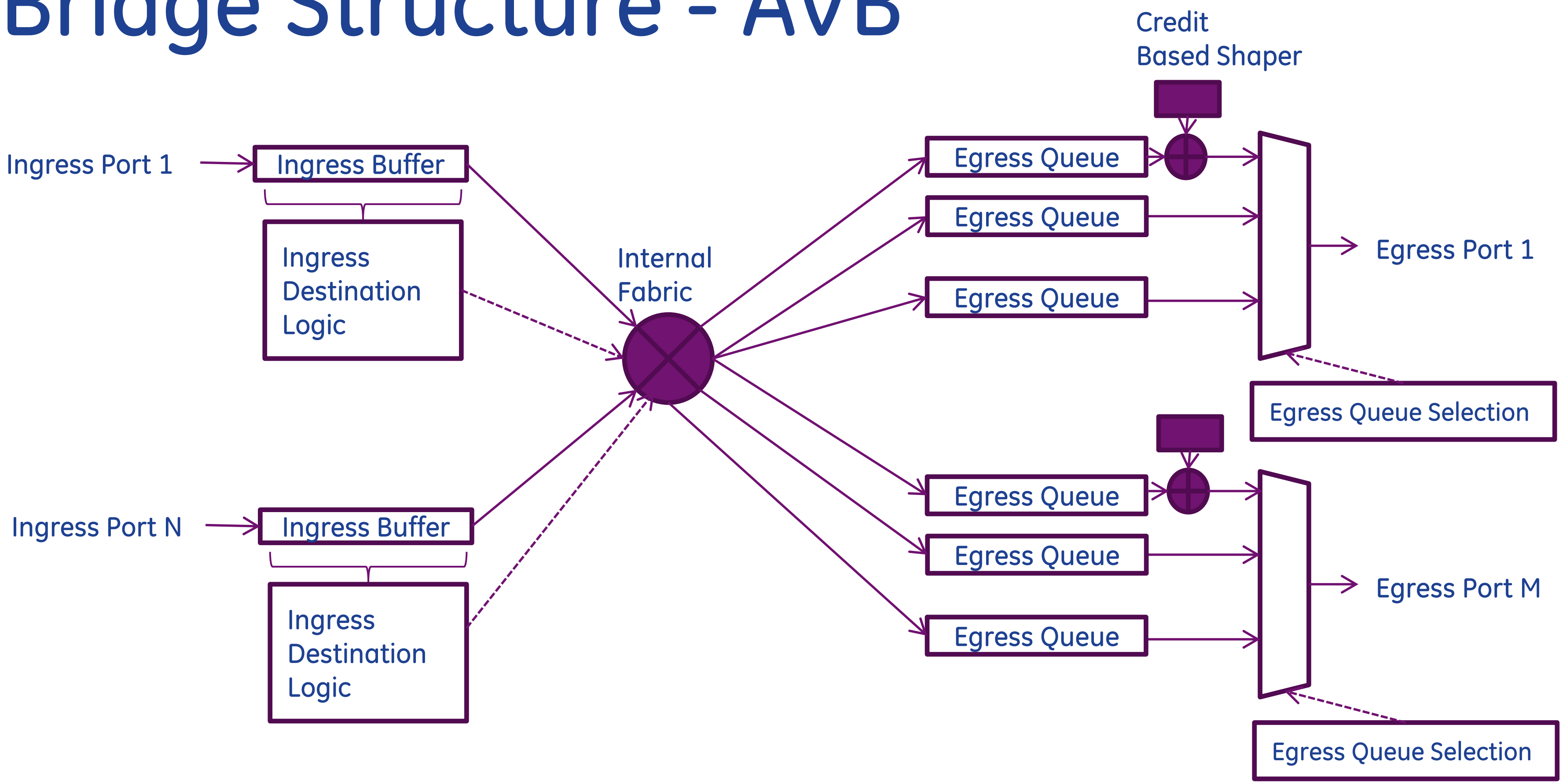
1. Review entitlements for latency and latency jitter for various time aware shapers.
2. Propose several options for time aware shapers and TSN operation.
3. Ask questions related to the implications of these various options.
4. Suggest an option that balances complexity with performance.

