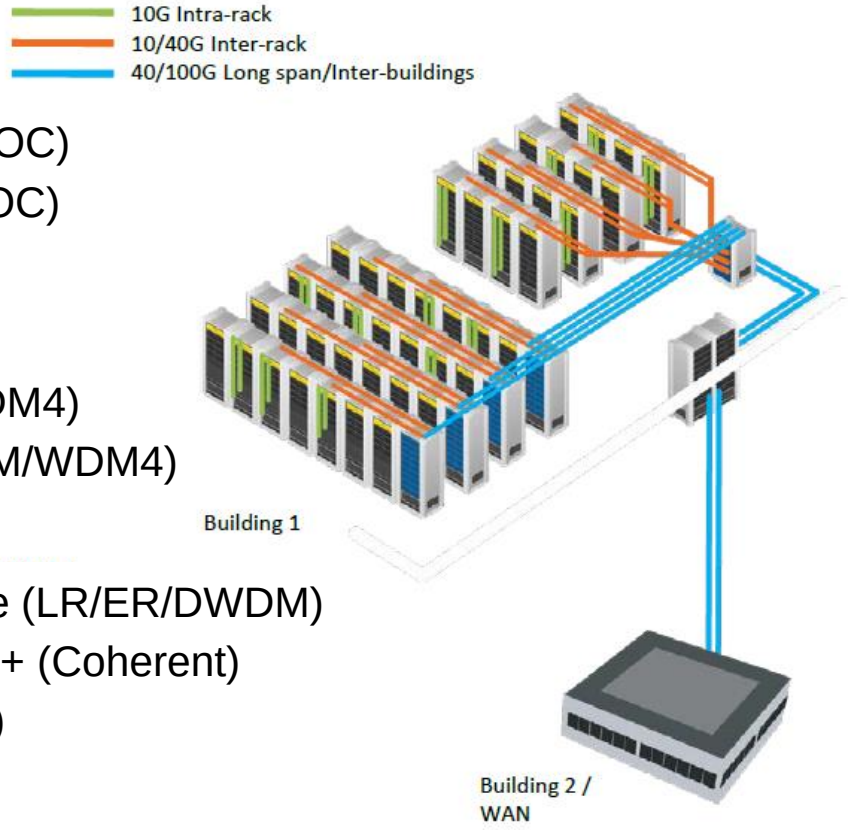
Inphi
Think fast.

100Gbit/s, 120km, PAM 4 Based Switch to Switch, Layer 2 Silicon Photonics based Optical Interconnects for Datacenters

Radhakrishnan Nagarajan, Sudeep Bhoja and Tom Issenhuth*
** Microsoft Corp, Redmond, WA*

