

Future Datacenter Interfaces

Based on Existing and Emerging Technologies

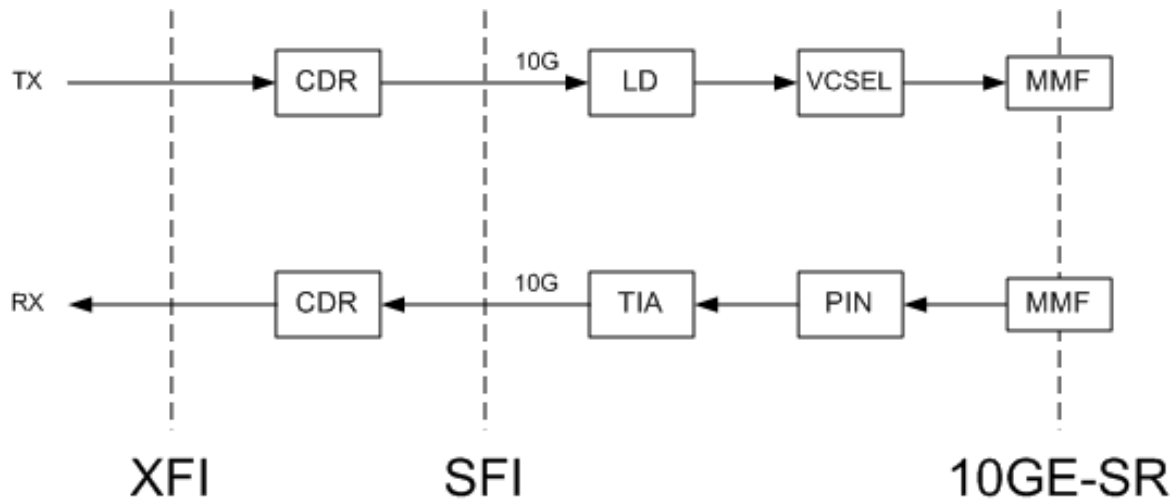
Summer Topicals
IEEE Photonics Society
Waikoloa, Hawaii
8-10 July 2013
Chris Cole



Outline

- 10G
- Multi-link 10G
- 40G
- Serial 40G
- 100G
- Higher Order Modulation
- 400G
- Board Mounted Optics
- Beyond 400G
- Summary