imagination at work

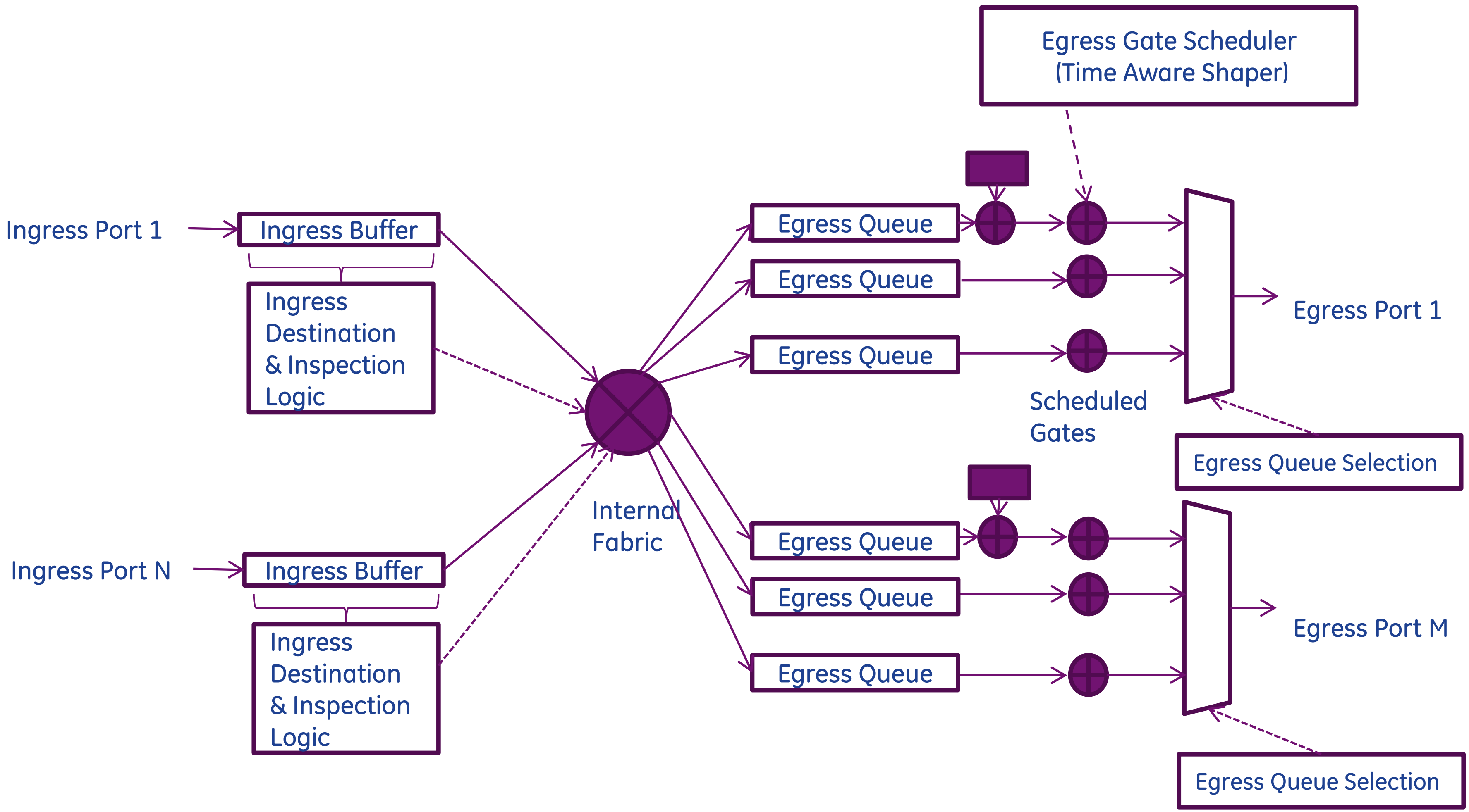
Bridge Structure - basic



Bridge Structure - AVB



Proposed Bridge Structure – time aware shaper



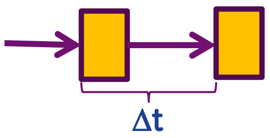
imagination at work

Timing analysis assumptions

Full duplex GE (1ns/b)

Cable v=1.85e8 m/s (100m = 540ns)

Times are from arrival of first bit to arrival of first bit in next network device.



Guard time: Example – Class A traffic: 1542 octets (includes IFG) = 12.336usec

imagination at work

Bridge Structure – Time aware egress Entitlement.

Assumptions:

1. Frame fully buffered and integrity checked before passed to proper egress queue.
2. 100m cable – 540ns
3. 1GE, 100 byte UDP packet: (170 byte times):1.360us
    1. IP Header, 20 bytes. UDP Header 8 bytes, Preamble through Ethertype including 802.1Q tag, 26 bytes, FCS, 4 bytes, IFG=12 bytes. L=26+20+8+100+4+12=170.
4. Assume no internal bridge delays
5. Egress MAC is free (all traffic has been suspended, guard time applied).
6. Egress MAC was freed early enough to account for slot time if required.
7. Best case added latency: 1.9usec (1.36us + 100m cable).

Note: Assumptions different than: new-pannell-latency-options-0311-v1, p3.

(Citation included slot time in latency, I have assumed that is not required or can be pipelined out).

Bridge Structure – Time aware egress Entitlement with Preemption.

