Optical Interconnects for Cloud Computing Data Centers: Recent Advances and Future Challenges

Dr. Muhammad Imran

muhammad.imran@ncp.edu.pk

National Centre for Physics
Islamabad, Pakistan.

Outline

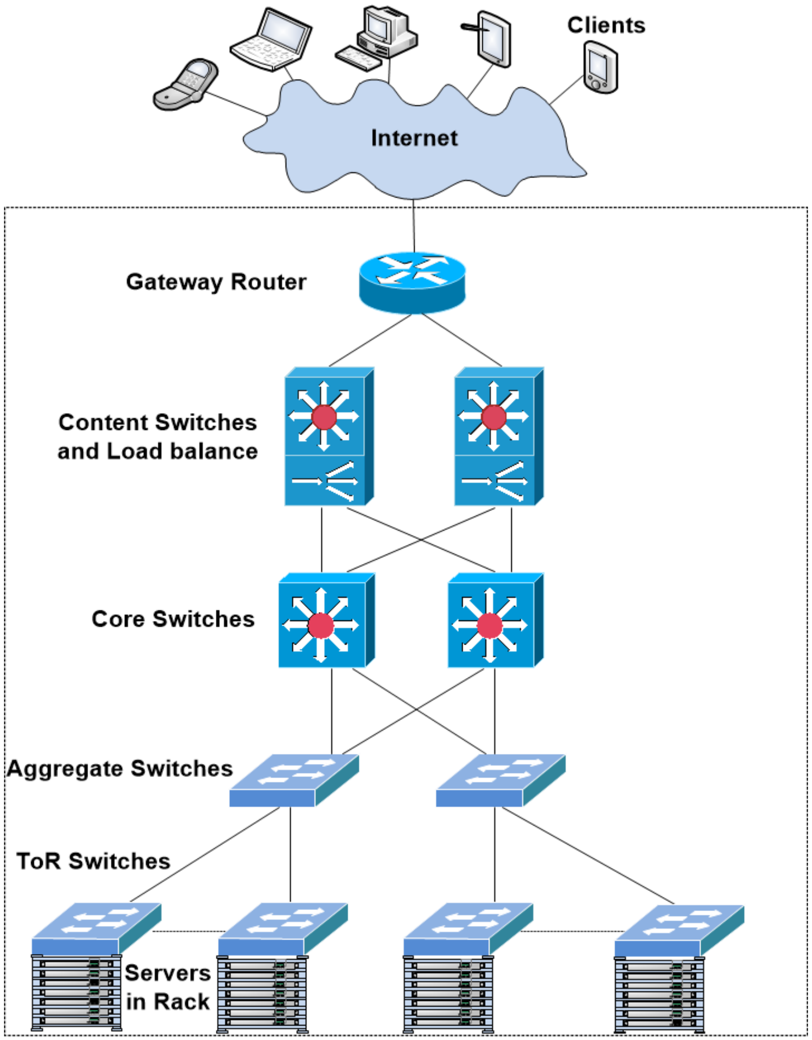
- Introduction
- Challenges in Traditional Data Center Network
- Optical Interconnects
  - Optical Switching Techniques
  - Optical Switches
- Recent Proposals of Optical Interconnects for DCN
  - Architectures based on MEMS:
  - Architectures based on SOAs
  - Architectures based on AWGRs
  - Architectures based on WSSs
  - Hybrid Architectures based on fast and slow optical switches.
- Comparison
- Conclusion and Future work.

Introduction

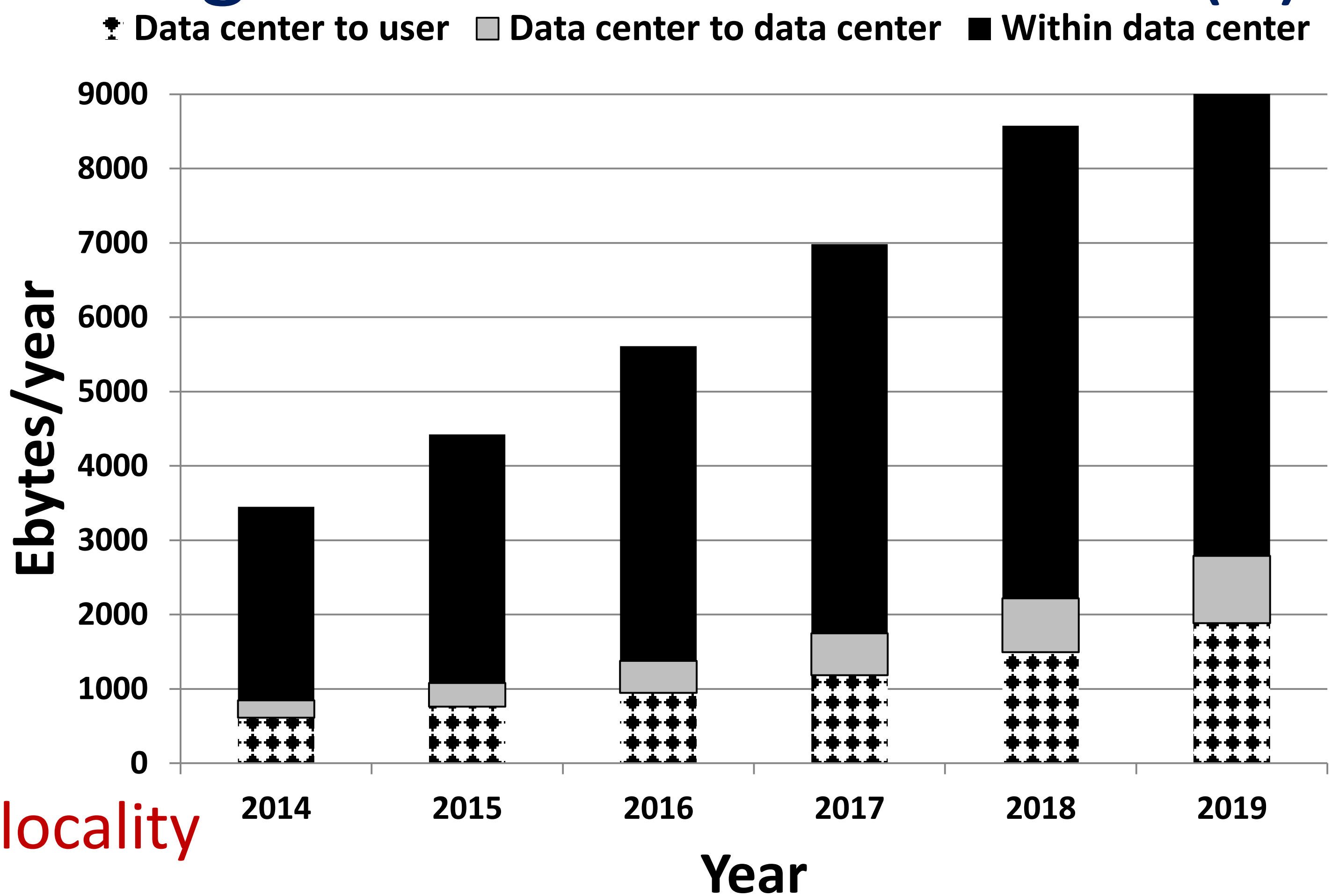
- Exponential increase of internet traffic:
  - emergence of cloud computing based applications.
  - applications run on the servers deployed in the data centres
  - require huge bandwidths.
- Importance of Data centres
  - Computation and storage away from desktop to large scale data centres.
  - IP traffic will keep on increasing due to smart devices and multimedia applications.
- Need of building fast and high bandwidth interconnect network in the data centre.

Challenges in Traditional DCNs

- Power consumption
  - ToR, aggregate/core switches and transceivers requiring O-E-O conversion are the main contributions to power consumption.
  - The power consumption of the interconnection network accounts for 23% of the total IT power consumption
    - It will increase significantly in future DCNs as data rates increase.
- Increasing bandwidth demand
- Latency
- Network Oversubscription
- Scalability



Challenges in Traditional DCNs…(2)



- Traffic locality
  - Cisco global cloud index reveals that majority of the traffic in future DCNs will remain within data centres while only a very small portion of the traffic will go the external network.

Optical Interconnects

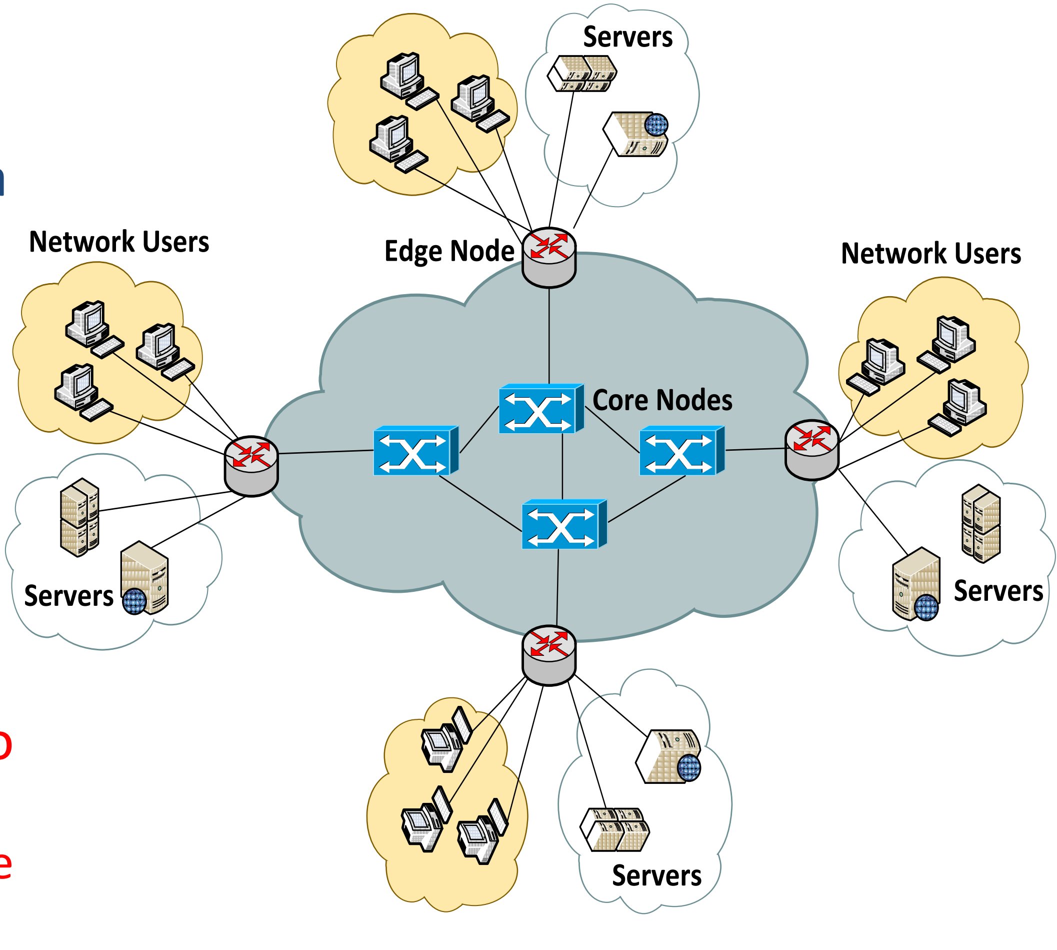
- Optical networks for data centre networks (DCNs):
  - Key requirement to achieve Internet-scale data centres
  - Provide huge bandwidths
  - Low latency
  - Power Efficient
- The performance of optical network is related with optical switching:
  - Optical Circuit Switching (OCS)
  - Optical Packet Switching (OPS)
  - Optical Burst Switching (OBS)

Optical Switching Techniques

- Optical Circuit Switching (OCS):
  - Connection Oriented
  - Suitable for long lasting connections
  - Large connection establishment time
  - Bandwidth underutilization in the case of low traffic load
- Optical Packet Switching (OPS):
  - Connectionless
  - O-E-O conversion of header for processing in electrical domain
  - Fiber delay lines (FDLs) for buffering during header processing
  - Lack of feasible optical buffer
  - Packet loss due to output port contention
  - O-E-O conversion increases energy consumption
  - Speed of header processing should be compatible with increasing data rate.