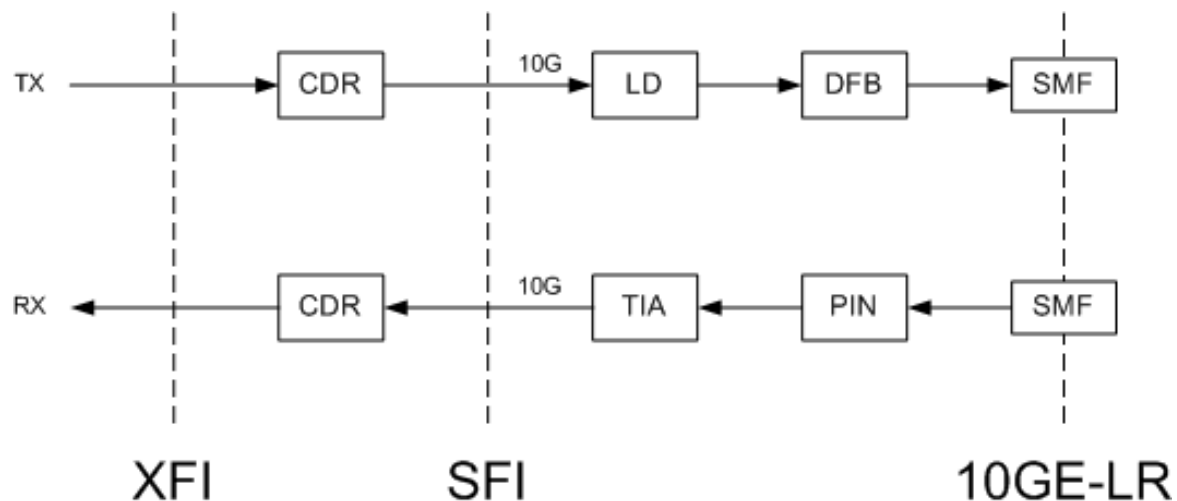
Datacenter 10G MMF: SR NRZ SFP+



Electric I/O		Optical I/O		
pin pair	Gb/s	fiber pair	λ	Gb/s
1	10	1	1	10
10		10		



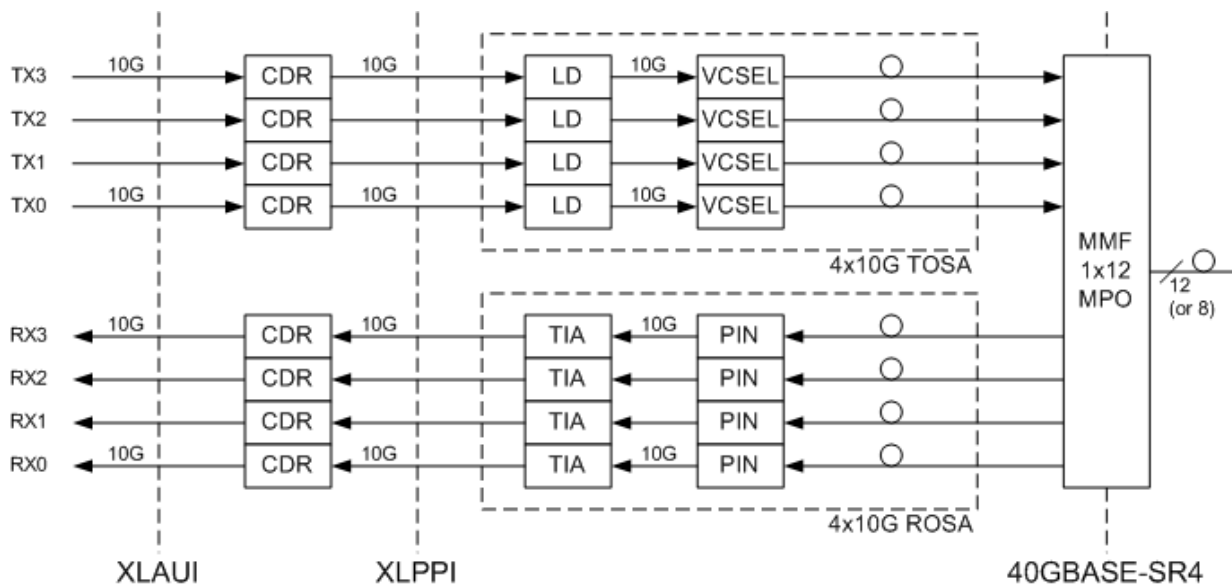
Datacenter 10G SMF: LR NRZ SFP+



Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
1	10	1	1	10
10		10		



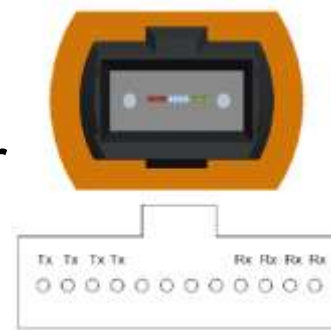
Datacenter 40G MMF: SR4 Parallel QSFP+



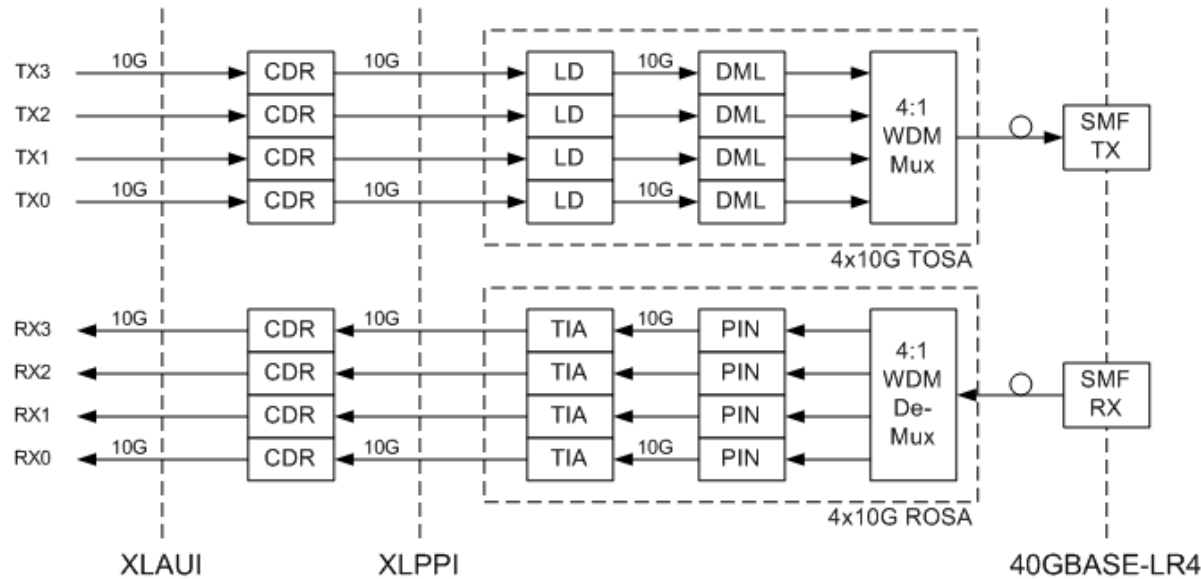
Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
	10		1	10
4		4		
40		40		

MPO parallel connector & MMF cable differs from 10GE-SR

MPO connector

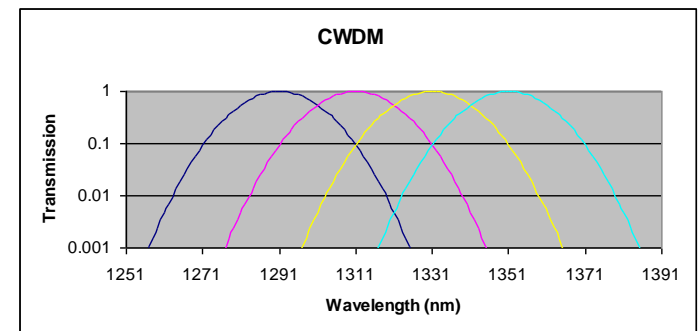


Datacenter 40G SMF: LR4 WDM QSFP+



Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
	10	1		10
4			4	
40		40		

duplex LC connector & SMF cable is same as 10GE-LR

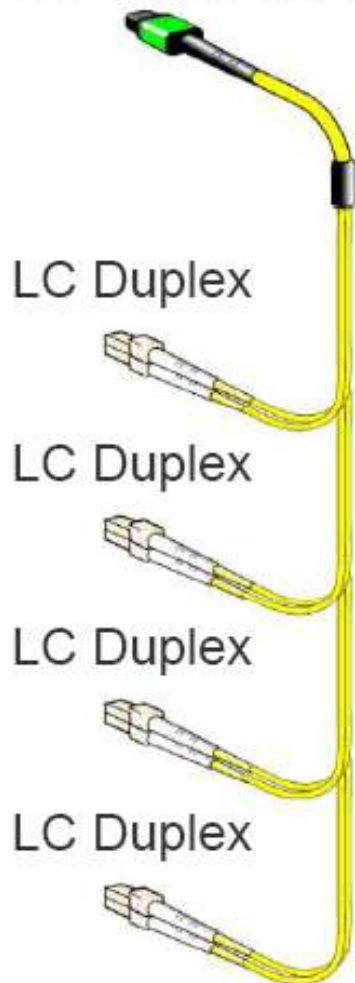


High Density 10G SMF: 4xLR Multi-link

Octopus break-out cable connects

QSFP+ \leftrightarrow 4x SFP+

1 x 12 (8 Used) MPO



Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
	10		1	10
4		4		
4x10		4x10		