Assumptions:

1. Frame fully buffered and integrity checked before passed to proper egress queue.
2. 100m cable – 540ns
3. 1GE, 100 byte UDP packet: (170 byte times):1.360us
    1. IP Header, 20 bytes. UDP Header 8 bytes, Preamble through Ethertype including 802.1Q tag, 26 bytes, FCS, 4 bytes, IFG=12 bytes. L=26+20+8+100+4+12=170.
4. Assume no internal bridge delays
5. Packet Preemption delay (interfering packet is already known to be lower priority) – τ_p
6. The system is full duplex, slot time is not required.
7. Best case added latency: 1.9usec (1.36us+ 100m cable), Worst case latency 1.9usec + τ_p.
8. Preemption overhead is unknown until operation is defined by 802.3.

imagination at work

Bridge Structure – Time aware egress Entitlement (cut through)

Assumptions:

1. Frame begins transmission after the 802.1 header & Ether type is received – egress MAC clear, guard time applied.
2. 100m cable – 540ns
3. Packet can start transmission after 26 byte times

Preamble(8)+MAC DA/SA(12) + 802.1Q(4) + Etype(2)

4. Assume no internal bridge delays
5. Best case added latency: 748ns (208ns + 100m cable)

imagination at work

Bridge Structure – Time aware egress Entitlement (cut through) with preemption

Assumptions:

1. Frame begins transmission after the 802.1 header & Ethertype is received.
2. 100m cable – 540ns
3. Packet can start transmission after 26 byte times if Egress MAC is free

Preamble(8)+MAC DA/SA(12) + 802.1Q(4) + Etype(2)

4. Assume no internal bridge delays
5. Best case added latency: 748ns (208ns + 100m cable),

   worst case latency = 748ns + τ_p

1. If Low latency packet is of absolute highest priority, this delay is not incurred.

What is the best we could possibly Do. Blind cut-through

Assumptions:

1. Frame begins transmission immediately after preamble, guard time applied
2. 100m cable – 540ns
3. Packet starts transmission after preamble
4. Egress is known as part of schedule
5. Assume no internal bridge delays
6. Best case added latency: 0.604us (64ns + 100m cable)

imagination at work

What is the best we could possibly Do with preemption? Blind cut-through.

Assumptions:

1. Frame begins transmission immediately after preamble, guard time applied.
2. 100m cable – 540ns
3. Packet starts transmission after preamble + start of frame delimiter (64ns)
4. Egress is known as part of schedule
5. Assume no internal bridge delays
6. Best case added latency: 0.604us (64ns + 100m cable),

