AVB Latency Math

v4 – Sept, 2010

802.1 AVB Face to Face – York, UK

Don Pannell - dpannell@marvell.com

Need

- Need to define standardized maximum latency calculations for AVB devices
- Then this standardized latency can be used in 802.1Qat's worst case latency it reports for each hop in the network
- Need equations for Talkers and Bridges
- Link Speed needs to be taken into account
- Bridge Fan-in needs to be included

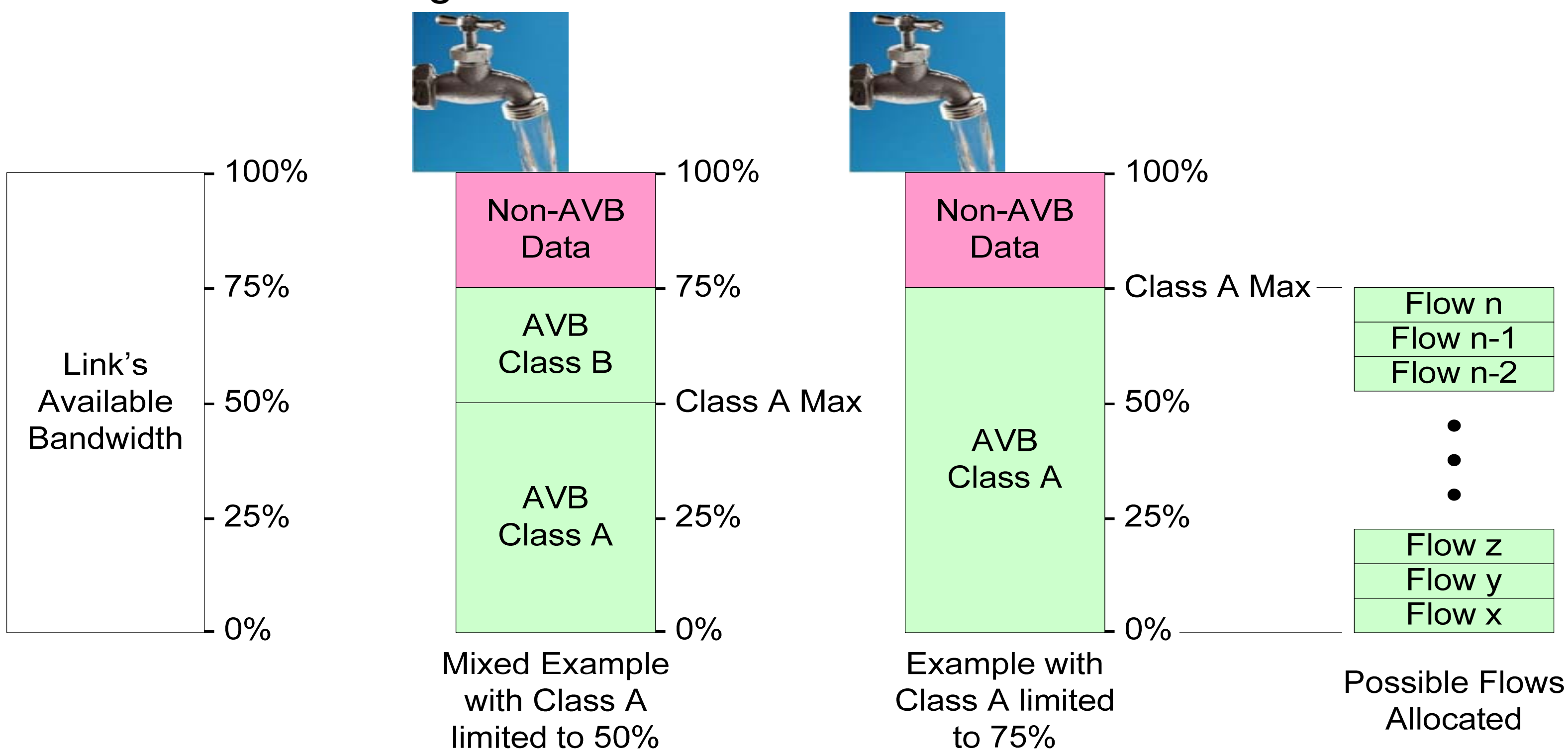
Problem Breakdown

- **Start with Class A equations**
- **Start with Talkers**
  - Create equations for Talkers 1st
  - Talkers are simpler to understand
  - We need to agree on what a Talker may generate (AVB flow wise) before we can drive this worst case into Bridges
- **Then move onto Bridges**
- **Then move onto Class B equations**

