Marker Protocol and DRNI

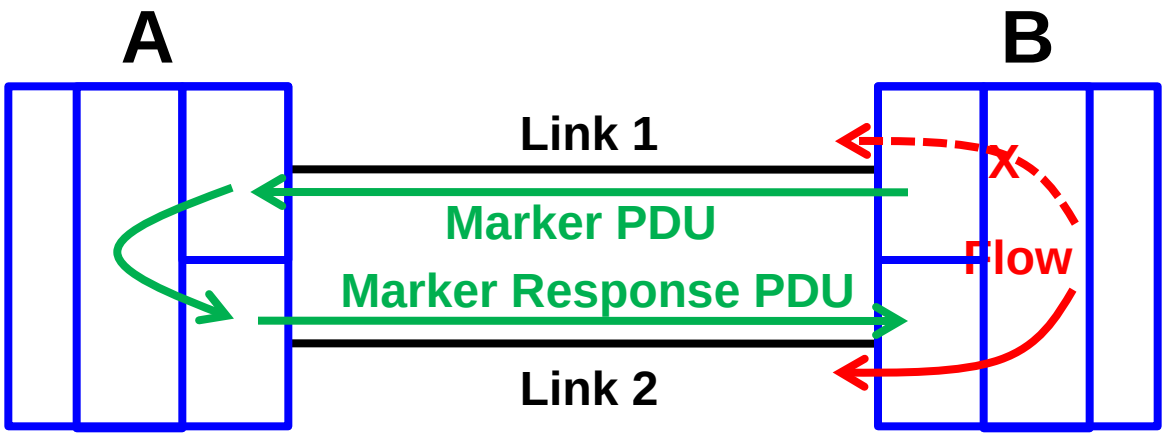
Rev. 1

Norman Finn

nfinn@cisco.com

Porting the Marker Protocol to DRNI

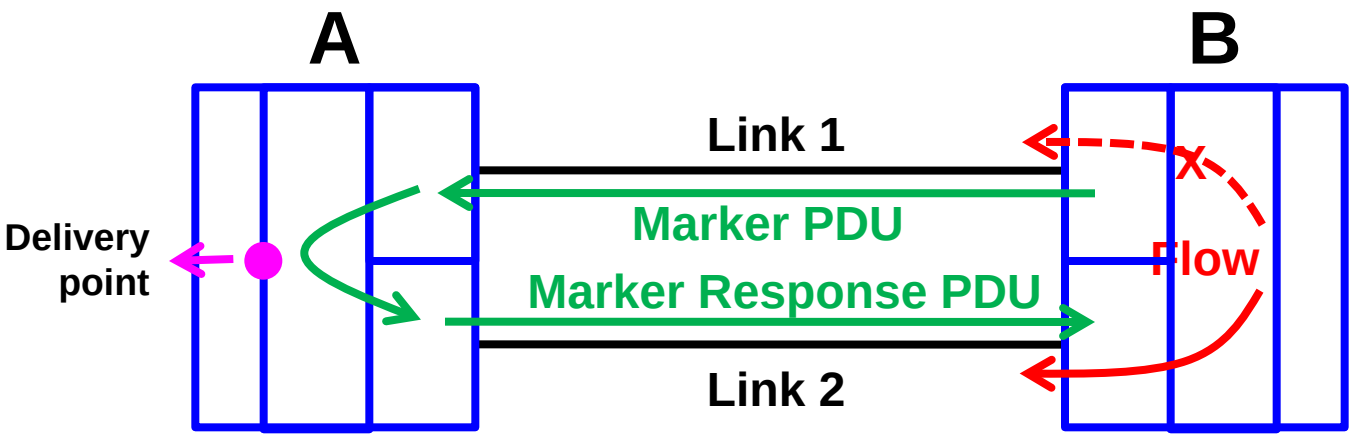
Marker Protocol in 802.1AX



B wants to shift a flow from Link 1 to Link 2

- Distributor **B** stops flow on Link 1.
- Marker Generator **B** sends Marker PDU on Link 1.
- Responder **A** returns Marker Response on any link.
- Receiver **B** receives Marker Response or times out.
- Distributer **B** starts flow on Link 2.