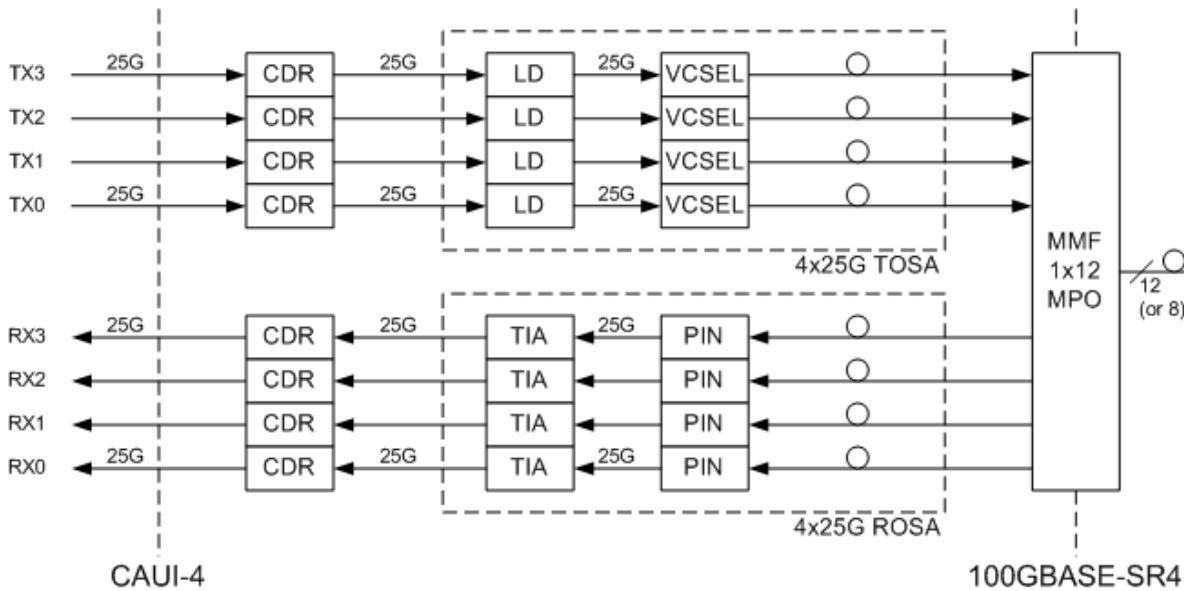
Same paradigm for 4xSR Parallel QSFP+

Multi-Link Example: 1UI Line Card Density

Form Factor	Electrical I/O	Rows	10GE Ports	40GE Ports	Power
SFP+	1x10G	Double	48	N.A.	~40W
QSFP+	4x10G	Double	160	40	~80W
CFP4 MLG	4x25G	Double	360	72	~120W

- Integer multiple density increase
- Thermal load increase
- Single channel failure requires bringing down multiple links to replace

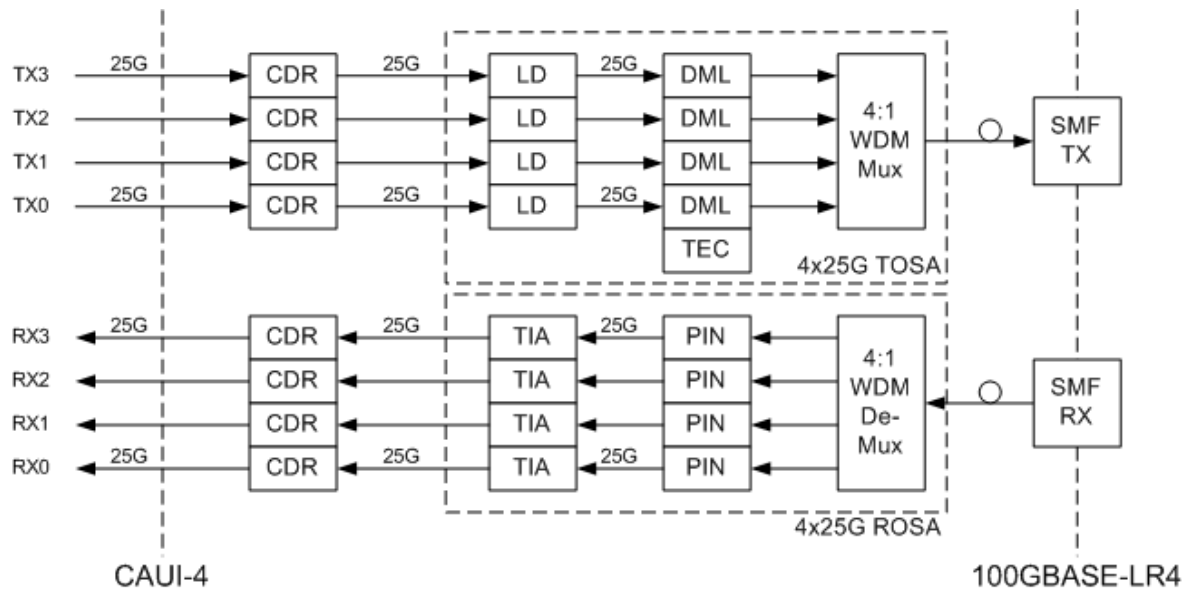
Datacenter 100G MMF: SR4 Parallel



Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
			1	
4		4		
	25			25
100		100		

