Flooding with Context Containment

Marc Holness, Nortel Networks
September 2002

Objective

- Demonstrate adherence to 802.1 packet reorder, duplication and loss requirements
- Summarize flooding technique

Terminology

- Clockwise (CW)
  - The RPR ringlet where packets travel in the clockwise direction
- Counter Clockwise (CCW)
  - The RPR ringlet where packets travel in the counter clockwise direction
- Flood
  - A transmission mechanism that ensures all RPR stations see a transmitted packet once for a given ringlet
- Flooding Scope (FS)
  - The number of hops a packet travels from a given source station to a destination station
- Context
  - The steering database used by a source station to steer traffic

Flooding Modes

Two modes of operation supported by this technique

- Strict mode: Adheres to 802.1 packet reorder, duplication, and loss requirements
    i. There is no guarantee that Service Data Units (SDUs) are delivered
    ii. Reordering of frames with a given user priority for a given combination of SA and DA is not permitted
    iii. Duplication of user data frames (to a client) is not permitted

- Relaxed mode: Adheres to 802.1 requirements under normal ring operation
    - In the advent of a protection event, an amount of reorder and/or duplication can be encountered

Relaxed Mode Flooding

- Uni-directional flooding uses SA or TTL to scope travel of packet
- Bi-directional flooding using TTL to scope the travel of CW packet and CCW packet
- MAC stripping rules outlined and described in RPR Draft 1.0, clause 6.8
- Supported by RPR frame structure described in RPR Draft 1.0, clause 8.0

Relaxed Mode Non Compliances

- Scenarios resulting in possible reorder/duplication
  - Ring (link or station) restoration event
  - Topology image of stations on the ring not being synchronized
  - Station failure resulting in pass-thru behavior (i.e., packets are sent through transit path unaffected)
  - Compound ring (link or station) failures resulting in segmented chains
  - Rapid cascading ring (link or station) failures

Strict Mode Requirements

- TTL and TTL_Base
- In-flight packets using stale context are killed upon protection switch detection
- Flooding type indication
- Wrap state indication
- Packets on secondary ringlet after unwrap gets deleted
- Station pass-thru can be supported

Strict Mode - Wrapping

Key elements of technique

- Add WrapState (WS) bits to RPR header
  - WS cleared when packet is launched
  - WE (WrapEligible) must be set in order for packet to be wrapped on secondary ringlet (along with WS being clear)
  - WS set when packet passes the source station on the secondary ringlet
  - Packet eligible for wrapping back on primary ringlet when WS and WE set
  - Once rewrapped, packet cannot be wrapped again
- All packets on the secondary ringlet is deleted from the ringlet following the healing of a protection event

Strict Mode - Steering

Key element of technique

- Context Containment: In-flight data packets are killed upon detection of protection switch event
    - Applicable to local unicast and multicast/broadcast/flood transmissions

Impacts

- Additional state checks required on data path to support wrapping systems
- If pass-thru supported one additional check required
  - `((TTL_Base-TTL)!=hops[SA])`
- No packet delivery on ring for ~15ms upon detection of protection switch event (in strict mode)