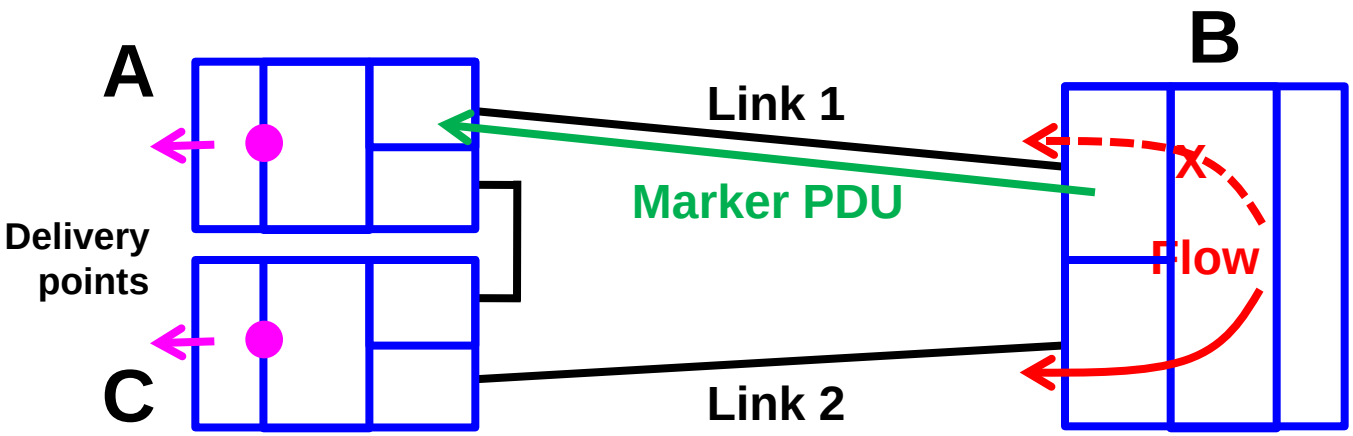
Marker Protocol in 802.1AX



How does this guarantee no out-of-order deliveries?

- Marker PDU has low priority in **B**'s queues on Link 1.
- Only one delivery point (●).
- Flow can start on Link 2 as soon as Marker PDU reaches the single delivery point.

