OPTICAL INTERCONNECTS IN DATA CENTER

Tanjila Ahmed

Challenges for Today's Data Centers

Challenges to be Addressed :

- Scalability
- Low latency
- Energy Efficiency
- Lower Cost

Challenges for Today's Data Center

- Cloud computing , social networking , video streaming
- Global Cloud Index from CISCO (Oct 2015) [1]

  In 2012 data center traffic 2.6 ZB per year

  In 2014 data center traffic 3.4 ZB per year

  In 2019 data center traffic 10.4 ZB per year
- Switching is still done in electronic domain : high power consumption and limited capacity upgrade
- Energy consumption : servers, network equipment, cooling and lighting
- Higher cost for capacity and energy consumption

Network Architecture of a Data Center

- Tree Topology [2] :

  Core : Layer 3 routers
  Aggregate : Layer 2/3 switches
  Edge/Access : Layer 2 switches

- Drawbacks :

1. All traffics are supposed to be North South traffic
2. Scaling : scale-up model

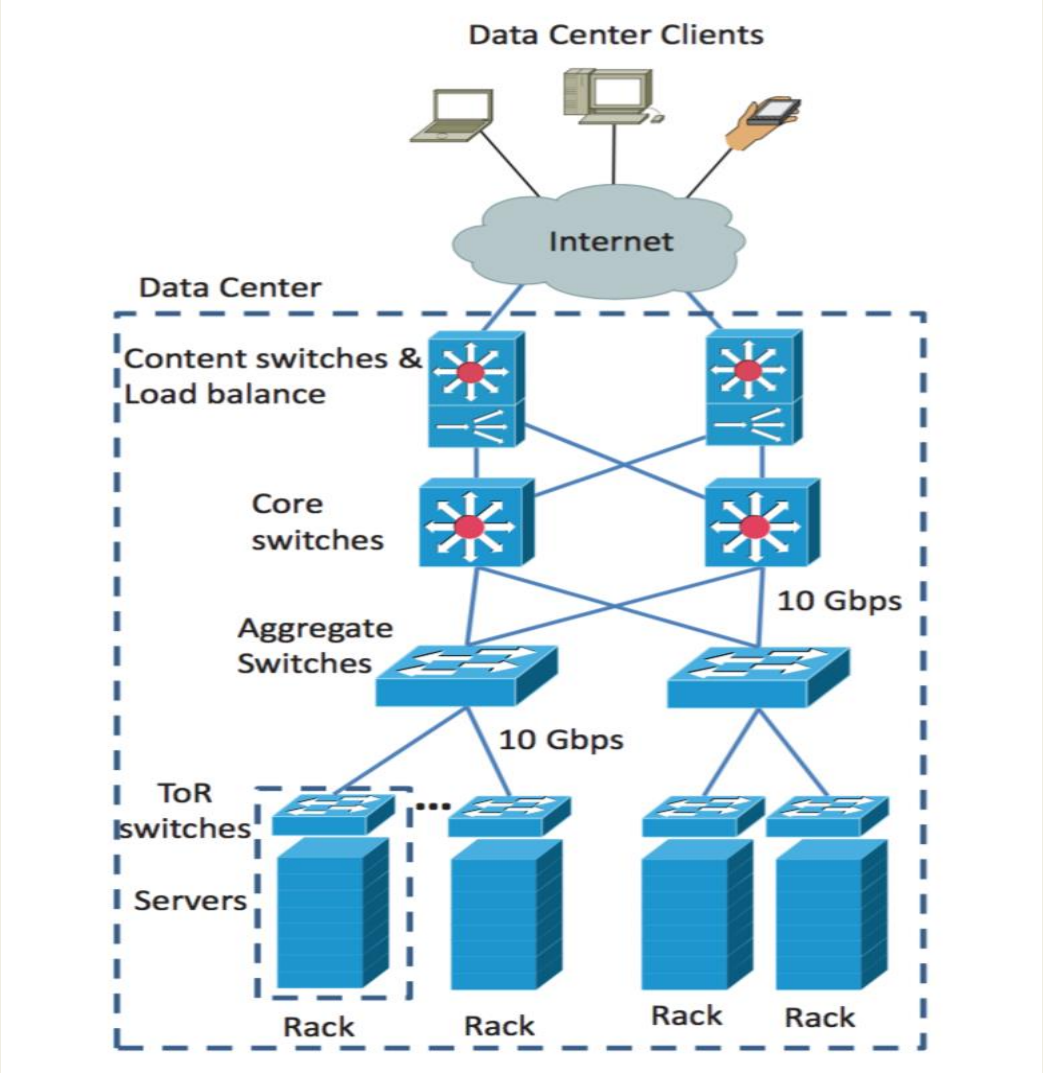


Image courtesy [2]

Network Architecture of a Data Center

- **Scale up model**[3] :

1. Fully meshed network
2. Connect every server with each other
3. Full bisectional bandwidth
4. Oversubscription required in every layer
5. Unable to scale beyond certain limit

because of the core switch limitation

6. Expensive equipment

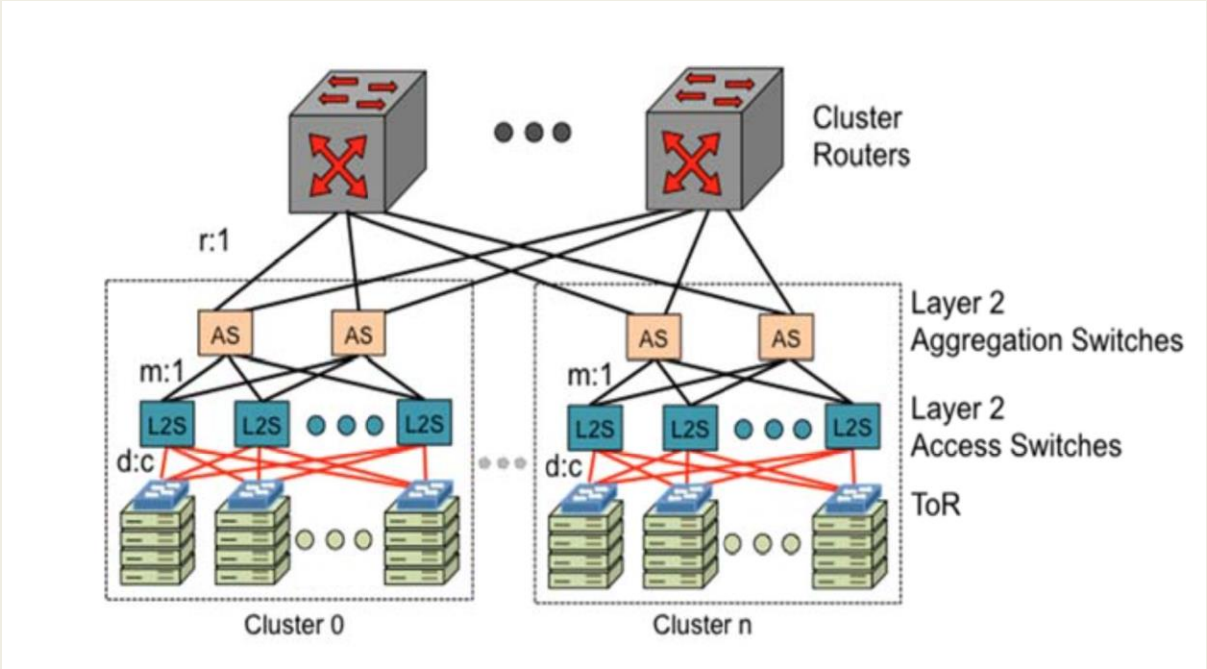


Image courtesy : [3]

Network Architecture of a Data Center

- Scale-out model

1. Small pods composed of identical switch
2. Extensive path diversity
3. Full bisectional bandwidth
4. Modular fashion scaling
5. Cost increases by constant cost/port
6. Supports east-west traffic