Talker: Class A Latency

Link Realities

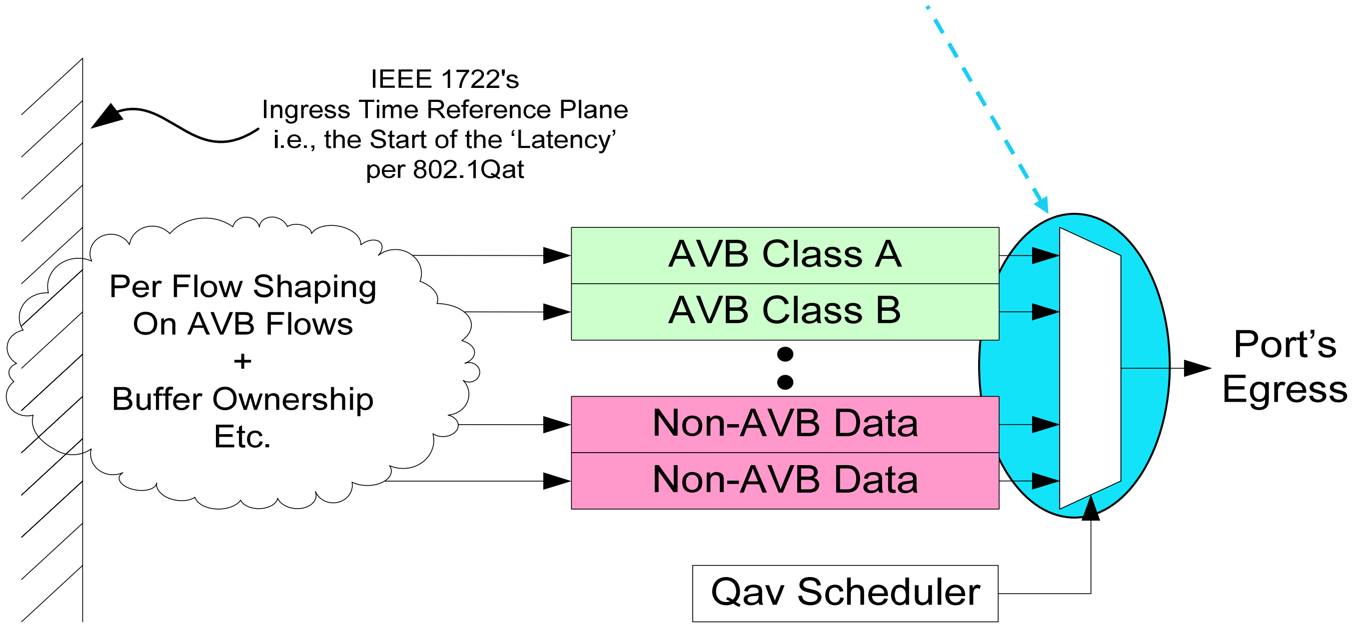
- **Limited Available Bandwidth determined by Link's speed**
- **Class A's Max setting**
  - 75% or less (software settable)
- **Position of a flow amongst its peers during any interval**
  - It could be 1st during one interval and then last on the next



Talker's Simplified Egress Model

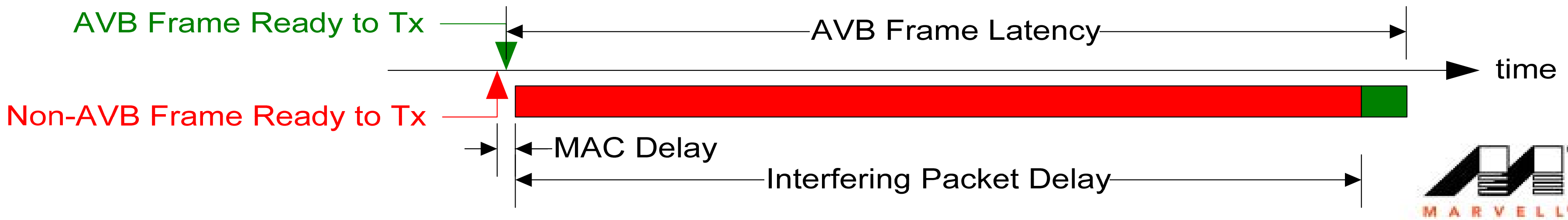
Worst Case Class A:

- Class A uses 75% of Link
- At t_0 all Class A flows need to be transmitted
- At $t_{0\ -1\ clk}$ a Max size Non-AVB frame starts out the port
- This Max size frame goes 1st then the Class A's



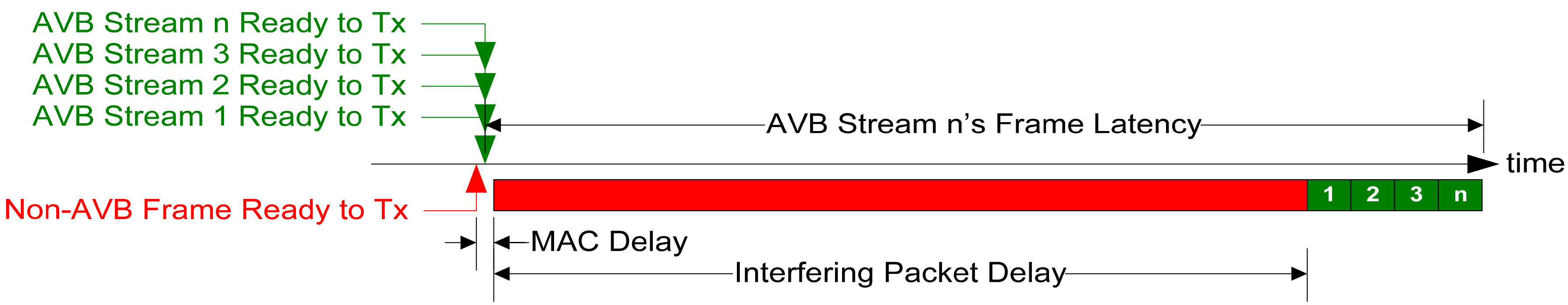
Latency Due to a Non-AVB Frame

- **Talker Latency = the delay from the Transfer of the Packet's Ownership (CPU to MAC) to the Complete Transmission of the Packet on the wire (otherwise called "Last bit in to Last bit out")**
- **Without Congestion this is the 'delay of the MAC' + Tx of the frame**
  - MAC Delay is VERY small for full wire speed MACs
  - It is delay from 'Ownership of Buffer to the MAC' until the MAC starts Tx
- **But with Congestion with another frame its the delay to finish transmitting the other frame + a min. IFG (assuming the MAC is full wire speed with back-to-back frames)**
  - Shown as the **Green** AVB Frame below
- **Worst case = Max size Non-AVB frame + a portion of MAC Delay**
- **MAC Delay ≤ 1 802.3 Slot Time or 512 bit times (propose to use this as a standard fudge unit)**
- **Therefore Max Latency due to a Non-AVB Frame = Transmission time of a Max Size Non-AVB Frame + 512 bit times**

