Title:	Liaison response to ITU-T SG15 LS-290
From:	IEEE 802.1
For:	Information
Contacts:	Glenn Parsons, Chair, IEEE 802.1 (glenn.parsons@ericsson.com)
To:	ITU-T Study Group 15 (tsbsg15@itu.int) Tom Huber, Rapporteur Q9/15 (tom.huber@coriant.com) Han Li, Associate Rapporteur Q9/15 (lihan@chinamobile.com)
Date:	January 21, 2015
Reference:	COM 15 – LS 290 – E

Thank you for your liaison COM 15 – LS 290 – E. We appreciate your continued interest in IEEE Std 802.1AX™-2014.
DRNI is able to meet a Transfer time of 50ms in the case of a single link failure as well as in the case of a single node failure. Please find details below.

The basic unit of processing that occurs each time the DRNI process is woken up is to process a received LACPDU (if any) and DRCPDU (if any), process a local state change (if any, such as a link down detection), and generate a LACPDU (if necessary) and DRCPDU (if necessary) for transmission. The actual processing time is very small (microseconds on any modern CPU). The limiting factor is how frequently the DRNI process can be woken up in a realistic system. We observe that CFM implementations are able to maintain CCM transmissions (which have a similar basic unit of processing) at 3.3ms intervals for a large number of MEPs. We can take this as a guideline and assume DRNI will take 3.3ms to be woken up and perform one unit of processing. Note that 802.1 typically assumes 1ms processing intervals, which is probably reasonable given that DRNI processing occurs per physical port whereas CCM is per service, however for this analysis we will use the 3.3ms interval as a guideline.

DRNI packets, LACPDUs and DRCPDUs, traverse individual links; they are not forwarded around the entire network diameter. Taking each link as ¼ the length of the entire network diameter is a sufficiently accurate estimate since the number of LACPDUs and DRCPDUs transmitted on an Aggregation Link or Intra-Portal Link respectively is approximately equal. Therefore, we can calculate the transmission delay on a 300km link as approximately 1.5ms. Since the actual DRNI processing takes a very small portion of the 3.3ms intervals, it is possible for the transmission to occur and to receive a response from the neighbor before the next DRNI processing interval. However, this is the best case. If we focus on the worst case, the response would not be received until the second processing interval after it was transmitted, so the time for round trip communication is 6.6ms.

Referring to the diagram in LS 290, to consider the case of an Aggregation Link failure assume that Aggregation Link 1 fails. The failure is detected by Gateway 1 (hopefully by an immediate indication from the physical layer, otherwise from CFM continuity check). Gateway 1 reacts to the local state change by sending a DRCPDU to Gateway 3 with the new state and resulting configuration. Gateway 3 responds with a DRCPDU confirming the new configuration. Gateway 1 processes the received DRCPDU in the second processing interval (i.e. 6.6ms) after its initial reaction to the state change, for a total Transfer Time comfortably less than 10ms. Gateway 2 and Gateway 4 would be undergoing the same sequence in parallel, and would be expected to conclude at roughly the same time.

To consider the case of an Intra-Portal Link failure, assume that IPL 1 fails. This is detected by Gateway 1 with the result of sending an LACPDU to Gateway 2 to communicate the change of state. This is followed by two round-trip DRCPDU exchanges between Gateway 2 and Gateway 4. (Other LACPDU exchanges on both Aggregation Links occur in parallel with the DRCPDU exchanges, but the critical time is when the DRCPDU exchanges complete.) The result is a total of five one way frame transmissions (15.5ms) for a Transfer Time comfortably under 20ms.

To consider the case of a node failure, assume that Gateway 3 fails. This is actually faster than the above case because both Gateway 1 and Gateway 4 can detect the failure directly. As above, Gateway 1 reacts by sending an LACPDU to Gateway 2. In parallel, Gateway 4 reacts by sending a DRCPDU to Gateway 2. Gateway 2 then sends a DRCPDU to Gateway 4 and a LACPDU to Gateway 1 (during the same unit of processing assuming that both frames had been received prior to the DRNI process waking up). Gateway 1 sends the final LACPDU to Gateway 2 for a total of three one way frame transmissions and a Transfer Time of approximately 10ms.

With respect to the text quoted from our March 2014 liaison, an example of "poor implementation choices" would be having a maximum frequency of waking up the DRNI process that is less than 100 times per second (i.e. 10ms intervals). An example of "slow transmission rates" and "long wires" would be if the physical media delay dominated the Transfer Time by having a physical transmission time per frame greater than 10ms.

IEEE 802.1 looks forward to maintain the good relationship with ITU-T Q9/15 and to further cooperation between our organizations.

IEEE 802.1 will be meeting next in Macau, 14-17 March 2016.