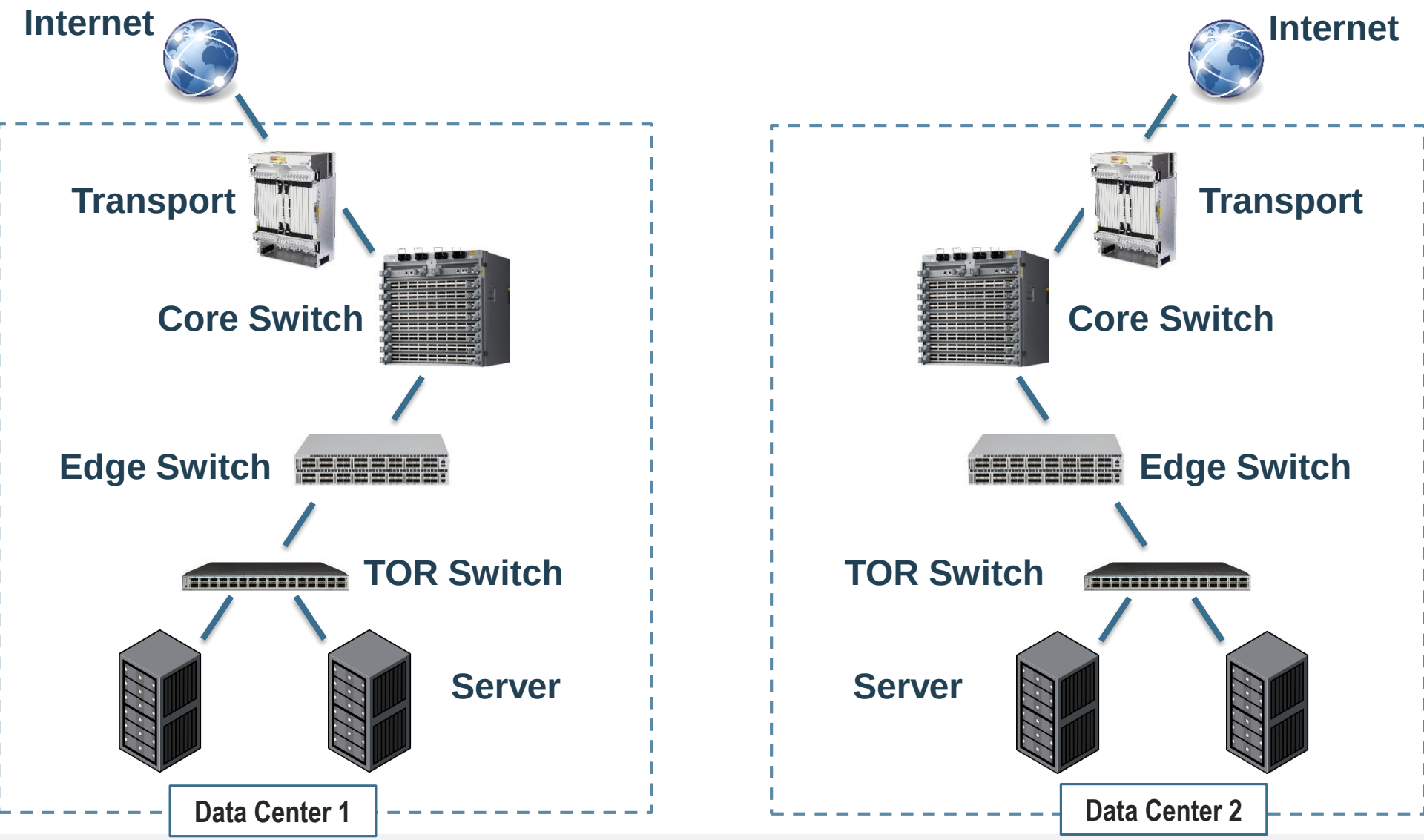
Data Center Connectivity Trends

- Within the Rack (Server to TOR)
  - 10GbE Deployed at Scale (DAC)
  - 25GbE Transition happening now (DAC/AOC)
  - 50GbE forecasted to start in 18/19 (SR/AOC)
- Between Racks
  - 40GbE Deployed at Scale (SR4/PSM)
  - 100GbE Transition in 2016 (SR4/PSM/WDM4)
  - Mix of 200GbE and 400GbE in 17/28 (PSM/WDM4)
- Inter Data Center
  - 10GbE/40GbE/100GbE Deployed at Scale (LR/ER/DWDM)
  - 100GbE/OTN Deployed at Scale for 80km+ (Coherent)
  - 200GbE Starting in 2017 (80km to 600km)
  - 400GbE Standardized now (DR8/FR8)

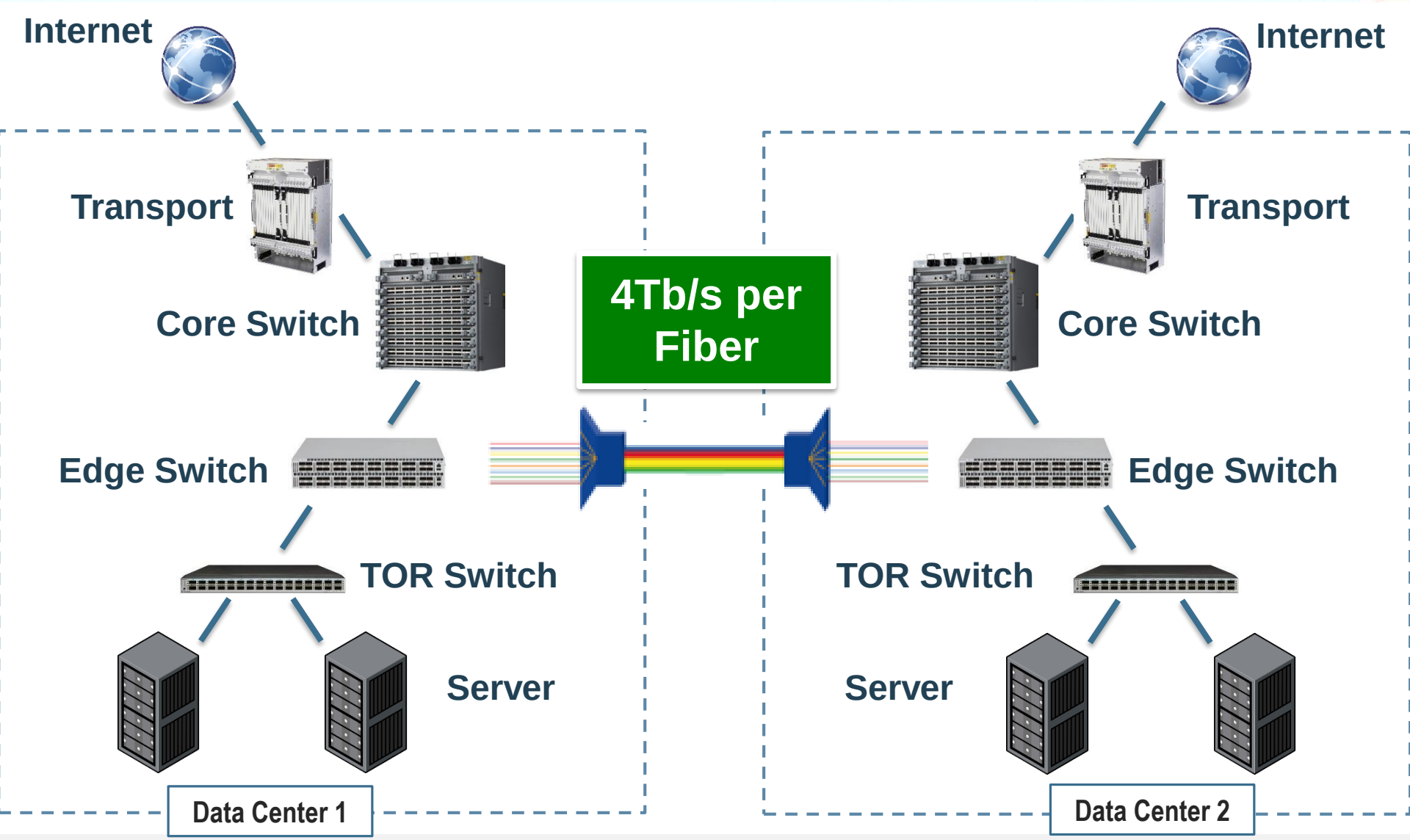


Increased Complexity, Speed and Volumes Driving Application Opportunities for Silicon Photonics

