OIF Flex Ethernet Liaison Statement Proposed Actions

Steve Trowbridge
Alcatel-Lucent

Supporters

- Mark Gustlin (Xilinx)
- Faisal Dada (Xilinx)
- Steve Gorshe (PMC-Sierra)
- Pete Anslow (Ciena)
- Sebastien Gareau (Ciena)
- David Ofelt (Juniper)
- Gary Nicholl (Cisco)
- Peter Stassar (Huawei)
- Juan-Carlos Calderon (Inphi)
- Jean-Michel Caia (Inphi)
- Brad Booth (Microsoft)

OIF Liaison Statements

- OIF_to_802d3_FlexE_Aug_2015.pdf was received before September interim meeting, advising of OIF work on Flex Ethernet and requesting an “O” code for a new kind of ordered set to be used as a deskew marker and indication of overhead position
- After request for clarification, OIF_to_802d3_FlexE_Oct_2015.pdf was received indicating that the first version of the Implementation agreement would use this ordered set on 100G PHYs, but they would expect to update the IA to use future higher-rate PHYs (including 400G) in the future. No plan to define Flex Ethernet or use the ordered set on currently standardized Ethernet rates <100G.

Pro/con of making assignment

- There isn't really any good alternative to what they are asking for what they are trying to do. Other "tricks" like sync header violations won't get past 257B transcoding. Values other than a new ordered set "O" code and D-bytes fall into the group of things the standard says "Shall not be transmitted and shall be considered an error if received"
- The supply of "O" codes isn't constrained. Only 2 of 16 possible values are assigned, and it has been more than a decade since the last one was assigned.

Choice of a number for assignment

- The two existing values 0x0 and 0xF have a Hamming distance of 4 between them. With what we now know about burst errors, these may not have been the smartest values to pick that have this property, but we are stuck with them with too many pre-existing implementations
- A consequence is that there are no remaining values with a Hamming distance of 4 or 3 from all existing codes.
- Several available codes with a Hamming distance of 2. If this is still an important criteria (along with thinking about burst errors), the next values to allocate would be 0101, 1010, 0011, 0110, 1100, followed by the codes with three zeros or three ones.
- If we apply this criteria, no reason not to assign OIF the value they have suggested (0x5=0101)

Other possible assignment choices

- If we decide that these are just codes (Hamming Distance and Burst Error tolerance are not an issue, many new PHYs using FEC), we could either:
  - Assign the remaining values in numerical order (IETF "FCFS"), in which case we would assign OIF the value 0x1
  - Assign new codes for new IEEE applications from the bottom up and for non-IEEE applications from the top down, in which case we would assign OIF the value 0xE (with 0x1 for the next IEEE-defined application)

What to call the ordered set

- OIF has suggested “FlexE ordered set”. The advantage is that it is simple, and clear what it is being used for
- If we are nervous that OIF might ask for a bunch of new codes for different applications (I don’t think this is likely), we could suggest something more generic like “OIF overhead ordered set” and leave it to OIF to figure out how to distinguish their uses of the ordered set if they have other needs in the future

Changes Needed to reflect in 802.3

Note that clause 119 refers to clause 82 for coding of ordered sets

Table 82–1—Control codes

Control character	Notation	XLGMII/ CGMII control code	40/100GBASE-R O code	40GBASE-R and 100GBASE-R control code
idle	/I/	0x07		0x00
LPI	/LI/	0x06		0x06
start	/S/	0xFB		Encoded by block type field
terminate	/T/	0xFD		Encoded by block type field
error	/E/	0xFE		0x1E
Sequence ordered set	/Q/	0x9C	0x0	Encoded by block type 0x4B plus O code, control codes are set to 0x00
Signal ordered set[a]	/Fsig/	0x5C	0xF	Encoded by block type 0x4B plus O code, control codes are set to 0x00
FlexE ordered set[b]	/Ffe/	NA	0x5	For 100GBASE-R and above only, Encoded by block type 0x4B plus O code, control codes are set as specified by OIF

[a]Reserved for INCITS T11 Fibre Channel use.

[b]Reserved for OIF Flex Ethernet use.

Last paragraph of 82.2.3.9

Sequence ordered sets may be deleted by the PCS to adapt between clock rates. Such deletion shall only occur when two consecutive sequence ordered sets have been received and only one of the two ordered sets may be deleted. Only idle control characters may be inserted for clock compensation. Signal ordered sets and FlexE ordered sets are not deleted for clock compensation.

Responsibility for Updates

- P802.3bs has scope to consider 400G aspects. Can submit a comment against D1.1.
- A maintenance request will be submitted concerning 100G aspects. The maintenance TF may assign this to P802.3bs (as it appears we may need to open clause 82 anyway), and this would also assure 400G and 100G are implemented in the same way.

THANKS!