Optical Switching Techniques...(2)

- Optical Burst Switching (OBS):
  - Packets are aggregated into a burst at the edge node
  - Control packet is sent ahead of the burst to reserve resources
  - The burst is sent a certain time after the control packet
  - Burst loss due to output port contention
  - Poor TCP performance due to burst loss
    - Contention induced losses are misinterpreted as congestion induced losses



OBS Network

Optical Switches

1. Slow Optical Switches:
   - MEMS
     - High port density
     - Transceivers are not required
     - Power efficiency due to passive switching
     - Low insertion loss and crosstalk
     - Data rate independent
     - Support bidirectional communications
     - Less expensive
     - **Switching time is in the order of tens of milliseconds**
   - WSSs
     - **Switching time is in the range of micro to milliseconds**
2. Fast Optical Switches
   - AWGR
     - Nanoseconds switching time.
     - Expensive and are not scalable.
   - SOA
     - Nanoseconds switching time.
     - Expensive

Recent Proposals of Optical Interconnects for Data Centers

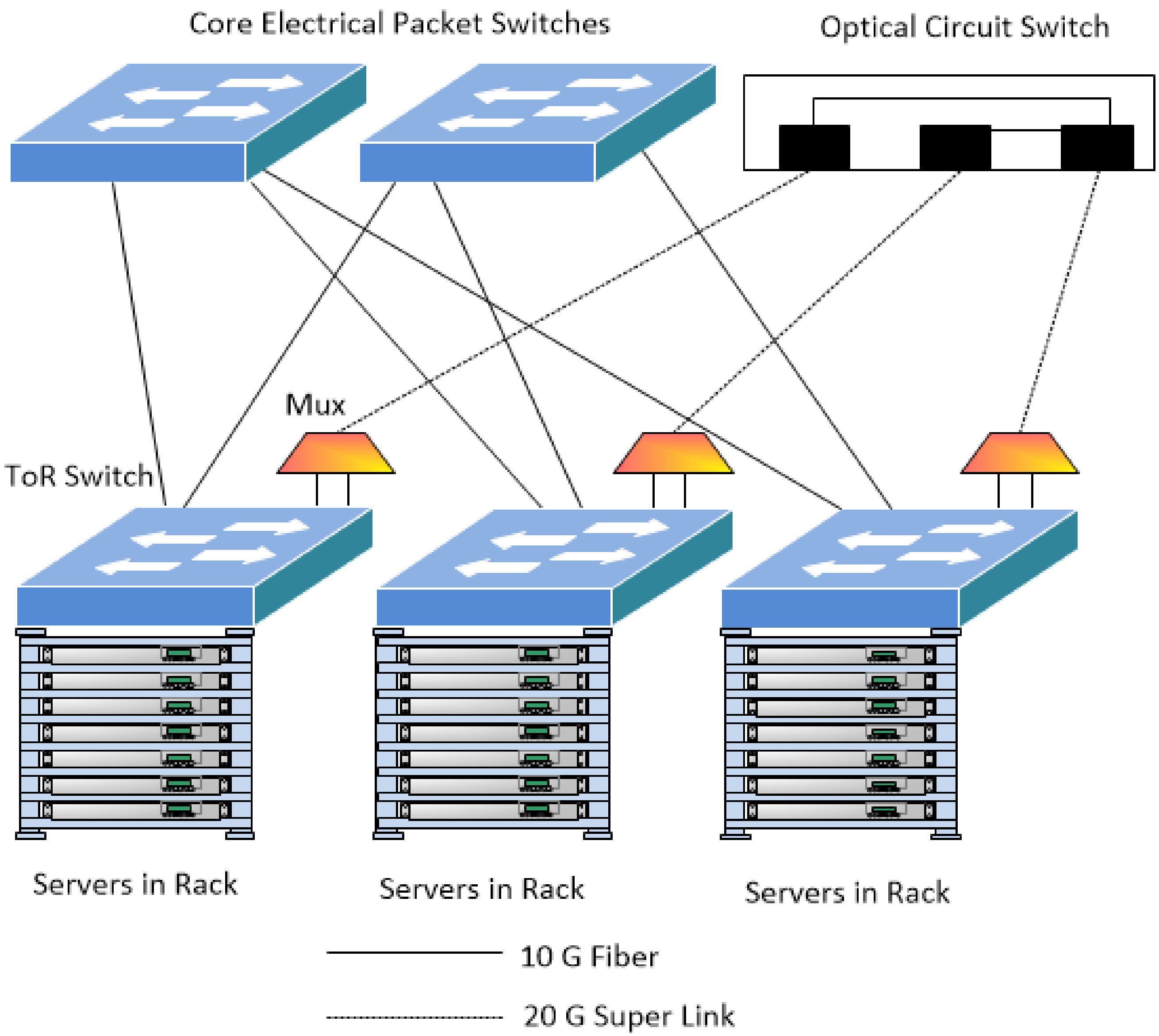
Optical Interconnects

1. Architectures based on MEMS:
2. Architectures based on SOAs
3. Architectures based on AWGRs
4. Architectures based on WSSs
5. Hybrid Architectures based on fast and slow optical switches.

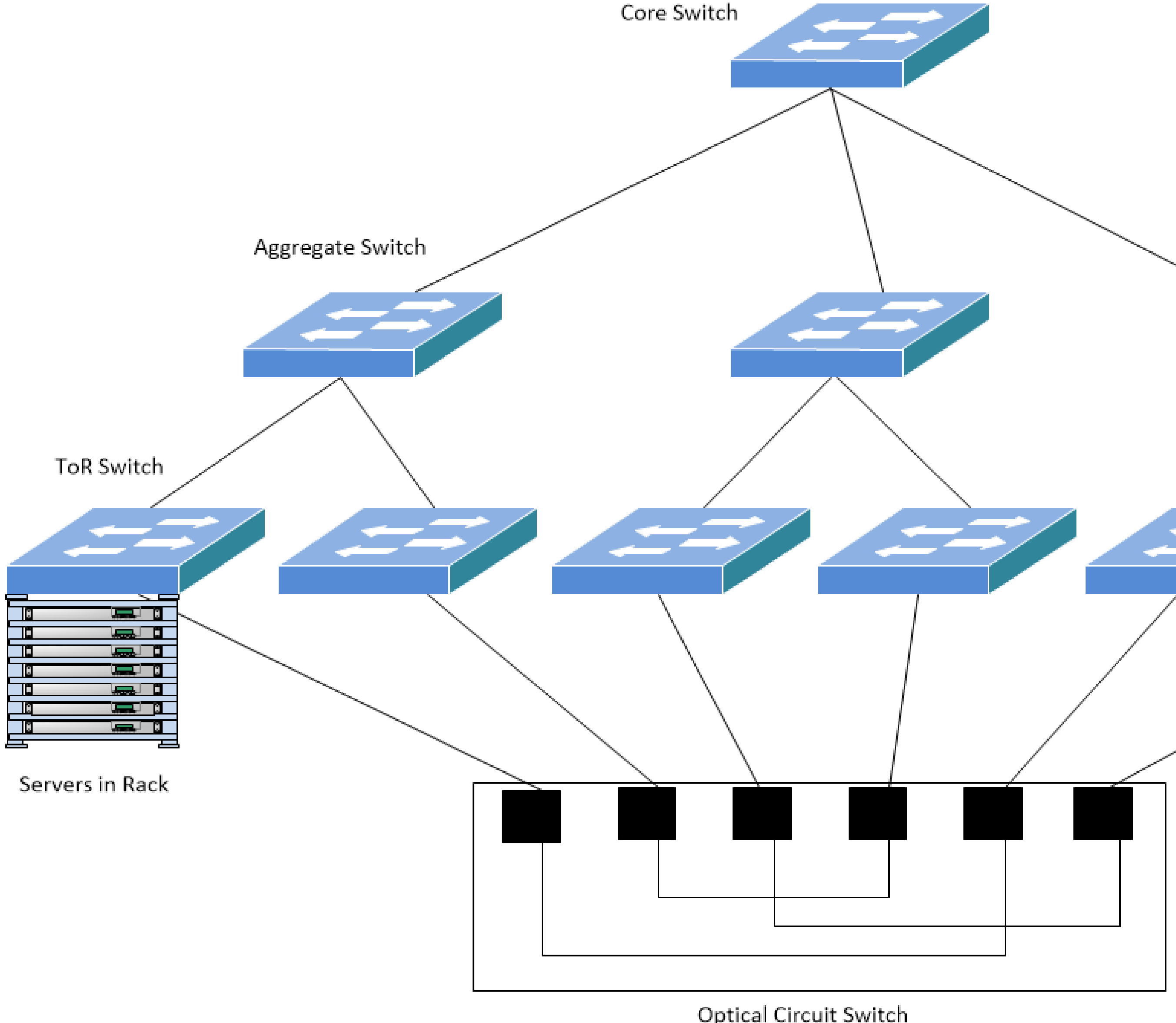
Architectures based on MEMS:

1. Helios
2. HyPaC
3. OSA
4. Reconfigurable Architecture
5. HydRA

Helios

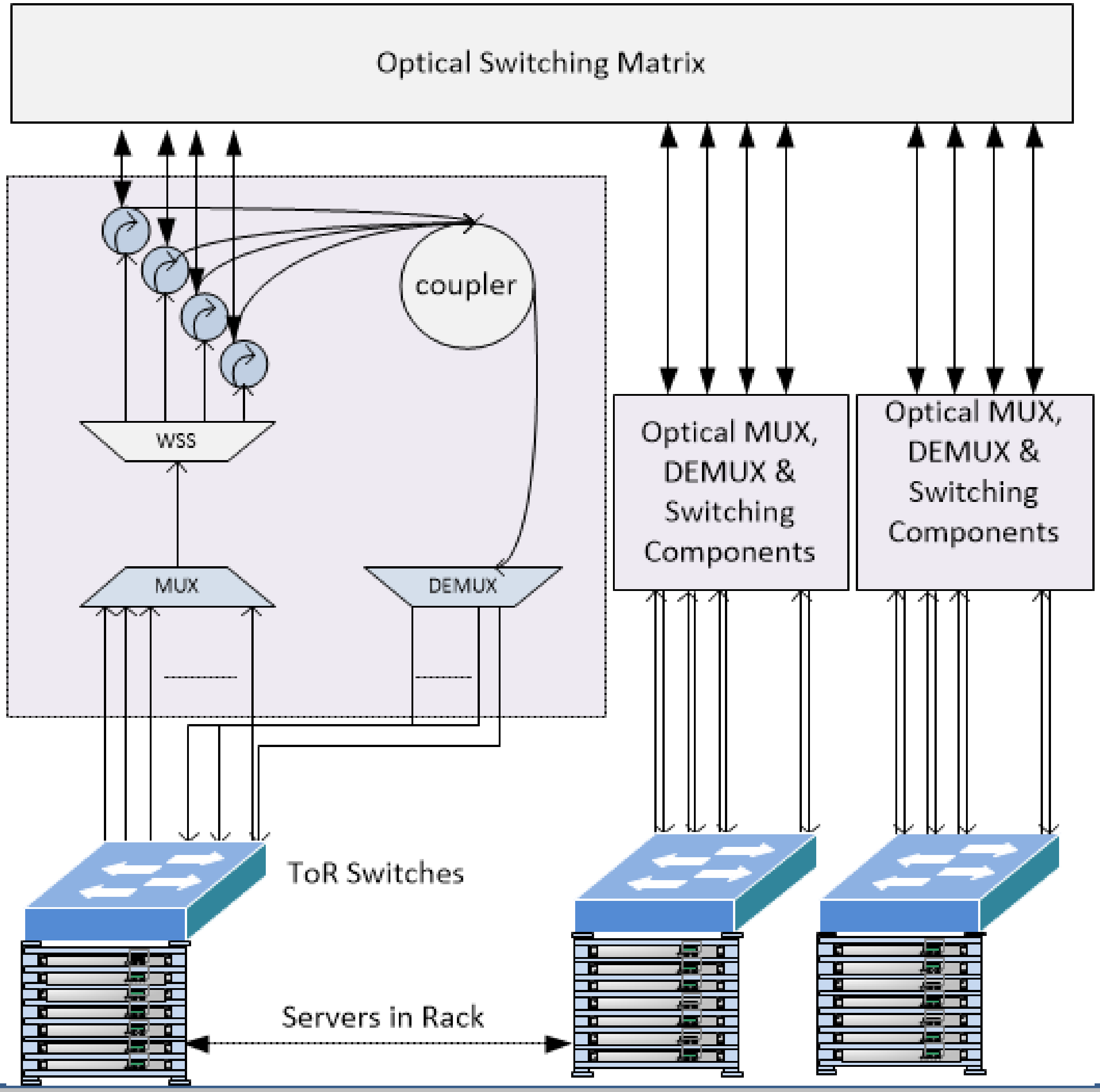

Core Electrical Packet Switches
Optical Circuit Switch
Mux
ToR Switch
Servers in Rack
Servers in Rack
Servers in Rack
10 G Fiber
20 G Super Link