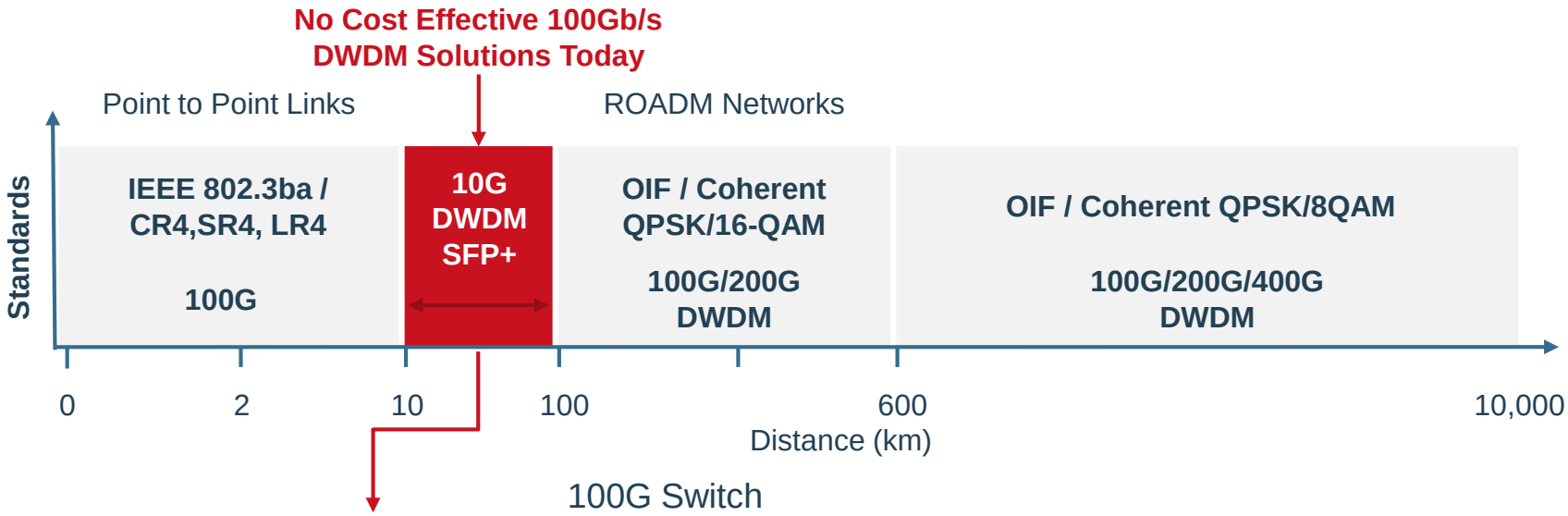
Virtual Datacenter Architecture: Latency Limited

