Thoughts on the Beyond 400 GbE Study Group

Mark Nowell – Cisco

Dave Ofelt - Juniper

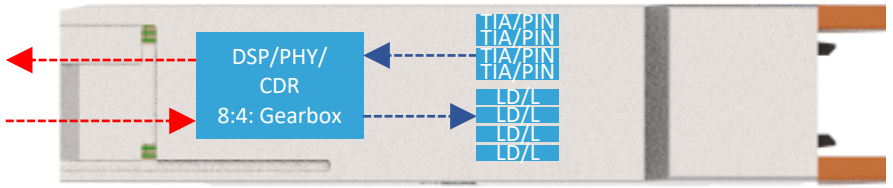
Topics:

- Looking back before looking forward
  - Changing use-cases – port flexibility
  - Lessons learned from past projects
- What does the market need?
- What should Study Group focus on?

How does the ability to flexibly configure interfaces affect things in IEEE?

- IEEE 802.3 specifies PHYs with the goal of being interoperable
- Flexible interface or port configuration is an implementation use-case that has become a more popular deployment option driven by:
  - Ethernet blocks in ASICs are typically capable of supporting multiple interfaces
  - Pluggable modules *can* be designed to support multiple independent PMDs

Example MAC/PCS block for 8 SERDES @ 50Gb/s
Many configuration variants



e.g. QSFP-DD Module capable of
supporting both 400GBASE-DR4 and Quad 100GBASE-DR

Port flexibility considerations

- While implementations may support multiple PHYs, from an IEEE 802.3 perspective only unique PHY specs are written – this flexibility doesn't appear in the specifications
- Where does Port Flexibility matter?
  - For defining market uses and needs. i.e., setting and justifying Study Group objectives
  - Minor consideration in a PHY spec. e.g., MDI spec (MPO vs LC) where knowledge of potential implementations is relevant

- More recently, IEEE 802.3 projects have been open to defining multiple PHYs specs at different MAC speeds
  - Supporting flexible use-cases properly means this should be a strong consideration for Study Group. Especially if a new "lane" rate is being defined.