HyPac



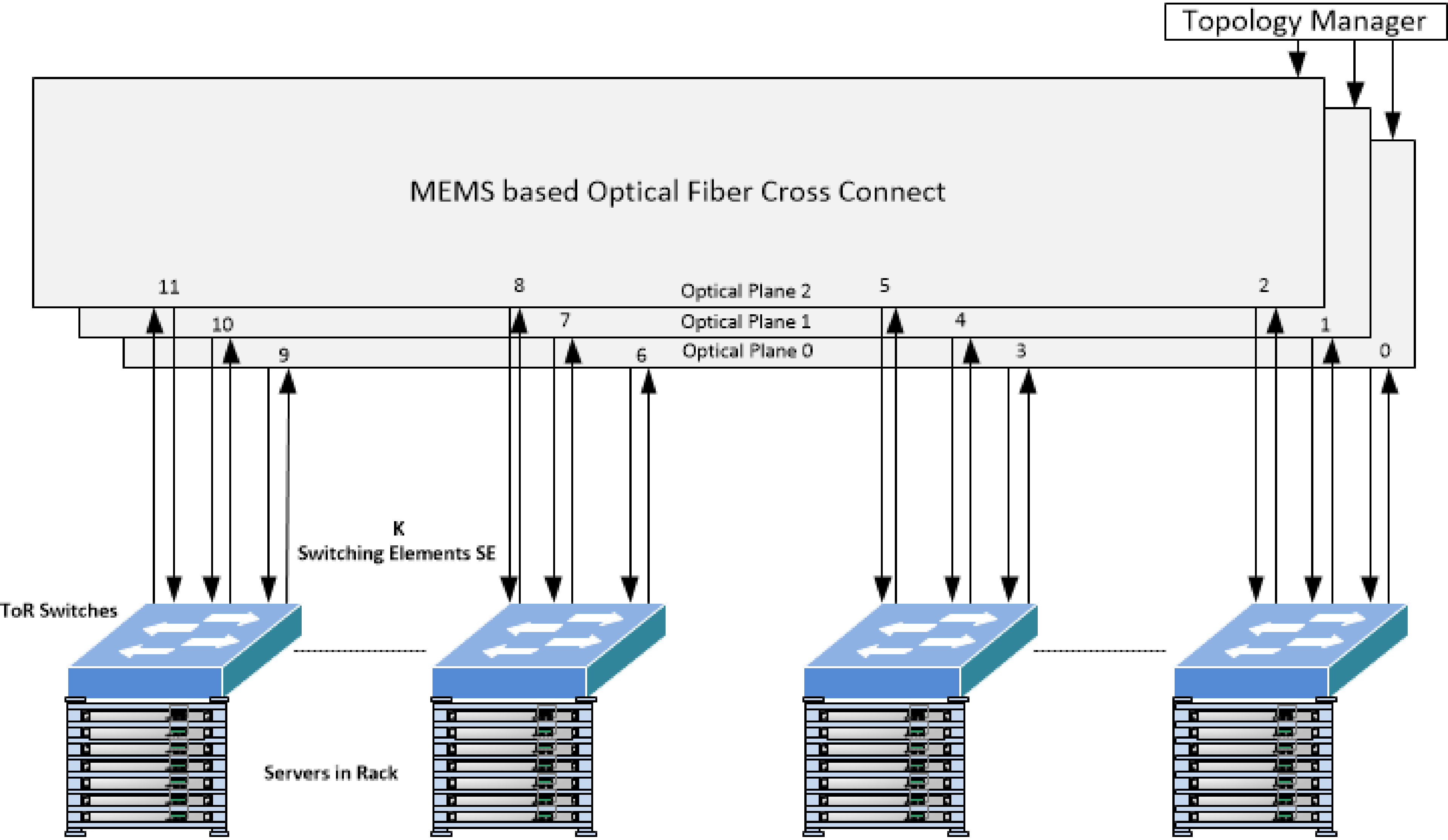
Core Switch
Aggregate Switch
ToR Switch
Servers in Rack
Optical Circuit Switch

OSA

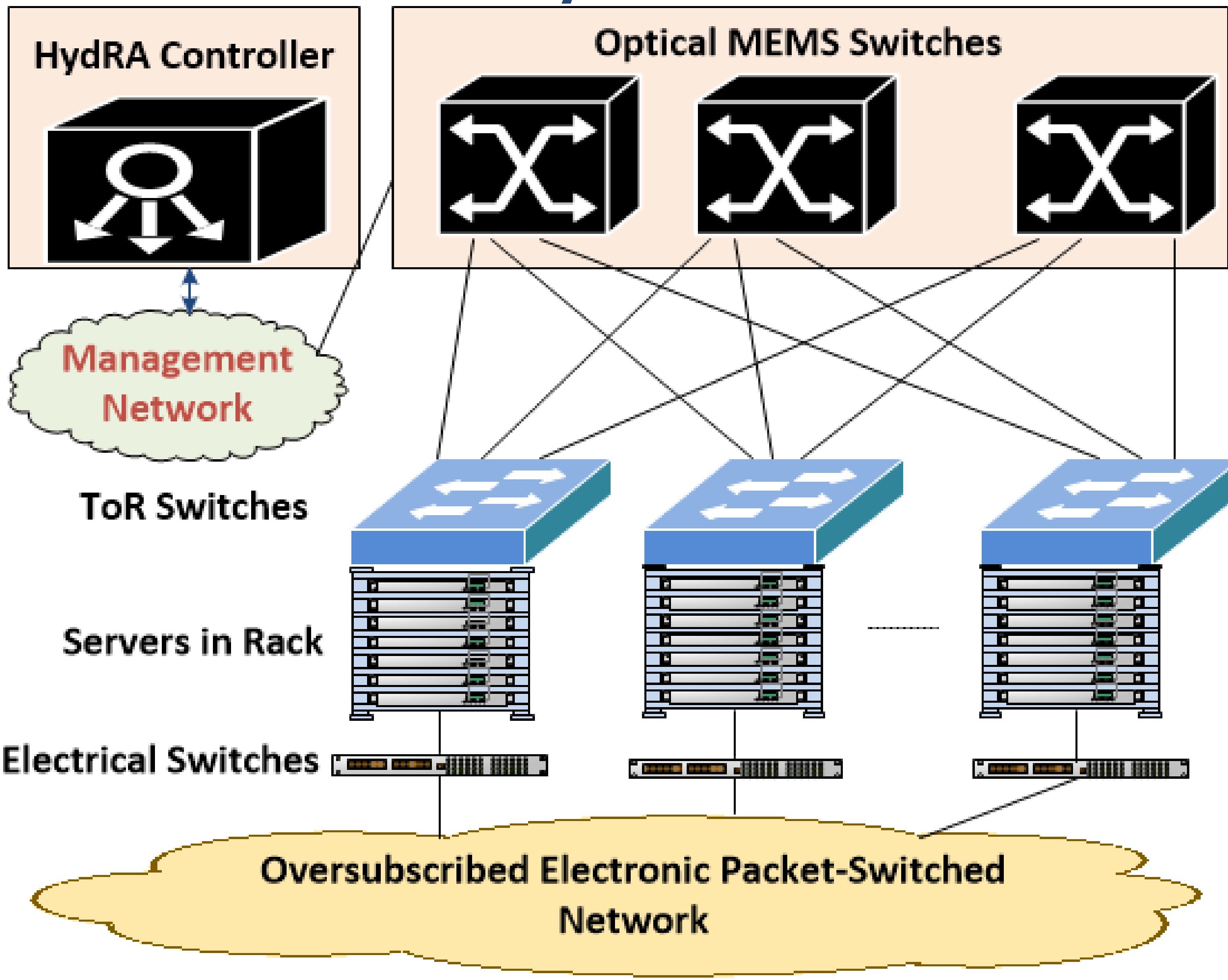


Optical Switching Matrix
coupler
WSS
MUX
DEMUX
Optical MUX, DEMUX & Switching Components
Optical MUX, DEMUX & Switching Components
ToR Switches
Servers in Rack

Reconfigurable Architecture



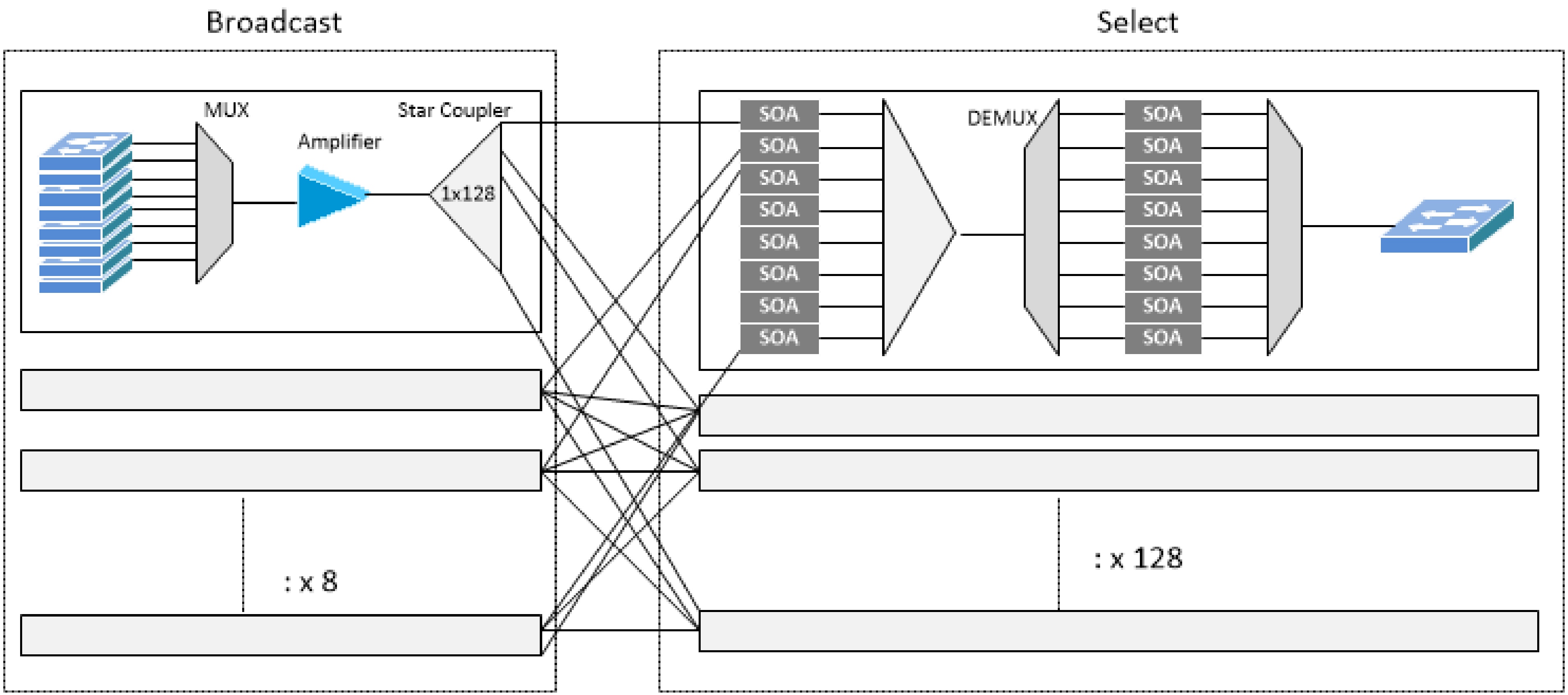
HydRA


HydRA Controller
Optical MEMS Switches
Management Network
ToR Switches
Servers in Rack
Electrical Switches
Oversubscribed Electronic Packet-Switched Network

