Enabling Software Defined Interconnection Layer for Application-driven Intra-Data Center Networks

Shan Zhong
Director of Technology and Product Development
CoAdna Photonics
OFC Market Watch, 2015
Agenda

- New Traffic Trend in Data Center
- Adoption of Optics and WDM for Data Center
- WDM & WSS to Enable Programmable Interconnection Layer for Large Scale DC
- Applications
- Component Readiness
- Summary
New Traffic Trend in Data Center

- Growth in server virtualization
- Exponential growth in mobile data
- Escalating growth in data traffic per server (cloud traffic growth at 32% CAGR)
- Massive East-West traffic capacity needed (75-80% of DC traffic)
- Need to support highly persistent high-capacity data flows for VM migration, data replications, and moving large data sets in addition to regular traffic
- The elephant-mouse flow challenge

*Cisco Global Cloud Index: Forecast and Methodology, 2013–2018*
Adoption Optics in Data Center along with Evolving Architecture

- Evolving architecture
  - Spine/leaf vs. 3 tier model: non-blocking links for optimal traffic & flow completion times

- Adoption of optics
  - High speed transceivers for high-bandwidth interconnection and front-panel density
  - Optical circuit switch (OCS) to provide low latency path

- Adoption of WDM
  - Efficient fiber bandwidth utilization, enabling simplified cabling system due to its flexibility
  - Mature wavelength switching technology to provide switchable low latency light-path
Adoption of WDM in Data Center

Early adoption of WDM to increase the bandwidth of fiber link. i.e. OpenOptics MSA (www.openopticsmsa.org)

Fixed OADM or NxN AWG to form full-mesh connection among all network nodes.

Wavelength switching and TDM scheme to provide bandwidth reconfigurable connection

To increase bi-section bandwidth:
1) Increase data rate per lambda, i.e. ALU Slot-Switching @100G per channel
2) Run parallel system
3) Do distributed wavelength cross-connect, i.e. CoAdna OvS

Reconfigurable WDM era

UCSD MORDIA proposal

CoAdna OvS
Implementation of Distributed Dynamic Wavelength Cross-Connect

CoAdna OvS (optical virtual switching) Platform

- Similar to degree-N ROADM system, it’s a broadcast-select style communication paradigm that allows any port to communicate with any other ports arbitrarily.
- Supporting *cast (unicast, multicast, all-cast etc.) communications optically
- Enabled by OpenFlow other SDN protocol

Tremendous bi-section BW:
16-node system: 32(lambda)x16(fiber)x10GbE ~5Tb/s

OvS to Inter-Connect Core-Switches or Spine Switches

- **OvS**: Distributed dynamic wavelength cross-connect
- **Benefit:**
  - Reconfigurable bandwidth among all switches
  - Distributed network core, a more resilient system
  - Expandable, pay-as-you-grow
2-D Implementation to Enable Flattened Butterfly Architecture for ToR to ToR Connection

- Enabling large scale (1000+ racks) inter-connections with minimized hopping
- Management & Control Plane monitor traffic matrix and move flows between fabric based on application & traffic pattern demand
WDM Brings Simplicity to Data Center

- Flattened network w/ fully programmable interconnection layer for application driven networking
- Dramatically reduced cable count
- Converged data/control/management networks

Smart Photonics for Optical Networks
Component Readiness

- Low cost WDM transceiver readiness
  - New DWDM laser chip designed for wireless front-haul
  - Silicon Photonics for integrated WDM MSA
  - New MSA: i.e. OpenOptics MSA

- WSS readiness
  - Significant cost reeducation due to performance/reliability relaxation

- Others
  - EDFA: ~10% cost down consistently over last 10 years and continue
  - MUX/DEMUX:
    - AWG Mux/DeMux already dirt cheap
  - Ribbon cable connector: 12/24/32 fiber cable and MPO connector are available now
Optical transceivers have been used for high bandwidth links inside data center where optical circuit switches show the capability to provide low-latency path for long-lasting elephant traffic among the large percentage east-west traffic.

WDM enables simple full-mesh connections while reconfigurable WDM brings data center networking to next level.

CoAdna’s OvS platform is the pioneer to implement such networking concept by providing fully programmable interconnections among spine-switches as well as ToRs.
Thank you!
Cost and Cable Comparison under Full Connectivity

**Total CAPEX Comparison**
- Fat Tree
- Our Solution

**Cabling Complexity Comparison**
- Fat Tree
- Our Solution

*Assuming ribbon cables used for all cases (QSFP or Instant Connect)*
Layered Structure of Future ToR Switch and the 2D Interconnection