Draft Individual Contribution P802.1Qcz

Paul Congdon

March 2019

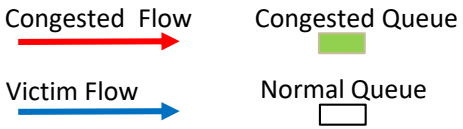
Agenda

- P802.1Qcz background and references
- Draft standard status and strategies
- Technical discussion of CI critical processes
- Next steps

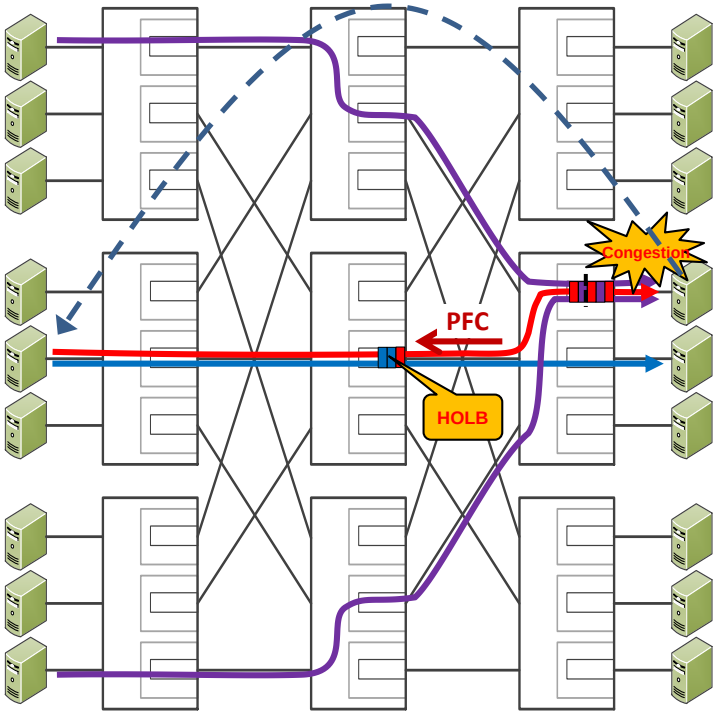
References

- This presentation supports a 2nd individual draft contribution
  - http://www.ieee802.org/1/files/public/docs2019/cz-congdon-draft-text-contribution-0319-v01.pdf
- Technical overview of CI
  - http://www.ieee802.org/1/files/public/docs2018/cz-congdon-congestion-isolation-review-0418-v1.pdf
- Possible changes to 802.1Q
  - http://www.ieee802.org/1/files/public/docs2018/cz-congdon-ci-Q-changes-0618-v1.pdf
- Objectives Discussion
  - http://www.ieee802.org/1/files/public/docs2018/new-dcb-congdon-ci-objectives-0118-v02.pdf