Architectures based on SOAs

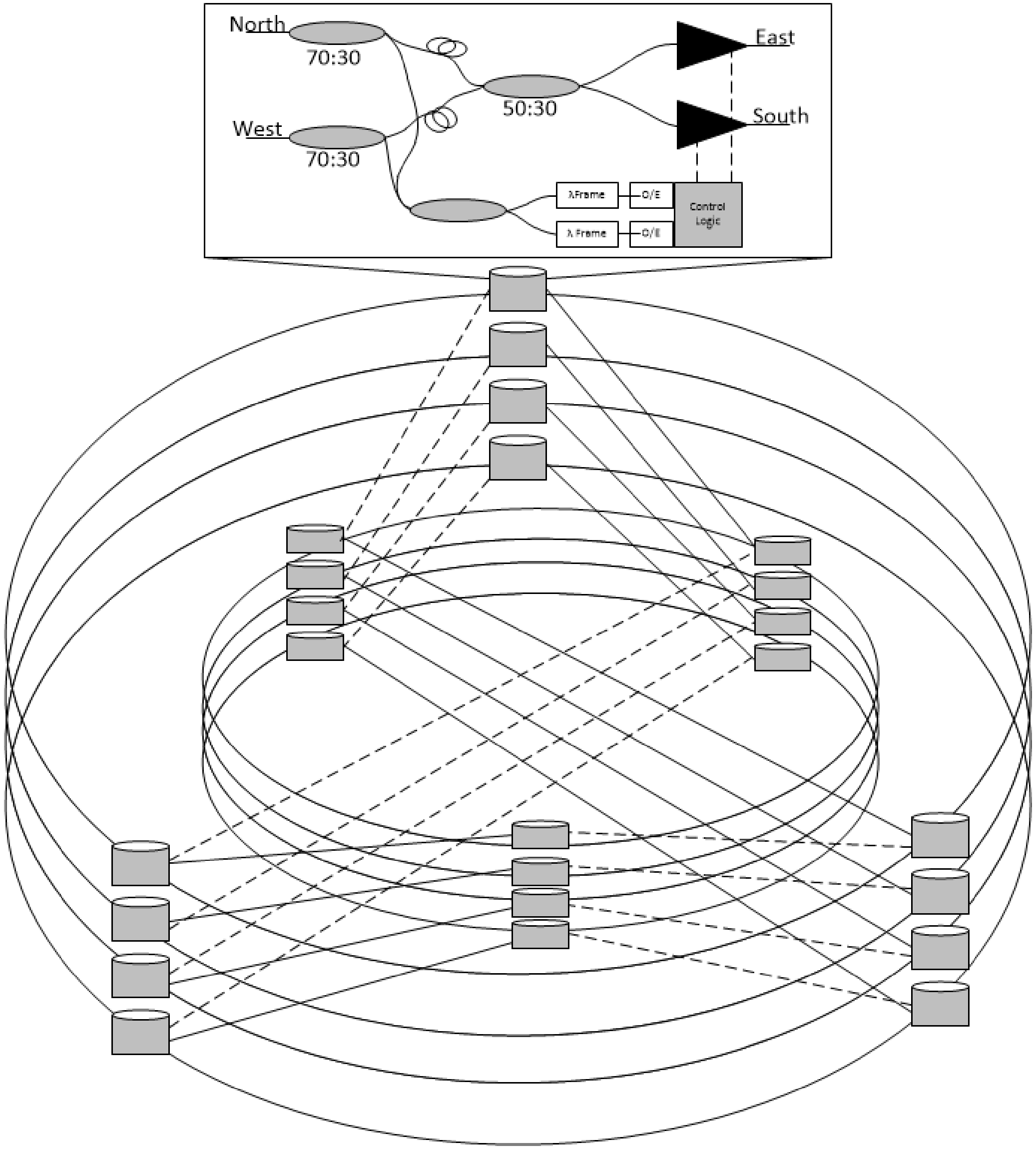
1. OSMOSIS
2. Data Vortex
3. Bidirectional Architecture
4. SW
5. STIA

OSMOSIS

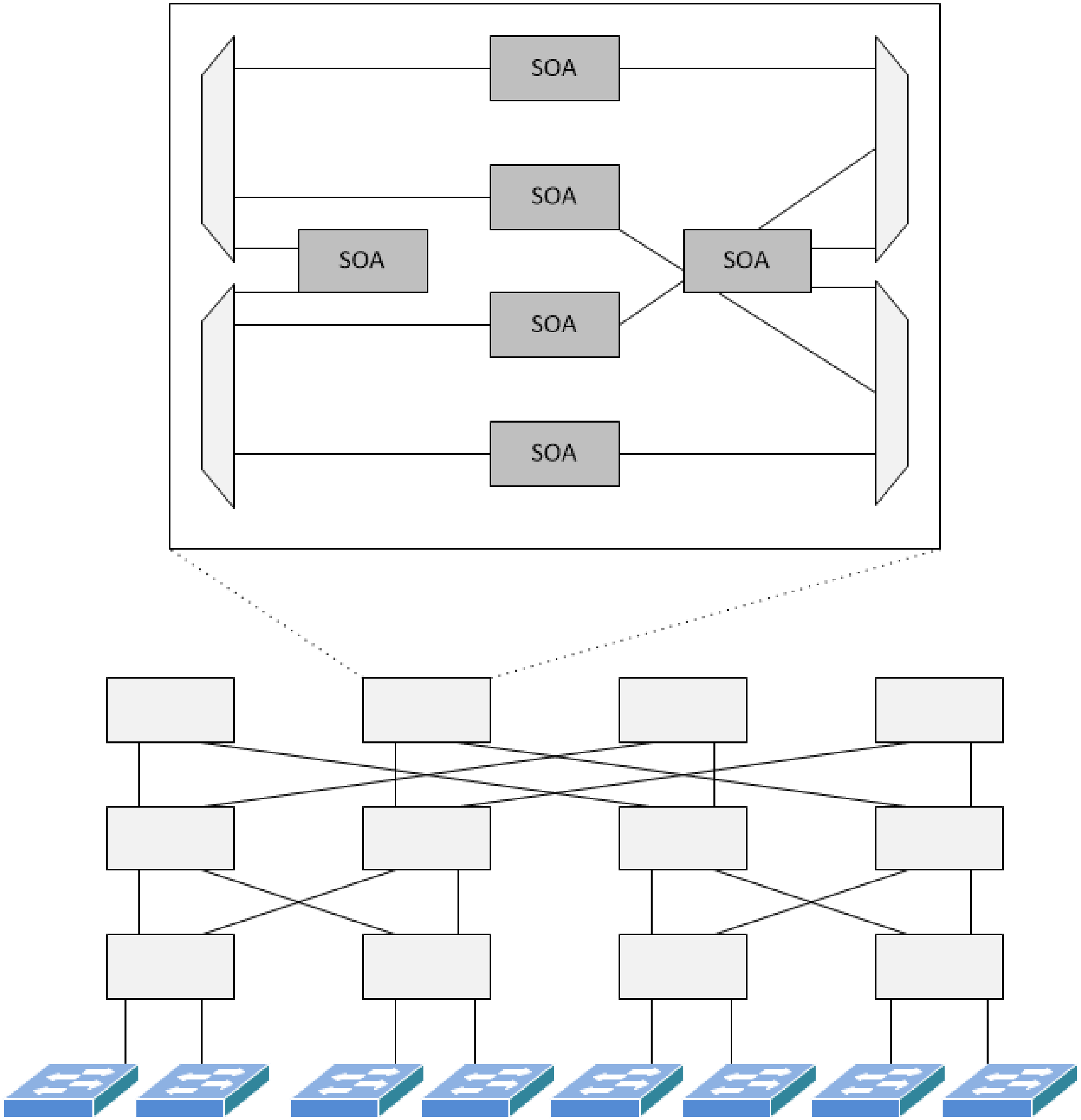


Broadcast
Select
MUX
Star Coupler
Amplifier
1x128
DEMUX
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
SOA
: x 8
: x 128

