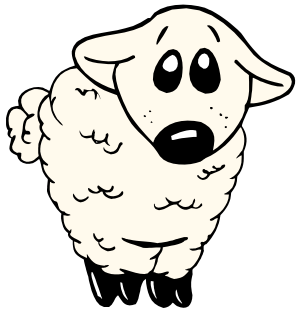
Flooding in 802.17 Networks

Anoop Ghanwani, Lantern
(anoop@lanterncom.com)
Li Mo, ZTE
(limo01@yahoo.com)

IEEE 802.17 Interim Meeting
Ottawa, ON
May 2002



Motivation

- There is more than one way to flood packets in RPR
- Need a comprehensive analysis of flooding techniques
- Need to evaluate each of the methods with respect to packet duplication, reordering, and loss at times when the ring is subject to a change
- The task was assigned to the bridging ad-hoc (BAH)
- This presentation is the outcome of that effort

MAC Service and the Issues of Loss, Reordering, and Packet Duplication

- The expectations of the MAC service are described in IEEE 802.1D-1998 (because bridges need to preserve the MAC service)
- Frames may be lost due to many reasons including corruption at the physical layer (Clause 6.3.2)
- Frame reordering is not allowed for a given user priority and a given source and destination MAC address (Clause 6.3.3)
- Frame duplication is not allowed under any circumstances (Clause 6.3.4)

Methods for Flooding In RPR

- Unidirectional
  - Strip based on source station identification only
    - TTL may be larger than number of stations on ring
    - Allows sending of traffic to/from stations even before topology has converged
    - Does not work for bridging with the current frame format
  - Strip based on TTL only (required for bridging)
    - TTL = N – 1, where N is the number of stations on the ring
  - Strip based on source and TTL
    - TTL limits propagation
    - Source stripping protects the packet in cases where a station on the ring dies
    - Mainly interesting when doing wrapping
- Bi-directional
  - Always strip based on the TTL
    - TTL_east + TTL_west = N – 1
  - May be symmetric or asymmetric
    - Symmetric: |TTL_east – TTL_west| <= 1
    - Asymmetric: |TTL_east – TTL_west| > 1

Interesting Scenarios

- Link failure
- Node failure (not source of flooded packets)
- Node failure (source of flooded packets)
- Node addition
- Node deletion
  - Protection event followed by a nodes' two neighbors getting directly connected
- Node in pass-through
  - Node goes into pass-through (no TTL decrementing)
  - A nodes' two neighbors get directly connected without any protection event
  - This may be caused by the actions of OXC or SONET Cross Connect
- Any others?

Frame Duplication

- Frames are prevented from circulating forever by the TTL
- However, depending on the TTL value, you could have packet duplication until the frame gets stripped
- This is usually not an issue when the source station has a way of identifying frames that it put on the ring

Frame Loss

- In some instances, frames may be “stripped too early”
  - Due to TTL too small when a new station is being added to the ring
  - Due a broken link in the ring for steering frames
- Selective loss is possible
  - Some stations may get the frame
  - However, other stations (possibly one or more of the intended recipients) may not get the frame

Frame Reordering

- Refers to reordering of frames *within* a service class
  - Frames are received in a order different than what they were when presented to the MAC
- It is possible for frames to get reordered
  - During times of protection
  - When changing from one method of flooding to another

Other 802 MACs and Frame Duplication/Loss/Reordering

- Most 802 MACs do not have issues with frame duplication, loss, or reordering
- These are problems that can sometimes happen in bridged network, but usually not on a single MAC
- 802.3
  - Under normal circumstances, no frame duplication, loss, or reordering at the MAC level
- 802.5
  - Ring monitor sets a monitor bit when it sees a frame so it can pull the frame off if the source dies
  - Damage is lower because only one outstanding frame on the ring at any time
  - The risk of duplication is no worse regardless of whether the source of the frame was local to the ring or not
  - No possibility of reordering

Unidirectional Source-based Stripping

SCENARIO	STEERING	WRAPPING
Link failure	Loss Reordering (1)	No issues
Node failure (not source)	Same as link failure	No issues
Node failure (source)	No issues	Duplication (2)
Recovery from failure (revert back to this mode)	Reordering (3)	Reordering (3)
Node Addition	Same as link failure	Same as link failure
Node Deletion	Same as node failure	Same as node failure
Node in pass-through	Duplication (4)	Duplication (4)

Unidirectional Source-based Stripping (Notes)

(1) Must switch to bi-directional flooding with TTL scoping and therefore possible reordering

(2) Duplication happens until TTL causes stripping for frames sourced by the dead node

(3) There may be a change in path length from source to destination

(4) Duplication happens if the node that is put into pass-through has sourced traffic, but hasn't stripped it and will happen until TTL expires for all of those packets; otherwise no issues