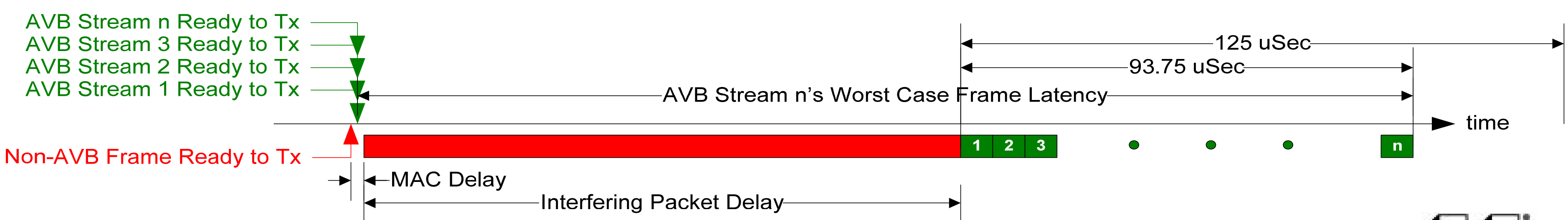


Latency Due to AVB Frames

- Max Latency due to this and Other AVB Frames = Transmission time of the Other AVB Frames + Tx time for this Frame
- But this is variable as the number of streams increase



- **So what is the worst case?**
- **Worst case is the Max % Delay of the stream's Class Measurement Interval**



Class A Talker's Latency

- **Worst Case Talker's Latency of a Class A Stream is =**

EQ 1: $$\text{Max Latency}_{\text{Class A Talker}} = t_{\text{MAC Delay}} + t_{\text{Tx Max Frame}} + t_{\text{Class A Max}}$$

- **In English it's the: MAC Delay (1 slot time or 512 bit times)**
- **Plus Time to Transmit a Max Non-AVB Frame**
  - Could be 1522 Bytes + IFG + Preamble or 2000 Bytes + IFG + Preamble
  - Want to manage this per encapsulation (MACSec, Provider) & management
- **Plus Time Allotted for Class A's Maximum Allocation**
  - Could be up to 75% of 125 uSec (or 93.75 uSec) or less

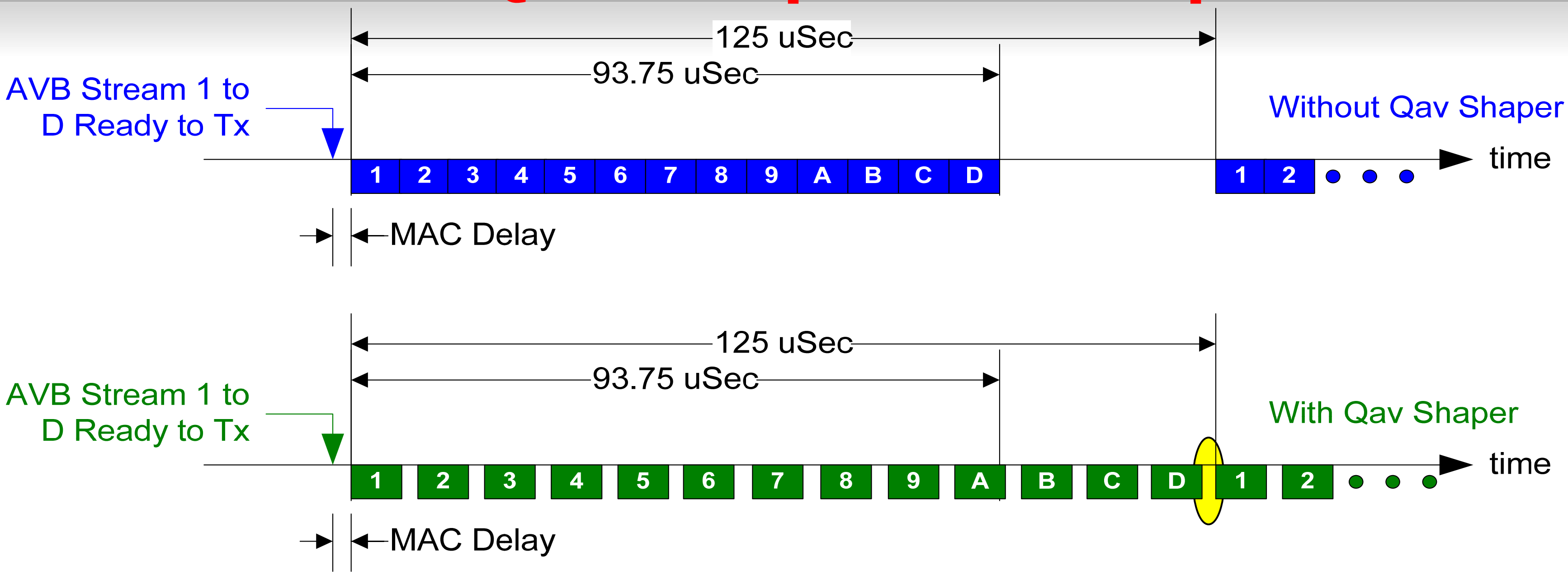
- **For Class A on FE (assume 1522 Max Non-AVB + 75% for Class A):**
  - = 5.12 uSec + 123.36 uSec + 93.75 uSec = 222.23 uSec
- **For Class A on GE (assume 1522 Max Non-AVB + 75% for Class A):**
  - = 0.512 uSec + 12.336 uSec + 93.75 uSec = 106.60 uSec