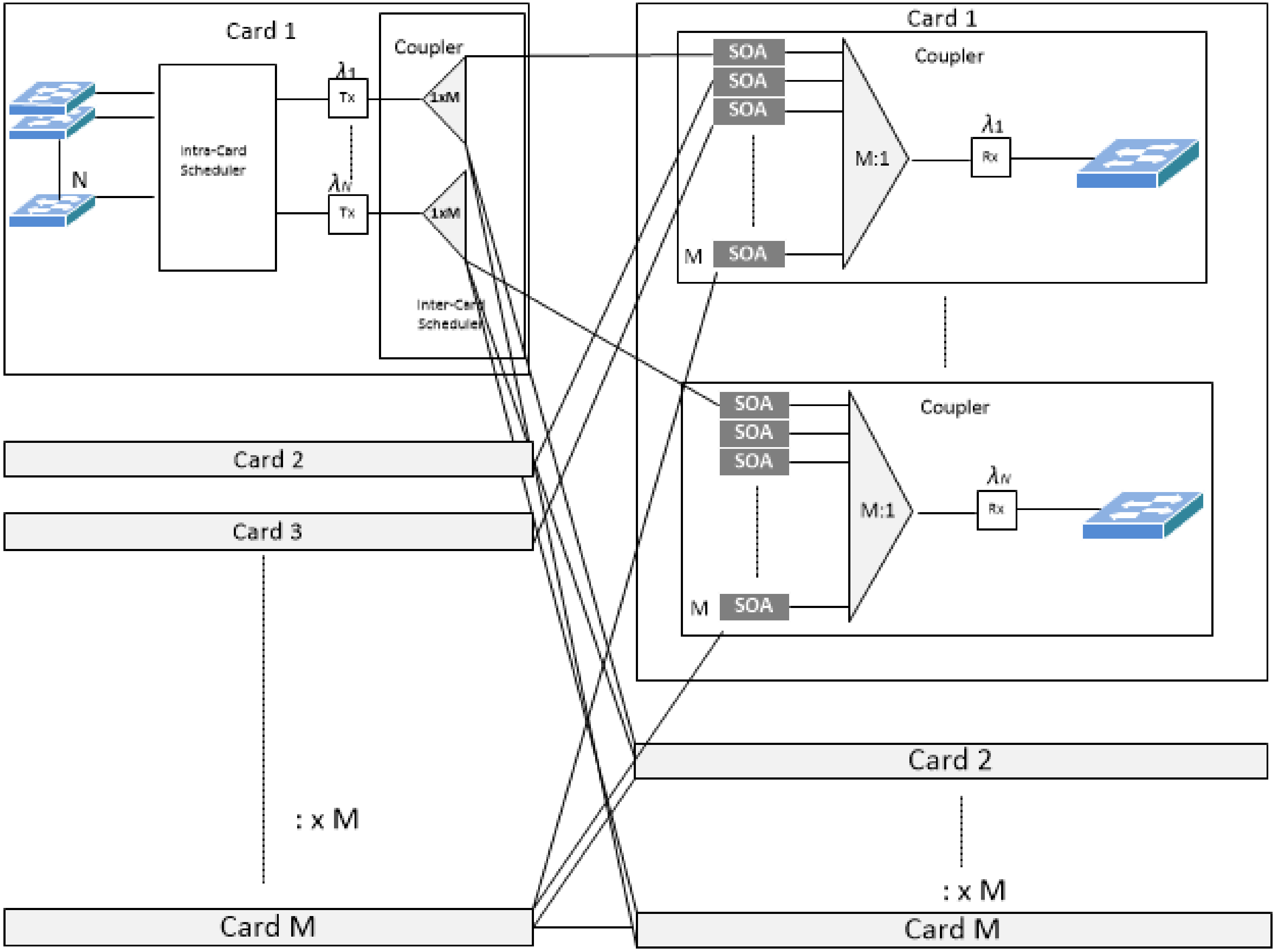
Data Vortex


North
70:30
West
70:30
50:30
East
South
λFrame
λ Frame
O/E
O/E
Control Logic

Bidirectional

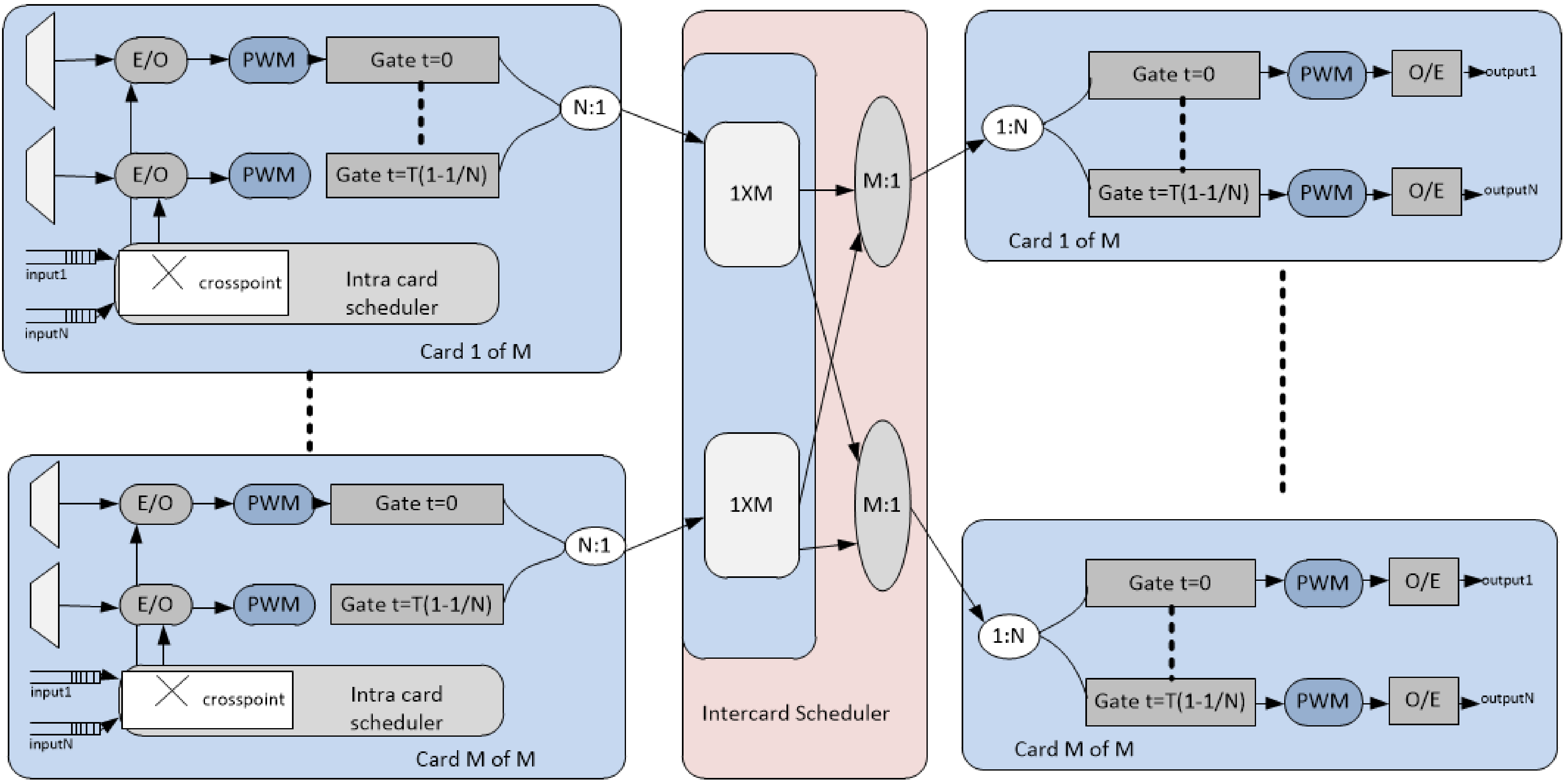


SW



Card 1
Coupler
λ_1
Tx
1xM
Intra-Card Scheduler
N
λ_N
Tx
1xM
Inter-Card Scheduler
Card 2
Card 3
: x M
Card M
Card 1
SOA
SOA
SOA
Coupler
M:1
λ_1
Rx
M
SOA
SOA
SOA
SOA
Coupler
M:1
λ_N
Rx
M
SOA
Card 2
: x M
Card M

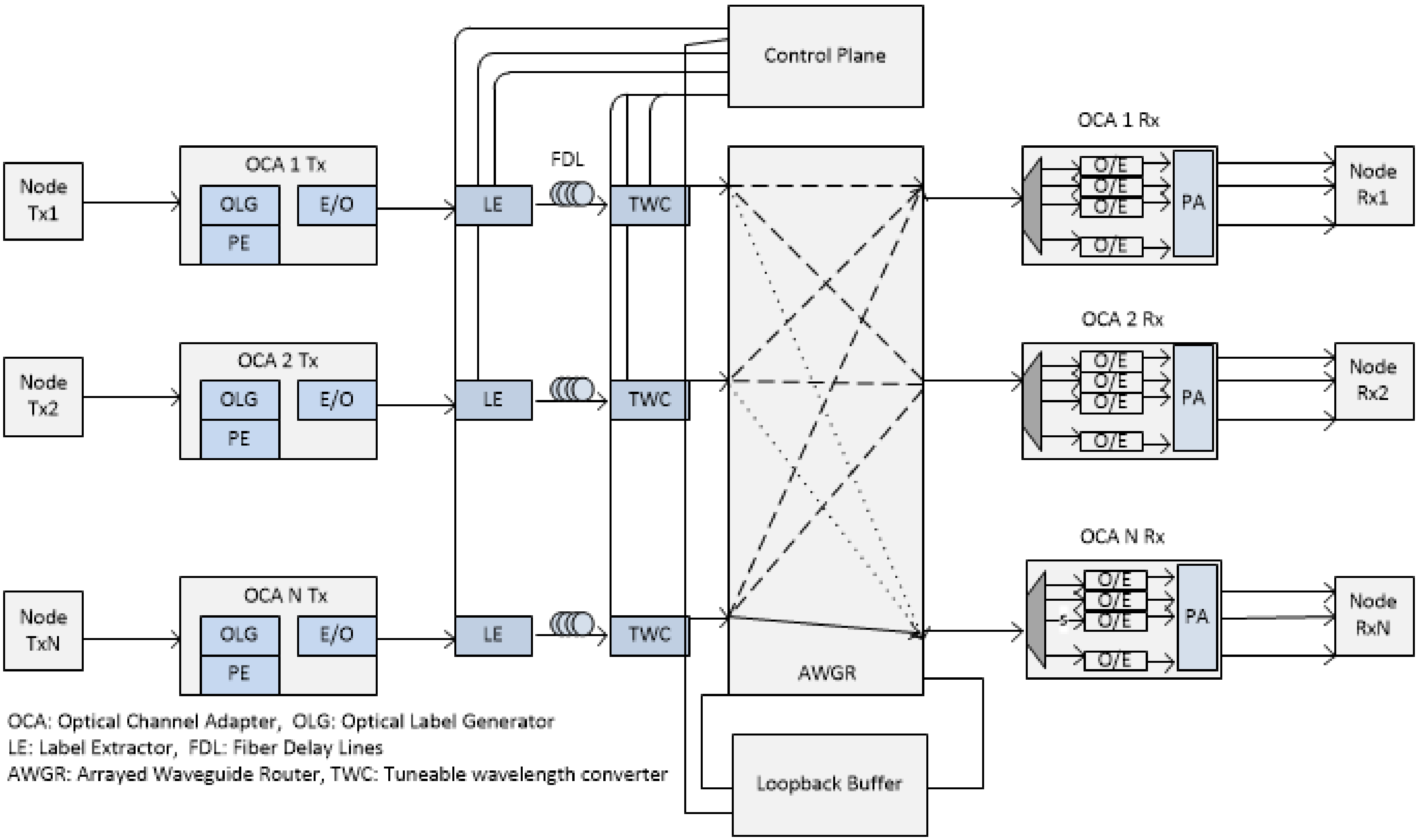
STIA


E/O
PWM
Gate t=0
Gate t=T(1-1/N)
N:1
input1
inputN
crosspoint
Intra card scheduler
Card 1 of M
Card M of M
1XM
M:1
Intercard Scheduler
1:N
O/E
output1
outputN

Architectures based on AWGRs

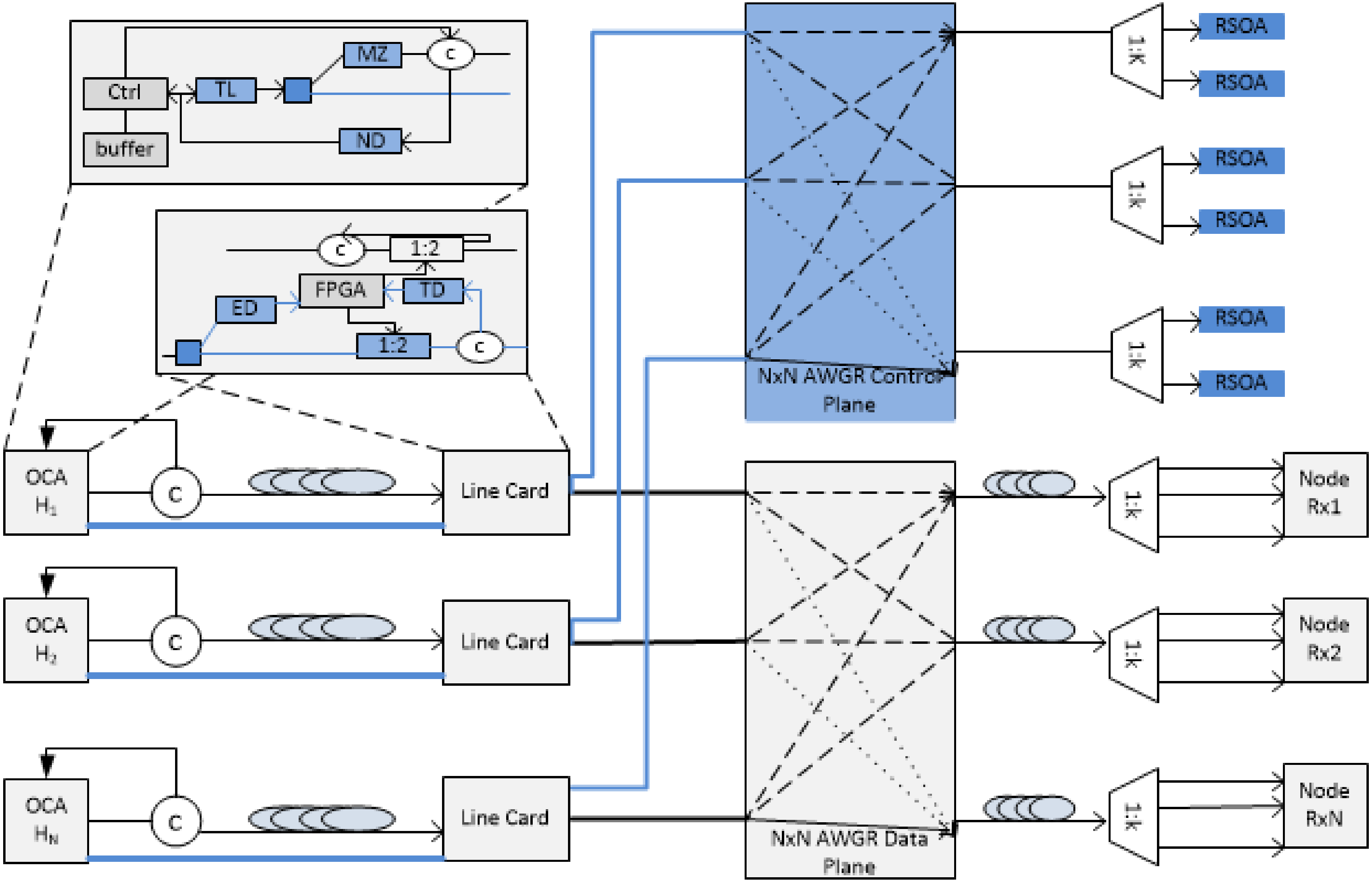
1. LIONS
2. TONAK-LION
3. Petabit
4. IRIS
5. OFDM-b ased

