

400G SR8 for Data Center Interconnect

January 20, 2018

Zuowei Shen

Google

Market Demand for 400G SR8 Optics

Cloud and AI applications calls for higher BW network.

- **Low cost:** SR8 is the cost optimized 400G end to end solution for greenfield/ brownfield application.
- **Flexibility:** Comparing with SWDM, SR8 Offers flexibility of fiber shuffling with 50G/100G/200G configurations. It also supports fanout at different I/O speeds for various applications: server, storage, GPU, TPU and flash.
- **High density:** 400G SR8 OSFP/QSFP DD transceiver can be used as 2x200G SR4, 4x100G SR2.
- **Backward compatibility:** Interoperable with 100G-SR4, 40G-SR4 with same 850nm wavelength.
- **Reach:** Same reach as 100G SR4: 70m OM3, 100m OM4.

Pros of 400G SR8 Transceiver

- Native serdes 50G PAM4, no gearbox required.
- Easy to deploy in both greenfield and brownfield with x16 ribbon fiber+ MPO16 connector.
- OSFP supports 32 per 1U and 15W thermal envelope
- Mature assembly process, very similar as 100G-SR4.
- Optical components are available NOW - faster time-to-market.
- For short reach (<50m), SWDM (SR4.2) fiber saving is less than module cost adder with customized VCSEL/optics.

Technical Feasibility of 400G SR8

- 100m OM3 reach with KP4 FEC ($2e-4$) has been demonstrated with PAM4 VCSEL and 25G NRZ VCSEL.
- VCSELs are available. More PAM4 VCSEL are being developed.
- Linear TIAs and PDs are available.
- VCSEL driver or CDR direct drive have demonstrated performance
- Stronger EQ helps bring down BER floor
 - Tx EQ: compensate VCSEL nonlinearity
 - Rx EQ: recover stressed eye over fiber.
- Use MPO16(1x16) APC to reduce MPI penalty. It is proven that lower link BER can be achieved with APC.
- Suggest to add optical return loss spec to mitigate MPI penalty

SR8 V.S. SR4.2

	SR8	SR4.2
VCSEL/Wavelength	VCSELs are available	Wavelength to be defined, Need new VCSEL development. Limited supply.
Assembly complexity	1x8 lens array, similar to 100G-SR4	More complex with mux and demux
Fiber ISI penalty over 100m OM3	OM3 has better modal BW 2000MHz · km at 850nm.	OM3 has smaller modal BW at longer wavelength. 1250MHz · km at 940nm Shorter reach or need new fiber for 100m reach
Link budget	Only fiber+connector loss	Link budget is challenging with 4.7dB penalty from PAM4. Mux/Demux introduces extra loss.
Optical connector	MPO16_APC Verified performance improvement with APC	MPO8 (APC ?) Need APC to reduce MPI penalty.

When?

- 50G PAM4 I/Os are ready. 400G MM optics needs to catch up.
- Flexibility, availability, fast to market and short lead time is key to enable 400G SR8 optics in data center/cloud. SR8 can be deployed in volume as early as 2018.
- Server BW increases as PCIe gen4 arrives. Cost competitive optical solution is a good fit for S1 to server connection >5m at 50G PAM4.

Summary

- 400GBASE-SR8 has distinct identity from SR16. It operates over a lower cost and more user-friendly MMF cable. It can breakout to support 50 G/100G/200G links.
- The previous slides have demonstrated technical feasibility, economic feasibility, and broad market potential for 400GBASE-SR8
- Recommend the Study Group to adopt an objective “Provide a physical layer specification which supports 400Gb/s operation over 8 pairs of MMF with channel lengths up to at least 100 m.”