Worst-case latency in 802.1Qav Ethernet bridges
(v2 – some early conclusions)

Michael Johas Teener

April 11, 2008

Change history

V	date	updates
1	31 jan 08	original version, class A only, no “observation interval”
2	11 may 08	validation of assumptions, where “class observation interval” is needed, more extensive conclusions

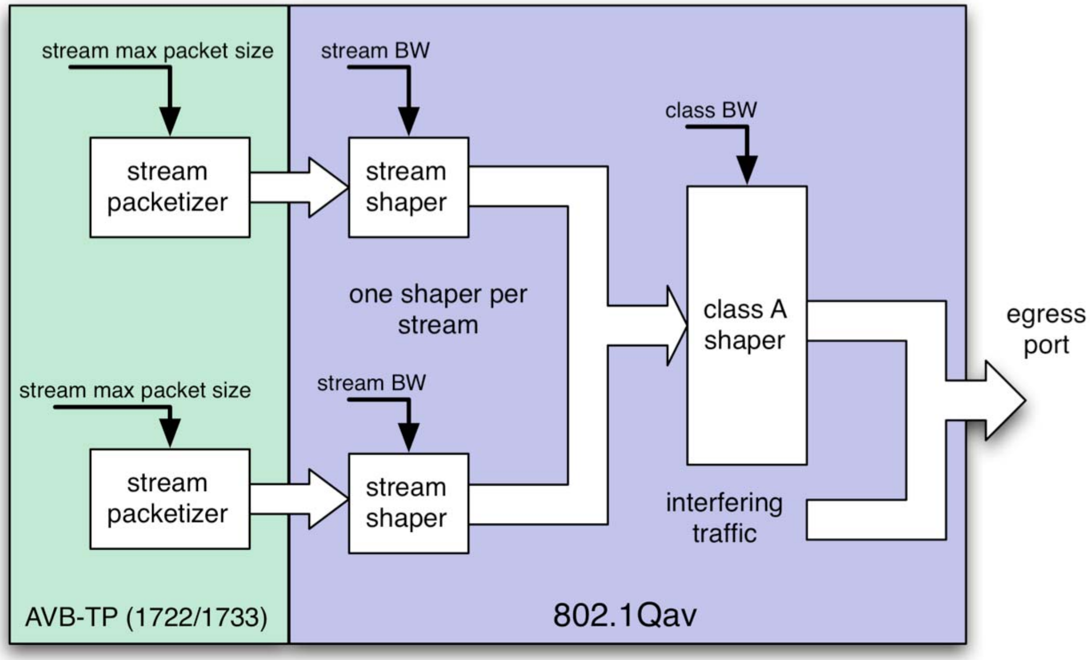
Notes

- This calculation is only for Class A
  - I want to make sure we understand the limits on a “2ms” latency network
  - Once we understand that, then I’ll add the Class B traffic to the analysis
- The parameters to be explored include:
  - Network topology (number of bridges and number of ports on each bridge)
  - Stream packet limitations (max packet size)
- All shapers are as described in Qav 0.3

Input, output & methodology

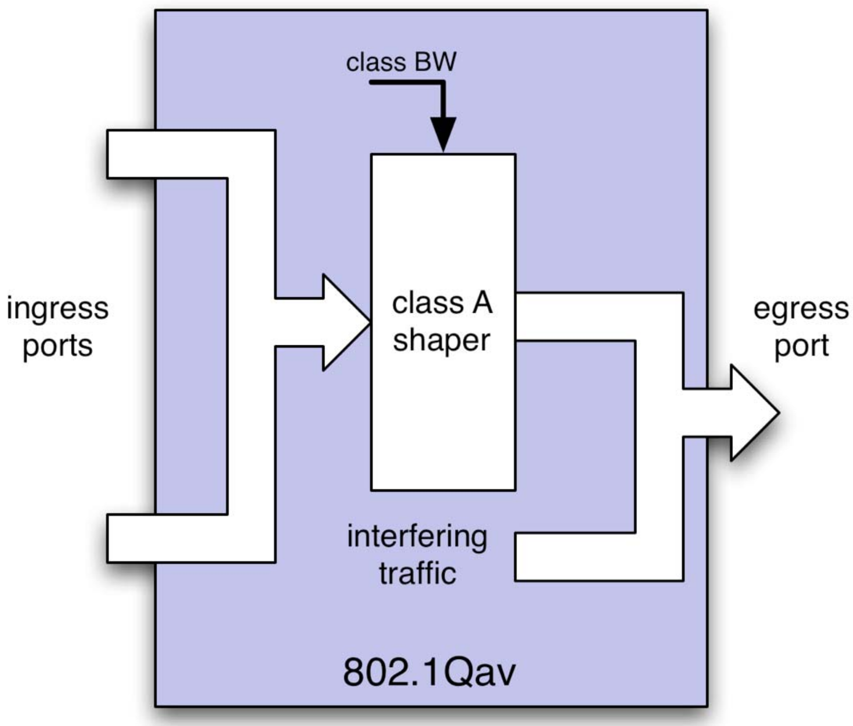
- The input parameters to be explored include:
  - Network topology (number of bridges and number of ports on each bridge)
  - Stream packet limitations (max packet size)
- Output is worst case delay
- Looking only at first order effects
  - mention will be made of 2^{nd} order effects that are being ignored for now