LIONS

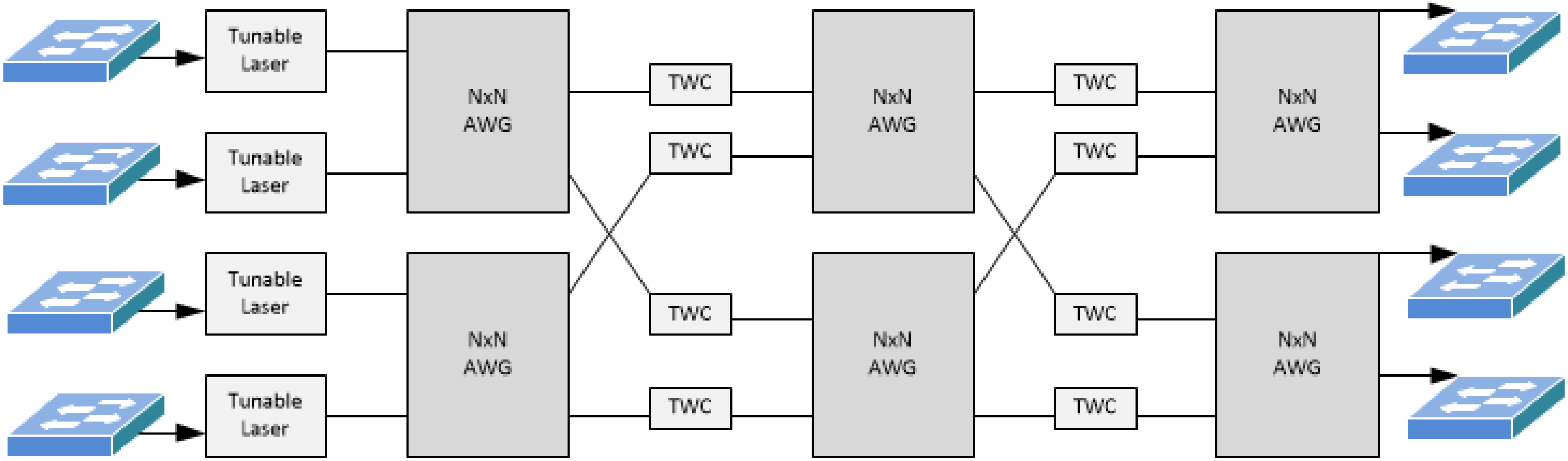


OCA: Optical Channel Adapter, OLG: Optical Label Generator
LE: Label Extractor, FDL: Fiber Delay Lines
AWGR: Arrayed Waveguide Router, TWC: Tuneable wavelength converter

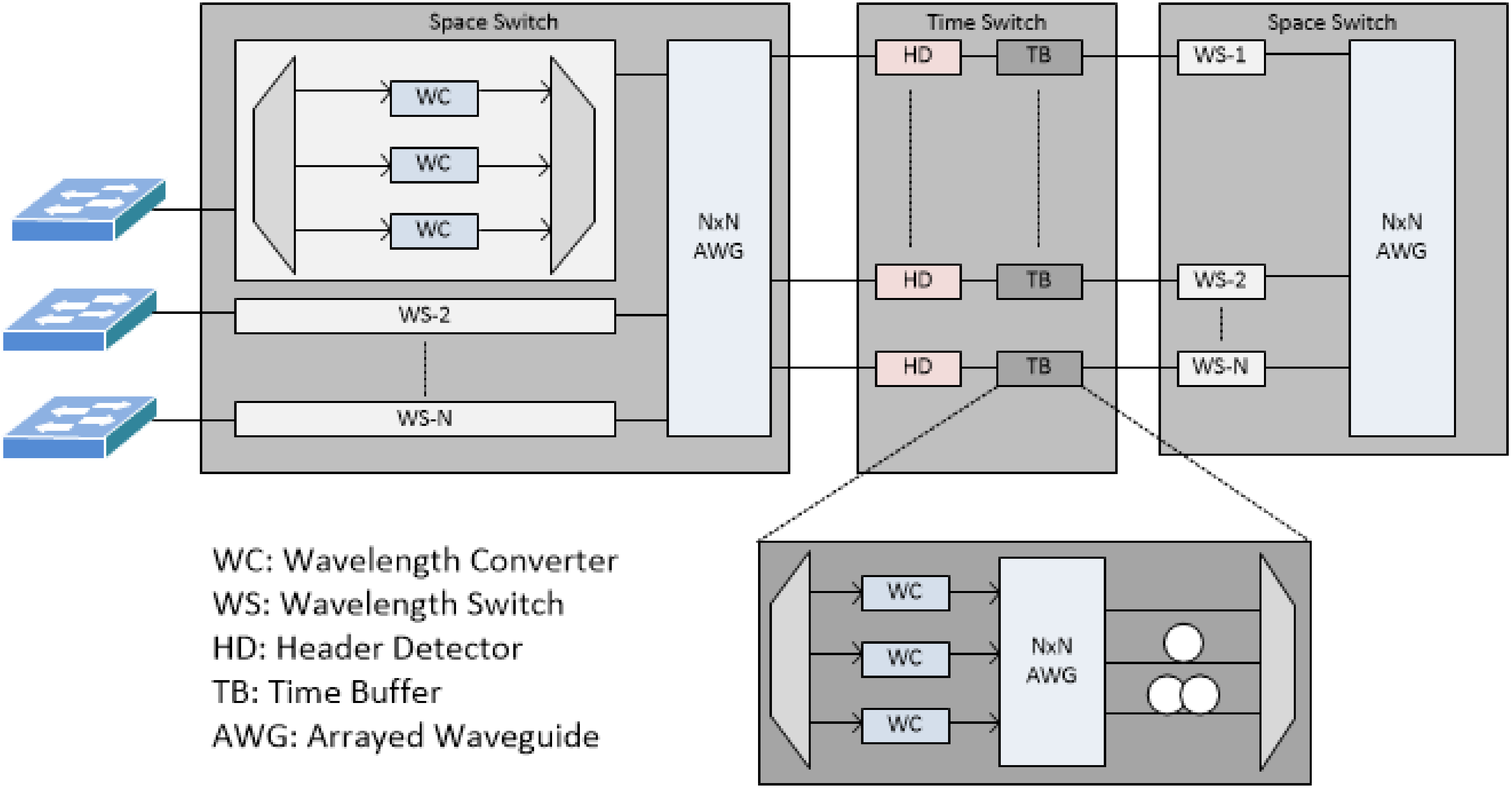
TONAK LIONS


Ctrl
TL
MZ
c
buffer
ND
c
1:2
FPGA
TD
ED
1:2
c
OCA H1
C
Line Card
OCA H2
C
Line Card
OCA HN
C
Line Card
NxN AWGR Control Plane
1:k
RSOA
RSOA
1:k
RSOA
RSOA
1:k
RSOA
RSOA
NxN AWGR Data Plane
1:k
Node Rx1
1:k
Node Rx2
1:k
Node RxN

Petabit

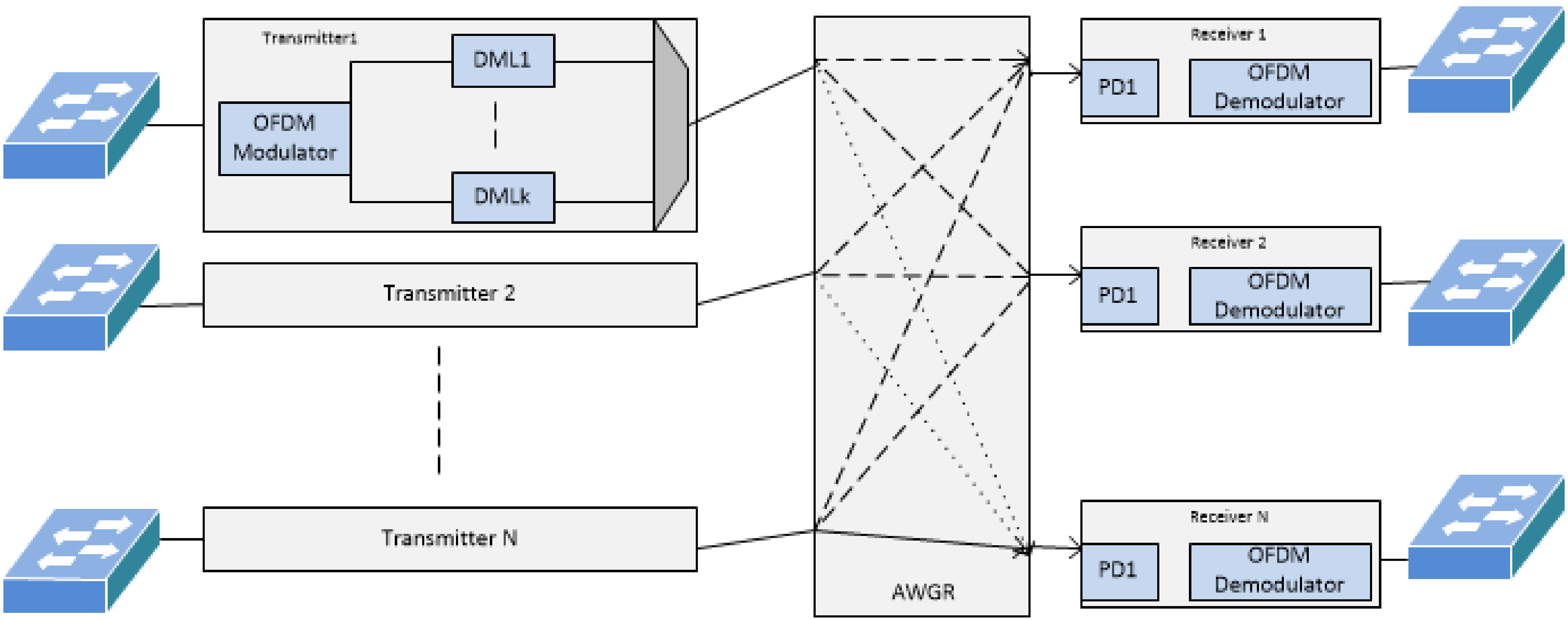

Tunable Laser
Tunable Laser
Tunable Laser
Tunable Laser
NxN AWG
NxN AWG
TWC
TWC
TWC
TWC
NxN AWG
NxN AWG
TWC
TWC
TWC
TWC
NxN AWG
NxN AWG

IRIS



WC: Wavelength Converter
WS: Wavelength Switch
HD: Header Detector
TB: Time Buffer
AWG: Arrayed Waveguide