- **% Class A Parameters don't change 10x with 10x the speed!**
  - But you get 10x number of streams!
  - And these parameters are manageable per interface
  - Some Talkers know how many streams they will support (microphone) and can define a pre-defined lower Max Class A Usage

Now Add the Qav Shaper . . .

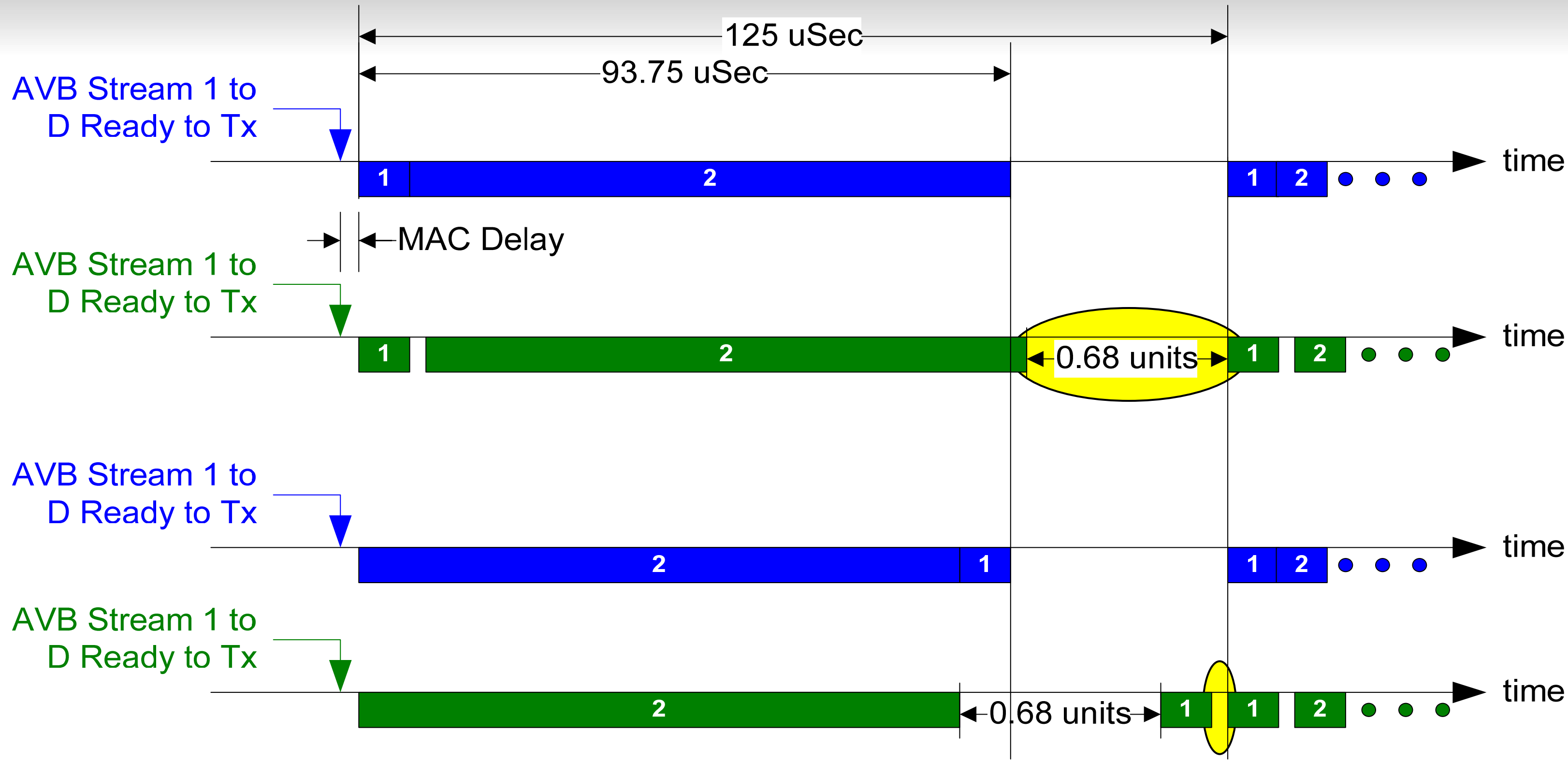
- **Need to look at the Shaper's effect without congestion**
- **Look at Largest Number of Min Size AVB Frames to fill the allowed Class Usage – the math is below:**
- **75% of 125 uSec = 93.75 uSec (Max AVB)**
- **93.75 uSec / 80 ns/byte = 1171.875 bytes**
  - Or 13.95 64 byte frames w/IFG & Preambles
- **Assuming 13 same size AVB frames**
  - 13 x 90 bytes = 1170 bytes
  - Therefore use 13 x 70 byte AVB frames w/20 overhead bytes each (12 IFG + 8 Preamble)
  - 1562.5 bytes (125 uSec) – 1170 bytes = 392.5 bytes / 13 = 30.19 extra shaping bytes between each frame
  - 13 * (90 + 30) = 1560 bytes or 124.800 uSec
  - AV Data + Qav Gaps have to fit under 125 uSec or the data will start pushing into the next Class Measurement Interval

