August 4, 2015

To: David Law and members of the IEEE 802.3 Working Group
CC: Pete Anslow, 802.3 WG Secretary

Subject: Flex Ethernet Implementation Agreement

From: Nathan Tracy, OIF Technical Committee Chair (ntracy@te.com)

Dear Mr. Law and members of IEEE 802.3,

We would like to inform you of our recent work to develop a Flex Ethernet implementation agreement. Agreement to start this project was reached at our 1Q2015 meeting, and at our 3Q2015 meeting in Ottawa our draft has reached the stage to issue for straw ballot. The project start proposal and current draft implementation agreement are attached.

This implementation agreement provides a bonding mechanism to create higher-rate interfaces out of multiple Ethernet PHYs, a mechanism to support smaller clients (Ethernet flows with lower effective MAC rates) over Ethernet PHYs, and a mechanism to multiplex multiple lower rate flows across a group of Ethernet PHYs. The first version of this implementation agreement is based on the bonding of 100GBASE-R Ethernet PHYs into a FlexE group. A future version is expected to support bonding of higher rate Ethernet PHYs such as 400G.

This work is an evolution of prior work on the Multi-link Gearbox (MLG), providing a more generalized way of carrying Ethernet flows at a variety of different rates over physical interfaces developed by IEEE 802.3. The general mechanism assumes that every client signal is encoded as a sequence of 66B blocks, encoded according to IEEE Std 802.3™-2012 clause 82. Different flows are interleaved on a 66B block basis according to a calendar that allocates 66B block positions to different flows, and the aggregate signal is then scrambled and distributed to PCS lanes using all of the lower layers of 100GBASE-R as specified by IEEE 802.3.

One area where we would request your guidance is the following: The definition of this mechanism requires the use of a 66B block to delineate the FlexE overhead information which can be distinguished from any 66B block which might be part of 66B encoded Ethernet information or lane marker. The most obvious choice for encoding of this block is as an ordered set. To distinguish this ordered set from the sequence ordered set used for link fault signaling in Ethernet (as well as to be distinct from the signal ordered set used by Fibre channel), we believe it would be safest to select a different "O" code than either of those currently in use for the IEEE 802.3 defined sequence ordered set or the INCITS T11 defined signal ordered set. The current FlexE draft uses the value 0x5, which is one of several possible values which would have a Hamming distance of 2 from either of the two 4-bit "O" codes currently in use by existing technologies. We would appreciate your advice as to whether this is an appropriate choice. Please advise us if there is a different "O" code you prefer that we use, or if you think there is a better way this information can be encoded which would still traverse the lower layers (including lane distribution, transcoding, RS-FEC in particular) of the 100GBASE-R stack as defined by IEEE 802.3.

As you review the draft, please consider that the solutions are in the process of iterative review and technical changes may still occur prior to finalization of these implementation agreements.

Sincerely,

Nathan Tracy,
OIF Technical Committee Chair (ntracy@te.com)

Attach:
oif2015.039.02 (Flex Ethernet Project Start Proposal)
oif2015.127.02 (FlexE Implementation Agreement Draft 1.1)