   worst case: 0.604us + $\tau_{p.}$

imagination at work

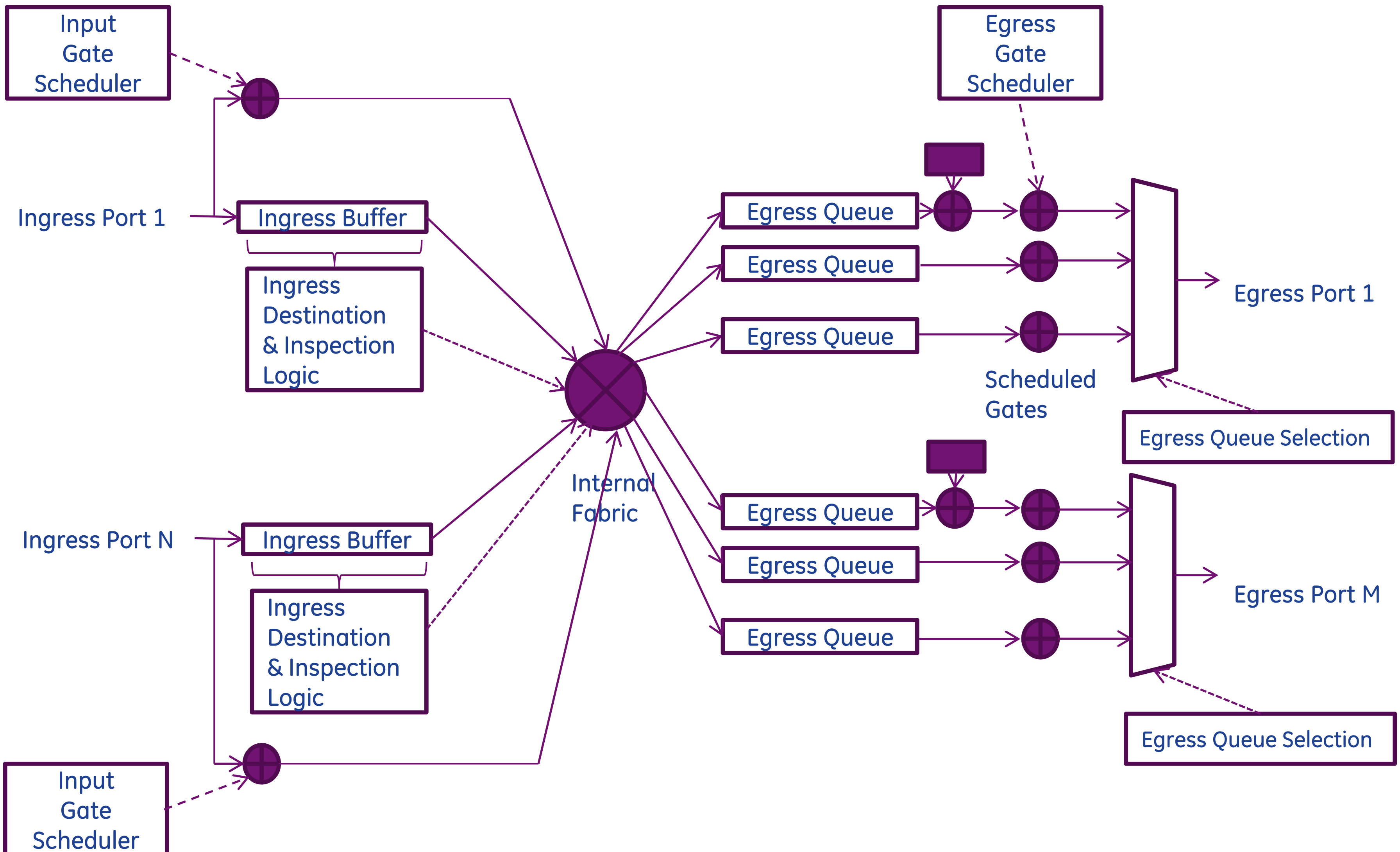
Single Bridge Delay Summary

	Cable Length				
Gen 2 Options	1M	10M	100M	Max Variability	Guard Time
802.1 VLAN[1]	74.9usec	75usec	75.3usec	72.5usec	0
AVB Gen 1	13.6usec	13.7usec	14.2usec	11.2usec	0
Buffered + Preemption	1.365usec $+\tau_p$	1.41usec$+\tau_p$	1.9usec$+\tau_p$	τ_p	0[2]
Buffered	1.365usec	1.41usec	1.9usec	0	12.336usec[3]
Cut through + preemption	0.213usec $+\tau_p$	0.262usec $+\tau_p$	0.748usec $+\tau_p$	τ_p	0[2]
Cut through	0.213usec	0.262usec	0.748usec	0	12.336usec[3]
Blind Cut Through + Preemption	0.069usec $+\tau_p$	0.118usec $+\tau_p$	0.604usec $+\tau_p$	τ_p	0[2]
Blind Cut Through	0.069usec	0.118usec	0.604usec	0	12.336usec[3]

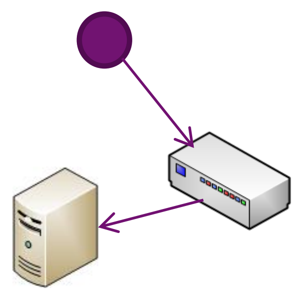
1. Assumes Highest Priority Frame with jumbo frames.
2. Lost network utilization of τ_p if a preemption occurs
3. This is lost network time, some may be recovered if a packet is queued that can fit in the window,

imagination at work

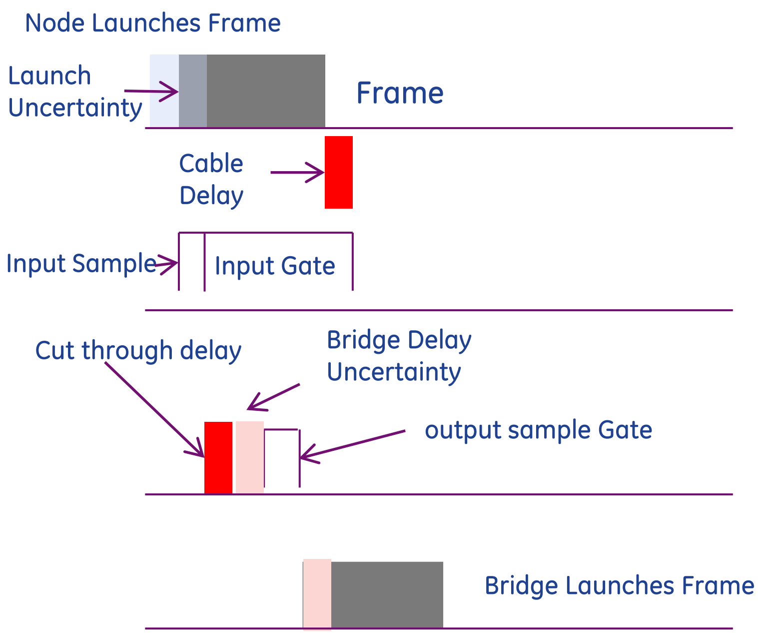
Proposed Bridge Structure – Blind cut through

