Real-time networks and preemption

More to it than latency

Rev. 3

Norman Finn

nfinn@cisco.com

http://www.ieee802.org/1/files/public/docs2011/new-avb-nfinn-real-time-networks-1111-v03.pdf

What is a real-time network?

- In a real sense, all networks are “real-time” except for simulations of networks.
- Video or voice data is certainly a kind of “real-time”
- Priority, resource reservation, and other methods work for many networks that have tight latency and/or jitter requirements.
- In this slide deck, “real-time” means a guaranteed response time to any given input or combination of inputs. **No excuses, no exceptions**.
- Typical examples are automatic automobile braking systems and robot control.

What do real-time networks lack?

- Some excellent presentations have been made this year on requirements from users and designers of real-time automotive and industrial networks.
- There are common threads that we can address:

  Topology

  Delivery

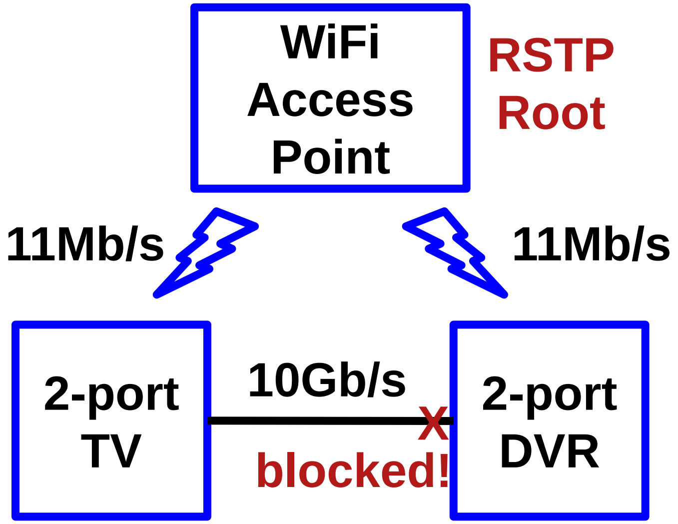
  Predictability

- But, we cannot address them in isolation, either from each other, or from more general uses of Ethernet networks.

Topology

Topology

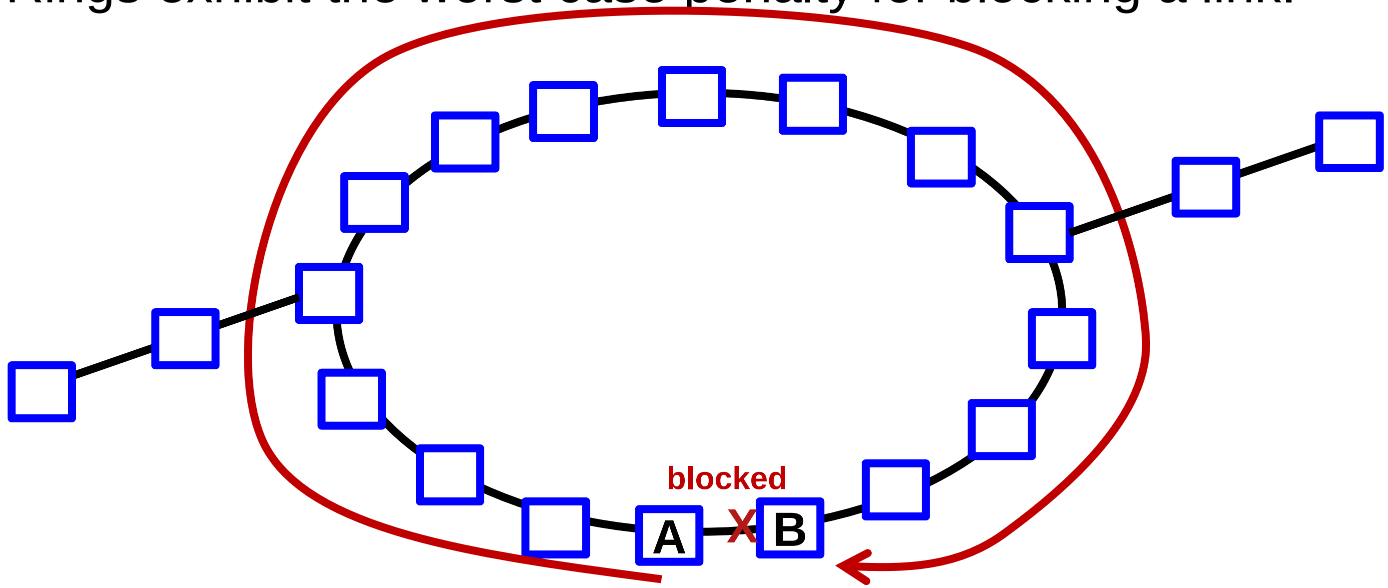
- As has been known for a long time, spanning tree has issues in simple networks with links of widely disparate data rates.