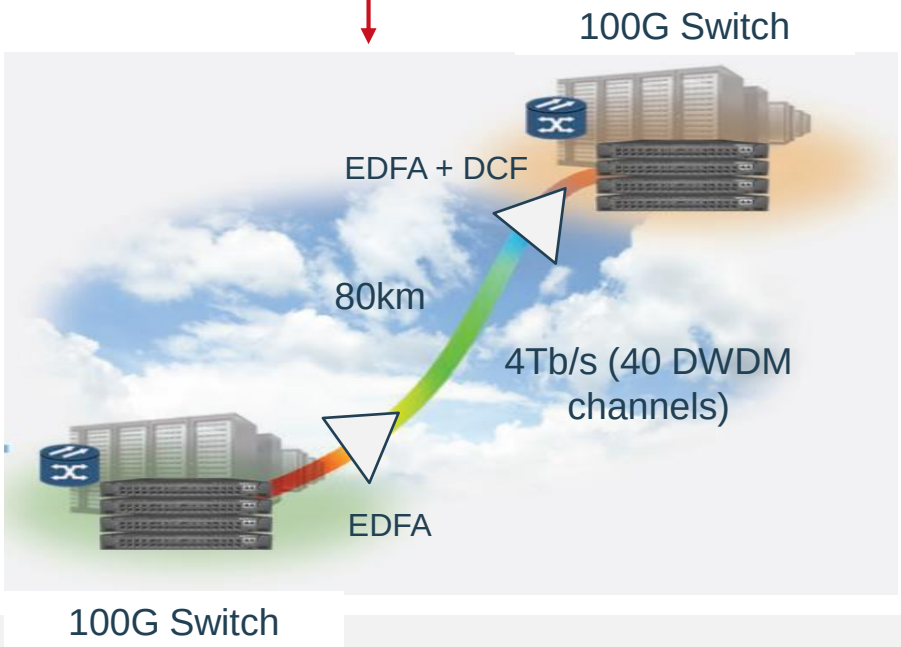


Virtual Datacenter Architecture: Latency Elimination



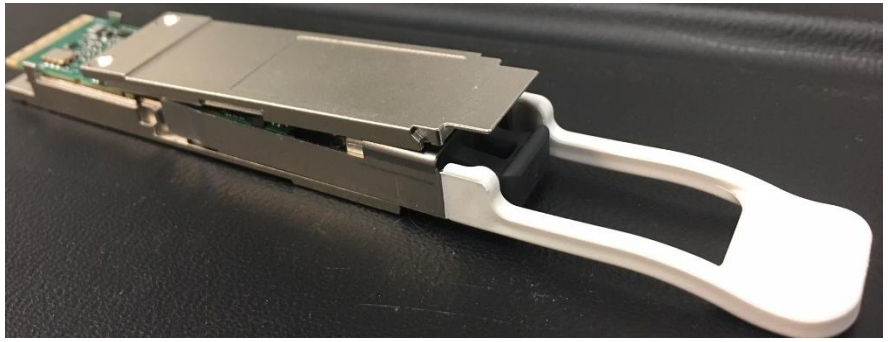
Metro Datacenter Interconnect (DCI) Gap



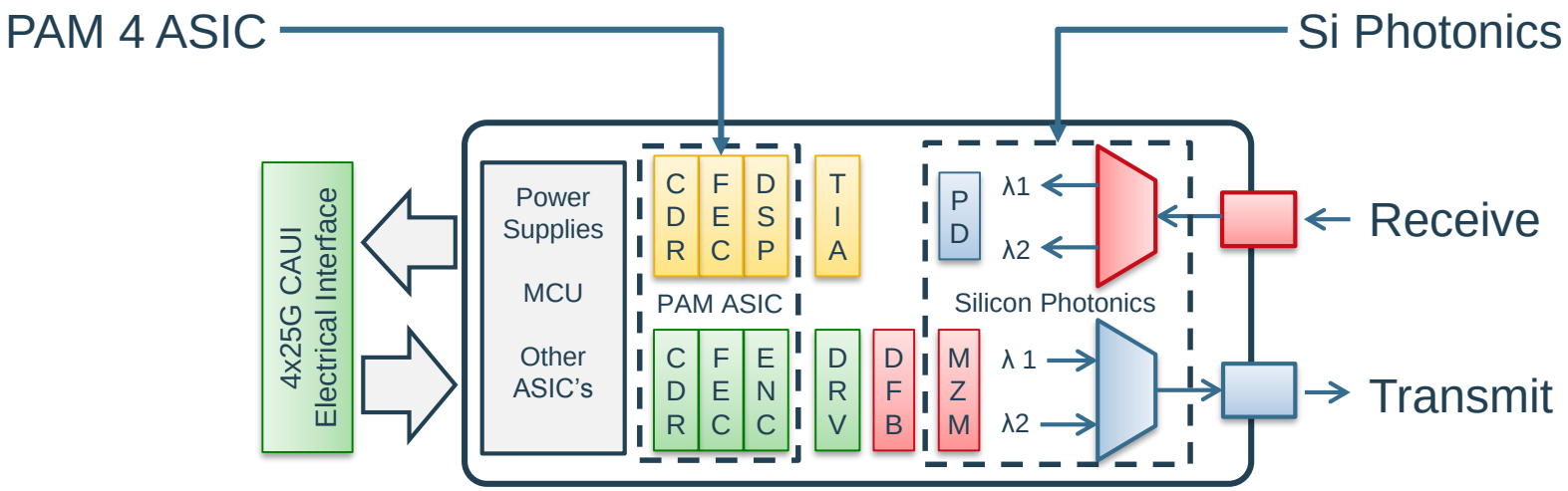


- Inter-datacenter links are being upgraded to 100G.
- Intra-datacenter links which enable large distributed architectures are the bottleneck.
- 100G DWDM QSFP-28: Cost effective approach to resolve the DCI bottlenecks