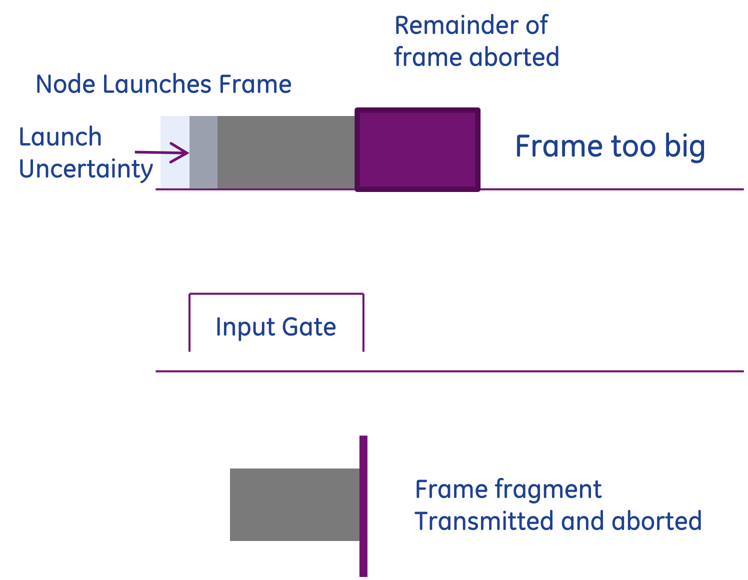


Blind cut-through frame timing
Scenario: Frame too large for time window



What happens if the frame doesn't fit?