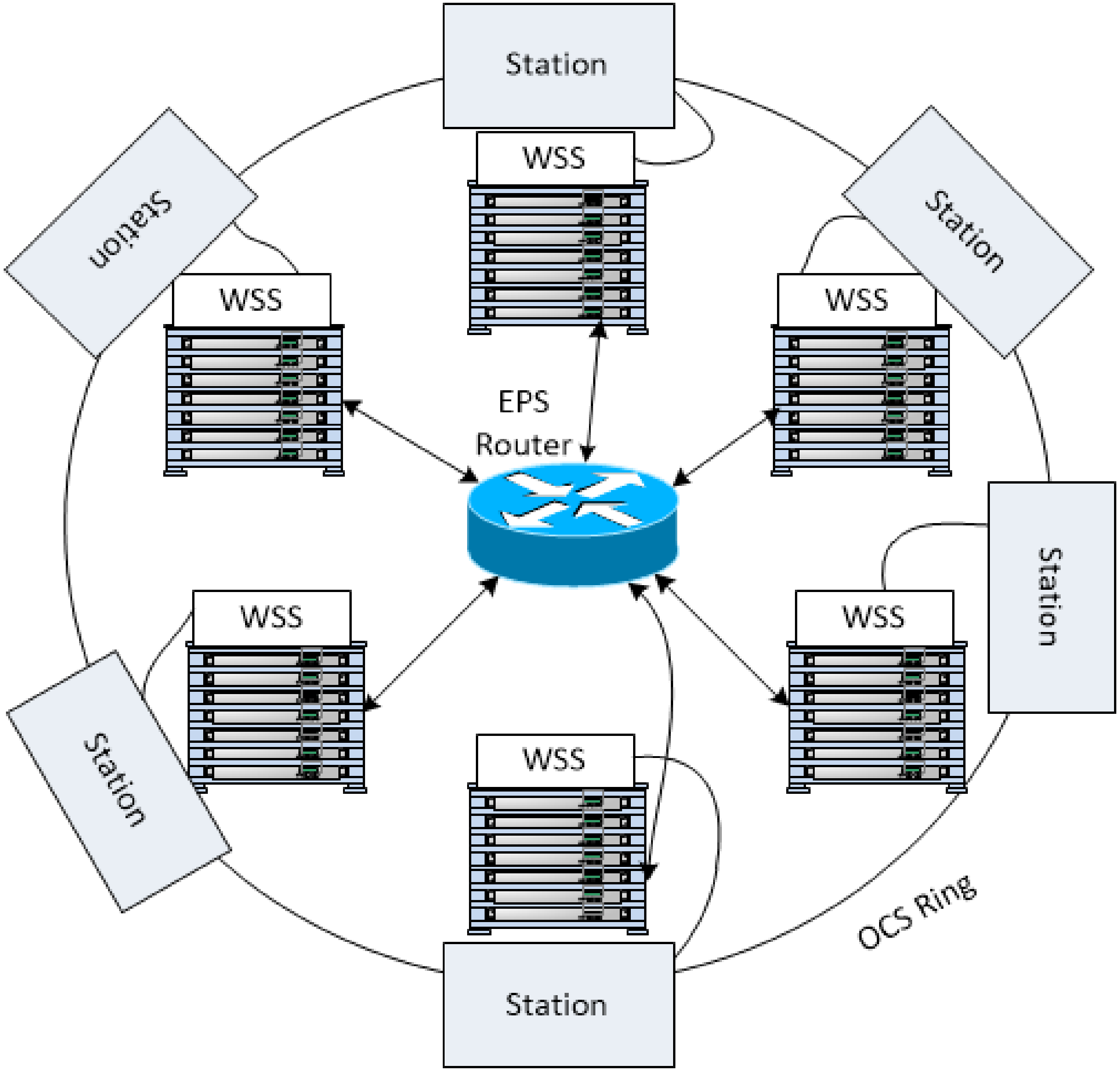
OFDM-based Architecture



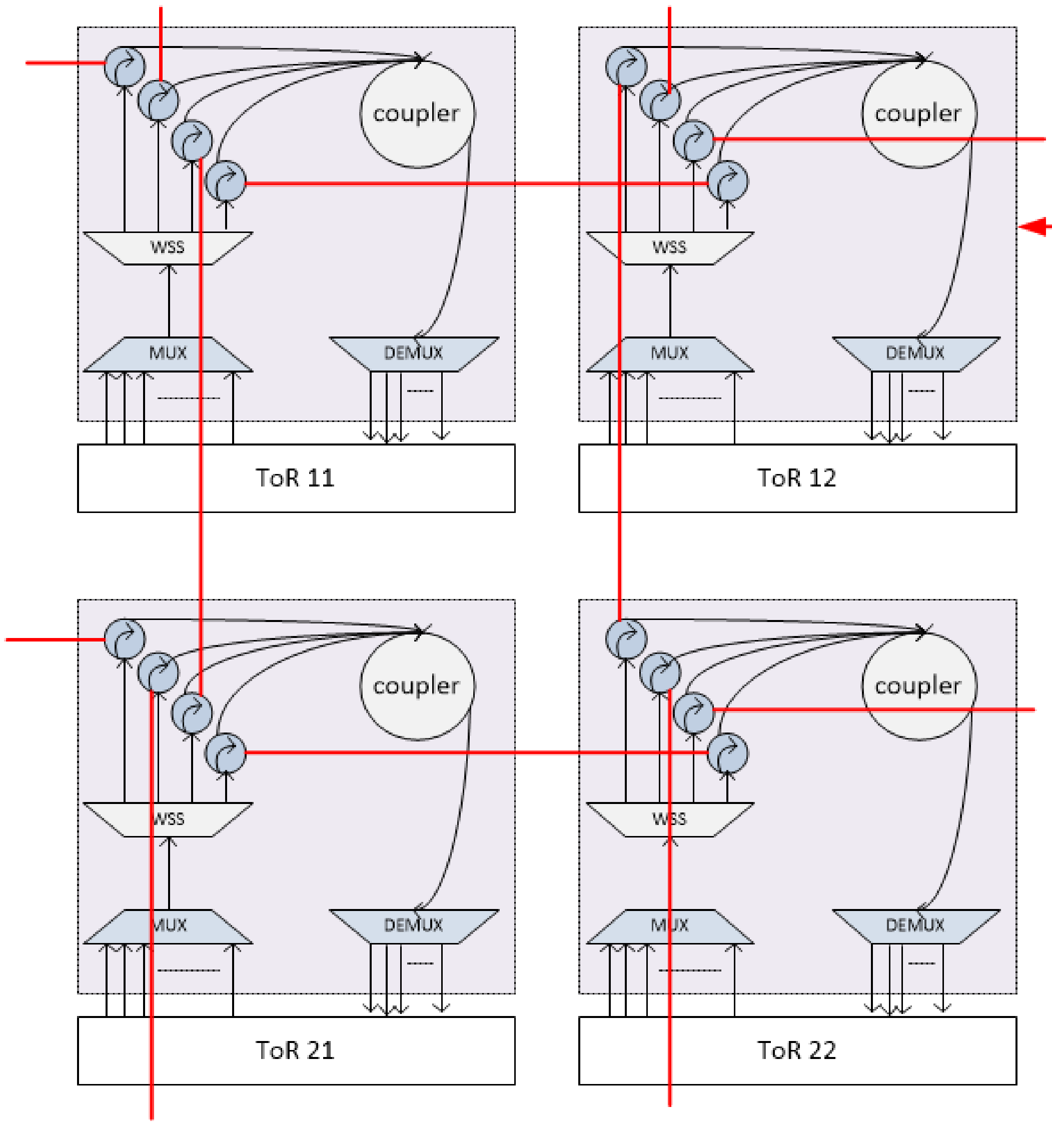
Architectures based on WSSs

1. Mordia
2. WaveCube
3. OPMDC

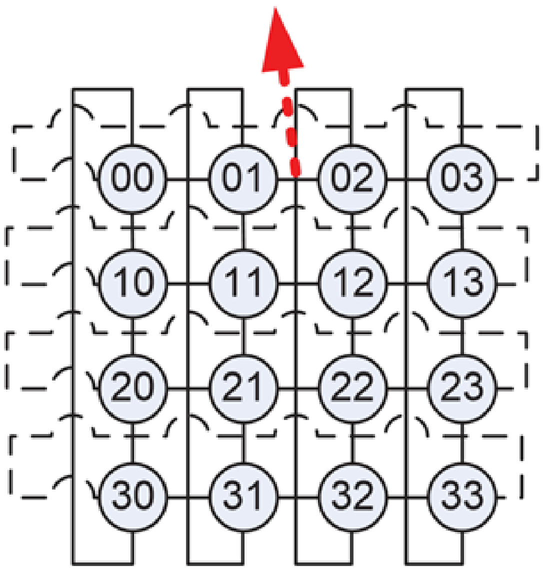
Mordia


Station
WSS
Station
WSS
Station
WSS
EPS
Router
Station
WSS
WSS
Station
WSS
OCS Ring
Station