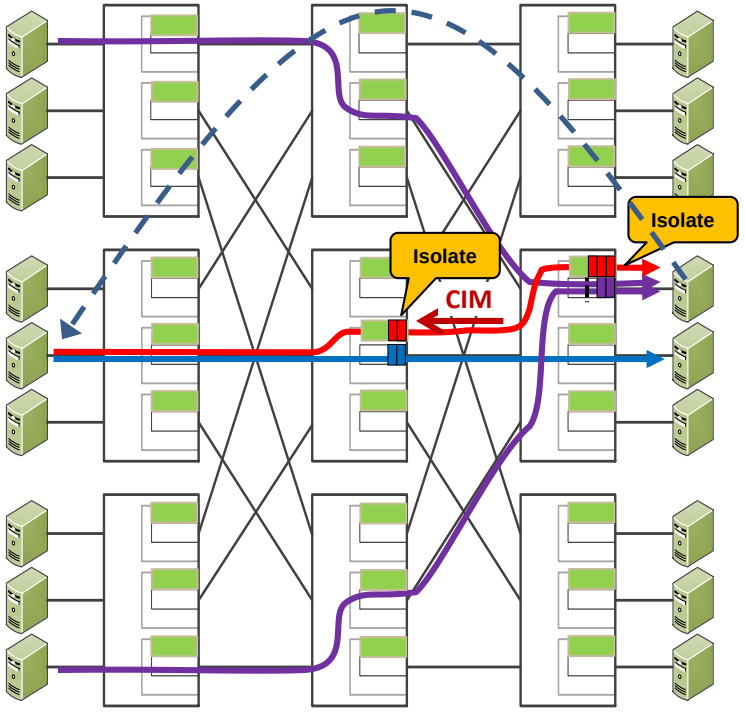
One Slide Congestion Isolation Refresh



Today – Without Congestion Isolation



Congestion Isolation



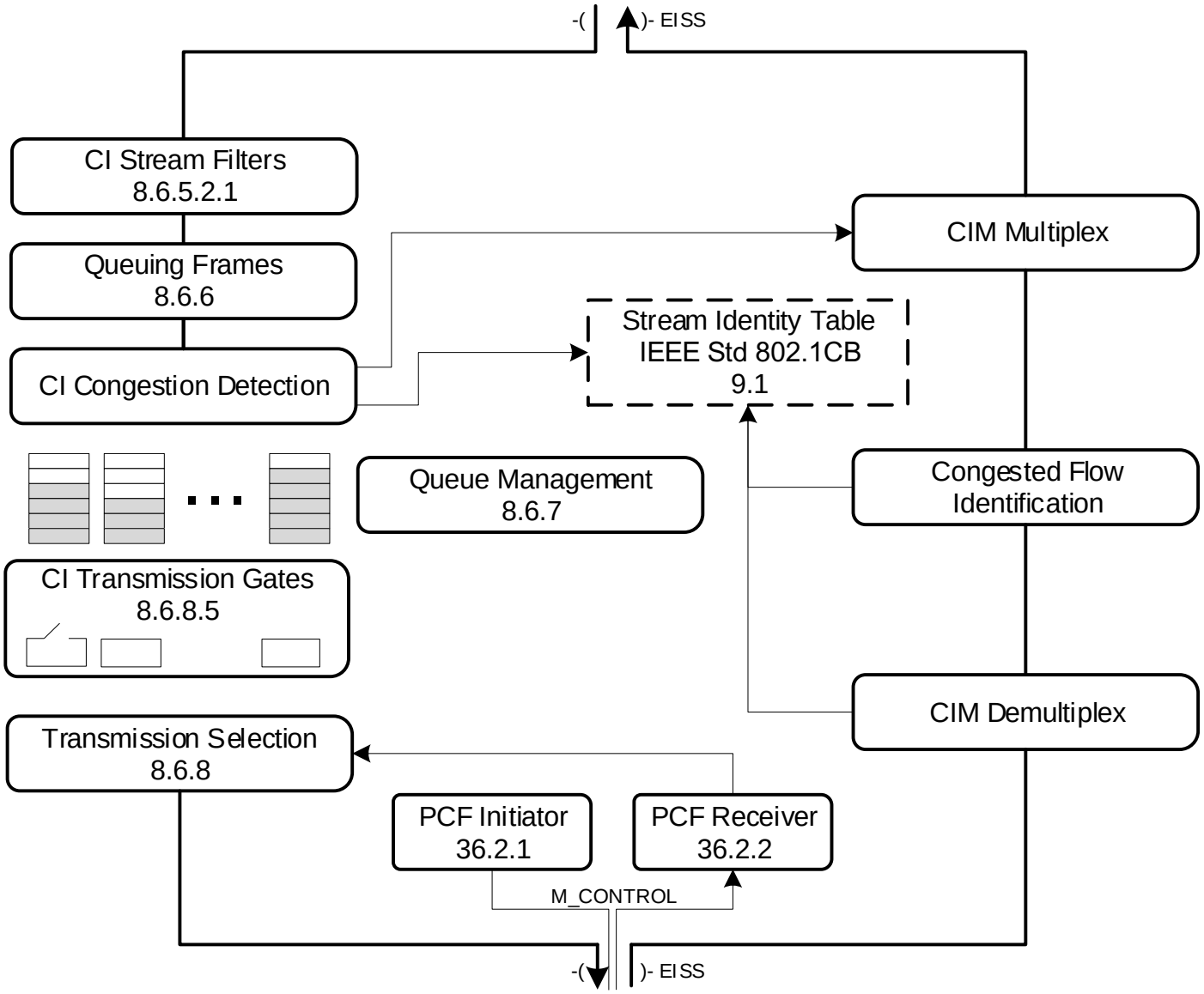
Summary of major draft changes

- Built using FrameMaker and standard 802.1Q master files
- Text related to congested flow identification
- Use of Stream Identification Function of 802.1CB instead of inventing separate Congested Flow Table
- Replaced ‘new’ transmission scheduling algorithm with re-use of PSFP architecture and simple transmission gates
- Updated reference diagram aligned with above
- Creation of Informative Annex X - Maintaining Packet Order with Congestion Isolation

Congestion Isolation Critical Processes

1. Detecting flows causing congestion
2. Creating flow entries in the congested flow table
3. Signaling congested flow identify to neighbors
4. Isolating congested flows without ordering issues
5. Interaction with PFC generation
6. Detecting when congested flows are no longer congested
7. Removing entries from congested flow table
8. Signaling congested to non-congested flow transitions to neighbors
9. Un-isolating previously congested flows without ordering issues