This policing could also be applied to egress queues and time aware shaper

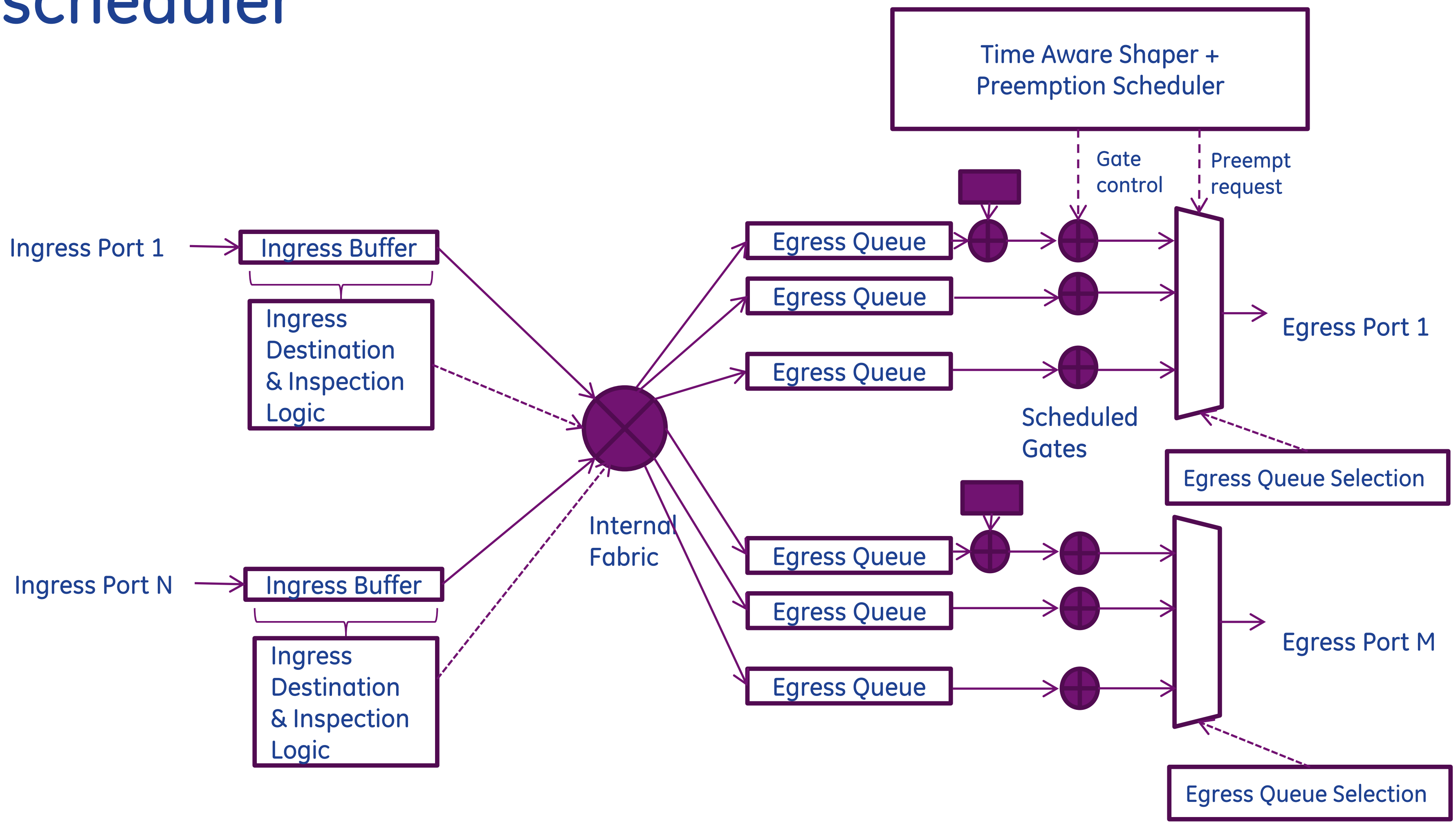
imagination at work

Modification – Preemption Scheduler

1. Same scenario as Time Aware Shaper
2. If a packet is being sent on the egress port it can be preempted on demand.

imagination at work

Proposed Bridge Structure – Preemption scheduler



imagination at work

Single Bridge Delay Summary

	Cable Length				
Gen 2 Options	1M	10M	100M	Max Variability	Guard Time
Buffered + Preemption	1.365usec $+\tau_p$	1.41usec$+\tau_p$	1.9usec$+\tau_p$	τ_p	0^2
Buffered	1.365usec	1.41usec	1.9usec	0	12.336usec
Buffered + Early Preemption	1.365usec	1.41usec	1.9usec	0	τ_p
Cut through + preemption	0.213usec $+\tau_p$	0.262usec $+\tau_p$	0.748usec $+\tau_p$	τ_p	0^2
Cut through	0.213 usec	0.262usec	0.748usec	0	12.336usec
Cut through + Early preemption	0.213 usec	0.262usec	0.748usec	0	τ_p
Blind Cut Through + Preemption	0.069usec $+\tau_p$	0.118usec $+\tau_p$	0.604usec $+\tau_p$	τ_p	0^2
Blind Cut Through	0.069usec	0.118usec	0.604usec	0	12.336usec[3]
Blind Cut Through Early Preemption	0.069usec	0.118usec	0.604usec	0	τ_p

Optimization 1 -Frame Stuffing

Purpose: To minimize unusable network bandwidth.

1. Inserting a frame which is currently in an ingress buffer into the unused guard time.
2. Length of frame is known
3. Frame must fit within the guard time
4. Short low priority frames can get ahead of longer higher priority frames (but not time scheduled frames)

imagination at work

Optimization 2 – Fixed Delay

Purpose: To minimize unusable network bandwidth at the expense of added latency.

1. Always delay outbound frame by maximum guard time.
   a. Eliminates guard times
   b. Eliminates variability
   c. Delays all frames *(From a controls perspective – not an optimization)*