CFP2 / CFP4 / QSFP28

Datacenter 100G SMF: LR4 WDM



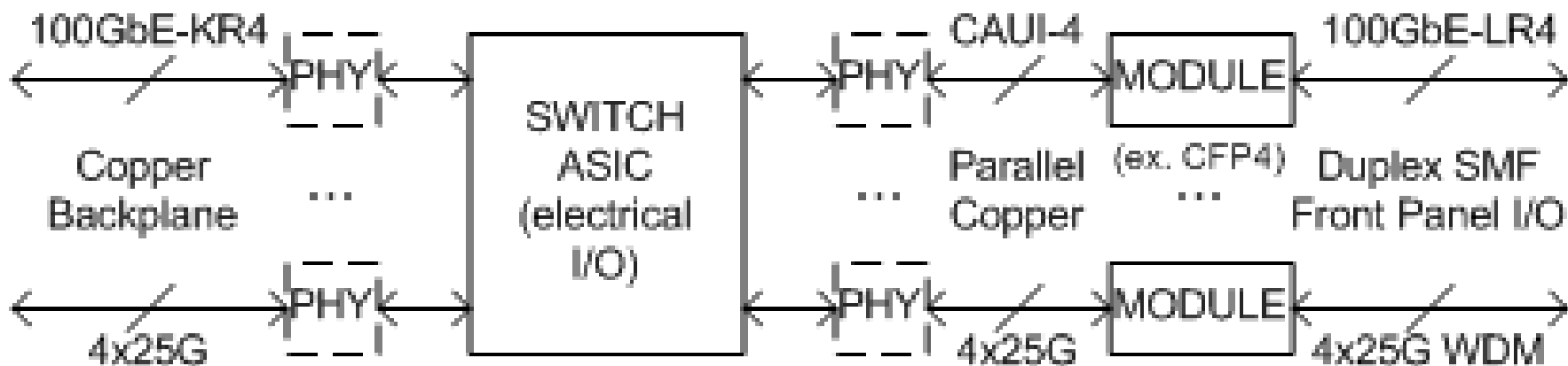
Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
		1		
4			4	
	25			25
100				100

CFP2 / CFP4 / QSFP28

Alternative to LR4:

CWDM eliminates TEC

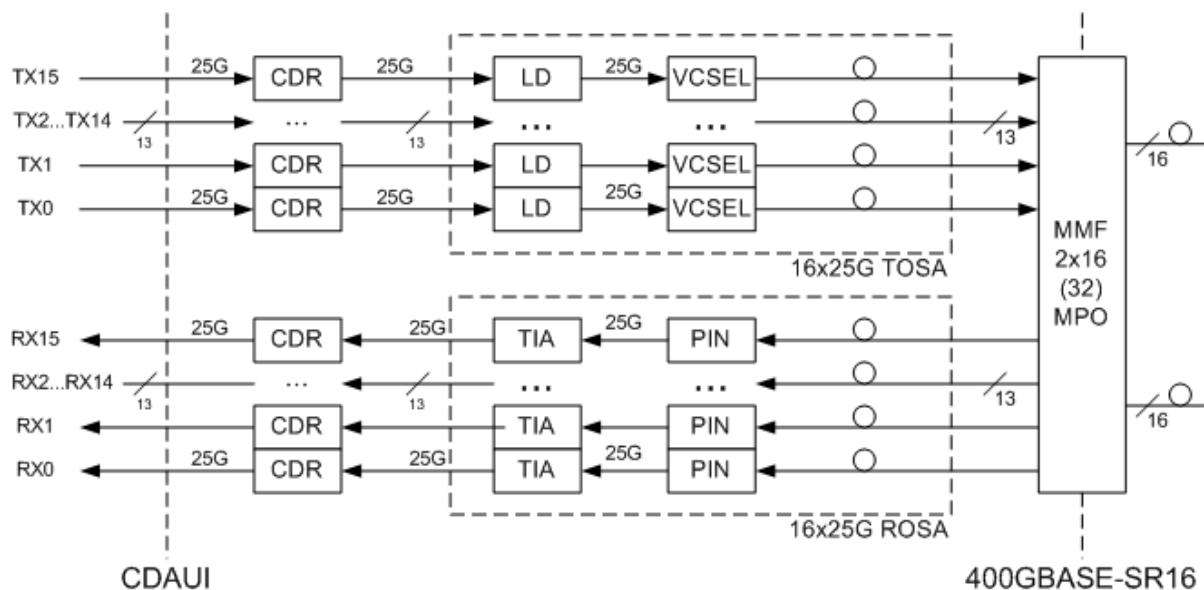
Datacenter 100G SMF: LR4 WDM



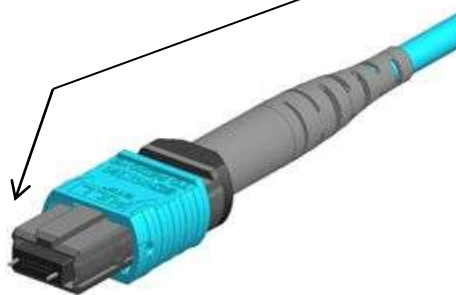
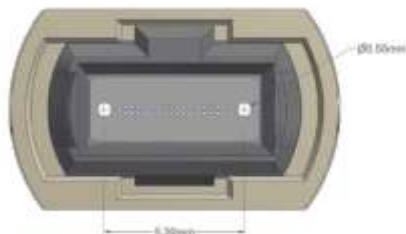
Lane Rate Increases

- 40Gb/s per lane (Serial)
 - 1x40G NRZ architecture (just like 10G NRZ SFP+)
 - AOC alternative to direct attach copper (DAC) for servers
 - 10Ms volume potential
- 50Gb/s per lane
 - 2x50G NRZ architecture for 100G
 - 8x50G NRZ architecture for 400G
 - 2x density improvement over 25G
 - 40-56Gb/s Electrical I/O Project started in the OIF
- 100Gb/s per lane
 - 1x100G Higher Order Modulation (HOM) architecture
 - 4x density improvement over 25G
 - Big question is when does it make sense