Effect of Qav Shaper at FE Speed



- **Blue is no Qav, Green is with Qav**
- **The last bit of the last AVB frame is at 125 uSec minus the IdleSlope time of the smallest possible frame**
  - Could combine IdleSlope with MAC Delay and just say its 125 uSec
  - Shaper rates must be set such that under no congestion the next set of frames can start without any added delay – but also where there is the maximum gap between AVB frames
- **This is true too for GE or 10 GE speeds!**
  - This is the worst case for GE and faster – not the previous equation!

Look at Two Mixed Size Flows



- **The last bit of the last AVB frame is <u>still</u> at 125 uSec minus the IdleSlope time of the smallest frame**
  - This isn't true when the smallest frame is first (top 2 examples)
  - But it is when the smallest frame is last! (bottom 2 examples)
  - Flow 1 could end up in this position if the 75% is allocated!

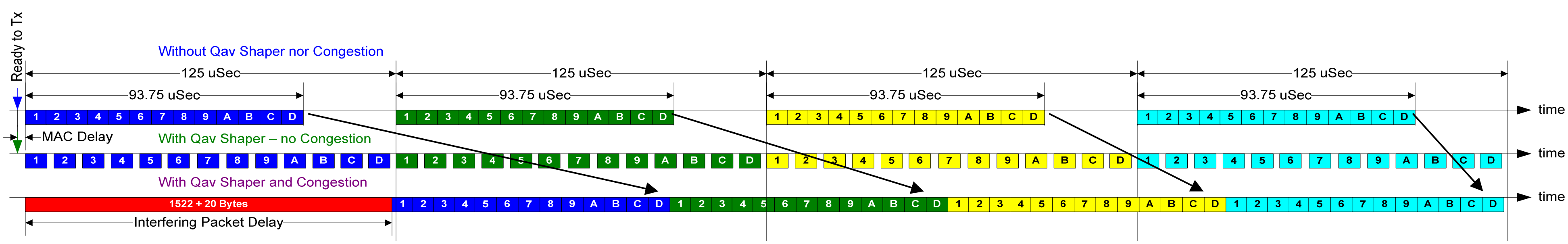
By Definition . . .

- **The entire idea of the Qav Shaper is to take the AV data and spread it out equally under the no congestion situation**
- **Therefore, the worst cast latency of a fully allocated link is at the extreme of the Class Measurement Interval**
- **Now need to add congestion back in . . .**

Effect of Qav + Congestion at FE Speed

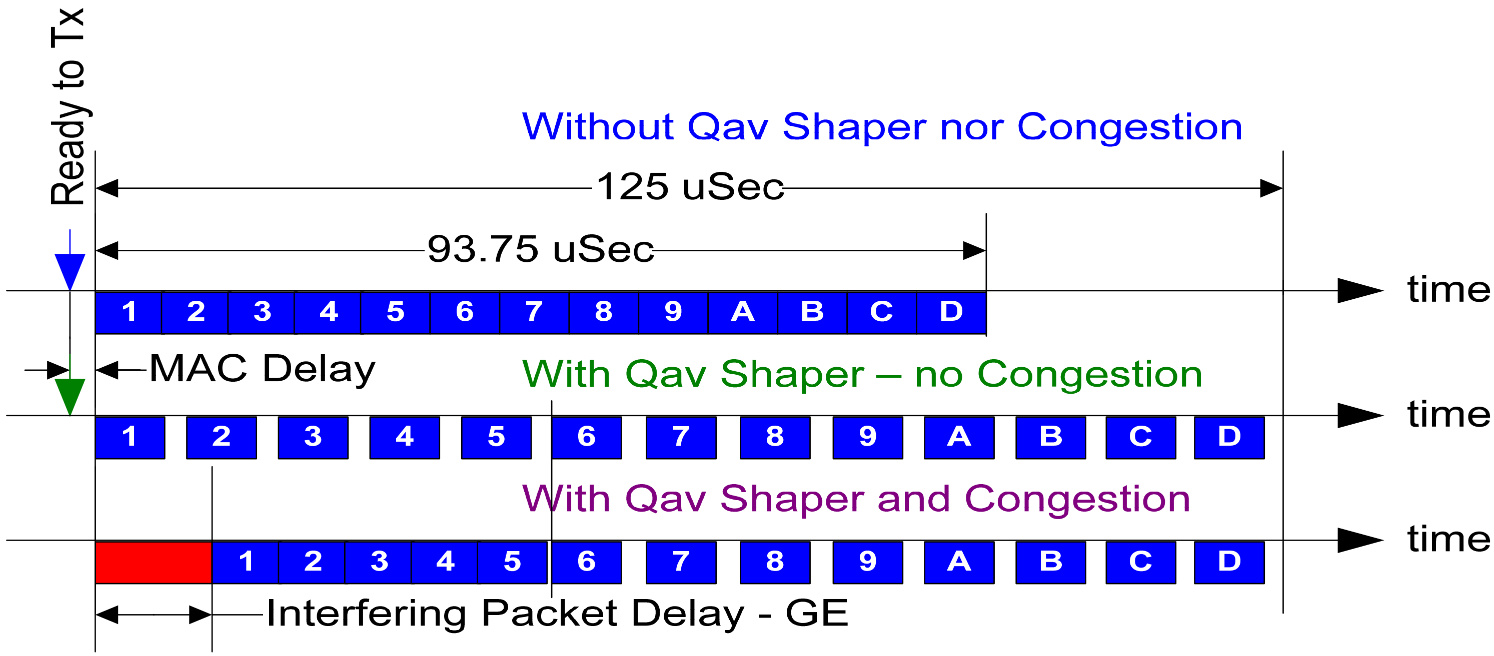
In the figure below:

- Top line is 75% AVB Flows with no Qav and no Congestion
  - It shows 4 x 125 uSec Measurement Intervals with 13 x 70 Byte AVB frames
- Middle line is same AVB Flows with no Congestion spaced out by Qav
- Bottom line is same AVB Flows with a 1522 byte interfering Frame
  - It takes 4 Measurement Intervals of AVB bursting to catch back up
  - At which point a new interfering frame can start
  - Or the AVB flows go back to looking like the Middle line
- Worst case Latency in the FE case is still EQ 1



Effect of Qav + Congestion at GE Speed

- **In the figure below:**
  - Top line is 75% AVB Flows with no Qav and no Congestion
    - It shows 1 x 125 uSec Measurement Intervals with 13 x 700 Byte AVB frames
  - Middle line is same AVB Flows with no Congestion spaced out by Qav
  - Bottom line is same AVB Flows with a 1522 byte interfering Frame
    - It takes 5 AVB frames bursting to catch back up (in this example)
    - At which point the AVB flows go back to looking like the Middle line
    - Or a new interfering frame can start – Lets look at this!



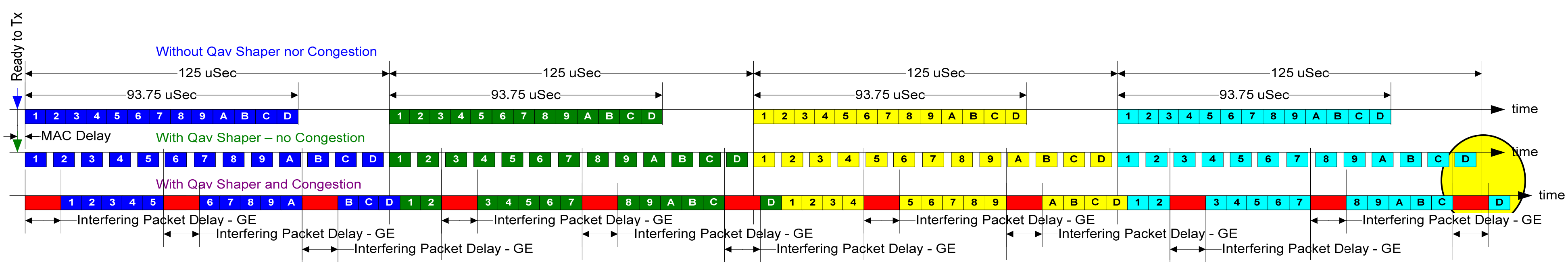
Effect of Qav + Congestion at GE Speed - pt2

In the figure below:

- Top line is 75% AVB Flows with no Qav and no Congestion
  - It shows 4 x 125 uSec Measurement Intervals with 13 x 700 Byte AVB frames
- Middle lines is same AVB Flows with no Congestion spaced out by Qav
- Bottom line is same AVB Flows with constant 1522 byte interfering Frames
  - Burst of Max size frames is quite realistic!

Worst Case is now a Max Size Interfering frame being transmitted just before the last frame in the Measurement Interval was going to be transmitted!

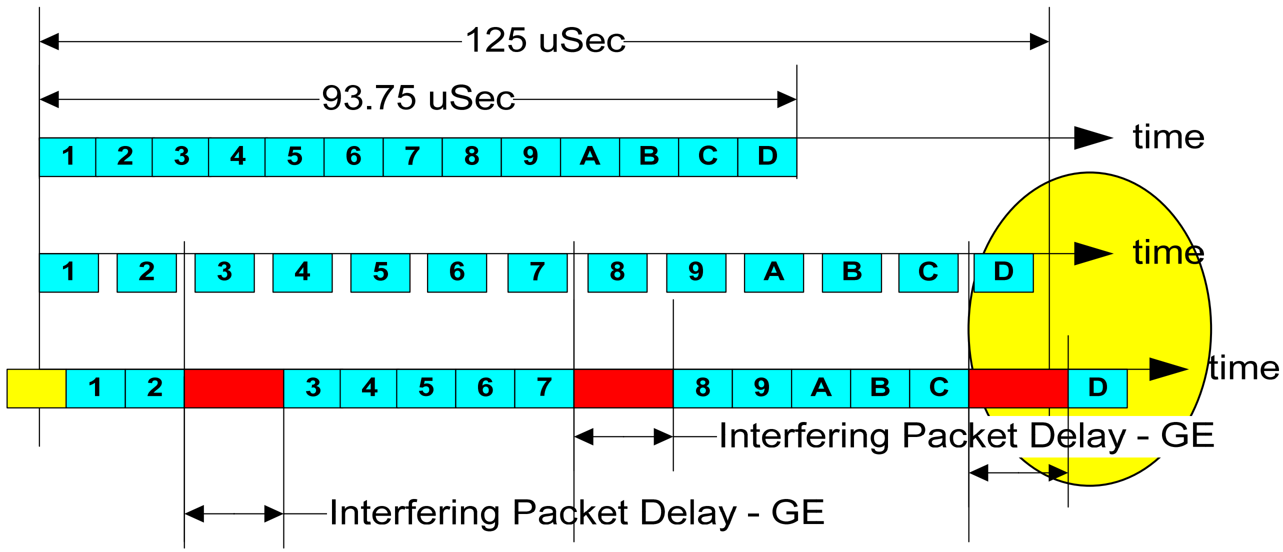
This needs to be looked at now!