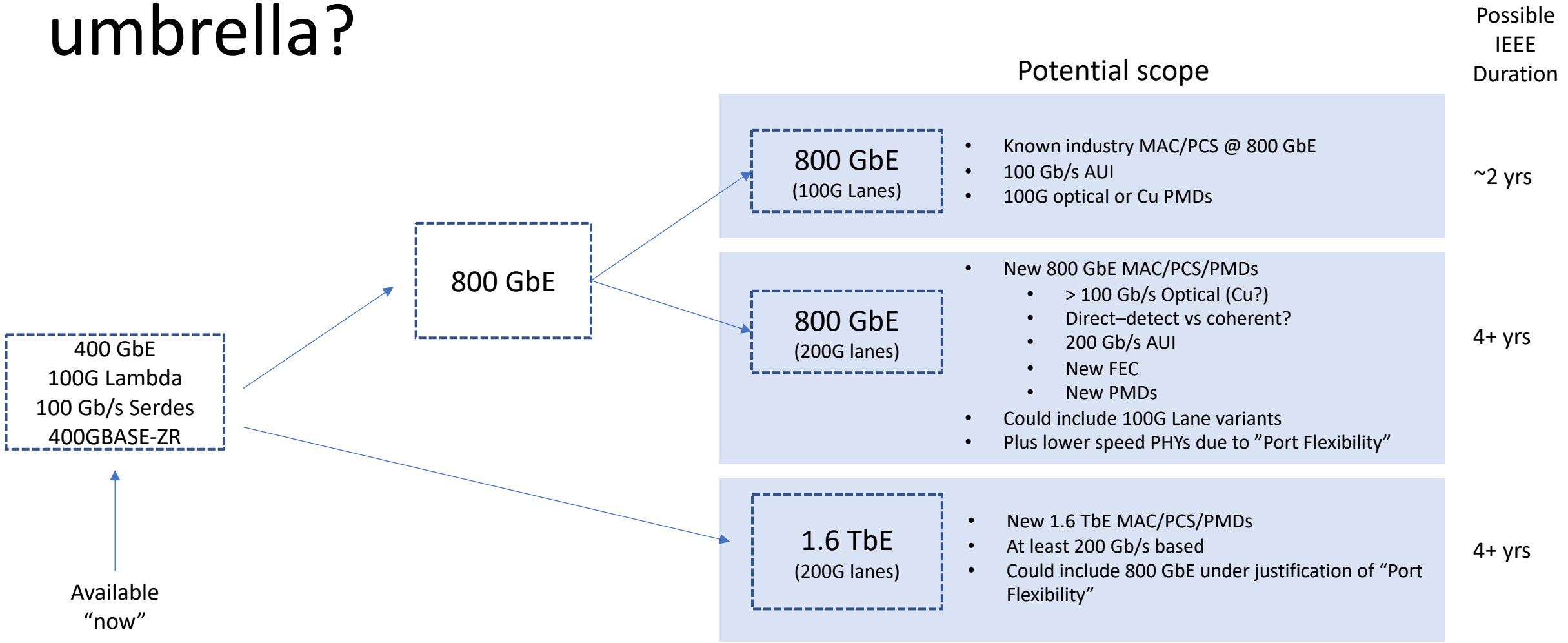
Lessons learned

- If port flexibility considerations were considered in past projects we might have saved a lot of effort
  - e.g. 100G optical lanes were initially defined across two projects
    - 802.3bs – 400GBASE-DR4
    - 802.3cd – 100GBASE-DR
  - Recommendation: where it makes sense to specify multiple PHYs that fall out of a flexible port use-case, include them in a common 802.3 project
- New lane rates are hard to define
  - Evolving industry learning and experience changed spec and test methodology for each new project based on the “same” technology
    - 802.3bs – 400GBASE-DR4
    - 802.3cd – 100GBASE-DR
    - 802.3cu – 400GBASE-FR4, 400GBASE-LR4-6, 100GBASE-FR1, 100GBASE-LR1
- However, the greater the consistency with past specs, the more the leverage
  - IEEE 802.3 now has good experience with writing PAM4 electrical and optical specs.

Lessons learned (2)

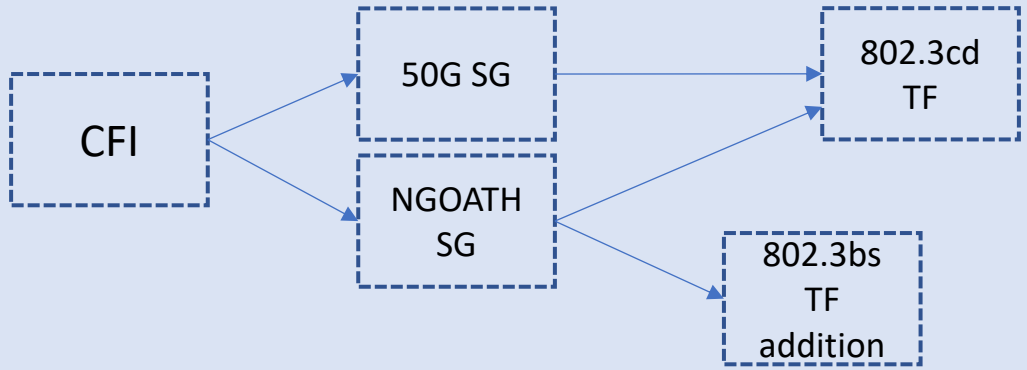
- Webscale deployments are driving the Ethernet switch and interconnect industry to develop solutions to allow them to continue to scale their networks.
  - Unfortunately, not a lot of consistency on how they build their networks
    - Radix, Fabric speed, Port speed, over-subscriptions, interconnect infrastructure
  - Therefore, focus on key building blocks is important, with knowledge that the range of implementations and usage may be broad
- For Webscale deployments, Multi-rate requirements needs to be considered (a.k.a. signaling backwards compatibility – not module backwards compatibility)
  - e.g. A host and module that supports 400GBASE-FR4, that can also be able to support 200GBASE-FR4, or even 100G-CDWM4

What could fall under a “Beyond 400 GbE” umbrella?