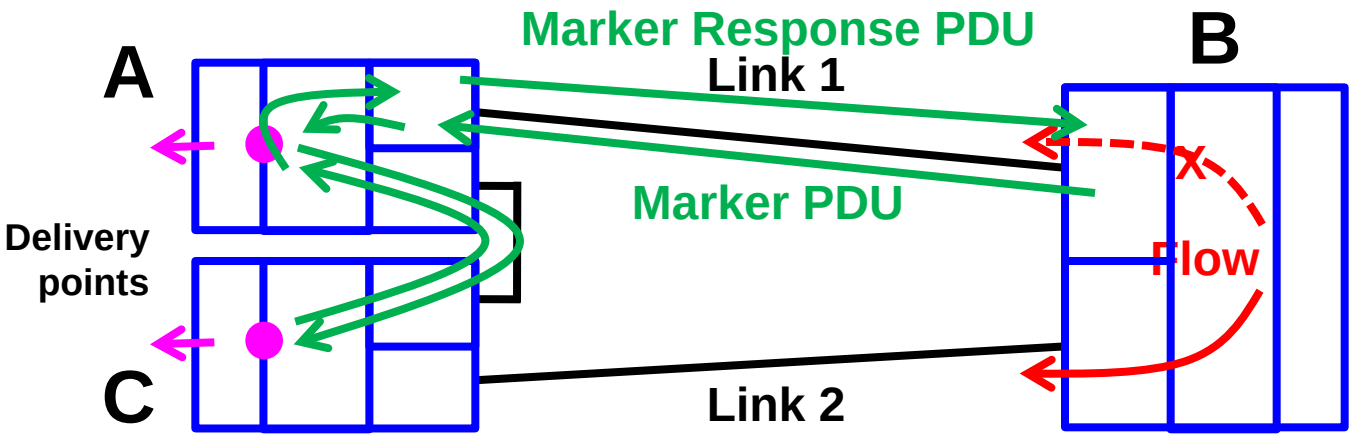
Marker Protocol in 802.1AXbq DRNI



But with DRNI

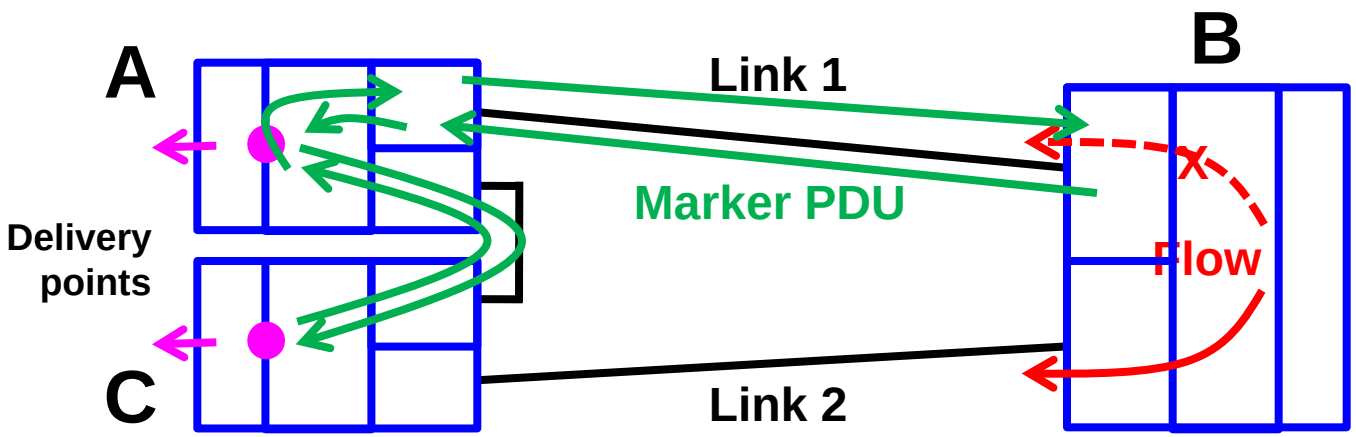
- We now have two delivery points.
- Data flow (or flows) taking Link 1 can spread to **C**.
- Marker cannot be returned until the path from Link 1 to **all** delivery points has been flushed.

Marker Protocol in 802.1AXbq DRNI



- **A**'s Marker Receiver/Generator must remember Marker PDU from **B**.
- A sends new Marker PDUs to **C** (and any other member of Distributed Aggregation Component [DAC]).
- When all have been returned to **A**, **A** can send Marker Response PDU back to **B**.