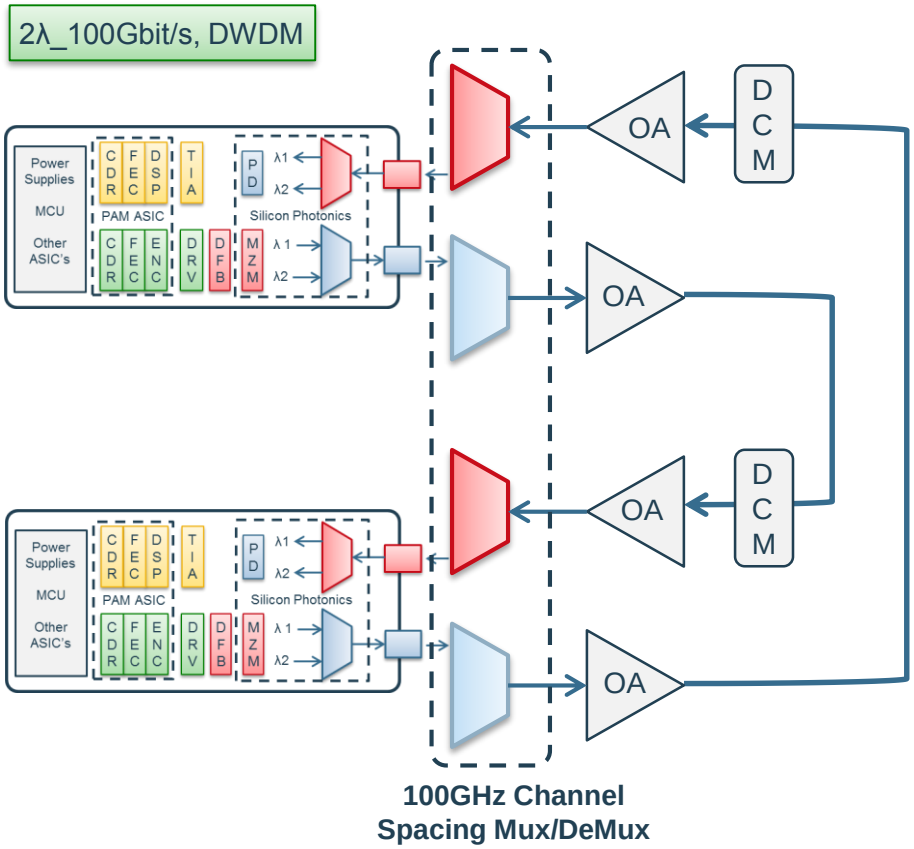
100G DWDM QSFP-28 Module



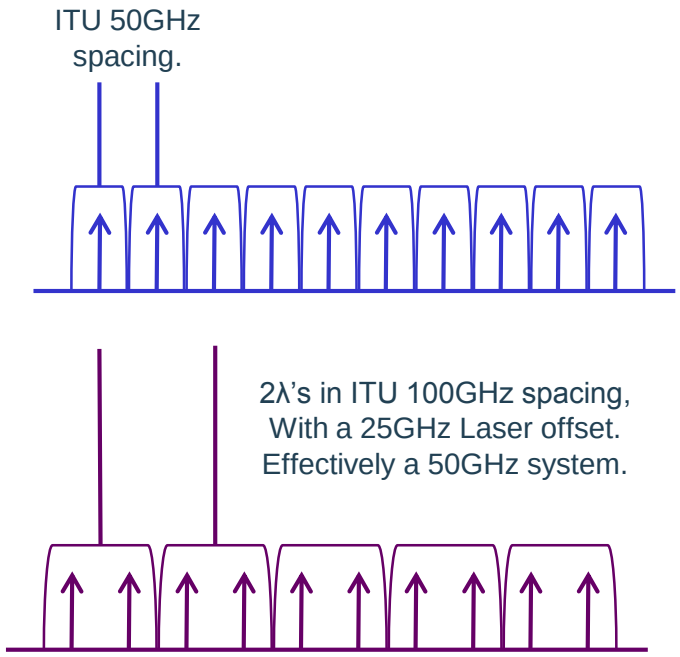
- Compatible with the QSFP28 MSA as described in SFF-8665
- Standard CAUI4 electrical interface
- Typical power consumption < 4.5W
- Electrical Input: 4 x 25.78125Gbit/s NRZ
- Optical Output: 2 x 28.125 Gbaud PAM4