Gen 1 400G MMF: SR16



2 x 16 MMF MT ferrule



Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
			1	
	25			25
16		16		
400		400		

Gen 1 400G MMF: SR16

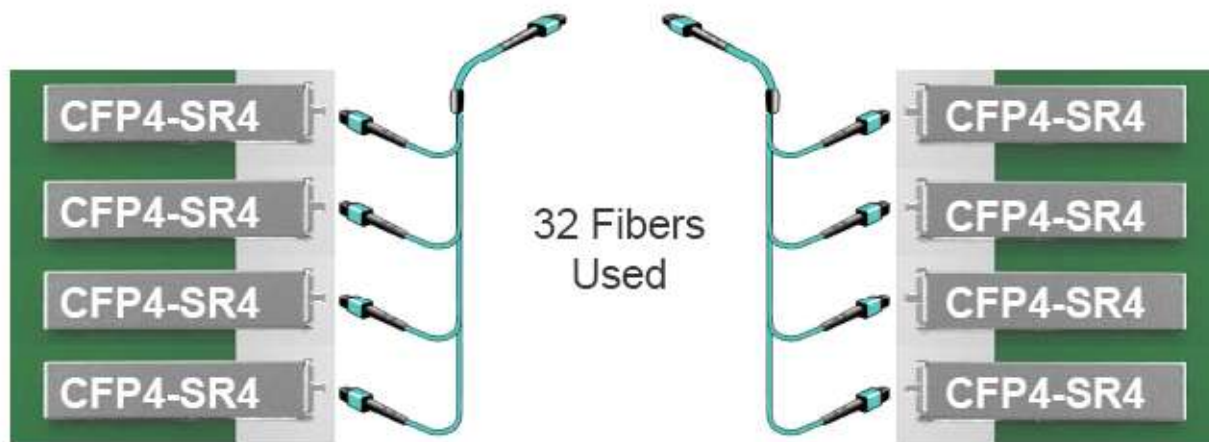
400GE over MMF by using the 100GBASE-SR4 PMD

100G Ethernet up to 100 m on OM4

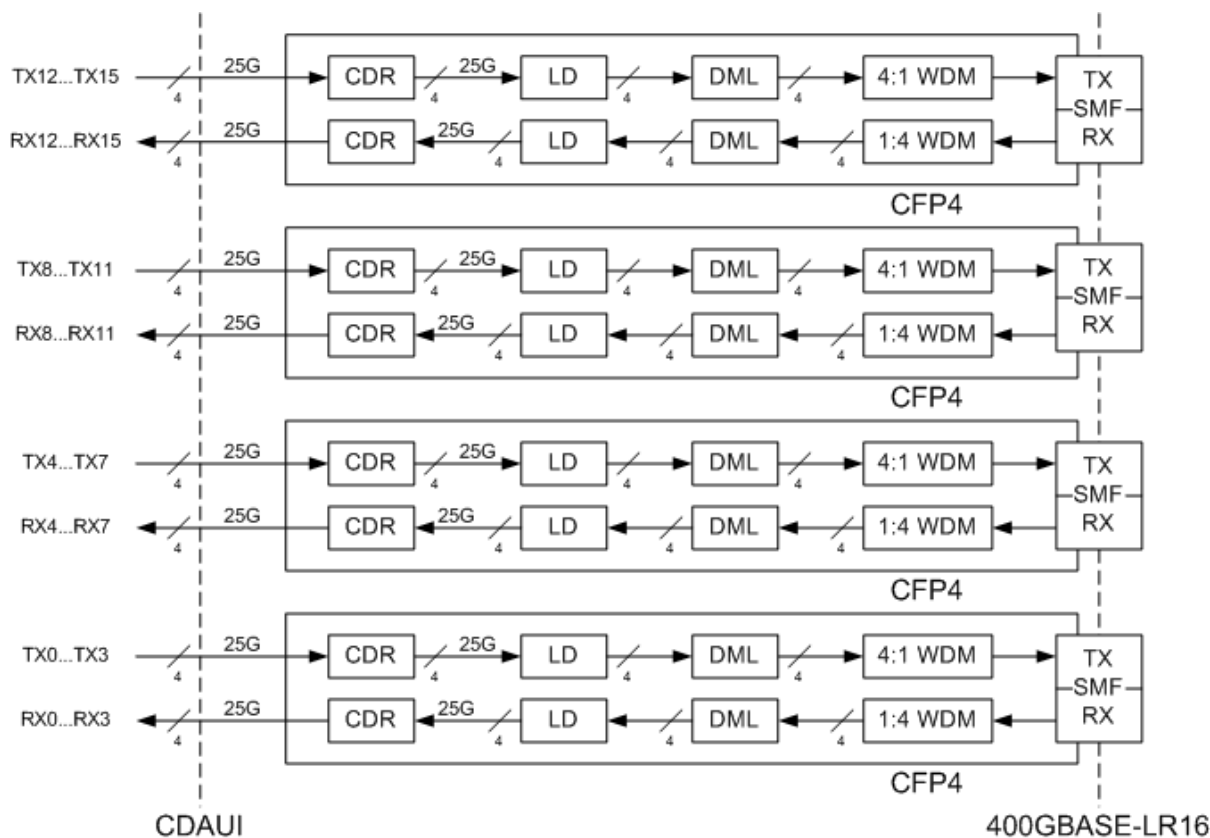


400G Ethernet up to 100 m on OM4

Parallel Multi-Mode Fiber Infrastructure



Gen1 400G SMF: LR4x4 4x CFP4



Electric I/O		Optical I/O		
pin pair	Gb/s	fiber pair	λ	Gb/s
		4	4	
	25			25
16				
400				400

