Frame Duplication and Misordering in RSTP Bridges

Tony Jeffree
Friday, March 03, 2000

This is a discussion paper for 802.1, addressing the potential issues of frame misordering and duplication in Bridged LANs containing Rapid Spanning Tree Bridges.

1. Background

In 802.1D:1998 (and earlier) Bridges using STP, the inherent delays imposed by STP before a Port is made Forwarding ensure that any frames of a conversation that were "in transit" prior to a reconfiguration event will have been delivered or discarded before any frames are transmitted on a Port that is made Forwarding as a result of the reconfiguration event. Hence, the risk of frame duplication and/or misordering on reconfiguration in LANs comprising legacy Bridges is small.

RSTP by its nature and intent can reduce the delay before a Port is made Forwarding to very small values; in the case of Root Port transitions, the delay can be as short as the hardware can manage it, or in the case of Designated Port transitions, as short as two frame transmission times on a single segment. Hence, it is conceivable that a reconfiguration event can take place, and a Port be made Forwarding as a result, while prior frames of a conversation are still in flight.

This paper attempts to examine these issues and evaluate the circumstances in which a problem might occur.

2. Frame Duplication

For a unicast frame, where the destination address of the frame has been learnt by the Bridges that may be required to forward it, that frame can only be buffered for transmission on one Port of one Bridge at any one time. Hence, it would appear to be impossible for frame duplication to occur in this situation.

If the unicast address has not been learnt, then the frame will be flooded (and therefore buffered) on all outbound Ports; there is therefore a possibility that a reconfiguration event could cause duplication of a flooded frame. However, as any subsequent response from the destination system would cause the address to be learned, the worst case for "normal" LAN protocols arising from this scenario would be two connect requests being seen by the destination machine, as all subsequent components of the conversation would be based on learned addresses.

Multicast frames, as these will in general be buffered at multiple outbound Ports, are potentially subject to duplication on reconfiguration events. However, of the multicast-based protocols that are in common use in LANs, the only one that would be sensitive to frame duplicatiojn is GARP, and GARP frames are not subject to frame forwarding in Bridges. Hence, this scenario seems not to be an issue.

3. Duplication – Conclusions

There is a small risk of duplication of unlearned unicast frames, and of multicast frames, on reconfiguration events. However, in both cases, it does not appear to result in any damage to known LAN protocols.

4. Frame Misordering

It is possible to conceive of a situation where, prior to a reconfiguration event, an end-to-end conversation was required to transit the maximum diameter of the Bridged LAN, and following reconfiguration, the same conversation only transits a single Bridge. This could happen, for example, with a Bridge that has a Root Port connected to one end of a LAN "diameter" and a Alternate Port connected to the other end of the "diameter". Block the Root Port and make the Alternate Port Forwarding, and an almost instantaneous switch of location occurs for end stations downstream of that Bridge, from one "end" of the LAN to the other. Clearly, as there could be frames in transit before the reconfiguration, frames transmitted immediately following the reconfiguration could arrive at their destination before the ones in transit, resulting in frame misordring as seen by the recipient.

How frequent an event is this likely to be in real LANs?

Are there any LAN protocols in common use for which this causes a problem? (Again, GARP is sensitive to misordering, but GARP frames will not be misordered under these circumstances, as they are not forwarded by Bridges.)