Fat Pipes Between DC's: DWDM QSFP 28



**Exact locations of OA's and DCM's in the link are subject to OSNR considerations.*

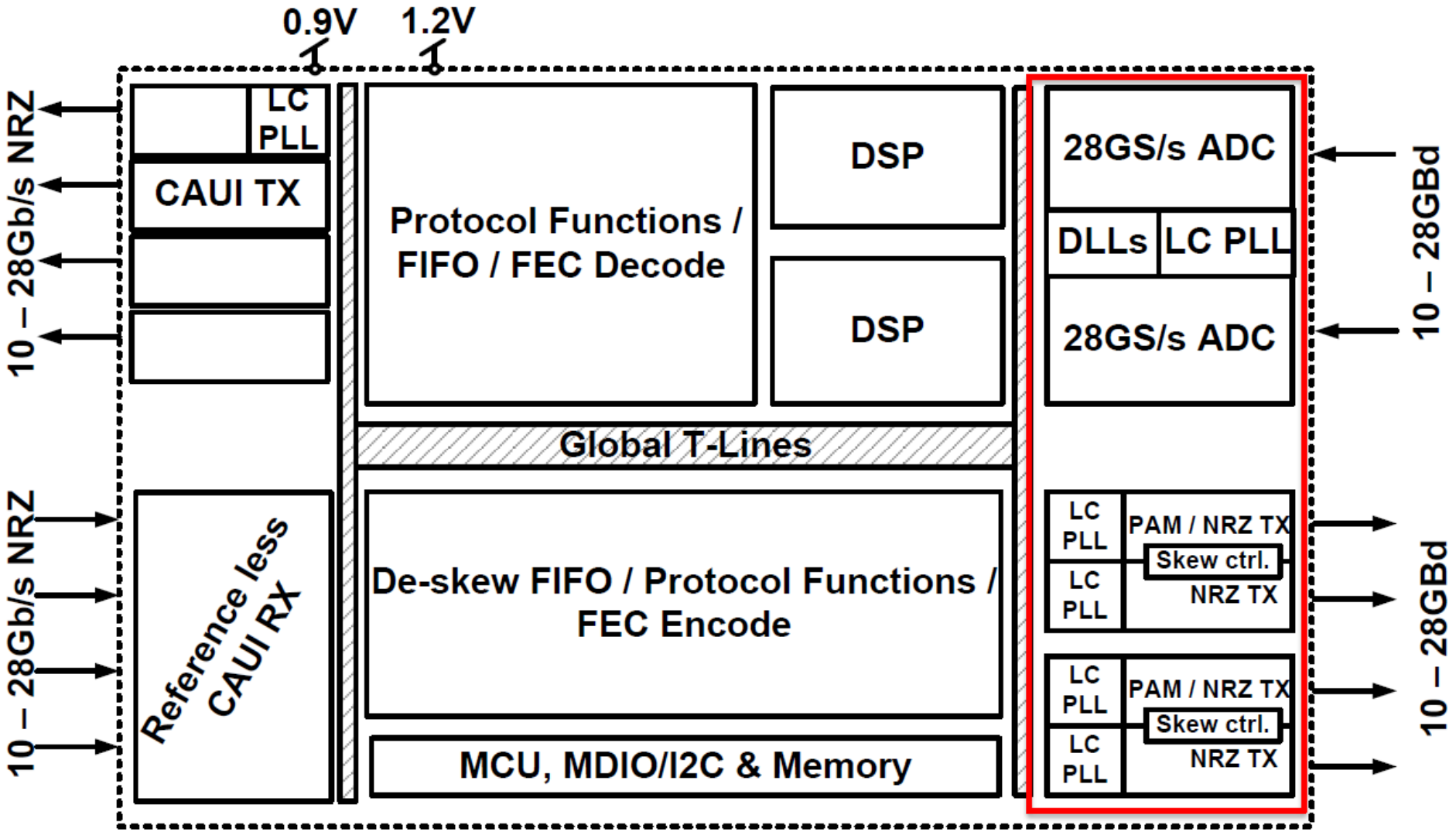


OA: Optical Amplifier

DCM: Dispersion Compensation Module

EDFA: Erbium Doped Fiber Amplifier

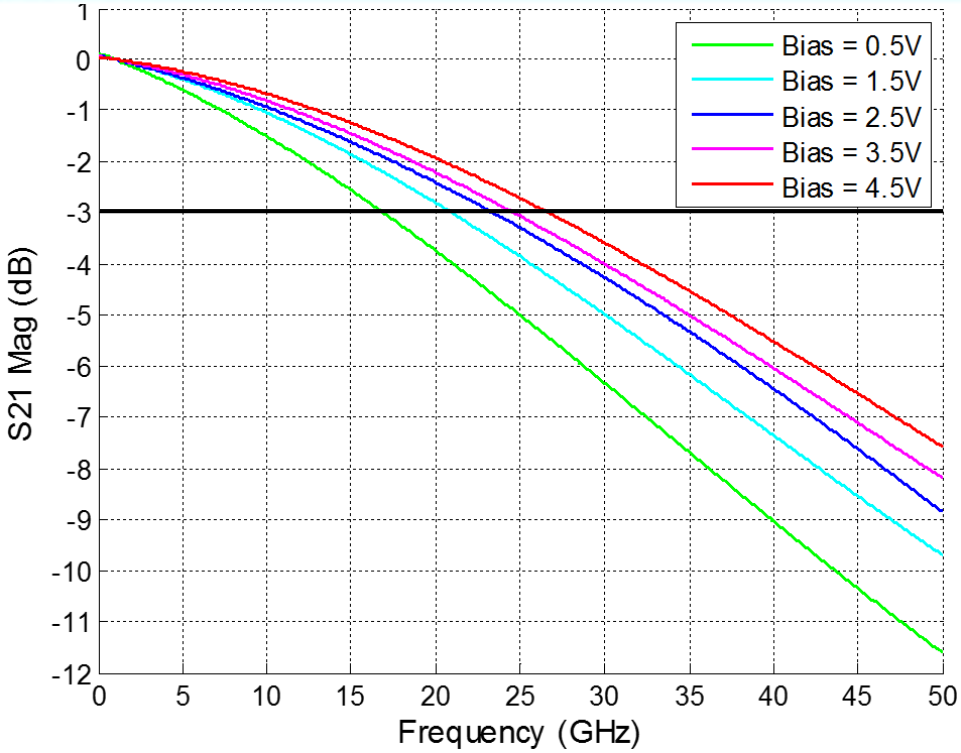
PAM 4 ASIC Block Diagram

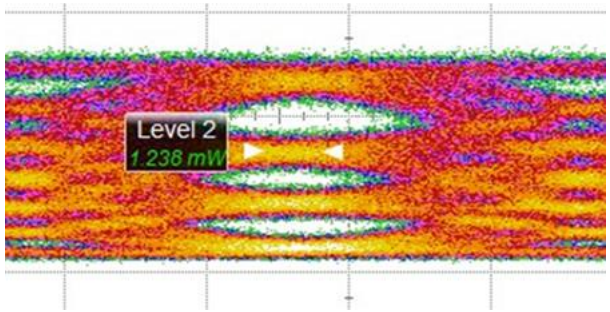


PAM 4 ASIC: Architecture

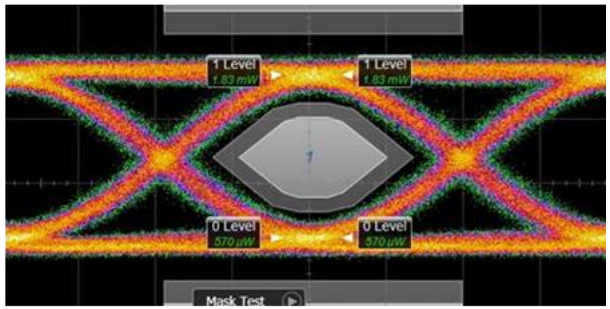
- CML Driver with CMOS backend
- Enables wide swing range and low power
- ½ rate TX Clocking and ¼ Rate RX Sampling
- 7-bit ADC-DSP based receiver with SAR core
- The clock path is CMOS based with regulators providing the required power rejection
- The data path are under independent regulator domains for proper isolation
- Multi-Tap FFE / DFE and Calibration in the DSP
- Reference-less, clock recovered from CAUI RX.