Gen1 400G SMF: LR4x4 4x CFP4

100G Ethernet up to 10 km

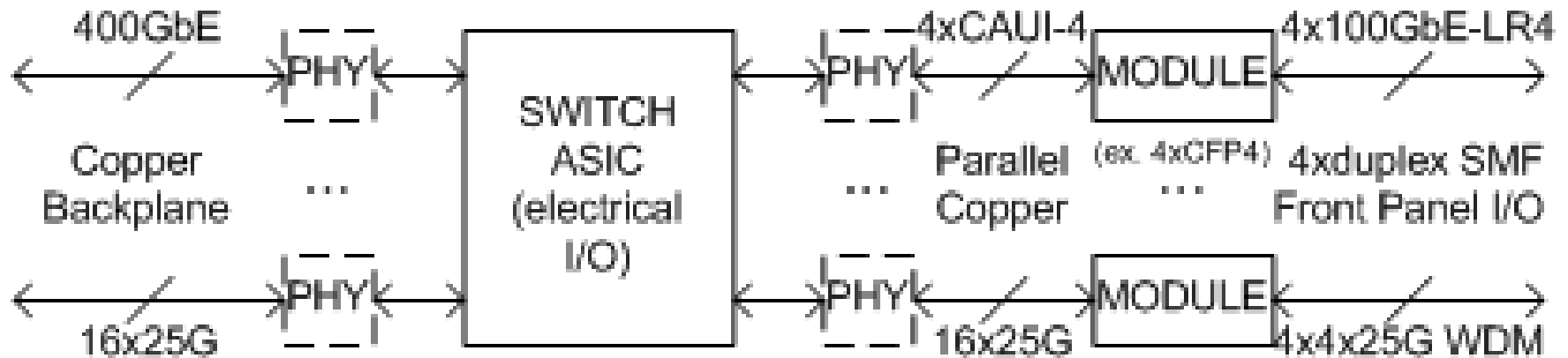


400G Ethernet up to 10 km

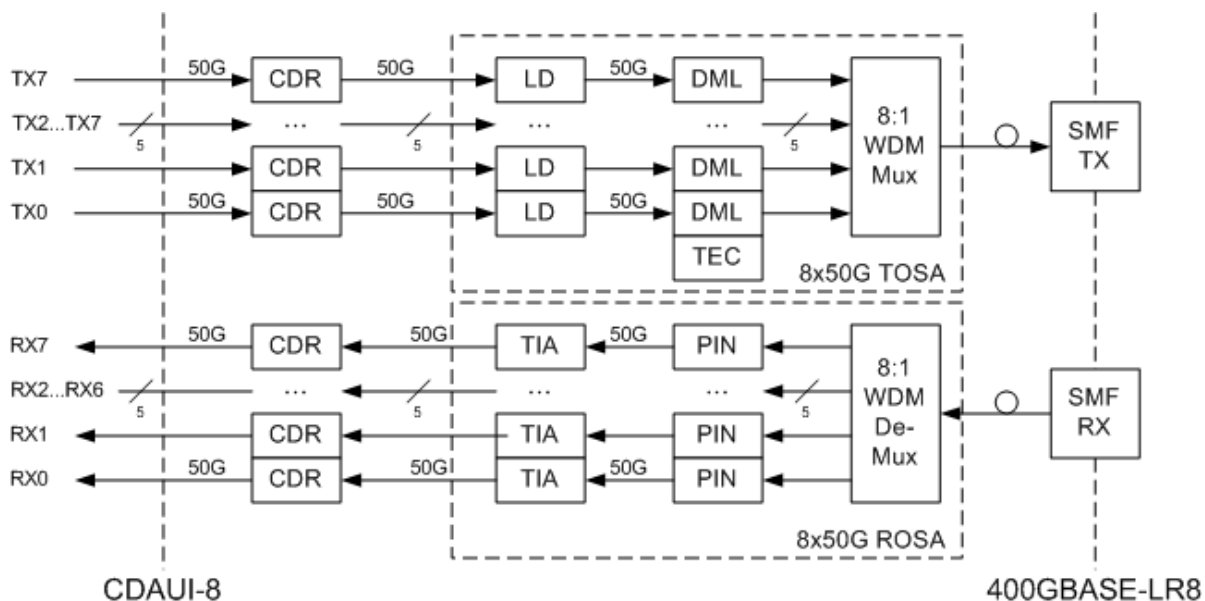
Parallel Single-Mode Fiber Infrastructure
(Actual distance limited by market adoption)



Gen1 400G SMF: LR4x4 4x CFP4



Gen2 400G SMF: LR8 CFP2



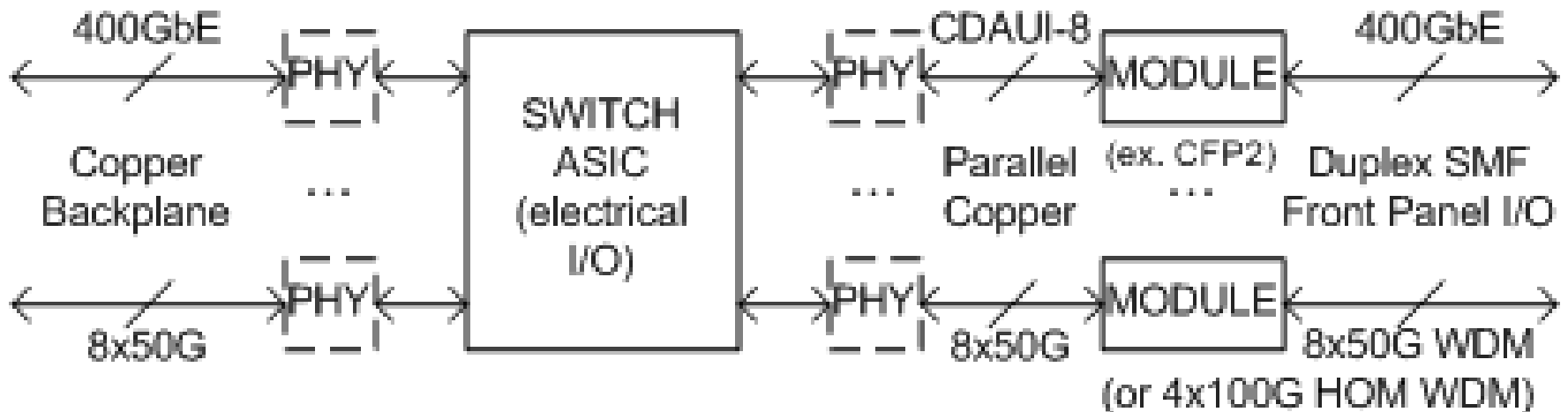
Electric I/O		Optical I/O		
pin pair	Gb /s	fiber pair	λ	Gb /s
		1		
8			8	
	50			50
400		400		

