Striving for Energy Efficient Networks in the Cloud

Brad Booth
Principal Network Architect
May 10, 2018
Microsoft’s Global Datacenter Network

DCs and network sites not exhaustive
Regional Architecture

- Distributed data center model
- Massively parallel and highly resilient
- Latency SLAs constrain maximum fiber distances
DC Ecosystem

1. Intra-DC
   - ToR-Tier1: ≤20m
   - Tier1-Tier2: ≤600 m

2. Inter-DC
   - DCI
   - Tier2-RNG ≤100km
   - Relatively fiber-rich

3. Inter-Region
   - Long-haul / subsea
   - ≥ 100 km
   - Fiber-poor
DC Ecosystem

1. Intra-DC
   - ToR-Tier1: ≤20m
   - Tier1-Tier2: ≤600 m

2. Inter-DC
   - DCI
   - Tier2-RNG ≤100km
   - relatively fiber-rich

3. Inter-Region
   - long-haul / subsea
   - ≥ 100 km
   - fiber-poor
Yesterday

- Sources: dedicated coherent transponders
- Line system:
  - Proprietary closed system compatible with sources
  - Proprietary NMS for control/monitoring
- Expensive and power-hungry!
  - 50-60W for coherent DSP
Inter-DC

Today

- Sources: 2-carrier 100G PAM4 (i.e. super-channel)
- Line system
  - Auto gain config and chromatic dispersion comp
  - Massive power reduction

Today

- to/from far-end RNG

Inter-DC

- tier 1
- tier 2

optical line system

≤100km fiber
Inter-DC

**Tomorrow (2020+)**
- Sources: single-carrier 400G 16QAM (OIF 400ZR)
- Line system: Same but simplified specifications due to coherent sources
- Power per bit continues to decrease
  - 4.5W @ 100G → 15W @ 400G

---

[Diagram of inter-data-center connections with tier 1 and tier 2, and labels for optical line system and ≤100km fiber].

[to/from far-end RNG]
Inside DC: Technology Barriers

400G Generation
$ Pluggable Optics ~2x > $ Switch Port (TCO)

ToR to Server: Copper DAC struggles
$ Optics ~10x > $DAC (Today)

Need cost effective replacement for DAC or new row architecture

Lane speed stalls
Coherent inside DC?

Switch Lane Speeds
- 25G Lanes: 2018
- 50G Lanes: 2020
- 100G Lanes: 2023
- 2025
Inside DC: Integration

Switch Lane Speeds:
- 50G Lanes (2020)
- 100G Lanes (2023)

Optics in Package
Optics on Die

On-board Optics
Summary

• Network power requirement is small compared to servers
• Continuous focus on energy efficiency per bit
• Microsoft’s Inter-DC approach offered massive power efficiency
• Integration is viewed as the North Star to reduce power inside the datacenter