Conclusion

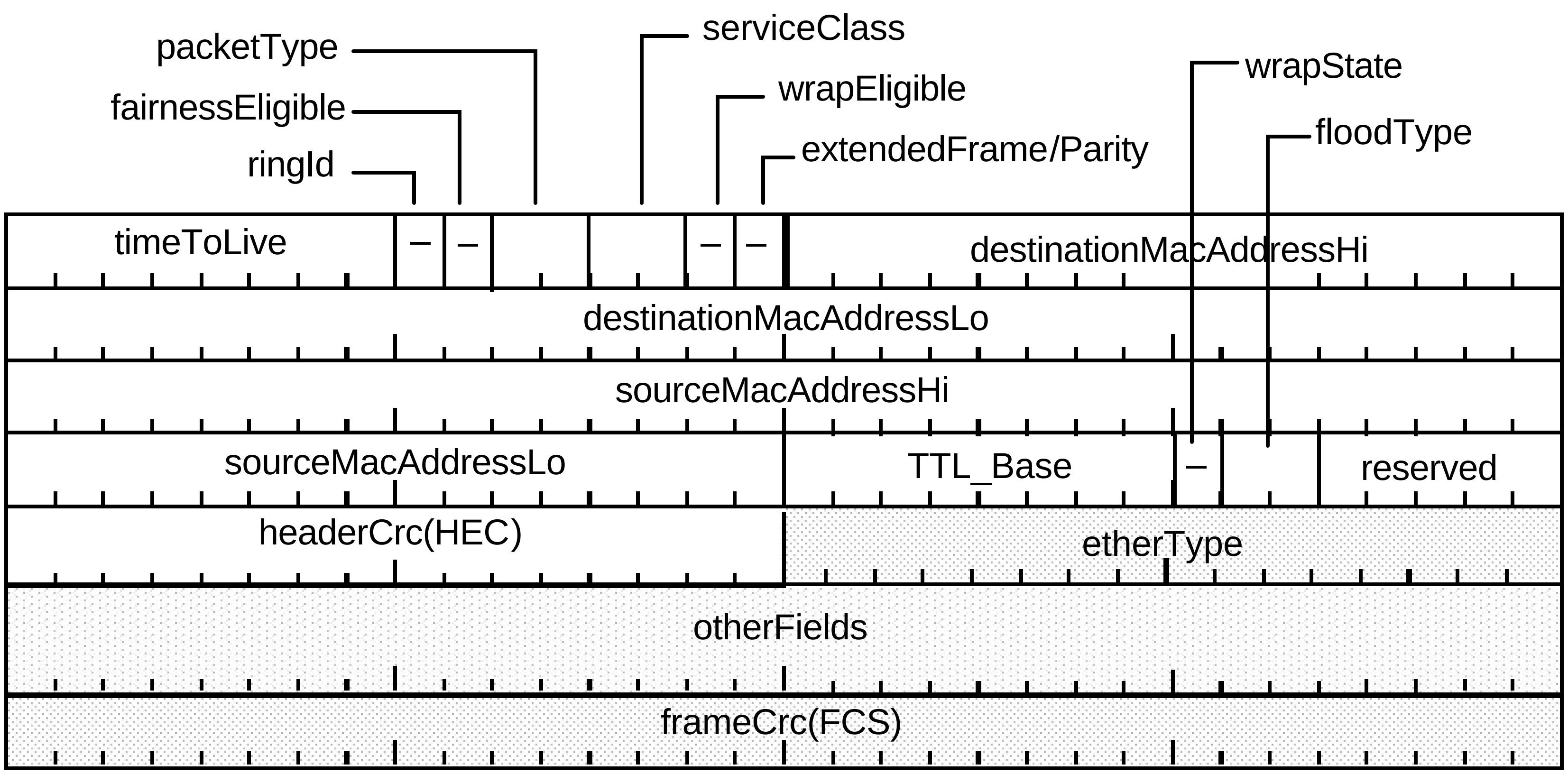
- Compliance to 802.1 can be achieved by this technique when operating in strict mode
- This flooding technique utilizes
  - Context containment for steering systems (in strict mode)
  - Wrap state information and data path stripping for un-directional floods with wrapping
  - Two modes of operation: strict and relaxed

Back Up

Context Containment

- Upon detection of a protection switch event, a node will remove in-flight data packets from the ring (that were launched using an old context)
  - Kills received data packets and data packets within TB(s)
  - Duration of kill is 15ms
    - Assume (max) 2000km ring span plus margin