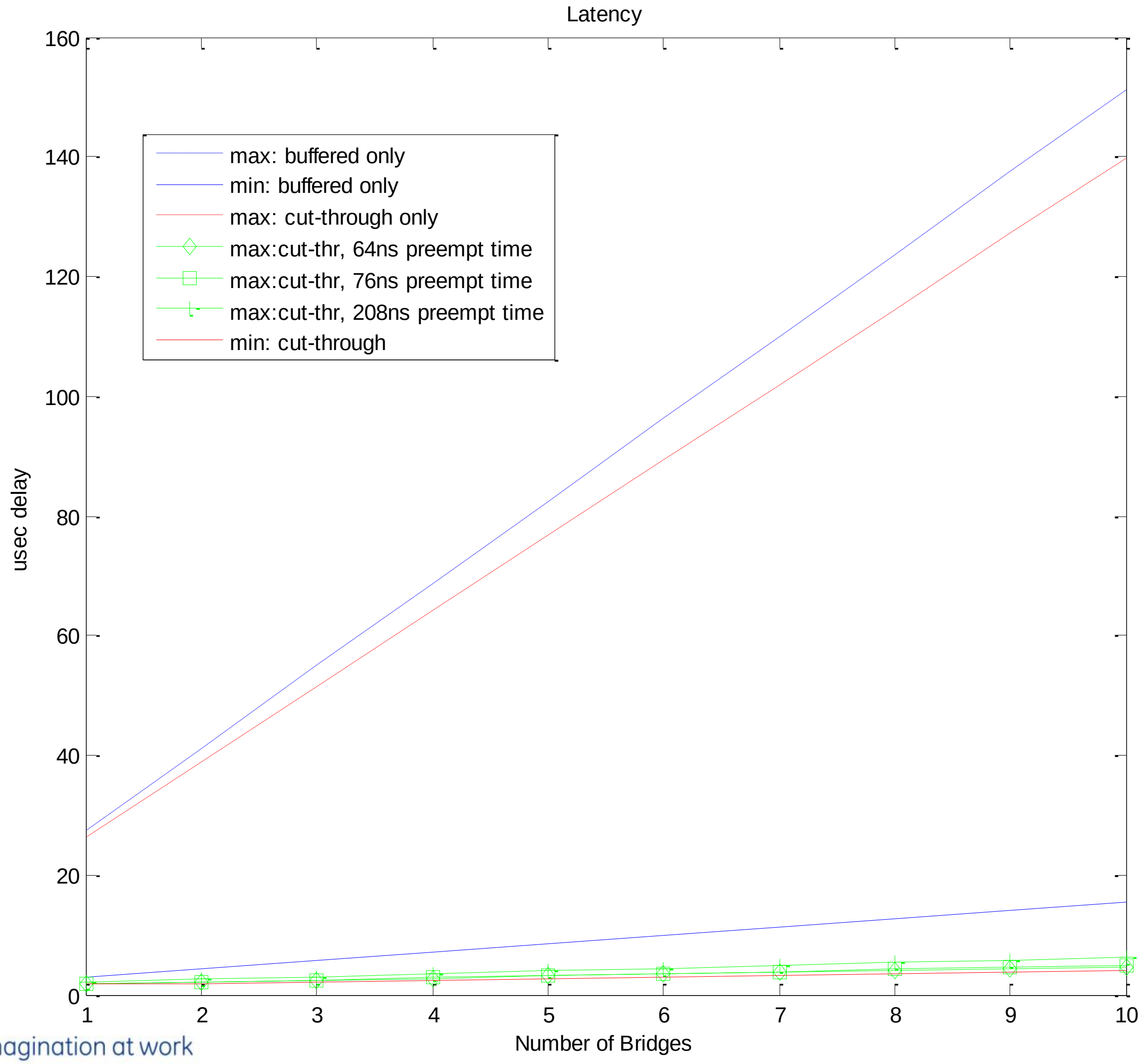
imagination at work

Single Bridge Delay Summary – Optimization 2

	Cable Length				
Gen 2 Options	1M	10M	100M	Variability	Guard Time
Buffered + Preemption	1.365usec $+\tau_p$	1.41usec$+\tau_p$	1.9usec$+\tau_p$	0	0
Buffered	13.7usec	13.8usec	14.2usec	0	0
Cut through + preemption	0.213usec $+\tau_p$	0.262usec $+\tau_p$	0.748usec $+\tau_p$	0	0
Cut through	12.549 usec	12.6usec	13.1usec	0	0
Blind Cut Through + Preemption	0.069usec $+\tau_p$	0.118usec $+\tau_p$	0.604usec $+\tau_p$	0	0
Blind Cut Through	12.4usec	12.45usec	12.9usec	0	0

imagination at work

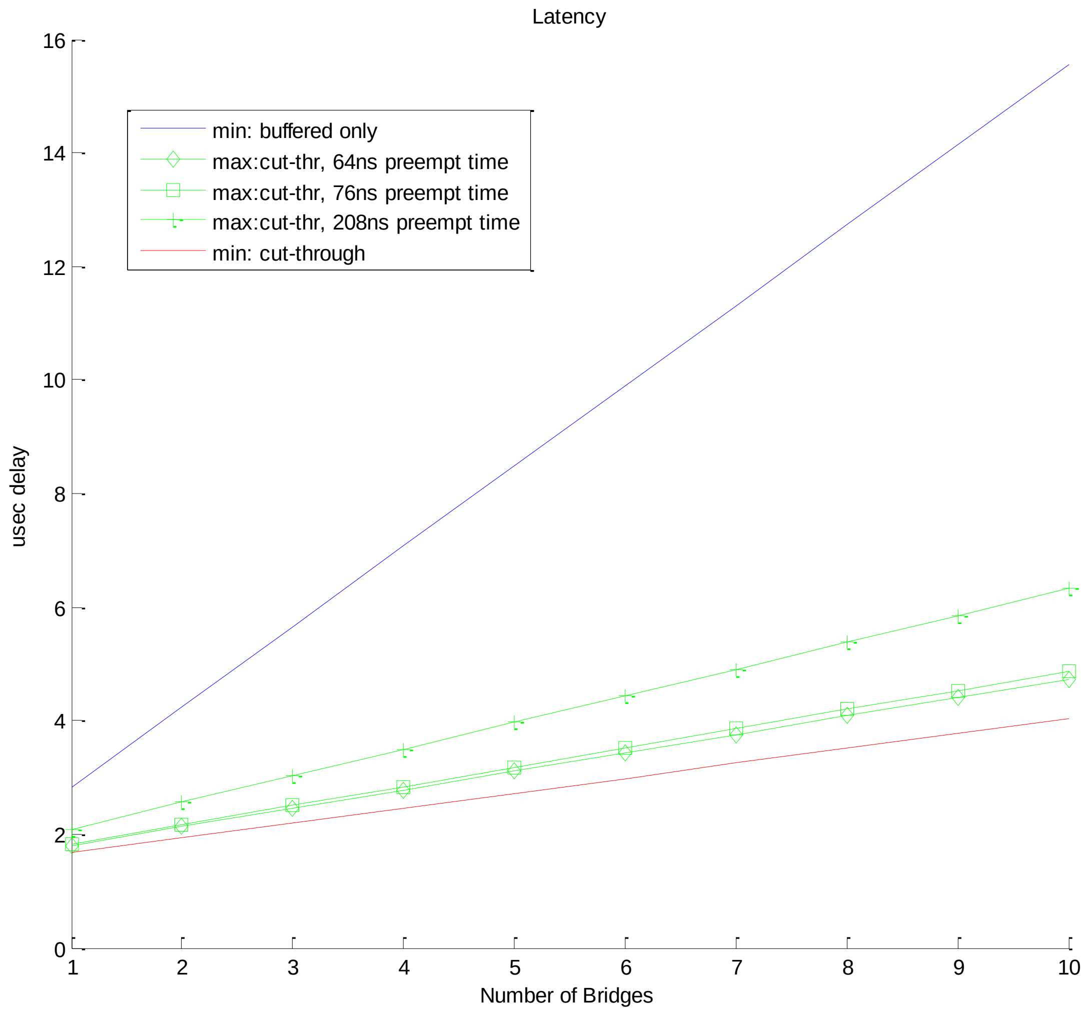
Frame Latency



From xmit bit on edge node to last recv bit on edge node.
Assumes:
100 byte packet
1GE
10m cable

