Follow-up on Peristaltic Proposal

Rodney Cummings
National Instruments
Introduction

• Follow-up on proposal for peristaltic transport

• Use 802.1Qbv scheduling to improve SRP streams
  • Deterministic delays for all streams
  • Delay scales with link speed

• New stuff: Bridges are cycle aware
  • Time-aware tags on ingress (for cycle count)
  • Two egress queues per traffic class (even/odd cycle)

• This presentation explores the new stuff
Assumptions

• All end-stations and bridges use same gating cycle
  • Cycle starts at same point in all
  • Cycle length has a default that can be changed
    • Use SRP to ensure that all agree

• Cycle has two windows: scheduled and best-effort
  • Scheduled window grows to fit all streams for that egress
    • Best-effort window shrinks
  • If new stream would cause best-effort window to be smaller than maximum frame, reservation fails
Reasons for Bridge-Aware Cycles

- Determinism (worst-case latency)
- Streams slower than frame-per-cycle
Determinism (1 of 2)

- If we enforce cycle-per-hop for each stream...

<table>
<thead>
<tr>
<th>Cycle count</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talker egress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bridge 1 egress</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

  - ...worst-case latency is linear function of #hops
  - If red received by bridge in schedule window 1 (odd), must hold back until schedule window 2 (even)
    - Requires cycle-aware bridges

  ![Diagram](image-url)
Determinism (2 of 2)

- If we allow egress earlier than cycle-per-hop...

  - ...simple formula for worst-case latency seems to remain: linear function of #hops
    - Worst that can happen in bridge: red pushed to cycle 2
    - If software tool has detailed topology & stream info (e.g. IS-IS), it can calculate precise worst-case latency
Slower Streams (1 of 2)

- Many control applications use multiple rates
  - Often harmonic (e.g. 250µs, 2ms, 8ms)
- Example of “cycle multiplexing”
  - 125µs cycle,
    talker 1 sends every 250µs, talker 2 sends every 250µs,
    talker 1 & 2 alternate: odd & even cycle count
- Network-wide cycle multiplexing can be complex
  - Various multiples (e.g. 8 1ms talkers, 13 2ms talkers)
  - Talkers share cycle across complex topology
  - Likely to require bridge data plane to be stream-aware
    - Not just cycle-aware
Slower Streams (2 of 2)

• Alternative: Limit cycle multiplexing to end-station
  • Bridge not required to be cycle-aware

• End-station reserves bandwidth in every cycle

• Multiple talkers in end-station alternate use of cycles
  • Specifics of cycle multiplexing outside 802.1 scope

• 802.1 feature: multiple destination addresses can share a bandwidth reservation
  • Not supported in 802.1 SRPv1
  • Supported in IETF RSVP
Conclusion

• 802.1Qbv scheduling viable for SRP streams
  • Cycle-aware bridges may not be required
  • Recommend support for multiple talkers per reservation

• Peristaltic proposal is Implicit Scheduling
  • Talker specifies bandwidth
  • 802.1 entities calculate schedule details

• Can be leveraged to provide Explicit Scheduling
  • Use traffic engineering concepts
  • Talker specifies schedule details
    • E.g. Distinct gating cycle per bridge
Thank you