Marker Protocol in 802.1AXbq DRNI



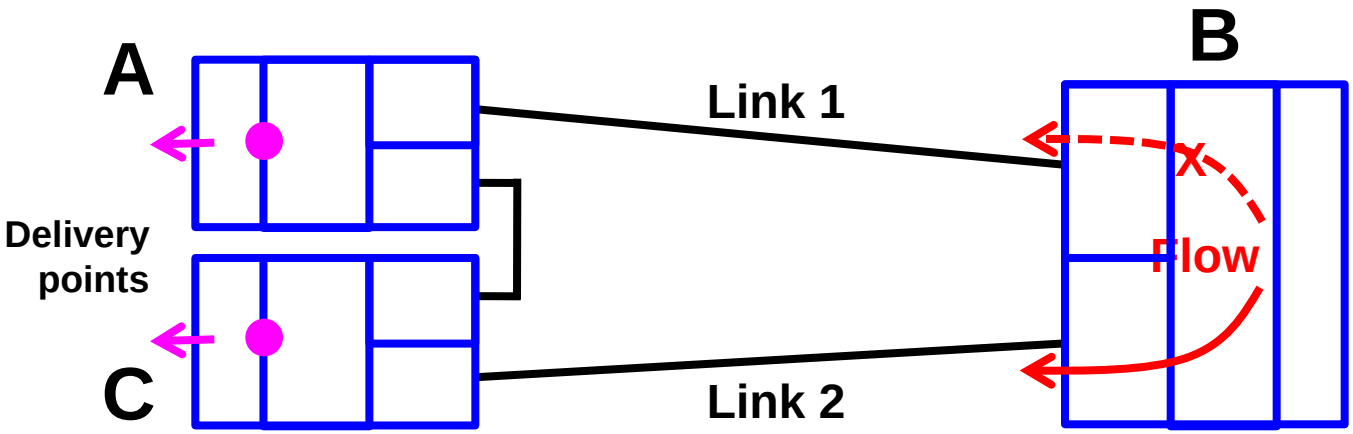
- This requires that **A** maintain a dynamic database to remember the Marker from **B** and track the responses from **C** and other members of **A**'s DAC, along with a timer to delete the entry if the responses are not returned.

Marker Protocol in 802.1AXbq DRNI

- What drives the need for the database is that the Marker PDU is not associated with any particular flow. If the Marker PDU were identified with a particular flow (e.g., a VLAN), **A** could forward the marker to just the relevant delivery point without state, and the Marker Response could be returned along any path.
- But, one of the advantages of the current Marker PDU is precisely that it is not associated with any particular flow, so that one Marker PDU can serve to transfer any number of flows (e.g., 100s of VLANs) to another Link.

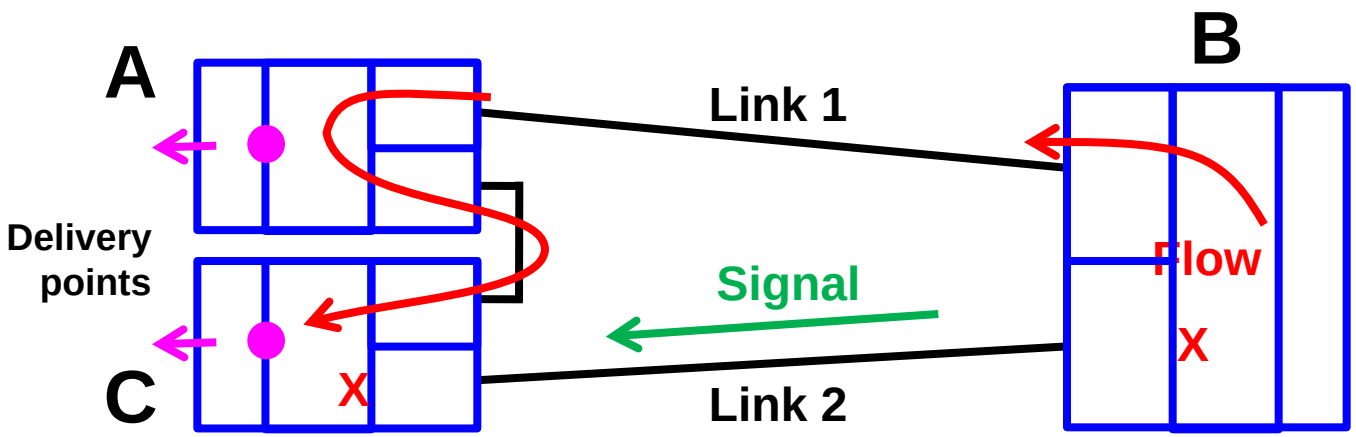
Another solution to avoid out-of-order delivery