RPR Frame Structure



Scenarios and Proofs

Packet duplication prevention

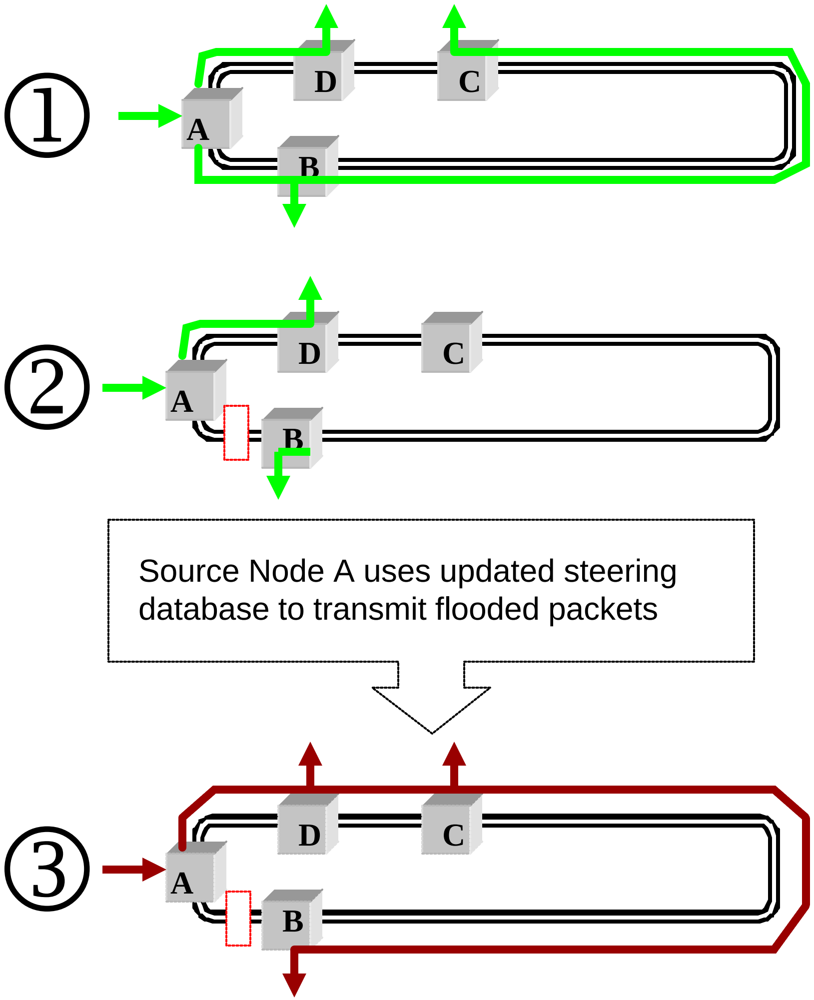
- If passthru is supported
  - Hop count consistency check: (`TTL_Base-TTL`)!=hops[SA]
- Wrapping systems
  - Once wrapped, a packet can not be rewrapped
- Steering systems
  - TTL scoping rules employed by source node will guarantee no duplication

Scenarios and Proofs

Packet reorder prevention

- Wrapping systems
  - Packets on secondary ringlet are killed following a healing of a protection event
  - Once wrapped, a packet can not be rewrapped
- Steering systems
  - Context containment: In-flight data packets are killed upon detection of protection switch event (i.e., packets on ring launched with old context are deleted)

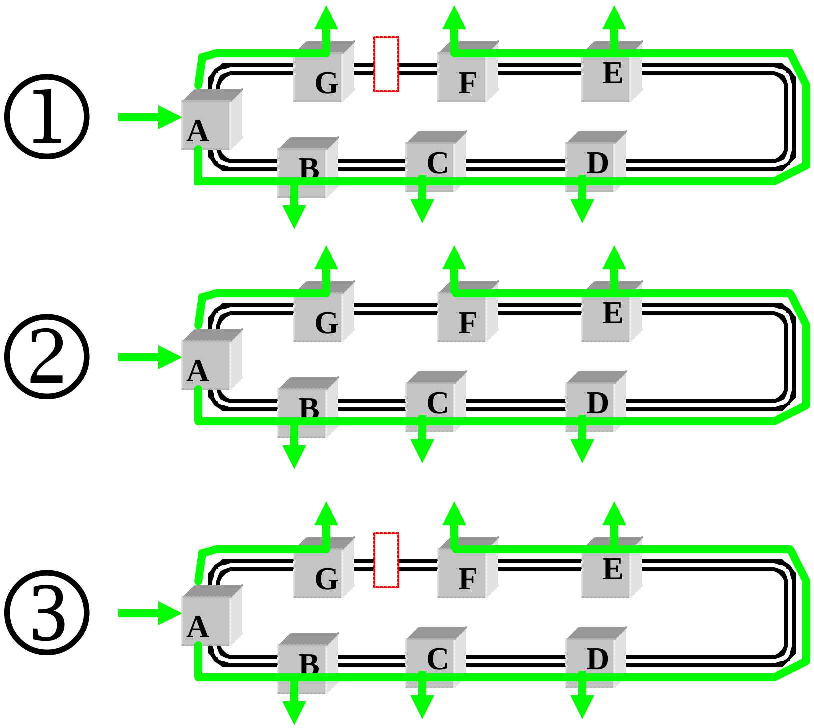
Scenario #1 – Protection Switch



Scenario Walk-thru

- At step 2, protection control packets launched
- At this point, every node that receives a protection control packet will
  - Update steering dB, and
  - Discard all data packets received and in TBs, for a duration of 15ms (i.e., kill all in-flight data packet launched using old context)
- At step 3, node A launches packets using new context
- No packet reorder can occurred!