Unidirectional TTL-based Stripping

SCENARIO	STEERING	WRAPPING
Link failure	Loss Reordering	No issues
Node failure (not source)	Loss Reordering	Duplication (1)
Node failure (source)	No issues	No issues
Recovery from failure (revert back to this mode)	Reordering	Reordering
Node Addition	Same as link failure (2)	Same as link failure (2)
Node Deletion	Same as node failure (2)	Same as node failure (2)
Node in pass-through	Duplication (3)	Duplication (3)

Unidirectional TTL-based Stripping (Notes)

(1) Because of wrapping the TTL = (N – 1) of frames already sent will cause the source to receive a second copy of the frame; duplication happens for all packets until topology re-converges

(2) Assumes that the system is in "protection" mode until the topology converges; otherwise you can get duplication, loss, reordering

(3) Putting a node in pass-through must be handled by first forcing a protection switch, and then putting the node into pass-through; if a node is just flipped into pass-through directly, we'll get packet duplication of all packets until topology re-converges

Unidirectional Source & TTL-based Stripping

SCENARIO	STEERING	WRAPPING
Link failure	Loss Reordering	No issues
Node failure (not source)	Loss Reordering	No issues
Node failure (source)	No issues	No issues
Recovery from failure (revert back to this mode)	Reordering	Reordering
Node Addition	Same as link failure (2)	Same as link failure (2)
Node Deletion	Same as node failure (2)	Same as node failure (2)
Node in pass-through	Duplication (3)	Duplication (3)

Unidirectional Source & TTL-based Stripping (Notes)

(1) Because of wrapping the TTL = (N – 1) of frames already sent will cause the source to receive a second copy of the frame; duplication happens until TTL causes stripping for frames sourced by the dead node

(2) Assumes that the system is in "protection" mode until the topology converges; otherwise you can get duplication, loss, reordering

(3) Duplication happens only if the source node is placed into pass-through before stripping all of its traffic, and it only happens for that traffic; otherwise, no issues

Bi-directional Flooding

SCENARIO	STEERING	WRAPPING
Link failure	Loss Reordering	No issues
Node failure (not source)	Loss Reordering	Duplication (1)
Node failure (source)	No issues	Duplication (1)
Recovery from failure (revert back to this mode)	Reordering	Reordering
Node Addition	Same as link failure (2)	Same as link failure (2)
Node Deletion	Same as node failure (2)	Same as node failure (2)
Node in pass-through	Duplication	Duplication

Bi-directional Flooding (Notes)

(1) Because of wrapping the TTL = (N – 1) of frames already sent will cause one of the nodes receive a second copy of the frame

(2) Assumes that the system is in “protection” mode until the topology converges; otherwise you can get duplication, loss, reordering

Flooding Requirements for 802.1D/Q

- The MAC must not reorder packets within a flow
  - A flow is defined as a DA, SA and user_priority
- The MAC must not duplicate packets

To satisfy these requirements, 802.17 must specify a mode of operation where there is no reordering or duplication

Avoiding Frame Reordering

- Causes of frame reordering
  - Change in the flooding method
    - Unidirectional to bi-directional or vice versa due to steering protection/recovery
    - Changing the TTL for each ringlet in the bi-directional method
  - Unwrapping after recovery from failure
- Reordering is always a transient condition
- When changing the flooding method, each station must
  - Stop sending (sourcing) any data
  - Start a timer approximately equal to a RTT
  - Send a packet to itself
  - If the packet arrives, or if the timer expires then it is OK to switch to the new method
  - Essentially, this requires a RTT of silence from a station before it can change the flooding method
- When unwrapping data must be dropped for a RTT

Avoiding Frame Duplication

- Frame duplication can be completely avoided by the doing the following:
  - Stations are never allowed to go into a pass-through mode
  - The flooding mode during protection is maintained even after recovery if a topology change is detected (node add or node delete)
  - If steering is used
    - For unidirectional, source stripping is sufficient if available
    - For bi-directional, TTL stripping is required
  - If wrapping is used
    - Flooding must always be unidirectional
    - Source + TTL must be used for stripping
- If TTL is used, flooding cannot be allowed until topology converges

Some Observations

- Allowing a failed node to go into pass-through is a bad idea
- Putting a node in pass-through requires
  - A way to identify the ring-local source station, and
  - The station does not source packets for a RTT before going into pass-through
- Once a node enters protection and starts steering, it should revert to the original flooding method only after topology has known to have converged
  - An exception is the case where source-based stripping is done
  - The node can revert to unidirectional flooding as soon as the ring is detected to have healed
  - The new node can start receiving traffic immediately

Conclusions

- All methods have problems with loss and reordering
- The most robust methods with respect to frame duplication are:
  - Unidirectional source-based stripping with steering, and
  - Unidirectional source & TTL-based stripping
  - In either, duplication happens only if the source is put in pass-through without a protection event, and without first having stripped all of its traffic
- The choice of flooding method to use is a local issue and need not be standardized as long as:
  - TTL is decremented properly by all stations which are not in the pass-through mode
  - When in pass-through, TTL is *not* decremented, if TTL stripping is used
  - The station that puts the frames on the ring knows how to strip it off reliably (or limit its scope)
  - A node is not put into pass-through without first signaling a protection event