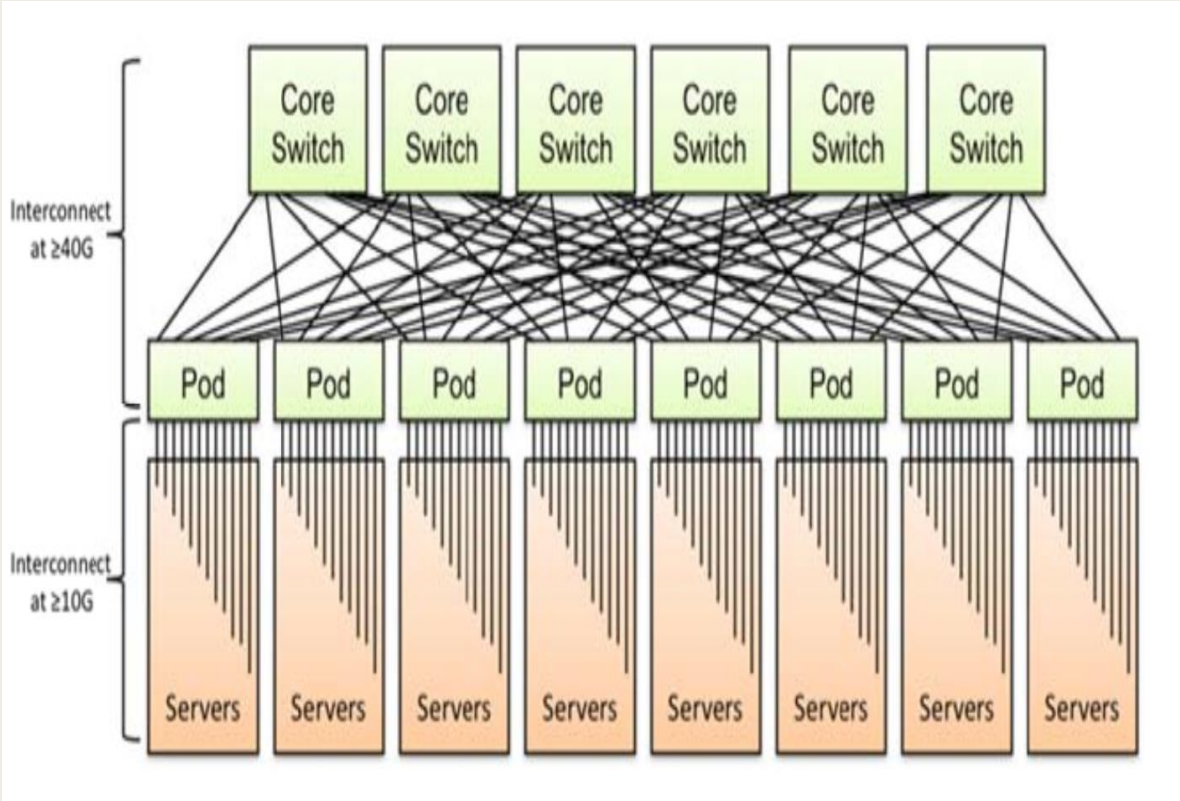


Image courtesy : [3]