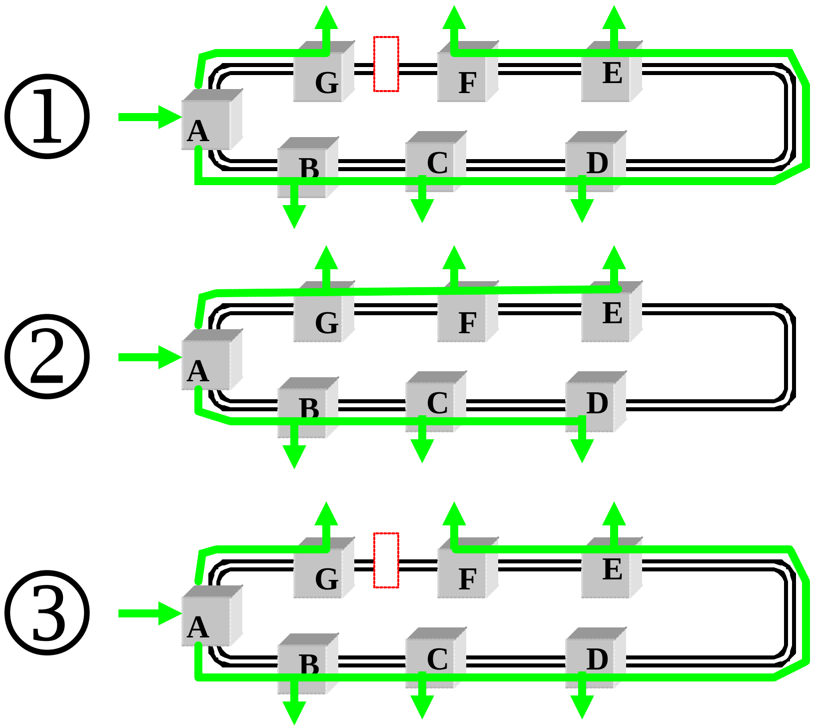
Scenario #2a – Cascading Failures



Scenario Walk-thru

- At step 1, consider packets in-flight on CCW ringlet (using context #1)
- At step 2, WTR timer set. Station A does not update its steering dB until WTR expires and protection control packets get broadcast
- Assume step 3 occurs prior to WTR expiry. Any protection control packets broadcasted do not change the steering dB associated with all the nodes (e.g., node A)
- No packet reorder can occurred!