imagination at work

Frame Latency



If preemption delay is 76ns or less, scheduled preemption provides little value
Better/simpler to add worst case preemption delay to all scheduled traffic.

Functional Proposal – Time Aware Shaper + Preemption Scheduler

1. Add egress gate that is used to make the egress queues visible to the output selection process to select which frame is being sent (already in proposed revision)
   a) Can be used to schedule the output of frames from an egress port. (Edge devices don't need precise timing)
   b) Can be used to delay the draining of queues.
   c) Multiple queues can be selected at the same time.
   d) Output frame can be delayed to create fixed deterministic delays (Optimization 2)
2. Allow a frame to begin to be forwarded to the egress queue once the 802.1 header + Ethertype has been received (cut-through).
3. Preemption occurs whenever a higher priority frame is available at the egress port and can also be scheduled as a timed event (if preemption overhead is >76nsec).
   a) Real time traffic needs to be at highest priority to be deterministic. If there is a higher priority frame available it can preempt a deterministic control frame.
   b) By scheduling a preemption event early we can eliminate the variable latency preemption might take.
4. If Preemption overhead is <76nsec, I believe the complexity the preemption scheduler adds outweighs it's value.

Q&A

1. What standards need to change?
2. Do we need a profile standard (other than 802.1BA Gen 2) to specifically cover highly engineered highly reliable control networks? – I would propose we do and that it has various levels for real time applications including safety critical applications.
3. Should cut-through be part of the standard? I would suggest that it is part of an application profile standard that would include latency specifications, Ex: 802.1?? Bridging for real time applications.
4. How should a switch describe itself? (Latency, shapers, etc, what protocol changes needed?)
5. What happens if a higher priority frame arrives at an egress queue while a control frame is being delivered to the egress port – will it preempt the control packet?
   a) Yes – but this behavior will break the determinism therefore the egress gates for high priority queues must be closed at least one max frame length ahead of time (as if we did not have preemption) or we can schedule automatic preemption of the egress queue.
6. What about Forking and Joining – Siemens proposal?
   a) I see no functional effect but I do see complexity in scheduling events.
7. What about a rate change, ingress is low rate, egress is high rate or ingress high and egress low or bridging to wireless?
   a) Cut-through won't work, frames must be fully buffered in the ingress queue before being loaded to the egress queue. We could make cut-through work in the high to low rate path but why? My recommendation is to limit cut-through to only to like media bridging.
8. How are gates programmed?
   a) Switch acts like a player piano playing the same tune over and over again.
9. What should the cycle time and granularity be?
   a) At least 100msec cycle time?
   b) Granularity ~ one "smallest frame" duration.
   c) Abilities beyond these simple requirements could be left to the vendors.

While I have your attention

Security:

1. Ingress Frame Verification:
   a. Ability to verify incoming frame headers and lengths to identify errors. Not required in all cases but may be part of a secure controls network profile.
2. What do we need/want from 802.1x?
   a. Should we require the use of EAPOL?
   b. Support for MAC Bypass?
3. TSNs should automatically isolate unauthenticated devices to a low bandwidth allocated stream. Traffic on unauthenticated ports get tagged to a low priority VLAN with limited time allocation based on egress schedule.
4. If MAC Bypass is used, the low bandwidth allocated stream should only flow over a VLAN to the authentication server.

   (The standard does not have to specify this, it can be done as an application standard elsewhere)
5. Frames that do not pass frame inspection (Described earlier) should not be forwarded, frames that are longer than expected should be aborted. The port where ingress frame errors occur should be quarantined – relocated to low bandwidth and limited access. Errors should be reported.

Conclusions

1. 802.1Q Proposal: *"avb-tj-peristaltic-shaper-in-clause-8-style-0313-v1.pdf" Tony Jeffree;* appears correct, will adequately cover the time aware shaper
2. Additional clause similar to above to cover preemption scheduling depending on the outcome of 802.3 TSN work group (may be a moot point if time to preempt is small)
3. Creation of a Controls based profile standard work group to create an application profile standard (Like 802.1BA – Audio Video Bridging (AVB) Systems).
4. Cut-through functionality need not be part of 802.1Q however minimum bridge latencies and latency variability specifications should be part of (3).
5. I currently see no changes needed in 802.1x (although I am not an expert). Security to the network fabric must be included in (3).