802.1Qcz Reference Diagram



Draft Position/Status on Critical Processes

1. Detecting flows causing congestion
   - Clause 98.2
   - No specific AQM algorithm required for interoperable implementations
   - Reference QCN/Congestion Notification Clause 30.2.1 as an example.
   - Reference IETF recommendations on AQM
   - Require implementation to assert ***ciCongestedFlow*** and provide first 64 bytes of sampled frame in order to generate CIM message

Draft Position/Status on Critical Processes

2. Creating flow entries in the congested flow table
   - Using the Stream Identification Function Clause 6 of IEEE Std 802.1CB-2017
     - Supports IPv4 and IPv6 flow specifications
     - Current open project to add more sophisticated flow matching / stream identification.
   - Referenced the same way as Per-Stream Policy and Filtering (PSFP) defined for TSN
   - Table is managed through 802.1CB managed objects

Draft Position/Status on Critical Processes

3. Signaling congested flow identify to neighbors
   - CIM message contains initial 64 bytes of sampled frame deemed to be causing congestion
     - NOTE: an optional feature would be for the participants to negotiate the desire to transmit more than 64 bytes in the CIM. This could be done using the LLDP TLV.
   - CIM is generated and transmitted similar to Congestion Notification Message (CNM) described in Clause 31.1.1 – Except to next hop peer
   - CIM uses ***Nearest Customer Bridge*** address as destination. Is this an issue?
     - NOTE: The CIM frame will traverse bridge relay, but must be transmitted out the port the congested flow was received, not flooded.
     - Alternatively a table of next hop peer addresses is needed for CIM generation.

Draft Position/Status on Critical Processes

4. Isolating congested flows without ordering issues
   - CI stream filters (8.6.5.2) will change traffic class based on stream_handle (in connection_identifier from 802.1CB stream identification function) and configured IPV.
     - NOTE: leveraged from PSFP for TSN
   - CI transmission gates (8.6.8.5) will make queues available for transmissions selection (i.e. blocking queues if out-of-order condition exists)
   - ***ciGateControl*** variable asserted by ordering algorithm. Informative Annex X describes a possible algorithm using markers and counters.
   - ***ciGateControl*** is always true if strict priority transmission selection algorithm. Minimum to implement requirement.

Draft Position/Status on Critical Processes

5. Interaction with PFC generation
   - From the standard specification point of view, CI and PFC are independent features.
   - To avoid ambiguity of which traffic class was used by an upstream neighbor, CI should require that the frame is VLAN/priority tagged.
   - When congested frames are in two traffic classes upstream, it is possible to block the non-congested queue.
   - See slides 7-13 of http://www.ieee802.org/1/files/public/docs2018/cz-congdon-ci-design-topics-1118-v01.pdf

Draft Position/Status on Critical Processes

6. Detecting when congested flows are no longer congested
   - Three approaches are discussed in: http://www.ieee802.org/1/files/public/docs2018/cz-sun-isolation-recovery-1118-v01.pdf
     - Define a new threshold for the congested queue
     - Keep accounting of packets in the congested queue
     - Use an inactivity timer to age congested flows
   - Solution needs to be compatible with leveraged stream identification function of 802.1CB
   - Likely need a ***ciUncongestedFlow*** signal to drive state machine behavior.
   - Specification strategy is TBD

Draft Position/Status on Critical Processes

7. Removing entries from congested flow table
   - Use the 802.1CB management interface for the stream identification function
   - Action triggered by ***ciUncongestedFlow*** signal
   - Potential out-of-order frame situation may exist requiring additional logic controlling the CI transmission gates.

Draft Position/Status on Critical Processes

8. Signaling congested to non-congested flow transitions to neighbors
   - A CIM ‘deallocation’ message is discussed in slides 14-16 of http://www.ieee802.org/1/files/public/docs2018/cz-congdon-ci-design-topics-1118-v01.pdf
   - Issue: the single ‘deallocation’ message that is subject to packet loss creates reliability concern
   - Specification work TBD

Draft Position/Status on Critical Processes

9. Un-isolating previously congested flows without ordering issues
   - Anytime the congested queue empties, the entire congested flow table can be flushed - no ordering issue.
   - Methods that can not assure that frames of a congested flow are vacant from the congested queue will need to block the non-congested queue to preserve order using CI transmission gates.
   - Minimum to implement strict priority transmission selection is perhaps the worst case scenario.
   - Specification work TBD

802.1Qcz Draft Next Steps

- Propose resolutions and positions for all critical processes
  - Resume design team discussions
- Obtain approval to product initial draft for Task Group ballot