Silicon Photonics: Mach Zehnder Modulator



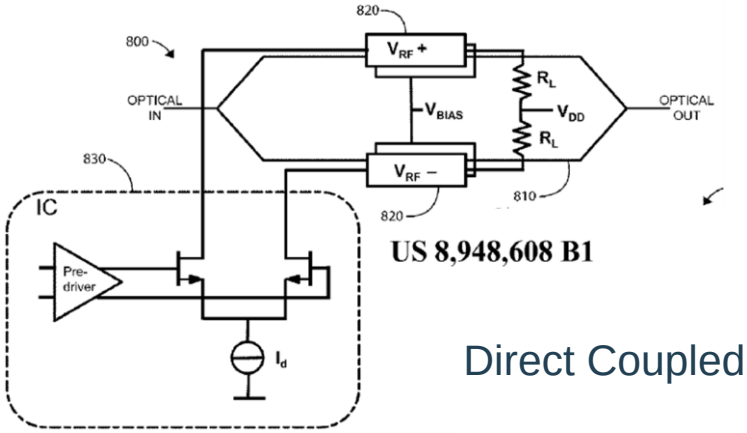


Optical Eye
PAM 4



Optical Eye
NRZ

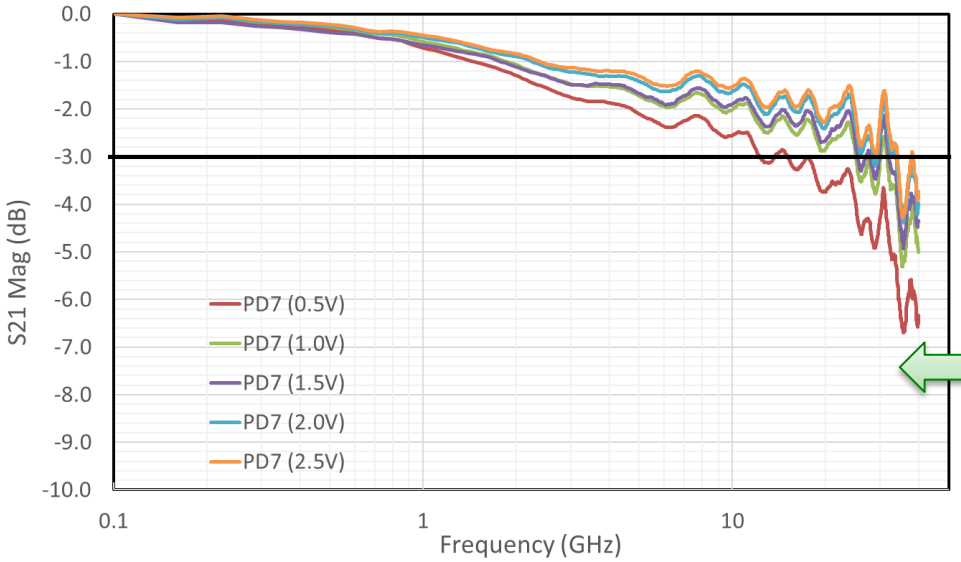
- Traveling Wave Mach Zehnder Modulator
- Optical-Electrical S21 Data