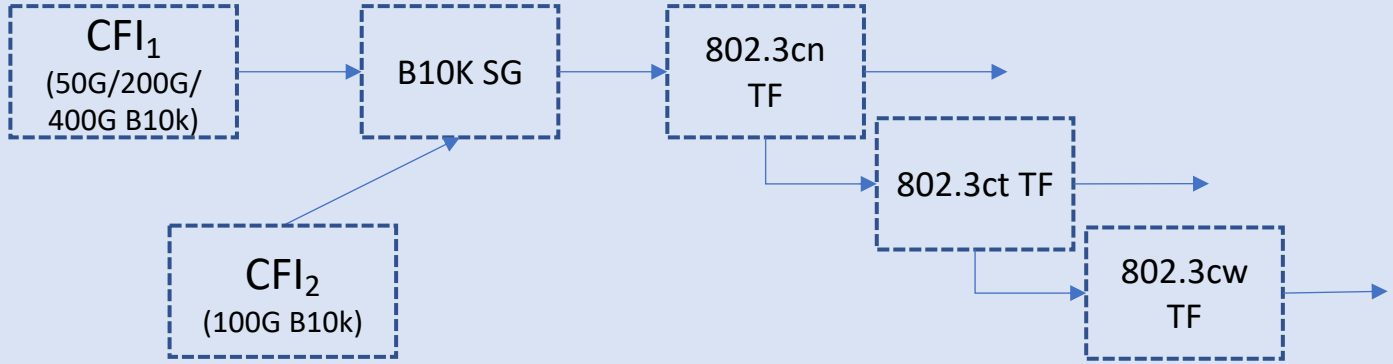


Any combination of these could fall under the scope of a “Beyond 400 GbE” study group

Past IEEE 802.3 project evolutions

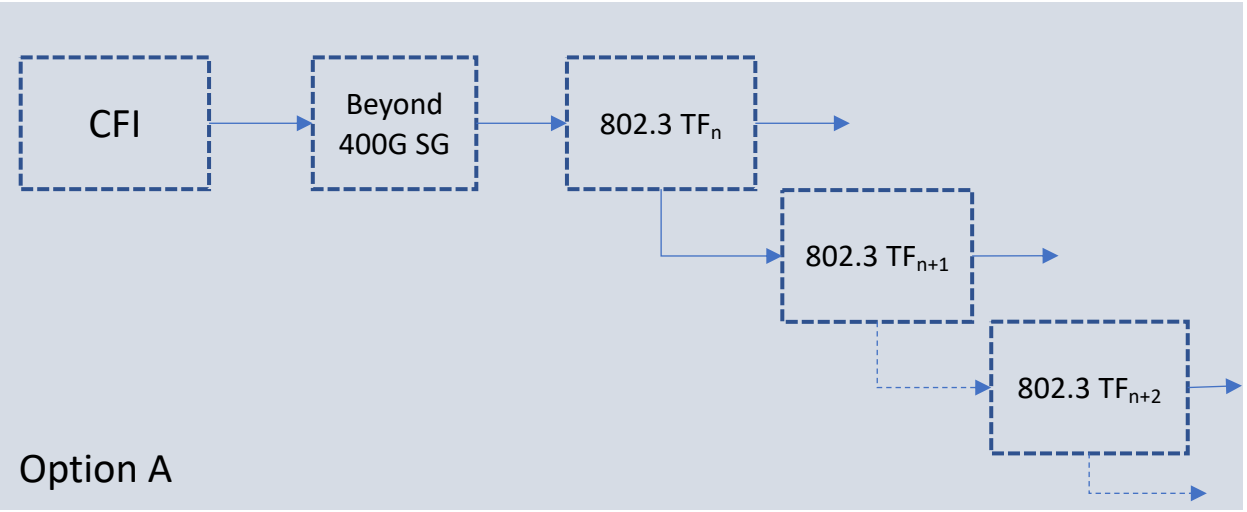


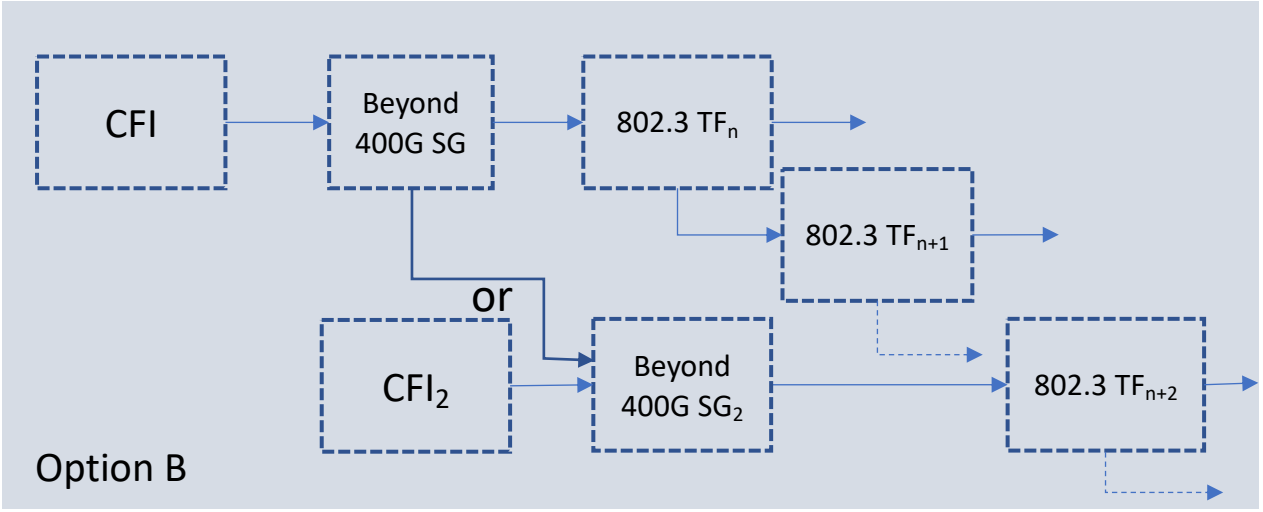
- CFI on 50 Gb/s, 100 Gb/s and 200 Gb/s Ethernet technologies
- 2 SGs originally since concern about different schedules
- Resulted in putting work into 2 TFs based on schedule considerations but NOT aligned with original SG assumptions
  - 802.3bs PAR/Scope modified to incorporate 200G SMF
  - 802.3cd formed with everything else