Alternative to LR8:

LR4 WDM with 4 λ s

Higher Order Modulation

Gen2 400G SMF: LR8 CFP2

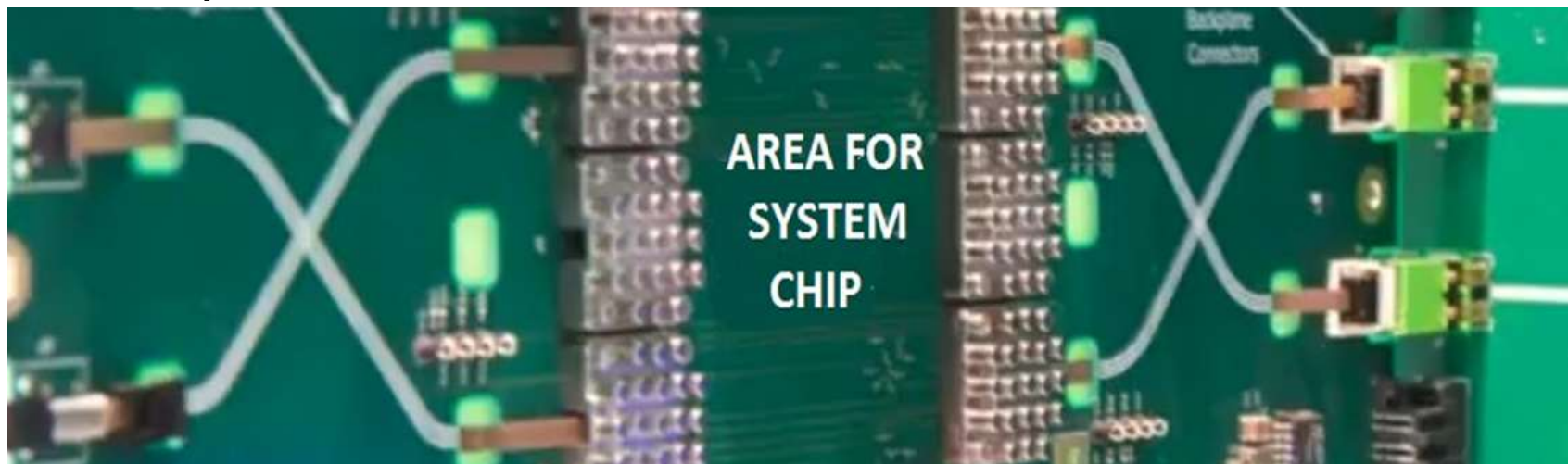


Alternative to LR8:
LR4 WDM with 4λs
Higher Order Modulation

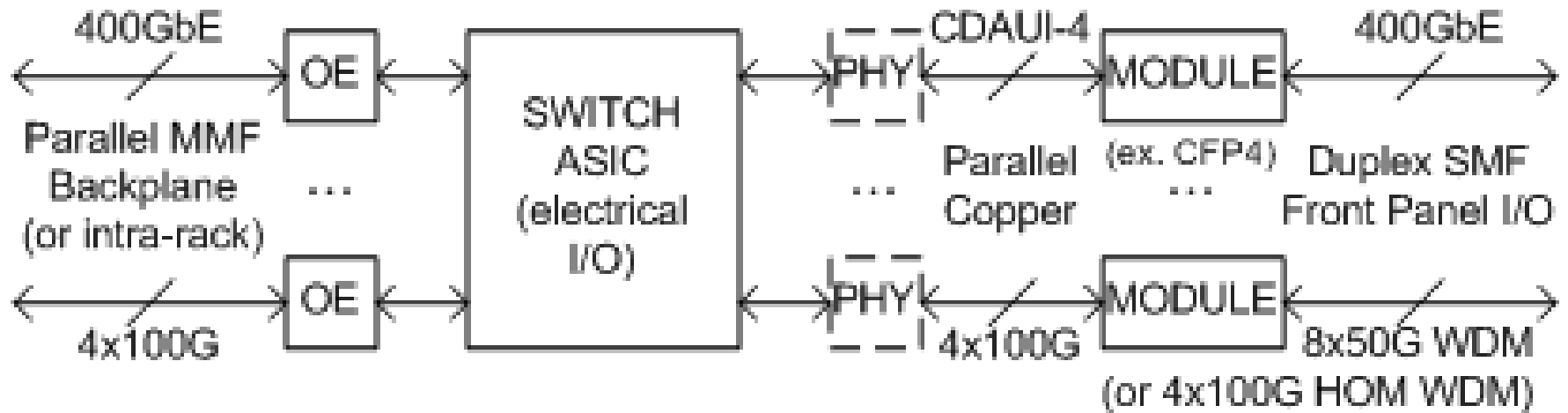


Board Mounted Optics

- Extensively used today for inter-rack I/O
 - POP4 (4x2.5G transceiver)
 - SNAP12 (12x2.5G, 12x10G transmit/receive pair)
- Advantages:
 - Minimizes S.I. due to short copper trace length
 - Distributes heat load
- Example:



Gen3 400G SMF: LR8 (or HOM LR4) CFP4



Will Board Mounted Optics Take Over?