Talker model



- Talker consists of data stream packetizers feeding into stream shapers feeding into class shaper
  - stream packetizer has a max packet size parameter
- Stream shapers have infinite "sendSlope"
  - has the effect of smoothing per-stream traffic, so there is no bunching within a stream
- Sum of all stream's "idleSlope" is the class "idleSlope"
  - SRP bandwidth allocation is "idleSlope"

Bridge model



- Same as a talker with no stream shapers
  – conversely, a talker can be thought of as a bunch of single stream sources each with an infinitely fast link to a bridge

Talker delays

- Talker has just started to transmit a best effort frame of b bytes
- There are m streams, each with a max packet size of s_j bytes
- Egress port rate is e bytes/sec
- Delay is

$$\frac{\left(b + \sum_{i=1}^{m} s_i\right)}{e}$$

Bridge delays

- Bridge has just started to transmit a best effort frame of b bytes
- There are m ports, each routing class A traffic with a max packet size of s_j bytes through the egress port
- Egress port rate is e bytes/sec
- Delay is

$$\frac{\left(b + \sum_{i=1}^{m} s_i\right)}{e}$$

Network delays

- There are n bridges
  - so there are n+1 devices for queuing delays
- For each hop between devices there is no common stream
  - so it's possible for a stream to always be delayed by new interfering packets on each hop
- Delay is $$\sum_{j=0}^{n}\left(\frac{\left(b_j + \sum_{i=1}^{m} s_{ij}\right)}{e_j}\right)$$

Simplification

- all links are 100 Mbit/sec (e_j)
- worst case best effort interfering packet of 2000 bytes (b_j)
- all other class A packets are the same size (s_{ij})
- the talker launches m streams and each bridge has $m+1$ ports
- so, delay is

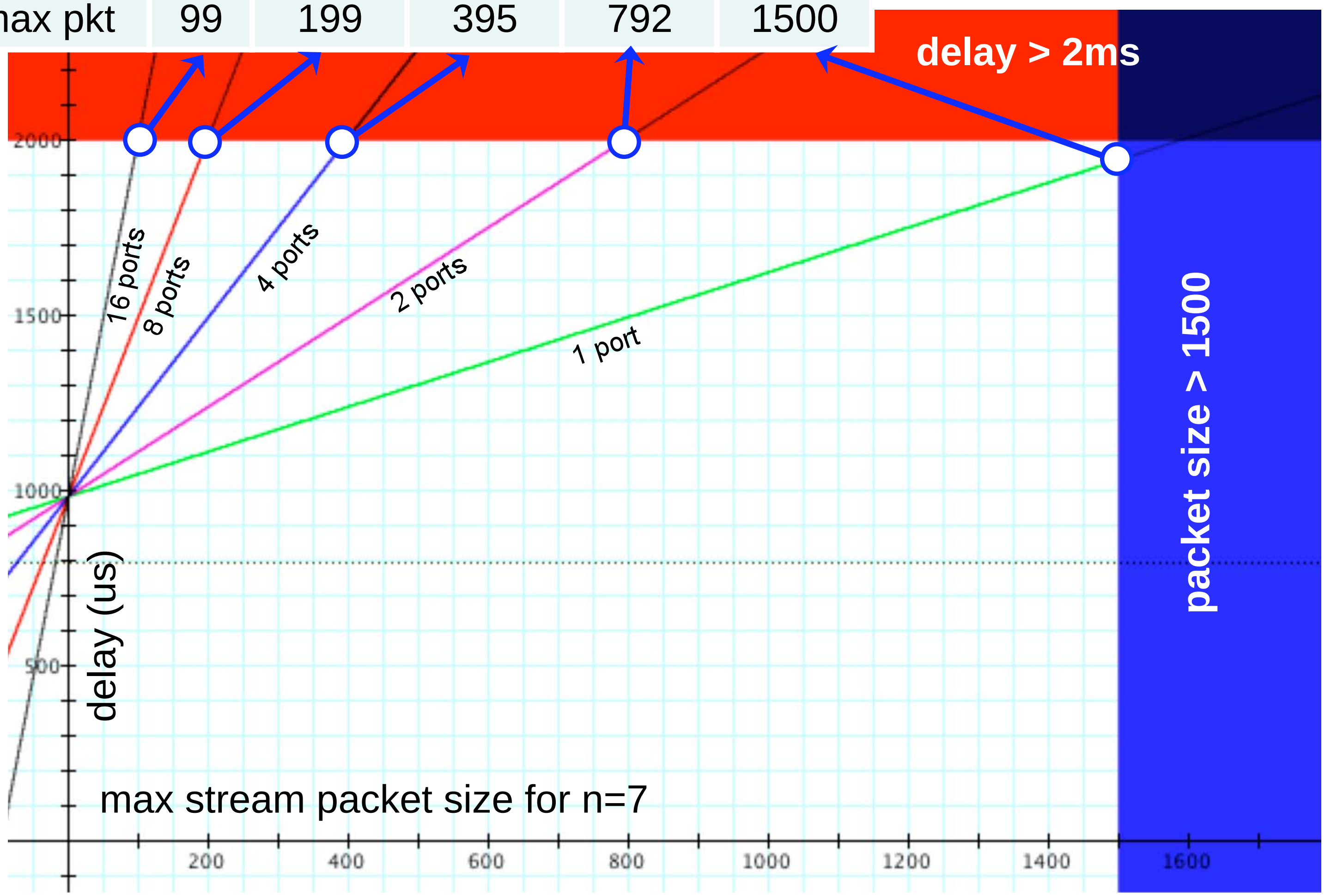
$$\frac{(n+1)(b+ms)}{e} = \frac{(n+1)(2048+ms)}{100}$$