Effect of a Late Interfering Frame

The figure below:

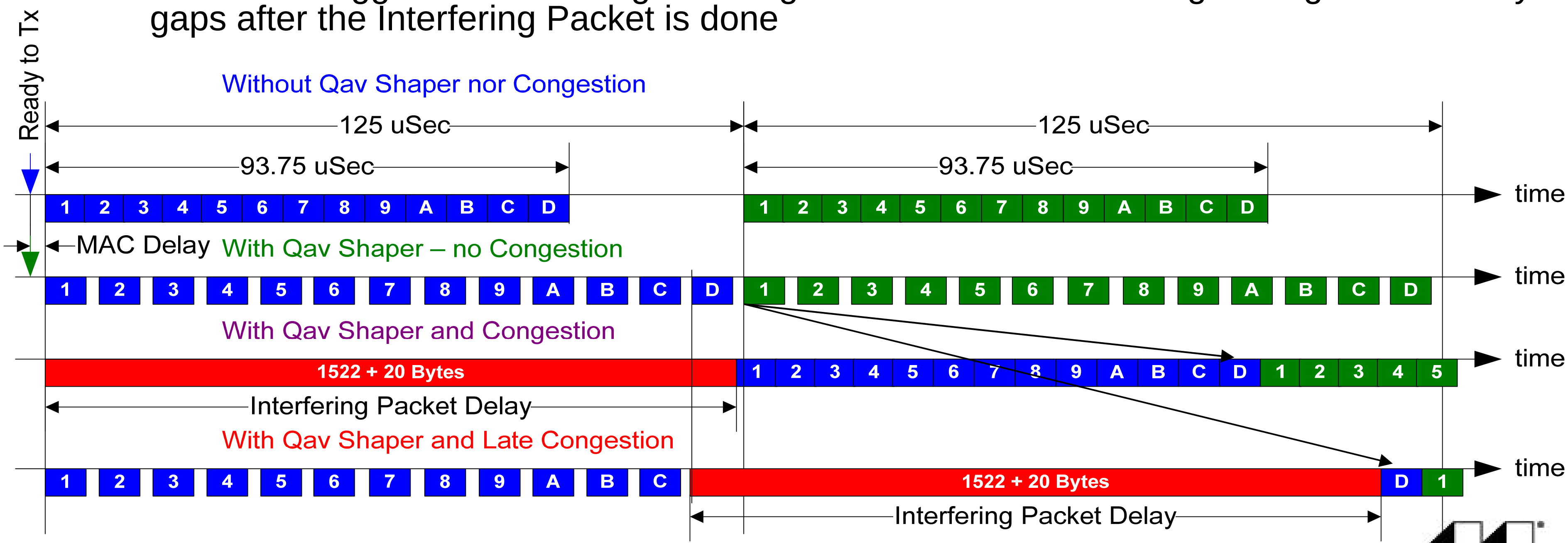
- Is a zoom in on the last part of the previous figure where the worst case can be seen
- AVB frame D is in the queue at the start of the 125 uSec interval
- It does not burst after frame C as the Shaper had caught up
- But just before D is scheduled to go, a Max size interfering frame gets to go instead
- Frame D is now delayed beyond the end of the 125 uSec interval



Example of a Late Interfering Frame

The below example is FE (small size):

