Thoughts on 25G cable/host configurations.

Mike Dudek   QLogic

11/18/14

Presented to 25GE architecture ad hoc 11/19/14.

Introduction.

- This is a short presentation that explores the implications of having different host trace losses, Cable lengths and FEC options.

- It is a discussion document.

- Proposals have been made for 3 different FEC options (No FEC, Clause 74 FEC (Firecode-FEC), and Clause 91 (RS(528,514)-FEC)(just called Clause 91 in the rest of this presentation), 3 different host losses, and two different cable lengths (3m and 5m).

- The total number of combinations are very large, but obviously some combinations are more interesting than others.

- The 802.3bj system will be used as the baseline with 5m cable, 6.81dB host trace loss (which I will call mid loss) and Clause 91 FEC.

Reason's for different options.

- **Cable length**
  - 3m cable length can allow either more host losses or lower latency (less powerful/or no)FEC
- **Host losses.**
  - More loss enables longer trace lengths particularly for the Switch, where it enables more front panel ports to be driven from a single IC. Alternatively it enables the use of lower cost PCB materials for the same trace length.
  - Servers generally don't have long traces and therefore could have less loss which could be used to allow more switch loss or lower latency.
- **FEC**
  - Clause 74 and no FEC significantly reduce latency. They also require less silicon and are lower power, but these may not be significant factors. They also provide maximum commonality with 10G ports.

Possible interesting configurations.

Configuration	Switch loss	Cable (m)	Server /switch loss	FEC	Comments
1	mid	5	mid	Clause 91	Baseline 100GBASE-CR4
2	high	5	low	Clause 91	Higher loss switch still 5m cable to server
3	high	3	low	Clause 74	Lower latency still high loss switch to server
4	mid	3	low	None	Lowest latency switch to server
5	high	3	high	Clause 91	Higher loss switch to switch link
6	mid	3	mid	Clause 74	(added based on ad hoc discussion) lower latency

- Based on assumptions on next page all these are possible using 100GBASE-KR4 like ASICs

Assumptions

- Options are based on loss budgets. Mellitz_25GE_01_1114 shows that this is not the whole story particularly for reducing FEC performance and more work is needed to understand this.

- 3m cable has 7dB less loss than 5m cable. (meets compliance spec of 15.48dB)

- High loss host is 3dB more loss (9.81dB)

- Low loss host is 3dB less loss (3.81dB)

- Clause 74 FEC requires <7dB less loss than Clause 91 FEC (compare theoretical coding gain difference of only approx 3dB)

- No FEC requires <10dB less loss than Clause 91 FEC (compare theoretical coding gain difference of only approx 5dB, but Mellitz_25GE_01_0114 indicates 14.5dB less loss might be needed).

Comments on Optical compatibility.

- **Compatibility with optical ports is desirable, particularly for switches.**
  - **Increasing losses beyond what is desirable for Chip to Module is an issue, however Fibre Channel has shown that the use of Tx FIR for Chip to module and more powerful Rx in the host ASIC for module to chip (along with FEC) can significantly increase the loss. (FC has informative host loss of 12.8dB for 32GFC with similar specs to CAUI4 at the specification points and a 10e-6 BER target). 9.81dB host loss even with only Clause 74 FEC is probably OK.**
  - **Although Clause 91 FEC is used for 100GBASE-SR4 and for 32GFC, Clause 74 FEC could be adequate for 25GBASE-SR.**
    - **Link budget calculations show that optical components that pass 32GFC specifications with Clause 91 FEC are likely to operate over 100m OM4 at the somewhat lower 25GBASE-SR symbol rate with just Clause 74 FEC.**

System implications.

- Cable information (5m v 3m) can be provided to the host in the same manner that 25G cable v 10G cable is provided.

- Auto-negotiation can be used to determine which FEC is to be used. However knowledge of the loss of the other end would be needed, which can be provided by the host through auto-negotiation advertisement.

Specification implications.

- ## Cables
  - Can use existing 5m cable specification
  - Need to create a COM specification for a 3m cable that works in all these applications. (probably can determine the worst case situations and it may be possible to only need two test cases as for the existing 802.3bj 5m cable spec.). (add extra column to the COM table for a 3m cable).

- ## Hosts
  - Probably can just add extra columns to the specification table for high and low loss hosts.

Backup

Attempt to auto-negotiate without knowledge of other end loss.

- **High loss, mid loss or low loss host with 5m cable requests Clause 91 FEC only.**
  - No issues
- **High loss host with 3m cable requests Clause 91 or Clause 74 FEC**
- **Mid loss host with 3m cable requests Clause 91, Clause 74 or No FEC**
- **Low loss host with 3m cable requests Clause 91, Clause 74 or No FEC**
- **If FEC is chosen in order of No FEC, Clause 74, Clause 91 then without knowledge of the other end's loss there is an issue (2 mid loss hosts would negotiate to No FEC which doesn't work.)**
- **Problem is that a mid loss host with 3m cable wants to work with No FEC with a low loss host, but Clause 74 FEC with a mid loss host so needs to advertise both, so a mid loss host on the other end will choose no FEC.**