- 2nd fast-follower related CFI.
- Chartering motion merged into common SG
- Objectives agreed but concerns raised about schedule harmony
  - New TF's spun out of original to allows subsets of specs to complete while further work of more challenging objectives
- One(+) CFI → 3 TF's

Possible “Beyond 400G” project evolutions





Key conclusion: The group has lots of options and possibly shouldn’t worry about it at this phase.

Primary near-term focus should be establishing clear objectives based on our knowledge of Technical and Economic Feasibility as well as known market demand/drivers and potential

This won’t be the only bus leaving the station

Thoughts on implementation

- The breadth of implementations is broadening
  - Modular chassis, Fixed,
  - Pluggable modules, Onboard optics, Co-packaged optics

- Need to be aware of, and can take advantage of, the new transitions

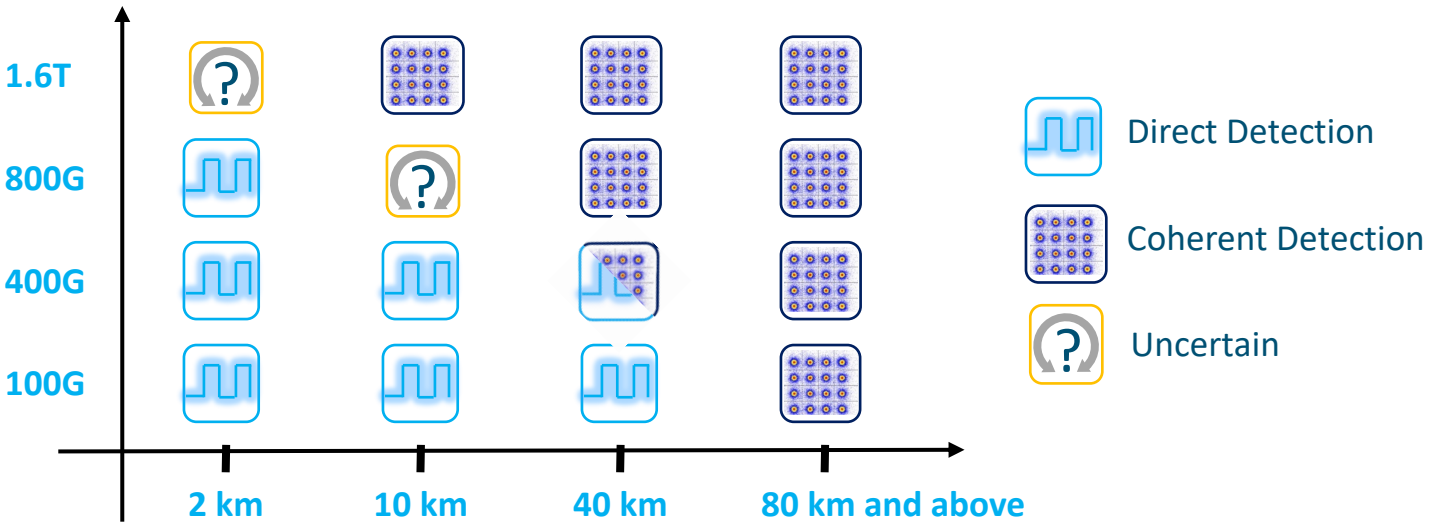
- Both risk and opportunity to Broad Market Potential may exist if we shift from current implementation or technology approaches.

Thoughts on 200G lanes

- Significant Study Group time will be needed to determine feasibility of component performance, module design and system design considerations as well as PCS/FEC analysis just to select objectives and answer CSD
- Potentially some hard decisions to make:
  - Viability of passive copper cables
  - Viability of current breadth of component design options
  - Consistency of coding between electrical and optical

- We have a stronger starting point on specification and testing methodology than before.
- Multi-rate compatibility needs to be a consideration as this has become a key deployment usage
- Port flexibility deployment considerations should mean inclusion of PHYs for 200 GbE within the “Beyond 400G SG” objectives (as long as there is a market justification).