Scale-up and Scale-out Approach

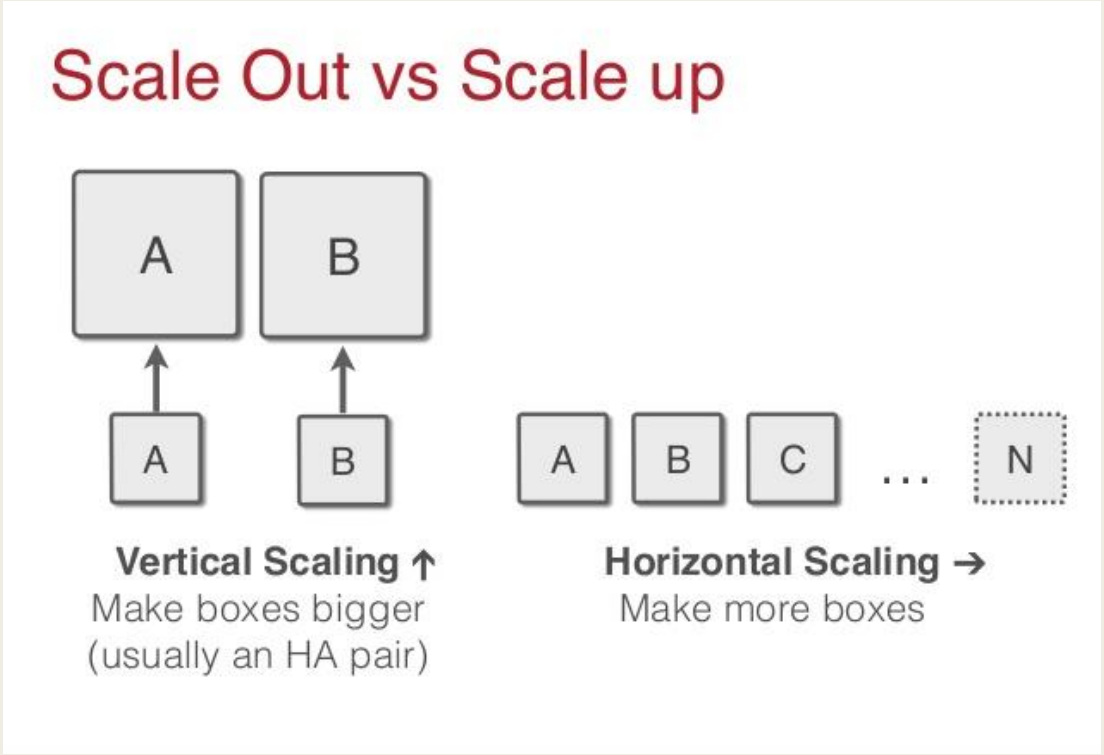


Image courtesy : [4]

Folded Clos Multi Rooted Tree Topology

- Advantages :

1. Does not require very powerful switches
2. All of the switches and servers are 10 /40 Gb/s
3. Multi pathing ensures the bandwidth utilization
4. Non blocking switching

- Disadvantages

1. 64 K servers will require 196,608 links
2. Links are expensive to manage
3. Copper cables cannot be used for its copper skin effect
4. Faster the data rate shorter the data link has to be
5. Optical fibers are used instead of copper cables
6. High number of transceivers are required