WaveCube

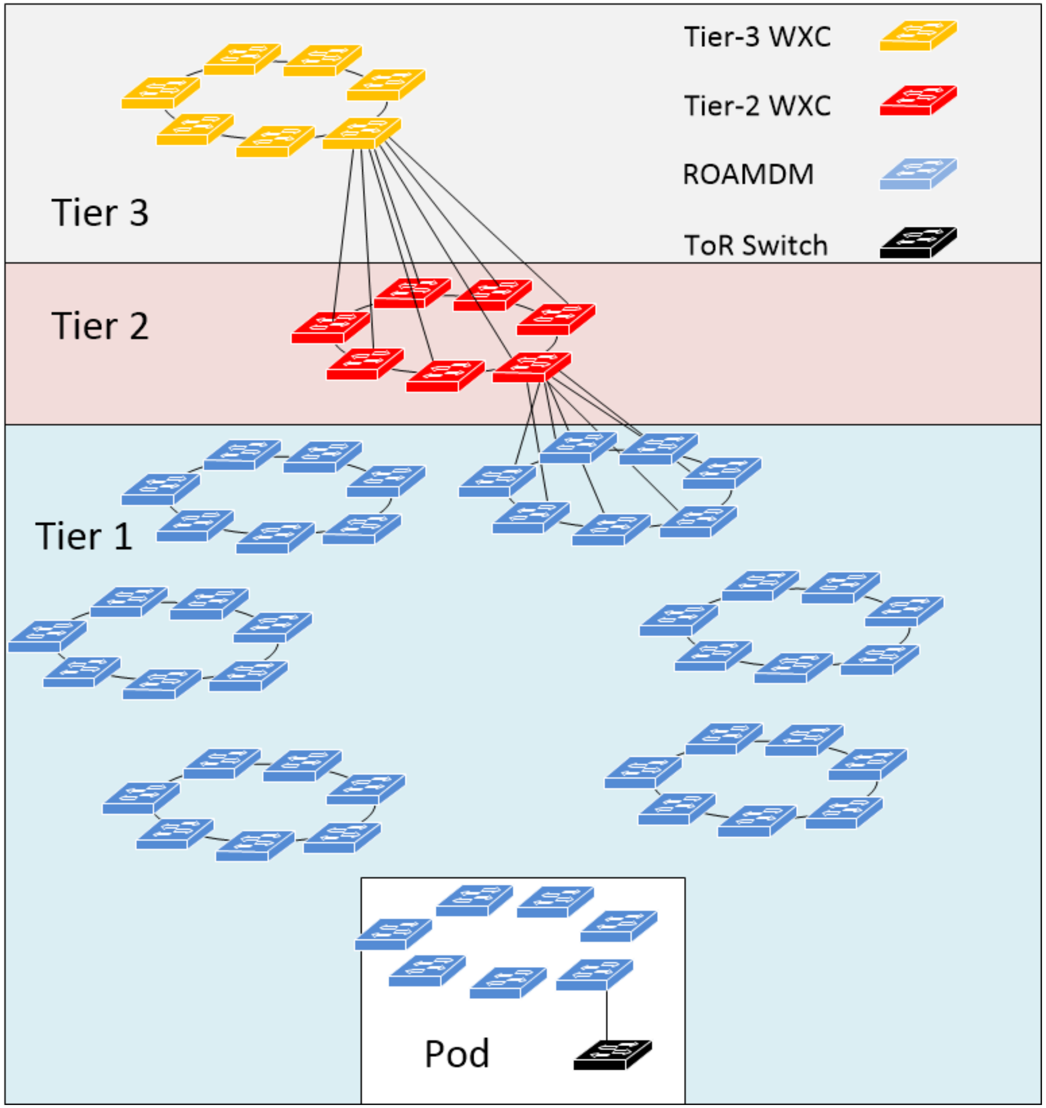


Physical Topology



ToR Level Logical Graph

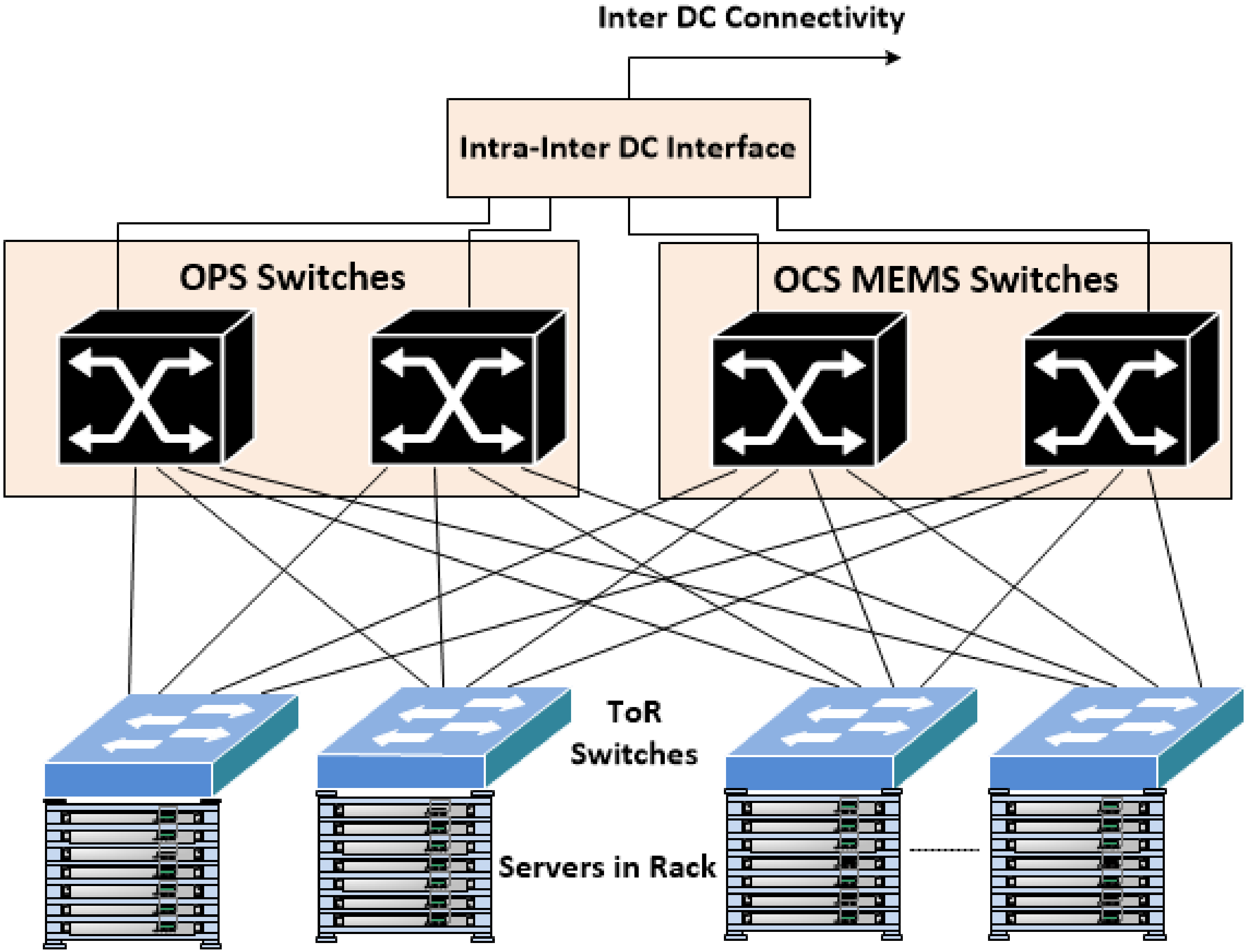
OPMDC



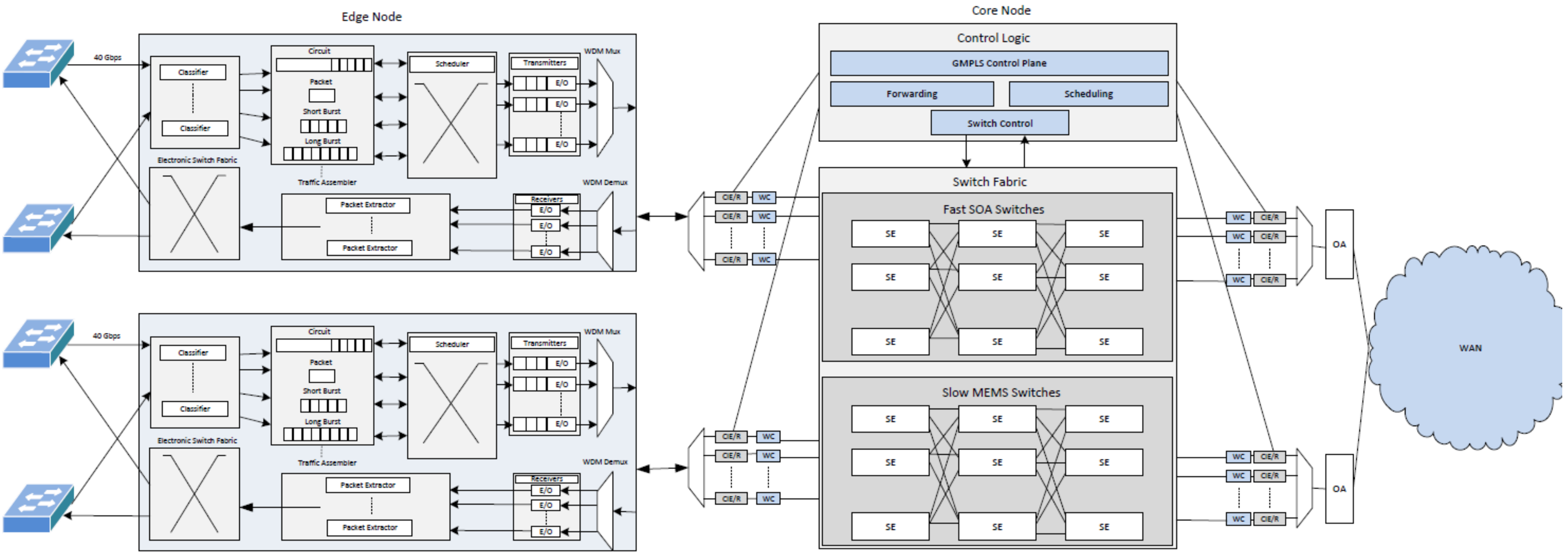
Hybrid Architectures based on fast and slow optical switches.

1. LIGHTNESS
2. Hybrid Optical Switching
3. HOSA

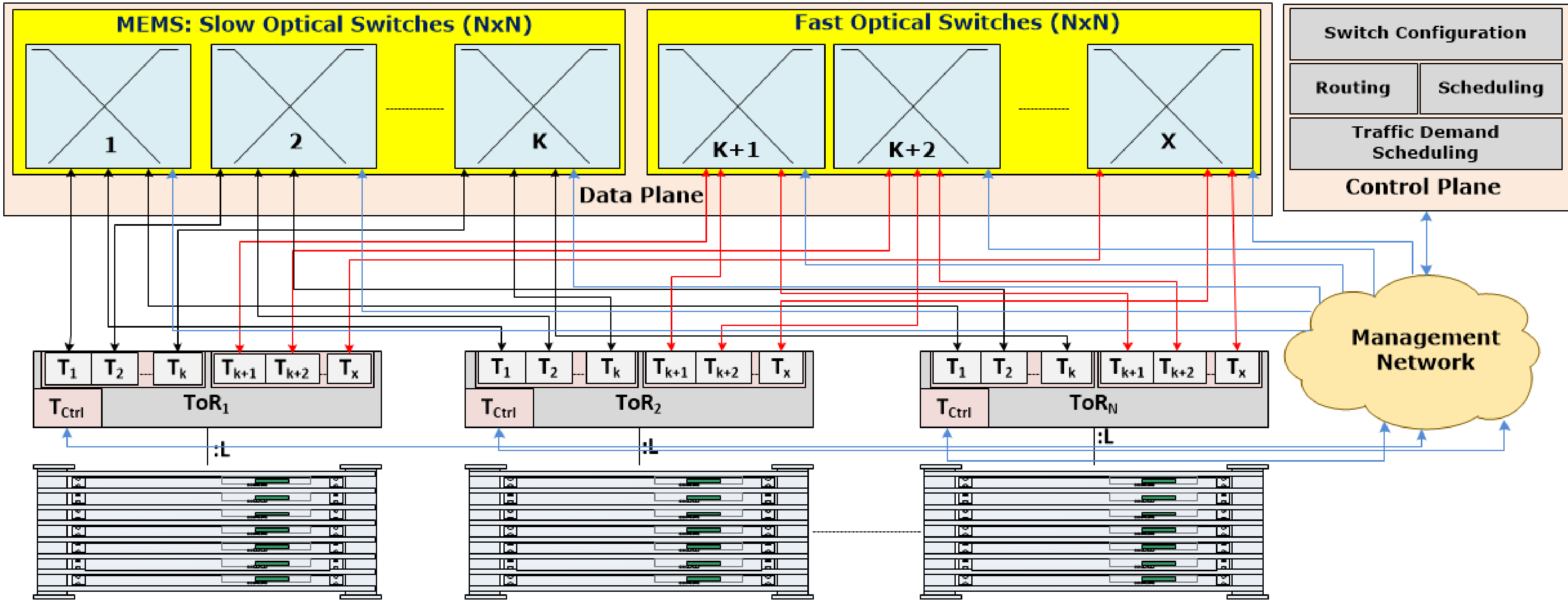
LIGHTNESS


Inter DC Connectivity
Intra-Inter DC Interface
OPS Switches
OCS MEMS Switches
ToR
Switches
Servers in Rack

Hybrid Optical Switching



HOSA


MEMS: Slow Optical Switches (NxN)
1
2
K
Fast Optical Switches (NxN)
K+1
K+2
X
Data Plane
Switch Configuration
Routing
Scheduling
Traffic Demand Scheduling
Control Plane
Management Network
T1 T2 Tk Tk+1 Tk+2 Tx
TCtrl
ToR1
ToR2
ToRN
:L

Comparison

Table 2: Comparison at a Glance

Architecture	Year	Switching Technique	Capacity Limitation	Scalability	Cost	Power Efficiency	Implementation Complexity	Prototype
Architectures based on MEMS								
Helios	2010	OCS + EPS	Transc.	Low	Low	Low	Low	●
HyPaC	2010	OCS + EPS	Transc.	Low	Low	Low	Low	●
OSA	2014	OCS + EPS	Transc.	Medium	Low	High	Medium	●
Reconfigurable	2012	OCS + EPS	Transc.	High	Low	High	Medium	●
HydRA	2015	OCS + EPS	Transc.	High	Low	High	Medium	●
Architectures based on SOAs								
OSMOSIS	2004	OPS	SOA	Low	High	Low	Medium	●
Data Vortex	2008	OPS	SOA	High	High	Medium	Medium	●
SW	2011	OCS	SOA	Low	High	Low	High	
STIA	2011	OCS	SOA	Low	High	Low	High	
Bidirectional	2011	OPS	SOA	High	High	Medium	High	
Architectures based on AWGRs								
DOS & LIONS	2010 & 2013	OPS	TWC	Low	High	High	High	●
TONAK-LION	2013	OPS	TWC	High	High	High	High	
Petabit	2010	OPS	TWC	High	High	Medium	High	
IRIS	2010	OPS	TWC	High	High	Medium	High	●
OFDM-based	2013	OCS	TL	Low	High	Medium	High	●
Architectures based on WSSs								
Mordia	2013	OCS	WSS	Low	Medium	Medium	High	●
WaveCube	2015	OCS	WSS	High	Medium	Medium	High	●
OPMDC	2015	OCS	WSS	High	Medium	Medium	High	●
Architectures based on Fast and Slow Optical Switches								
LIGHTNESS	2013	OCS + OPS	Transc. + SOA	High	Medium	High	High	●
HOS	2014	OCS + OPS + OBS	Transc. + SOA	High	Medium	High	High	
HOSA	2015	OBS	Transc. + SOA	High	Medium	High	High	

Conclusion & Future Work

- Every architecture has some pros and cons.
- The more feasible architecture would be the one which is:
  - Scalable
  - Power efficient
  - Cost effective
  - Low latency
  - High throughput.
- Software Defined Optical Networking (SDON)
  - Inter and Intra Data Centers Interconnect
- Elastic Optical Network
- Photonic Integration
  - Provide fast optical switching
  - Increase capacity

Thanks