POTENTIAL LINE CARD IMPLEMENTATIONS FOR INCREASED BW



Courtesy:
Nick Ilyadis

- Pluggable module with VSR/CAUI-4 interface may require 2 retimers



- SiP/optical engine connected to switch through PCB may require retiming



- SiP/optical Engine mounted on the package does not require retiming



- SiP connected to switch through TSV does not require retiming



Why Pluggable Module?

- The good (il buono)
 - multiple applications supported
 - pay as you go
 - confined, replaceable failures
 - common market
 - specialized R&D & production
- The bad (il cattivo)
 - increased component count
 - SI complicated by I/O connector
 - power increased by I/O ICs
 - placement limited to the front
- The ugly (il brutto)
 - poor thermal interface
 - heat localized at host front



Why Active Optical Cable (AOC)?

- The good (il buono)
 - lowest cost optical interconnect
 - no optical connectors
 - no interop specification
 - least if any optical testing
 - used like copper cables
- The bad (il cattivo)
 - cost advantage goes away with bulkhead connections
 - stocking of all needed reaches
 - S.I. same as pluggable modules
- The ugly (il brutto)
 - uncontrolled thermals
 - cooling features not practical



Optics Types Comparisons

- Existing optics types and reaches:
 - Submarine: 10,000km (5,000x)
 - Long Haul: 2,000km (1,000x)
 - Datacenter: 2km (1x)
- Why are existing optics types different?
 - Longer reaches require more complex/expensive optics
 - For shorter reaches this is wasted performance/cost
- Emerging optics types and reaches
 - Datacenter: 2,000m (10,000x)
 - Inter-Rack: 20m (100x)
 - Backplane: 1m (5x)
 - PCB: 0.2m (1x)
- Why will emerging optics types be different?
 - Same as above; Swiss Army knife optics are wasteful

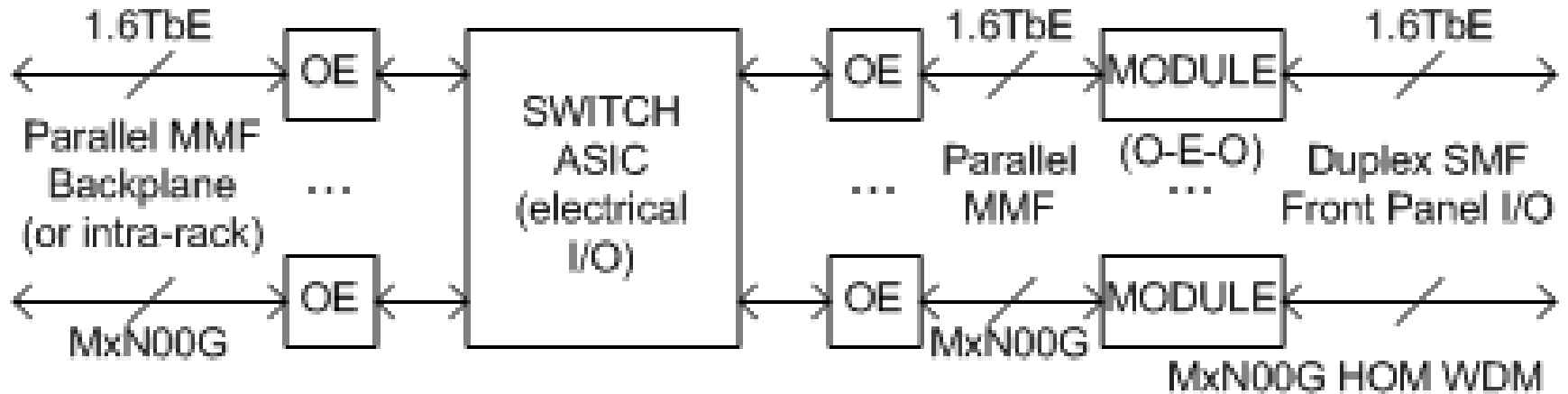
Emerging Optics Types Characteristics

- Datacenter:
 - Duplex SMF
 - WDM
 - Higher Order Modulation
 - FEC
- PCB / Backplane / Inter-rack
 - Short, direct electrical connection to switch I/O
 - High lane count Parallel MMF (or SMF)
 - High Baud NRZ
 - No FEC
- How to interconnect the two optics types?
 - Same as today between different existing optics types
 - Optical \Leftrightarrow Electrical \Leftrightarrow Optical (O-E-O) conversion
 - Pluggable module is the ideal platform

What's After 400G?

- 1Tb/s Ethernet
 - Has been extensively discussed
 - Vestige of 10x historical Ethernet speed jumps
 - Will require huge R&D investment
 - 2.5x speed increase from 400G is not compelling
- 1.6Tb/s Ethernet
 - 4x speed increase reasonable return on R&D \$
 - 4x is more likely for future speed increases
 - Similar to historical 4x Transport speed jumps
 - Gen1 can use 4x 400G architecture

1.6T Switch Example



Ex. 1: 16x100G

Ex. 2: 8x200G

Ex. 3: 4x400G

Summary

- Today's interconnect
 - Pluggable 10/40/100G optics modules
 - Board mounted optics
 - Direct attach copper (cable)
 - PCB copper traces
- Future interconnect
 - Pluggable multi-link 10/40/100G modules
 - Pluggable 400G & 1.6T modules
 - Active optical cables
 - Board mounted optics
 - Short PCB copper traces