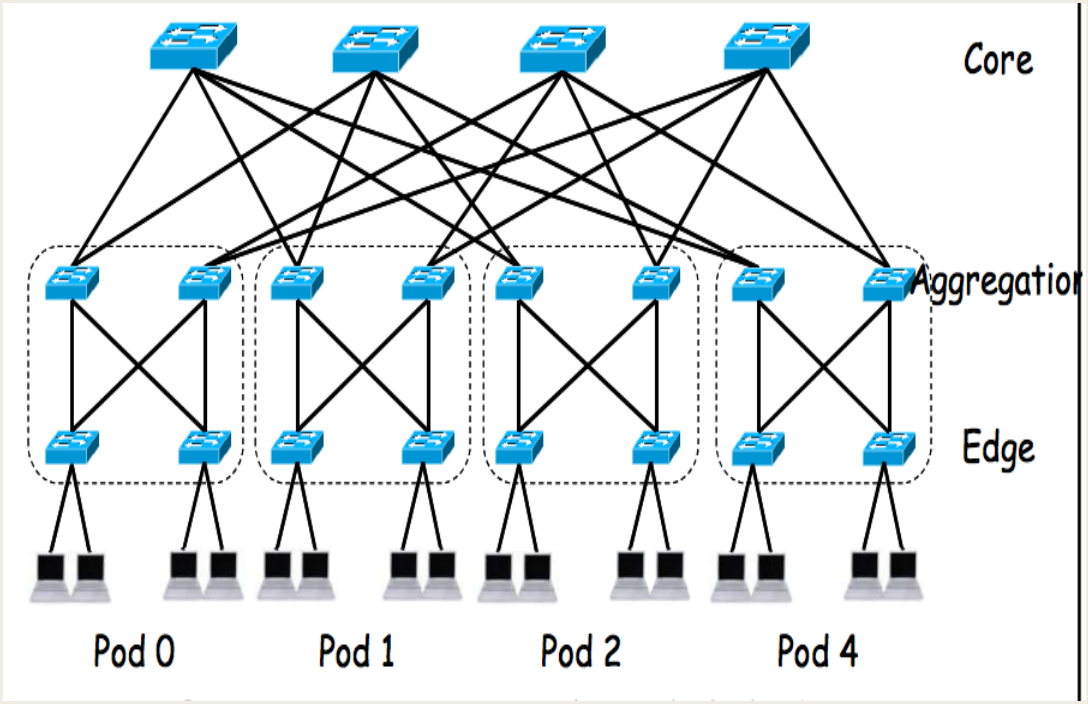


Image courtesy : [5]

Optical Switches [6]

1. **Hybrid electronic/ optical switching**

- Packet and circuit switching
- Electronic switch gives fine granularity
- Optical circuit switching (OCS) gives high capacity
- Upgrade limitation in electronic part

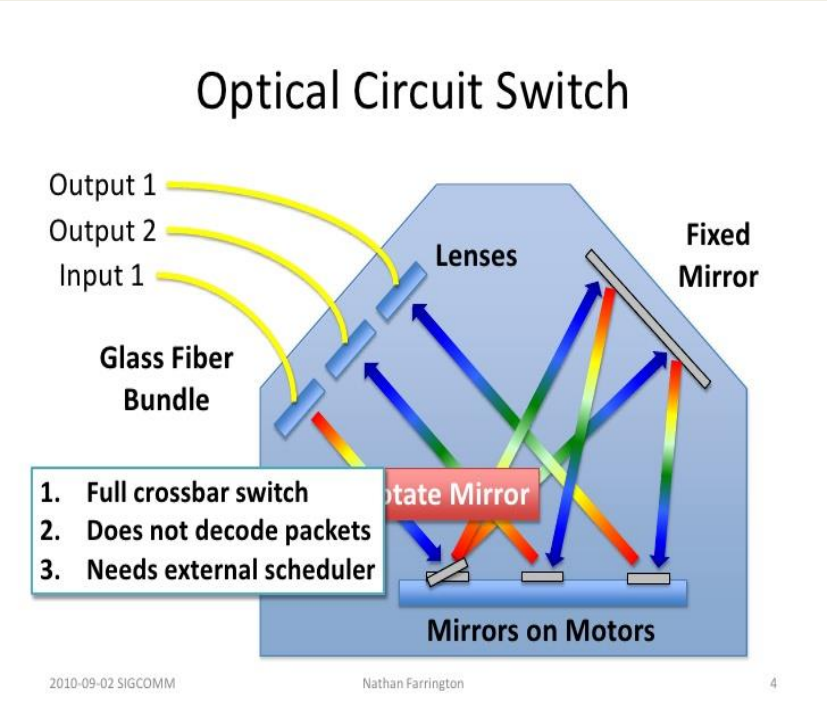


Image courtesy : [7]

2. **All optical switching :**

- OCS gives less granularity for switching
- Optical packet switching (OPS) gives higher granularity for switching
- Lack of optical buffering(optical RAM) and complex computing , logical operation using optics
- Extra OE and EO conversion required

J. Chen, Y. Gong, M. Fiorani and S.Aleksic, " Optical Interconnects at the Top of the Rack for Energy-Efficient Data Centers", *IEEE Communications Magazine*, vol.53, issue. 8, pp. 140-148, Aug.2015

Passive Optical Interconnects(POI)

- Passive components : splitter/combiner, arrayed waveguide gratings
- Access network solution for telecommunication network
- Products for 100 Gb/s PON is under development

Two traffic scenarios in ToR:

- Traffic to/form outside of the rack
- Intra rack communication (multi point to multi point)

Passive Optical Interconnects

Requirements :

- Optical Interface sends traffic to different destination
- Rack controller to solve the contention issues
- Media access protocol for resource allocation algorithm