2^{nd} order effects

- Cumulative “bunching”?
  - Does not appear to be an effect, at least for networks carrying a single AVB class
- Remember, no analysis given to multiple classes!

Max network delays

#ports	16	8	4	2	1
max pkt	99	199	395	792	1500



Conclusions (part 1)

- Class 5 max packet size directly effects the latency, as does the number of bridges in a path from talker to listener, as does the number of ports on those bridges
  - We have been assuming 7 hops is a good limit for class A at 2ms max delay.
  - So we need to assume limits for the number of ports on the bridges and the max packet size
- For a 7 hop 100 Mbit/sec Ethernet configuration, we should perhaps assume 8 port bridges are a maximum
  - If so, then class 5 packets need to be no larger than about 200 bytes
- SRP *can* allow larger packets, but it will have to be ready to deny requests even when there is bandwidth available on a egress port
  - because the latency budget of “250 usec/bridge” is used up

Changing the simplifications

- This is all a bit too restricting
  - We end up with very large latencies unless we restrict everyone's packet sizes
- It's also not necessary
  - The important value is the amount of interfering traffic:

$$\left(b_j + \sum_{i=1}^{m} s_{ij}\right)$$

  - We can't limit the best effort packet length, but the sum of the stream packet lengths could be under control

So ….

Using a measurement interval (part 1)

- "Class measurement interval" is useful to determine the max packet size parameter of the stream packetizer
  - packet size for stream $s_{ij} \leq a_{ij} c$ where
    - a_{ij} is the allocated bandwidth for a stream
    - c is the class measurement interval
- Since the total allocated bandwidth must be less than 75% of the egress rate: $\sum_{i=1}^{m} a_{ij} \leq 0.75 e_j$
- We then know that $\sum_{i=1}^{m} s_{ij} \leq 0.75 e_j c$ or about 1180 bytes for 100Mb/s

Using a measurement interval (part 2)

- Using a class measurement interval allows stream bandwidth (available to SRP) to be used directly as an analog for stream-induced delays so that SRP can relate max delay to max bandwidth

Alternate simplification

- all links are 100 Mbit/sec (e_j)
- worst case best effort interfering packet of 2000 bytes (b_j)
- all class A packets uses the same class measurement interval of 125us
- the talker launches m streams and each bridge has $m+1$ ports
- so, delay is $\dfrac{(n+1)(b+\sum_{i=1}^{m} s_{ij})}{e}$

$$= \frac{(n+1)(2000 \times 8 + .75 \bullet 100 \times 10^{6} \bullet 125 \times 10^{-6})}{100 \times 10^{6}}$$

$$\cong (n+1)0.000254us = 2.03ms \;\; for\ n = 7$$

Conclusions

- If a class measurement interval is used to force frame sizes to smaller size, then the max delay is not dependent on the number of ports on a bridge, just on the number of hops
- A 125us measurement period for class A will guarantee ~ 2ms worst case delay over 7 hops