- This diagram illustrates the problem in the home.

Topology

- Similarly, large rings, as are common in automobiles and industrial networks, are the least-favored topology for spanning tree.
  - Rings (with tails) exhibit the worst case reconfiguration times.
  - Rings exhibit the worst case penalty for blocking a link.



Topology

- We could build on spanning tree. But ...
  - Bridges running MSTP lack a view of the whole network, and this may useful information to applications.
  - Using MSTP requires that MSRP or similar protocols must converge *after* MSTP converges, instead of simultaneously.
- For these reasons, and because the blocked-link problems in the previous slides are solved, this author believes that a **link-state protocol should be the basis for real-time networks**.

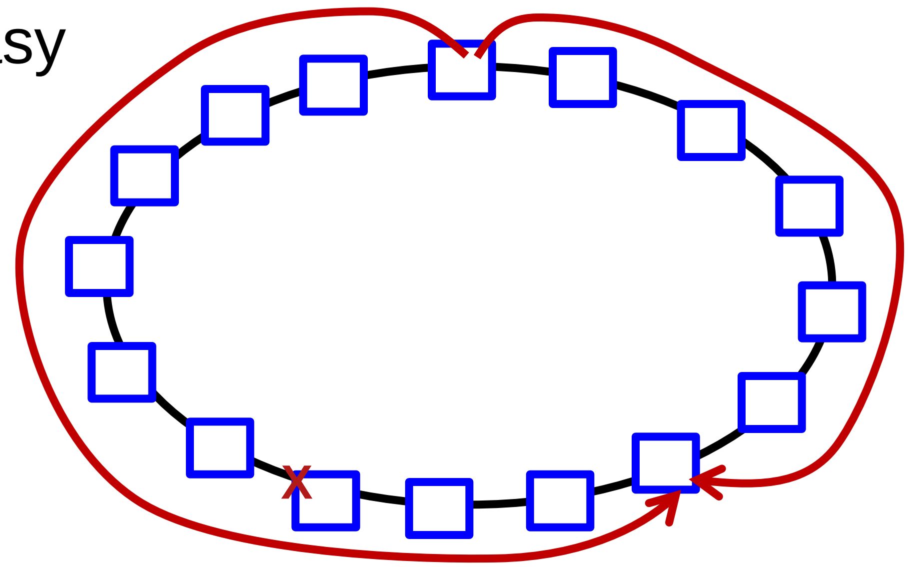
Shortest Path Bridging

- Coincidentally, SPBV (VLAN-mode Shortest Path Bridging) can be made plug-and-play for networks in the size range we're interested in.
- Some work would still be needed:
  - We must balance the number of VLANs against number of bridges ([number of bridges] * [number of VLANs] < 4096).
  - Learning MAC address can preclude the use of two paths between two stations.
- It is true that SPBV is more complex than alternatives that are based on a fixed topology. But, not all real-time networks are rings, and one must ask whether the topology is *really* fixed.

Delivery

Delivery

- For ultra-reliable communications between consenting stations, delivery of frames along two paths would be very helpful, and there are documented methods for it.
- This cannot be easily done by current bridging/routing protocols: paths are not equal cost, overriding the topology to slip past blocked links breaks address learning, and it is not easy to discovery maximally-disparate paths.
- But, if we can do it, the value will be significant!



Delivery

- It is worth pointing out that P802.1Qbf Segment Protection can route frames **outside** the spanning tree or SBP framework, including simultaneous delivery along multiple paths.

Predictability

Predictability

- Preemption reduces queue size, and thus latency, by **exactly one frame**.
- Improving one flow's latency makes all other flows' latency and jitter even worse.
- As soon as you have two preemptive flows, collisions between those flows put you in the same place you were in before you introduced preemption.
- Cut-through forwarding of preemptive frames would improve best-case latency. Is this improvement also necessary?
- (Cut-through forwarding of preemptable frames results in an intractable fragmentation problem.)
- So, there is more to the predictability problem than latency or preemption.