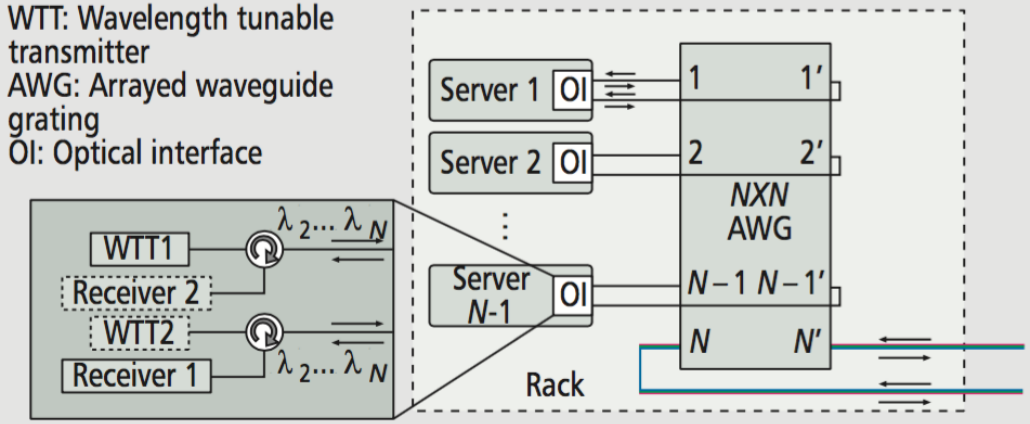
Scheme 1 : N x N AWG based POI

1. N ports of AWG corresponds to N wavelength
2. Optical interface should be able to deal with multiple wavelengths
3. Cyclic property of AWG ensures no conflicts

Passive Optical Interconnects

Scheme 1 : N x N AWG based POI

4. Ports of AWG can be distributed among intra rack and inter/outside data center traffic



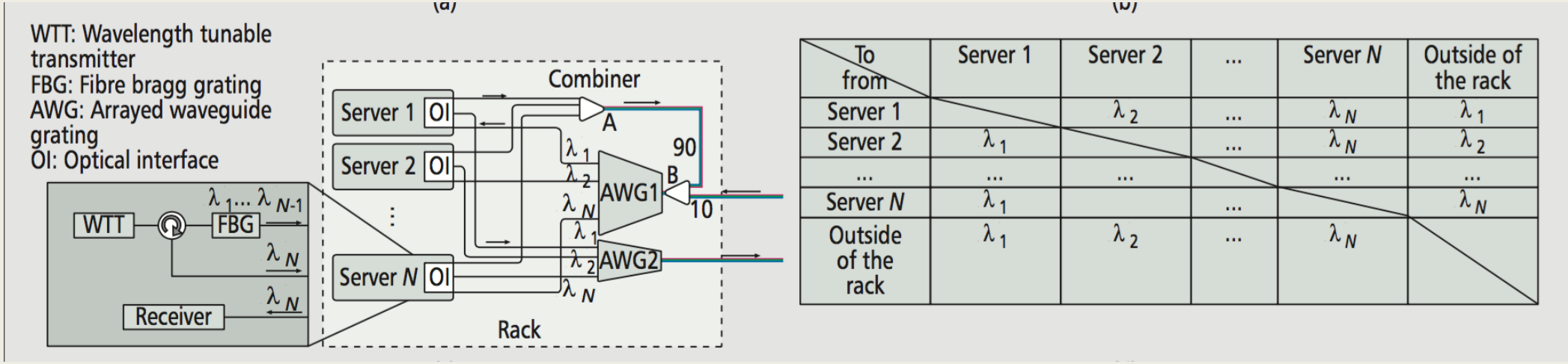
To / from	Server 1	Server 2	…	Server N-1	Outside of the rack
Server 1		λ_2	…	λ_{N-1}	λ_N
Server 2	λ_N		…	λ_{N-2}	λ_{N-1}
…	…	…	…	…	…
Server N-1	λ_3	λ_4	…		λ_2
Outside of the rack	λ_2	λ_3	…	λ_N	

AWG based POI and corresponding wavelength plan[6]

Passive Optical Interconnects

Scheme 2 : AWG + Coupler Based POI

1. A wavelength is assigned to each server
2. Total number of wavelength in a rack = total number of servers inside
3. Wavelength dependent coupler is used as a combiner/splitter