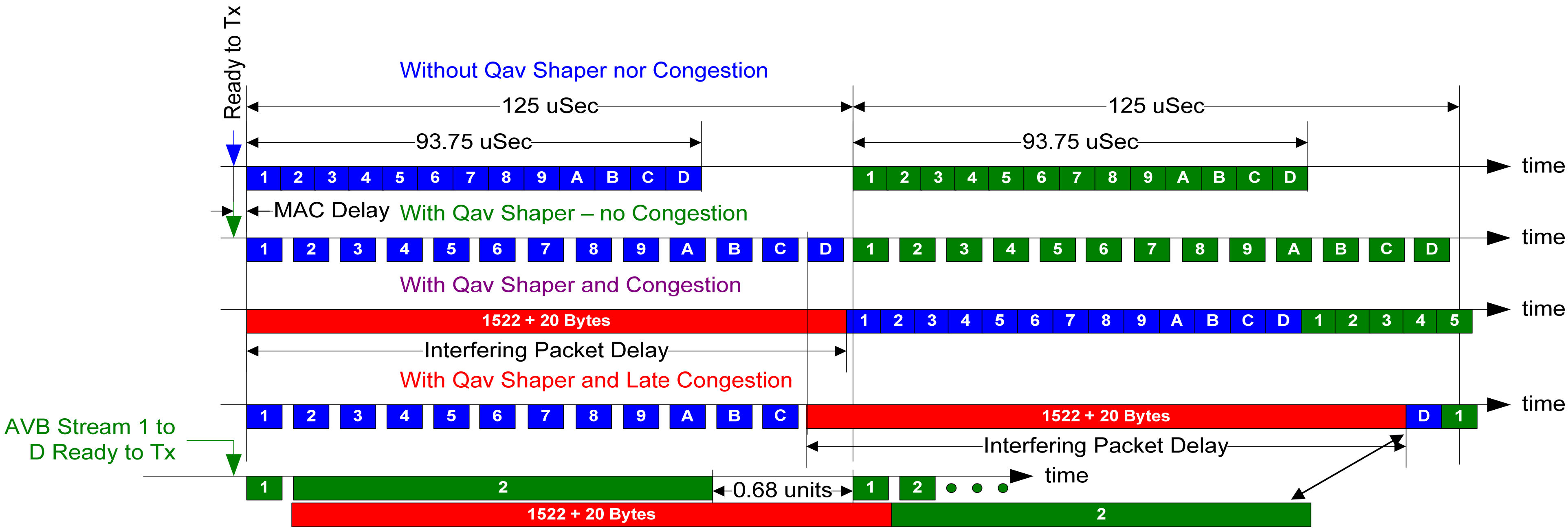
- It can be seen that the Late Interfering Frame (just before frame D is to egress) moves frame D much further out
- EQ 1 is no longer worst case even for FE!
- This is with small (70) byte AV Frames
- The worst case gets better with increasing frame sizes for D
  - When D is bigger the Interfering Packet gets to start earlier and D gets to go without any gaps after the Interfering Packet is done



Example 2 of a Late Interfering Frame

The below example is FE (large size):

- It can be seen that the Late Interfering Frame (just before frame 2 is to egress) moves frame 2 must further out (shown below the previous example)
- This example is not as 'worst case' so smallest AVB frame size appears to be the worst case
- GE is not nearly as bad as the Interfering frame is much shorter in time



A New Equation is Needed

- **With the discovery of the Late Interfering frame, EQ1 is not the worst case for GE nor FE!**
- **In the previous examples I used 90 bytes of streams with 30 bytes of pacing gap which correlates to 75% and 25% as expected**
- **The min size of a stream frame is 64 bytes + 20 overhead bytes or 84 bytes**
- **Using the same 75%/25% ratio, 84 bytes of stream data needs 28 bytes of pacing gap with the two added together being 112 bytes**
- **Therefore the worst case start of a Late Interfering frame is 112 bytes before the end of the Class Measurement Interval**
  - Or 896 nSec for GE, or 8.96 uSec for FE
- **So the equation is looking like (for a 64 byte stream):**

EQ 2: $\text{Max Latency}_{\text{Class A Talker}} = t_{\text{MAC Delay}} + t_{\text{Class A Interval}} - t_{\text{Tx 112 Bytes}} + t_{\text{Tx Max Frame}} + t_{\text{Tx 64 Bytes}}$

A New Equation is Needed – pt 2

EQ 2: $\text{Max Latency}_{\text{Class A Talker}} = t_{\text{MAC Delay}} + t_{\text{Class A Interval}} - t_{\text{Tx 112 Bytes}} + t_{\text{Tx Max Frame}} + t_{\text{Tx 64 Bytes}}$

- **EQ2 is for 64 bytes streams which is worst case as increasing streams frame sizes cause a larger subtraction vs. a larger addition**
- **This is because the Tx 112 bytes = (Frame Size + 20) * 1.3333 due to the 75% ratio (i.e., the 1.3333 = 100%/75%)**
- **So in terms of stream frame size EQ2 becomes:**

EQ 3: $\text{Max Latency}_{\text{Class A Talker}} = t_{\text{MAC Delay}} + t_{\text{Class A Interval}} - t_{\text{Tx (Stream's Frame + 20) * 1.333}} + t_{\text{Tx Max Frame}} + t_{\text{Tx (Stream's Frame)}}$

- **Using the smallest stream frame size the worst case latency becomes:**
- **For Class A on FE (assume 1522 Max Non-AVB + 75% for Class A):**
  - = 5.12 uSec + 125 uSec – 8.96 uSec + 123.36 uSec + 5.12 uSec = 249.64 uSec
  - For 7 FE hops this is 1,747.48 uSec – which is below our 2.0 mSec target goal!
- **For Class A on GE (assume 1522 Max Non-AVB + 75% for Class A):**
  - = 0.512 uSec + 125 uSec – 0.896 uSec + 12.336 uSec + 0.512 uSec = 137.46 uSec
  - For 7 GE hops this is 962.22 uSec – which is better than FE, but not 10x better

Current Observations

- With the discovery of the Late Interfering frame, EQ1 is not the worst case for GE nor FE, EQ3 is!
- Unfortunately, the worst case gets better with larger stream frame sizes when the network works better with smaller stream frame sizes
- GE does not support 10x lower latencies. It is not quite even 2x!
- Reducing the link's bandwidth percent that Class A streams can use does not change the latency as the Qav Shaper still spreads the data out over the entire Class Measurement Interval.
- The only way to reduce Class A's latency is to add Class B steams which is not desirable either (if they are simply 'dummy' streams)
- I believe these equations will be the same for bridges as the same time start of 'n' frames in a bridge is its fan-in
- Talkers that support a limited # of flows can take this into account
- This what we have today. Can any improvements be addressed in 802.1BA? Or in 802.1Qav ver 2?

Thank You