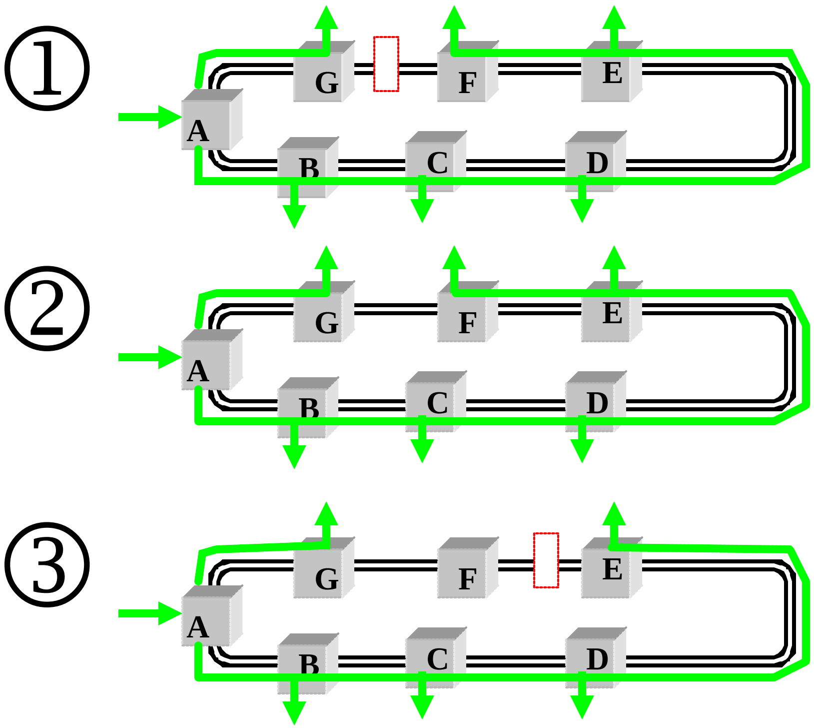
Scenario #2b – Cascading Failures



Scenario Walk-thru

- At step 1, consider packets in-flight on CCW ringlet (using context #1)
- At step 2, WTR timer set. Station A does not update its steering dB until WTR expires and protection control packets get broadcast
- Assume WTR timer expires
- At this point, every node that receives a protection control packet will
  - Update steering dB, and
  - Discard all data packets received and in TBs, for a duration of 15ms (i.e., context #1 packets are killed)
- At step 3, another protection event is detected and in-flight data packets launched using old context is deleted
- No packet reorder can occurred!

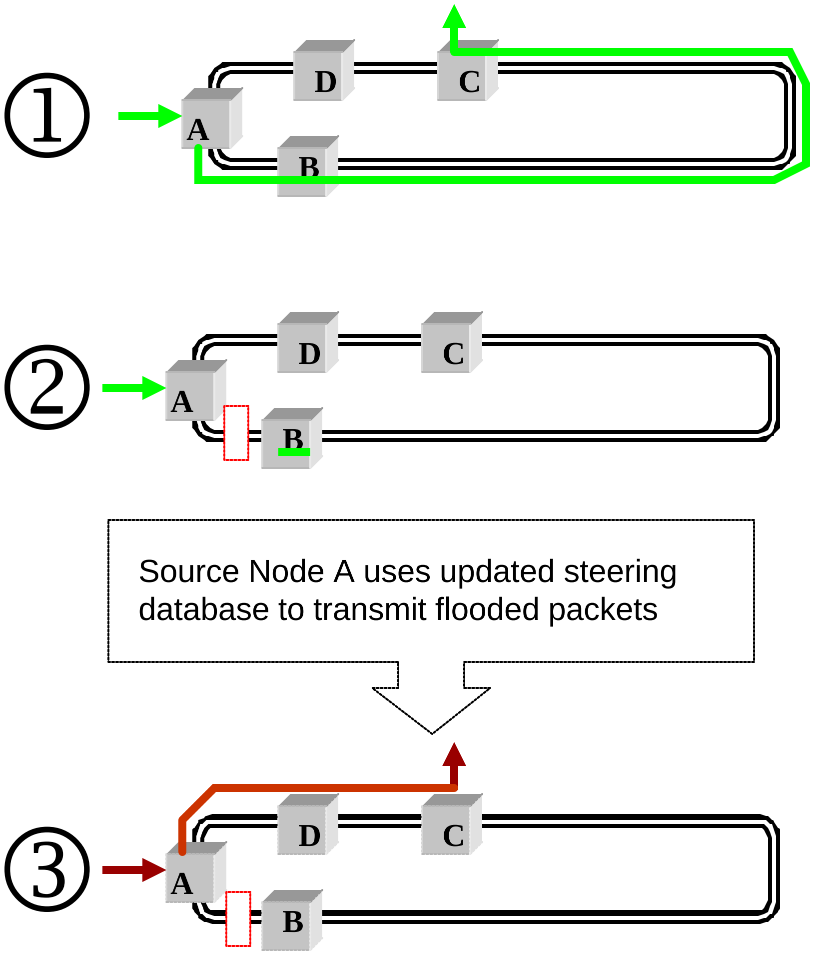
Scenario #3 – Cascading Failures



Scenario Walk-thru

- At step 1, consider packets in-flight on CCW ringlet (using context #1)
- At step 2, WTR timer set. Station A does not update its steering dB until WTR expires and protection control packets get broadcast
- Assume step 3 occurs prior to WTR expiry. Protection control packets get broadcasted
- At this point, every node that receives a protection control packet will
  - Update steering dB, and
  - Discard all data packets received and in TBs, for a duration of 15ms (i.e., context #1 packets are killed)
- Node A launches packets using new context (i.e., context #3).
- No packet reorder can occur!