To / from	Server 1	Server 2	...	Server *N*	Outside of the rack
Server 1		λ_2	...	λ_N	λ_1
Server 2	λ_1		...	λ_N	λ_2
...	...	...		...	...
Server *N*	λ_1		...		λ_N
Outside of the rack	λ_1	λ_2	...	λ_N	

AWG + coupler based POI and corresponding wavelength plan[6]

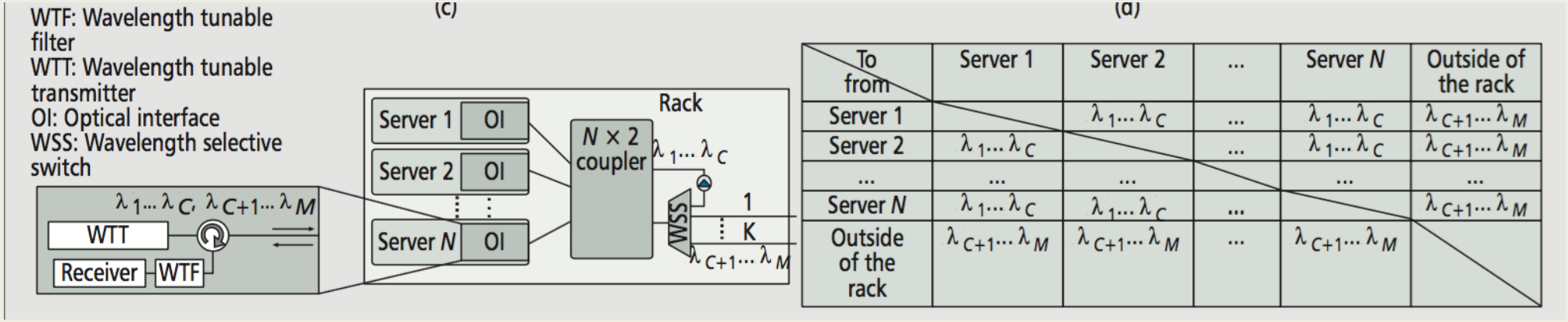
Passive Optical Interconnects

4. All signals from different servers coming to same rack is combined
5. Scalability depends on splitting ratio of couplers

Scheme 3 : Coupler Based POI :

1. A wavelength Selective Switch (WSS) gives flexible channel allocation
2. Nx2 optical coupler is used
3. Broadcast and select

Passive Optical Interconnects



Coupler based POI and corresponding wavelength plan[6]

Performance Comparison

1. Large power saving using POI
2. Scheme 1 (w backup) consumes most power among 3 for using two optical transceiver for protection
3. Scheme 3 is consuming more power due to Wavelength tunable filter and WSS.
4. Passive components are not sensitive to data rate so the cost is lower than Electrical ToR in each case
5. Scheme 1(w backup) always costly among 3 Because of the expensive WTT.
6. All optical architecture reduces overall power Consumption.

Image Courtesy: [6]

