CABLE REACH OBJECTIVE CONSIDERATIONS

Adee Ran, Intel
David Chalupsky, Intel
Richard Mellitz, Intel

Goals of presentation

- Consider implications of having only one cable-oriented Physical layer specification (or PHY, or port type), with either 5 meter or 3 meter reach
- Consider two reach objectives, two port types and two channel specs – can everything work together?
- Focus on PMD and cost implications
  - Ignore latency for now
- Recommend objectives

Implications of having only “5 meters” objective

- Will likely lead to specifying a PMD with electrical specs at the MDI similar to those of Clause 92
- 5 m reach will enable aggregation of many servers into large switches spanning several racks, saving OPEX (and possibly CAPEX too)
- Mandatory support 5 meters essentially means PMD design challenges similar to those of 100GBASE-CR4
  - Same host loss budget requires low-loss material and/or short traces (possibly, separated PMD)
- For servers/switches that need only 3 meters:
  - No CAPEX saving opportunity
  - Insignificant OPEX saving, if at all (RX power saving?)

Implications of having only “3 meters” objective

- Will likely lead to specifying a PMD with relaxed electrical specs at the MDI, compared to clause 92
- Relaxed specifications likely mean cost saving compared to 100GBASE-CR4
  - … 100GBASE-CR4 ports meet stricter specs, so can still be compliant (e.g. re-use PMD with break-out cables)
- Having only 3 m reach will prevent aggregation of many servers into large inter-rack switches.

Can we have the cake and eat it?

- It is desirable to…
  - Enable use cases that need 5 meter reach (e.g. switches spanning several racks)
  - Enable cost saving in servers (and switches) that can do with 3 meter reach
  - Enable choice, but maintain compatibility
- Suggested solution:
  - Minimum PMD specifications for mandatory support of 3 meters.
  - Optional specifications for “extended reach”.
- The following slides describe one way to do it. Other ways may also be possible.

Two PHY specifications

- Mandatory minimum spec (3 meter support):
  - Relaxed TX specifications at MDI (e.g. linear fit pulse peak)
  - Relaxed RX specifications at MDI (e.g. lower loss test channel)
- Optional “extended reach” spec:
  - TX and RX specifications at MDI consistent with clause 92
  - May have a separate port type, e.g. “25GBASE-CRX” vs. “25GBASE-CR”
- Both specs **technically feasible** with same silicon, and possibly different trace length or PCB material
- Two “minimum spec” devices support up to 3 meter cable
- Two “extended reach” devices support up to 5 meter cable
  - May allow lower-latency mode with shorter cable
- The two device types **can interoperate** over 3 meter cables (possibly even support a 4 meter cable).

Two cable assembly specifications

- Cable for “up to 5 meters”
  - TP0-TP1 and TP4-TP5 paths as in clause 92
  - Enables connecting two “extended reach” devices
- Cable for “up to 3 meters”
  - TP0-TP1 and TP4-TP5 paths longer than clause 92, consistent with minimum PHY specification at MDI
  - Loss/noise budget will align with “minimum spec” devices
  - Compatible with “extended reach” devices too
    - “Extended reach” devices may be able to use reduced latency solutions

Use cases and cost saving

- Large switch spanning several racks:
  - Can be designed to meet “extended reach” specifications on all ports
    - If we had only 5 meter specs: the exact same requirements would apply ➔ suggested solution has no implications
  - But may also mix and match ports
    - Extended reach for center ports or breakout of 100G ports; minimum spec for end-of-panel ports
    - ➔ Cost saving opportunity
- ToR switch:
  - If we had only 5 meter specs: would have to meet lower loss budget, likely higher cost
  - Assuming it needs only 3 meters, can be designed to meet only minimum specs
  - ➔ Cost saving opportunity
- Server/NIC:
  - Two separate types are possible – minimum spec (lower cost) and extended reach (higher cost)
  - If we had only 5 meter specs: all servers/NICs would need to be equivalent to extended reach (likely higher cost)
  - ➔ Cost saving opportunity
- Cables:
  - The same cable lengths will be used, regardless of what PHY/cable specs we specify
  - ➔ having two PHY specs has no unexpected implications

CSD responses

- Broad marker potential
  - The two main use cases discussed in study group are addressed
- Compatibility
  - All PHYs will interoperate over supported channels
  - 100GBASE-CR4 port breakout is possible
- Distinct identity
  - Two separate solutions for two channel specifications
- Technical feasibility
  - Based on 100GBASE-CR4
  - More restrictive specs for cables are met by cable models contributed for 802.3bj
- Economic feasibility
  - Comparable to 100GBASE-CR4
  - Enables cost reduction in most ports

Recommendation

- Adopt two objectives:
  - “Define a single lane 25 Gb/s PHY for operation over links consistent with copper twin axial cables, with lengths up to at least 3m”
  - “Define a single lane 25 Gb/s PHY for operation over links consistent with copper twin axial cables, with lengths up to at least 5m”
- Exact specifications, port types, identification etc. can be dealt with in the task force.