Avoiding out-of-order without Markers



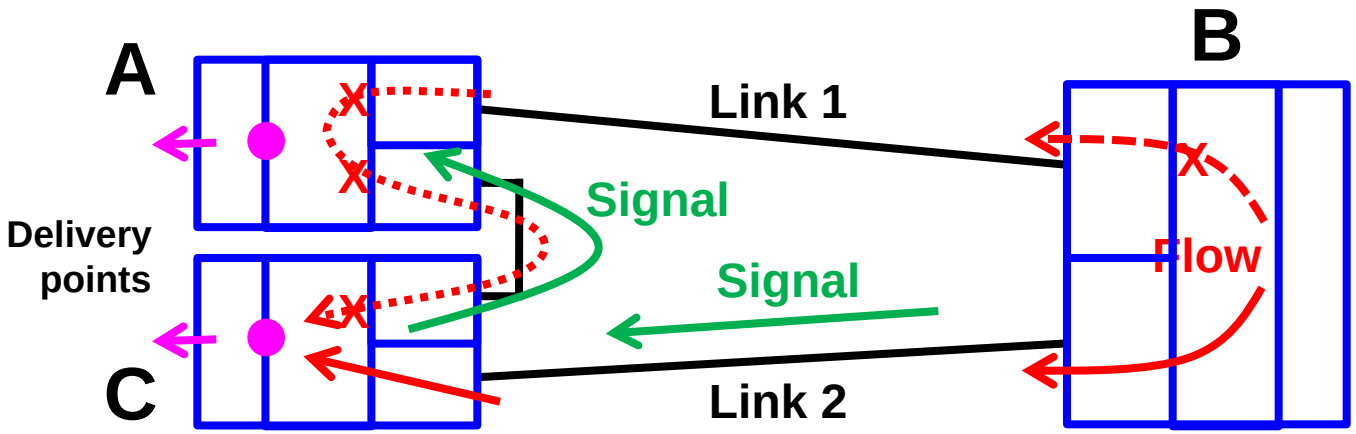
- For any given flow, exactly one of the delivery points in a DRNI is active.
- We would like the flows to be symmetrical – a given flow uses the same Link in both directions.
- **If DAC A-C discards frames arriving on an unexpected Link, no out-of-order delivery can occur.**

Avoiding out-of-order without Markers



- Initial state: **B** is sending some flow(s) over Link 1.
- **C** has the delivery point for these flows.
- **A** and **C** enable these flows to pass through some ports, but block them from others (e.g. **C**'s Link 2 port).

Avoiding out-of-order without Markers



- **B** signals its intention to move some number of flows from Link 2 to Link 2, and moves the flows.
- **C** relays this signal to **A**.
- **C** stops accepting frames for those flows from the **A-C** Link, and starts accepting them from Link 2.
- No matter whether **C** or **B** switches first, no out-of-order deliveries can occur; only lost frames can occur.

Avoiding out-of-order without Markers

- Note that fewer frames are lost than is typical for the Marker Protocol, because the time taken to return the Marker Response is wasted.
- This idea only works if we have a mechanism for signaling flow choices. This is easy for 4094 VLANs, but difficult for 5-tuples.
- This idea only works if the receiving end enforces the transmitting end's decisions. This is easiest if decisions are bi-directional with respect to the data direction.
- Some kind of signal relaying and handshaking is required. Details To Be Determined.

Conclusions

Conclusions and questions.

- Responding to a current, existing system's Marker PDUs **requires** a state machine, a dynamic database, and a multiplication of Marker PDUs through the DAC.
- Ensuring out-of-order delivery using a common means of distribution (e.g. per-VLAN) plus enforcement on data reception gives us what we want without Markers.
- Do we:
  1. Define and/or require dynamic Marker multiplication?
  2. Define and/or require reception enforcement of distribution choices?
  3. Some combination of the above?