Reach



800 GbE:

- Building blocks are “known”
- What will be used for reaches is “likely known”

1.6 TbE

- Building blocks are “sort of known”
- What will be used for reaches is TBD

- As speeds increase, reach becomes a challenge
- Mitigations:
  - Change Media: DAC→ Active Cu → MMF → SMF
  - Change Media: Duplex → Parallel
  - Change technology: Direct Detect → Coherent Detect

Options (opinions) for priority of objectives

800 GbE
(200G lanes)

- New 800 GbE MAC/PCS/PMDs
  - > 100 Gb/s Optical (Cu?)
  - Direct–detect vs coherent?
  - 200 Gb/s AUI
  - New FEC
  - New PMDs
- Plus lower speed PHYs due to ”Port Flexibility”

- Area of highest interest from a building block perspective.
- 200G per lane technologies are key for an optimized 800 GbE and necessary for a 1.6 TbE solution

800 GbE
(100G Lanes)

- Known industry MAC/PCS @ 800 GbE
- Known 100 Gb/s AUI
- Mostly known 100G optical or Cu PMDs

- Industry not gated by lack of IEEE specifications here.
- Likely won’t delay anything if there is interest in adopting as objectives

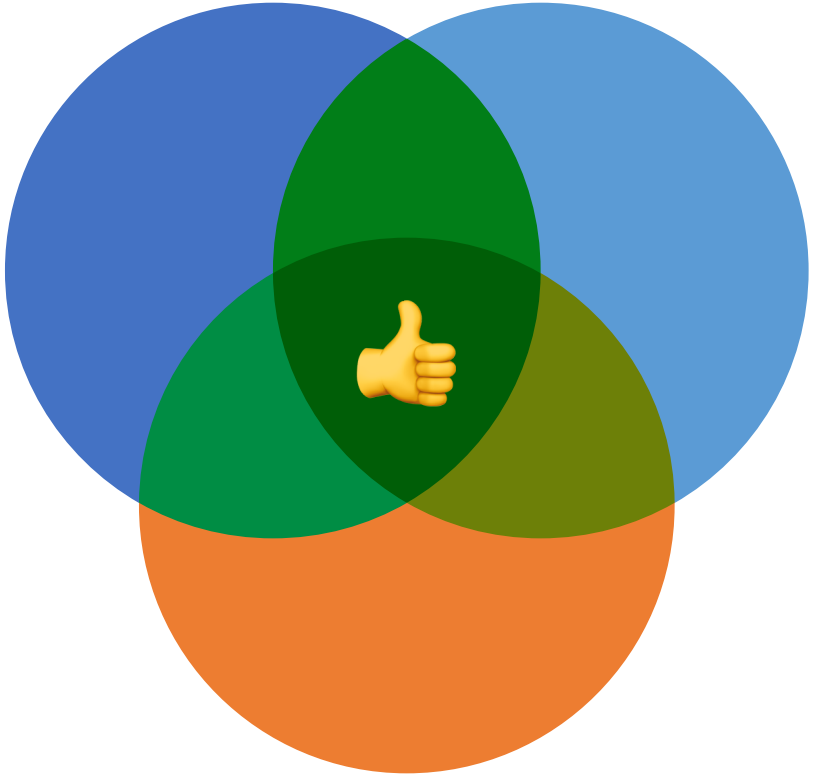
1.6 TbE
(200G lanes)

- New MAC/PCS/PMDs
- At least 200 Gb/s based
- Could include 800 GbE under justification of port flexibility

- Significant leverage on any 200G/lane definitions
- A lot of uncertainty on network drivers and technology outcomes to start defining specific 1.6 TbE PHYs at this time.
  - Eg coherent vs direct detect – reach PMDs
- Timing is important - do not want to work to specify and build solutions too soon for market adoption.

Importance of the CSDs

Economic Feasibility



Technical Feasibility

Broad Market Potential

Takeaways

- The crystal ball for 1.6 TbE is very murky…but clearer for 800 GbE
- The building blocks needed for 800 GbE will likely be useful when we need to figure out 1.6 TbE
  - But will also accelerate optimization of lower speeds
- Lack of 1.6 TbE doesn’t preclude 1.6T pluggable modules (e.g. 2x 800 GbE)
- Focus on objectives around 800 GbE and understanding feasibility and market
- Time will help improve the clarity.