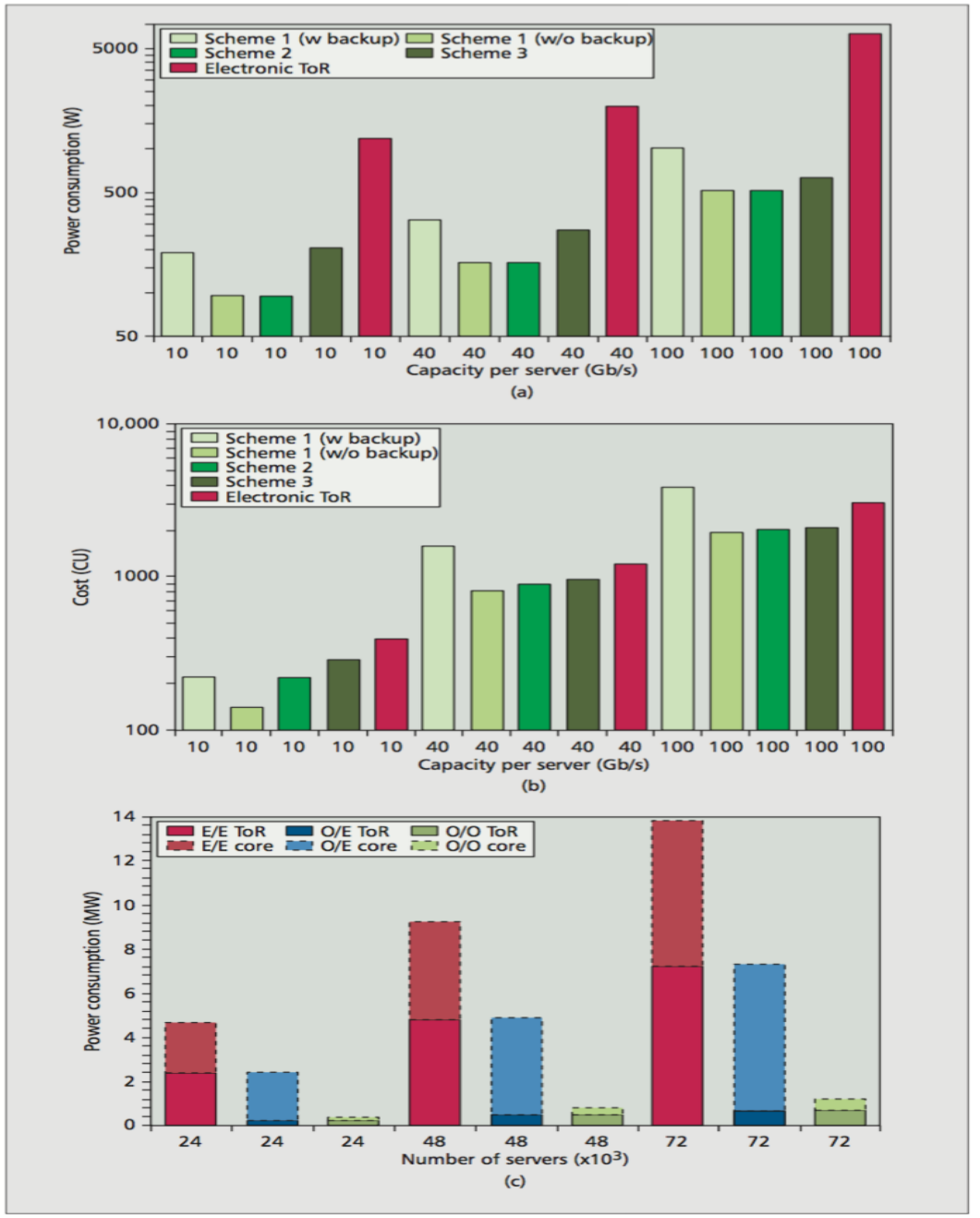


Figure 4. a) Power consumption; b) cost of the three presented schemes and a conventional commodity switch at ToR; c) total power consumed by the overall data center considering three different configurations.

Reference:

1. http://www.cisco.com/c/en/us/solutions/service-provider/global-cloud-index-gci/index.html
2. C. Kachris , K. Bergman and I. Tomkos, " Introduction to Optical Interconnets in Data", *IEEE Communications Surveys & Tutorials*, vol.14, no.4, pp. 1021-1036,2012
3. H. Liu, R. Urata and A. Vahdat ,"Optical Interconnects for Future Data Center Networks", Springer, Avenel, NJ(2013), pp. 17-31
4. http://www.slideshare.net/randybias/pets-vs-cattle-the-elastic-cloud-story
5. M. Veeraraghavan ,"Data Center Networking",Tutorial at IEEE ICC 2014 June 14, 2014.
6. J. Chen, Y. Gong, M. Fiorani and S.Aleksic, " Optical Interconnects at the Top of the Rack for Energy-Efficient Data Centers", *IEEE Communications Magazine*, vol.53, issue. 8, pp. 140-148, Aug.2015
7. http://www.slideshare.net/nfarring/helios-a-hybrid-electricaloptical-switch-architecture-for-modular-data-centers