Improving one hurts all others

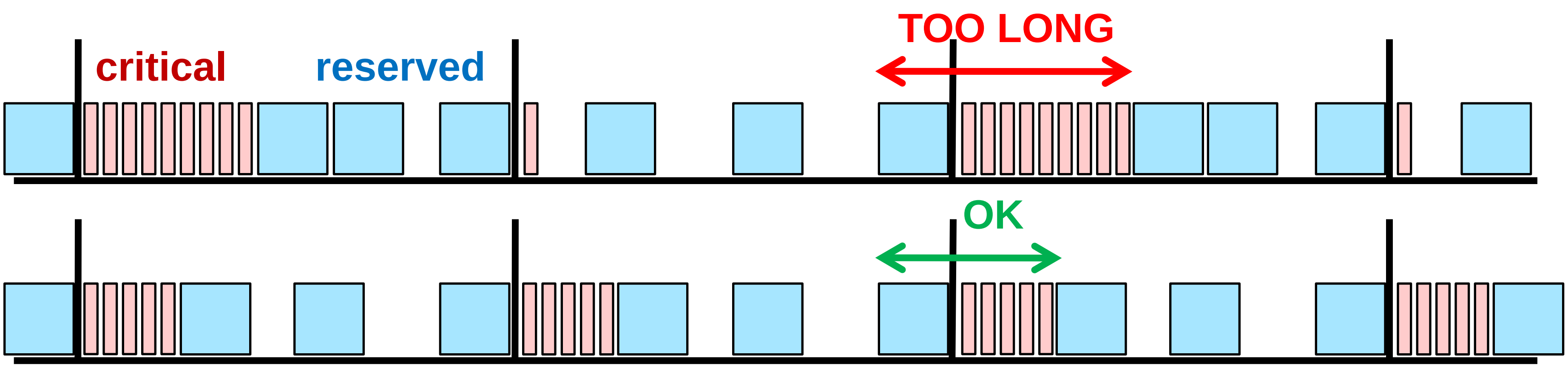
- The work on 802.1Qat showed that the biggest impacts on the latency and jitter of a reserved stream (or the highest-priority non-reserved stream) are, in increasing importance:
  1. The largest conflicting lower-priority frame.
  2. The fan-in at each bridge to a controlled queue.
  3. The percentage of bandwidth reserved for this level of controlled traffic.
  4. The percentage of bandwidth reserved for higher-priority controlled traffic.
- Preemption eliminates 1. 2 is not a concern for networks that are mostly rings. Preemption has a major impact on other, lower-priority bandwidth-reserved flows.

Time synchronization

- There is a long history of real-time networking, especially in the aerospace industry.
- In this world, “real time” does not mean interrupts and preemptive process scheduling. It does not mean “best effort delivery.”
- “Real-time” means **scheduling**: scheduling processes within a station, scheduling communications between stations, and coordinating the stations’ schedules.
- Scheduling guarantees that all processing and communications happen within the required time limits.
- Even network recovery is accounted for by scheduling alternatives.

But ...

- Critical traffic must live with bandwidth reserved traffic, also.
- If scheduled critical traffic takes a high enough percentage of the bandwidth for a long enough time, it will starve the bandwidth reserved (audio or video) traffic.



- These requirements **can be incompatible**.

Non-real-time networks

Preemption in non-real-time networks

- In the typical enterprise or data center network, there is a high degree of connectivity in order to minimize congestion.
- A higher degree of connectivity implies more opportunities for flows to collide.
- Preemption reduces queue size, and thus latency, by **exactly one frame**.
- In a highly connected network with many transmitters, preemption makes a difference only in benchmark tests.
- It would be unfortunate if we allow a class of unrepresentative benchmark tests skew real requirements.

Summary

Real-time networks: 3 networks in 1

- Three levels of service: **Critical**, **Reserved**, and **Best-Effort**.
- **Critical** traffic uses preemption (and maybe cut-through forwarding) so that other classes do not disturb it.
- **Critical** traffic uses time synchronized transmissions to ensure that 1) critical flows do not interfere with each other, and 2) critical flows do not overly disrupt Reserved traffic.
- **Reserved** traffic uses bandwidth reservation and shaping to guarantee audio/video requirements.
- **Best-effort** traffic gets what’s left.

Real-time networks: 3 networks in 1

- **Preemption** is required to ensure that non-critical traffic cannot interfere significantly with the scheduled operation of critical traffic.
- **Cut-through forwarding** (of critical traffic only) **may** be needed to minimize latency and the impact of critical traffic on other traffic classes.
- **Scheduling** of critical traffic is required both to meet application requirements and to avoid disrupting bandwidth reserved traffic.
- Existing **bandwidth reservation** and **shaping** are required to meet audio / video requirements.
- Existing **priorities** support best-effort service.