Inter Data Center Networking: Is the time right for optics to move beyond point-to-point connectivity?

Benn Thomsen
Microsoft Research
From Mega to Distributed “Regional” DC’s

- Site Availability
- Power, cooling, water
- Resiliency: Provides geographically separated failure domains

Distributed data centers across an 70km ($\leq 2.0\text{ms}$ RTT) region to keep pace with cloud growth
Can we improve performance

Packet Switched Hub & spoke
- 2*B*n ports
- simple design
but, latency inflation

Full Mesh Optimal for latency but,
- B * n² ports
- port exhaustion for small DCs
- hard to scale

(B ports per DC, n number of DCs)
A perfect region

- Intuitive design
- Ports per DC independent of the number of DCs
- Easy scaling
- Latency optimal
- No hub
- Exploit fiber availability
Slow Optical Switching to Provide Partial Mesh on Demand

- Number of ports per DC as in hub & spoke model (i.e. B ports)
- Emulate a distributed hub through dynamic paths for DC-DC traffic
- Benefits
  - Improves latency over hub and spoke model. Does it matter??
  - Distributed network eases site selection
  - Graceful bandwidth scaling. OXC independent of line-rate
Network Reconfiguration time (secs-hours)

Colored optics

Regional network (80 km links)

Gray optics

Mega data center

Regional architecture
Physical layer challenges

Control Plane
• Distributed control plane / Synchronisation
• Network management changes: Hybrid Packet/Circuit network
• Predictive dynamic traffic engineering to set DC to DC bandwidth

Physical layer
• Line system stability?
  • Fully load line (32-40 Wavelengths) At all times
  • 1-16 optically amplified spans (In distributed scenario)
• OXC requirements (<100ms, Wavelength granularity switching)
  • 80 rings 40 wavelengths
  • Ideally: 80x80 fibre ports, 40λ waveplane switch
  • Grouped Fibre: 16x (6x6 fibre ports, 40λ waveplane switches)
• Fast locking (<10ms) colorless coherent transceivers