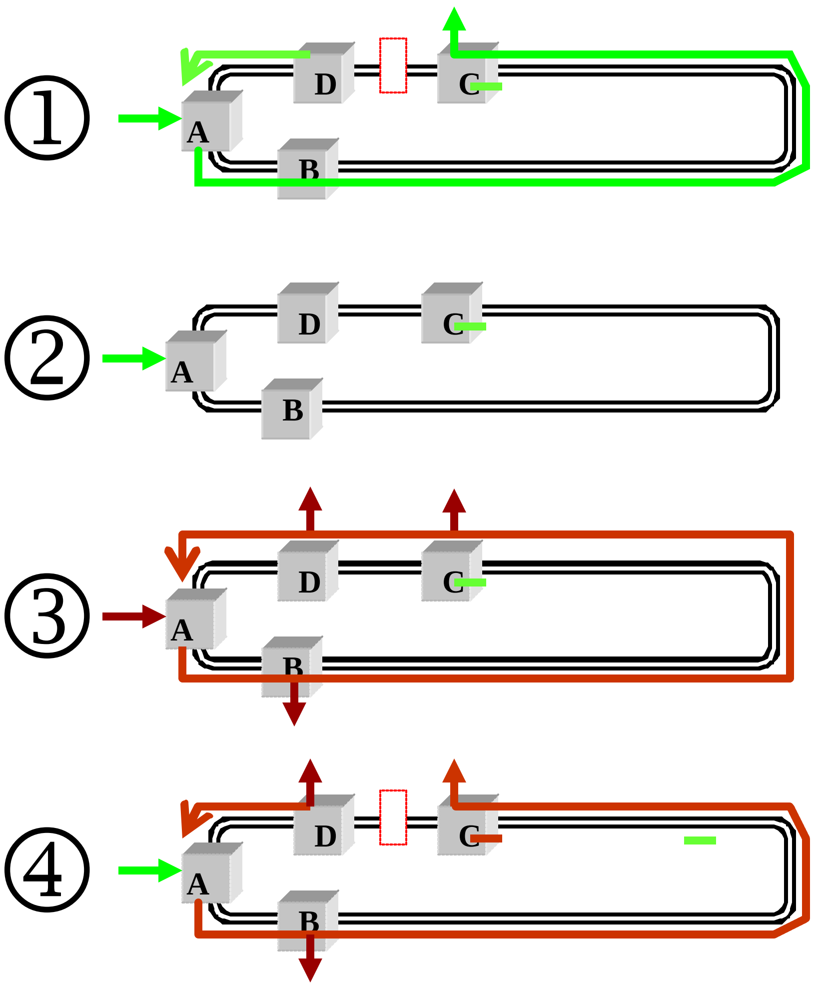
Scenario #4 – Protection Switch



Scenario Walk-thru

- At step 1, station A is sending unicast traffic destined to station C
- At step 2, protection event is detected and protection control packets are launched
- At this point, every node that receives a protection control packet will
  - Update steering dB, and
  - Discard all data packets received and in TBs, for a duration of 15ms (i.e., kill all in-flight data packet launched using old context)
- At step 3, node A launches packets using new context
- No packet reorder can occurred!

Scenario #5 – Protection Switch



Scenario Walk-thru

- Consider a wrapping system, where there is a link failure at link DC
- At step 1, station A is flooding (unidirectional) packets on the CCW ringlet
- At step 2, link DC heals
  - Protection control packets are launched by station D and station C
  - Packets on secondary ringlet circulate (until TTL expires)
- At step 3, station A is flooding (unidirectional) packets on the CCW ringlet
- Stations detect protection (heal) event and discard all data packets where packet.RI is not equal to ringlet, for a 15ms duration
- At step 4, link DC experiences a failure during 15ms timer
  - Stations D and C wrap
  - Data packets with wrong RI still get discarded until 15ms timer expires
- Packet reorder is prevented!