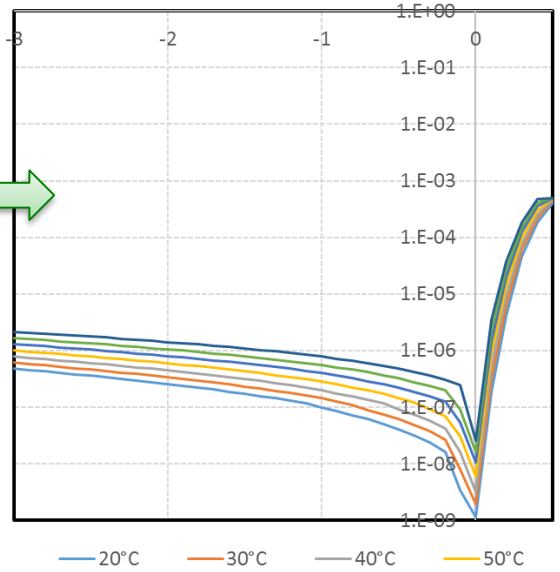
US 8,948,608 B1

Direct Coupled Driver

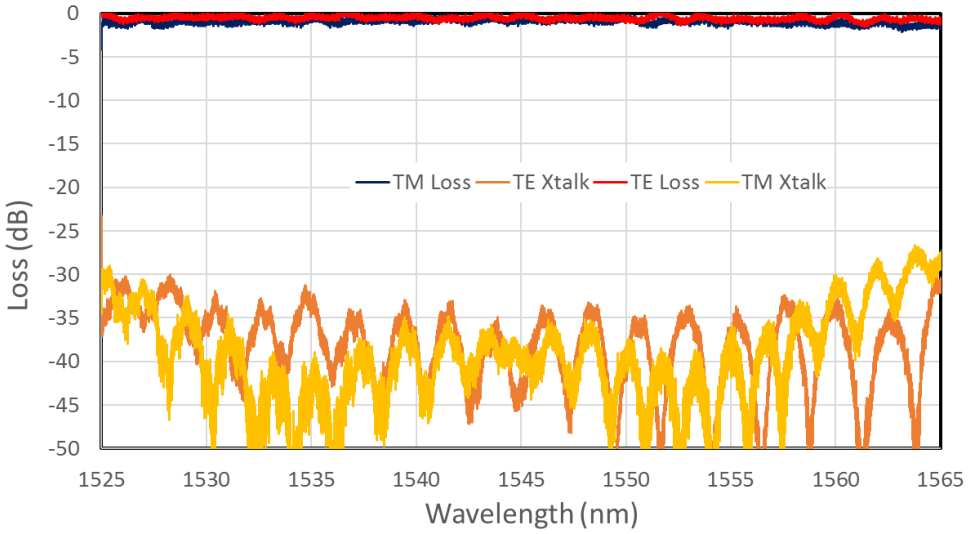
Silicon Photonics: Receiver Path



- High Speed Ge Photodetector
- Optical-Electrical S21 Data

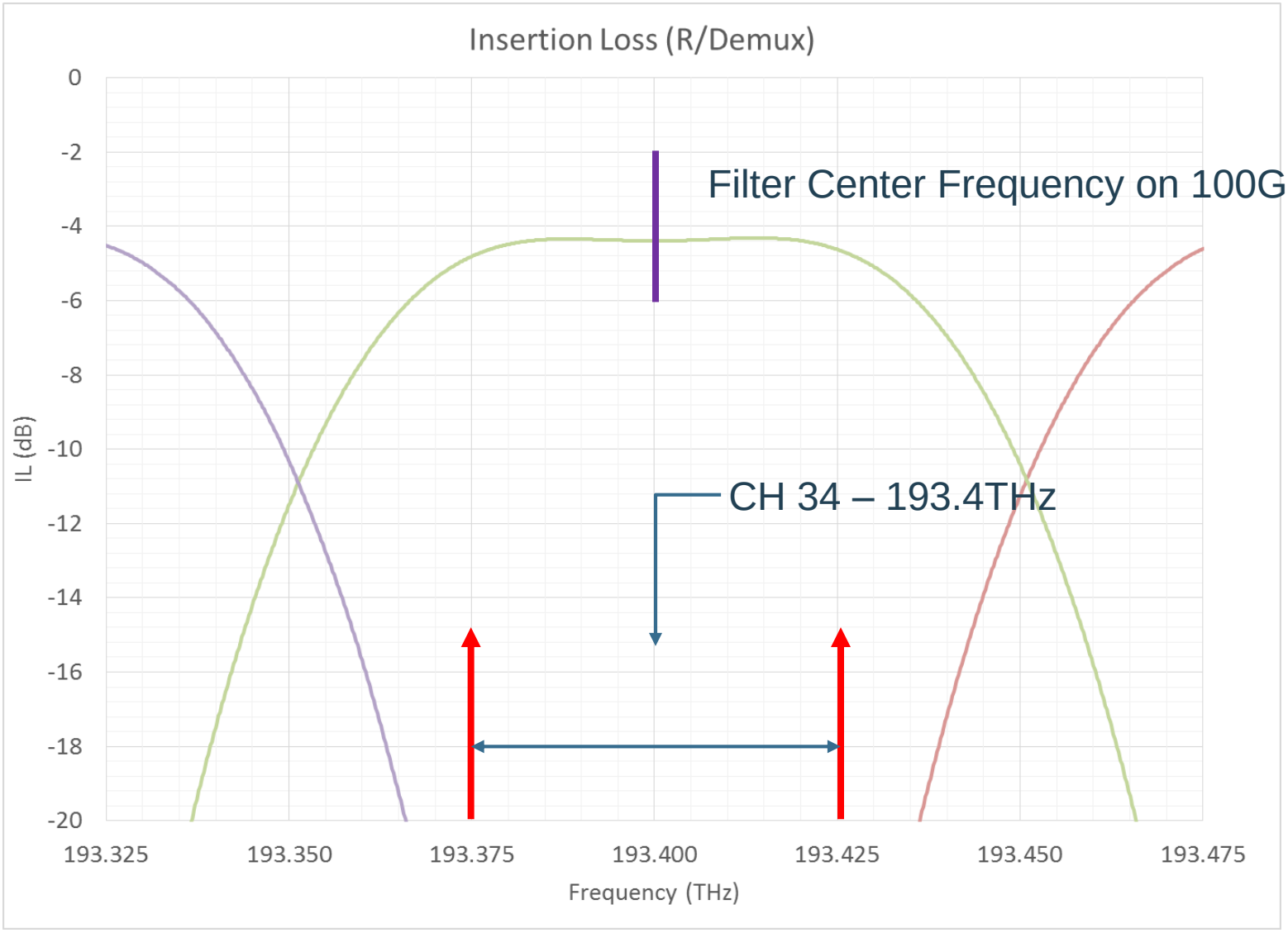


PD Dark Current

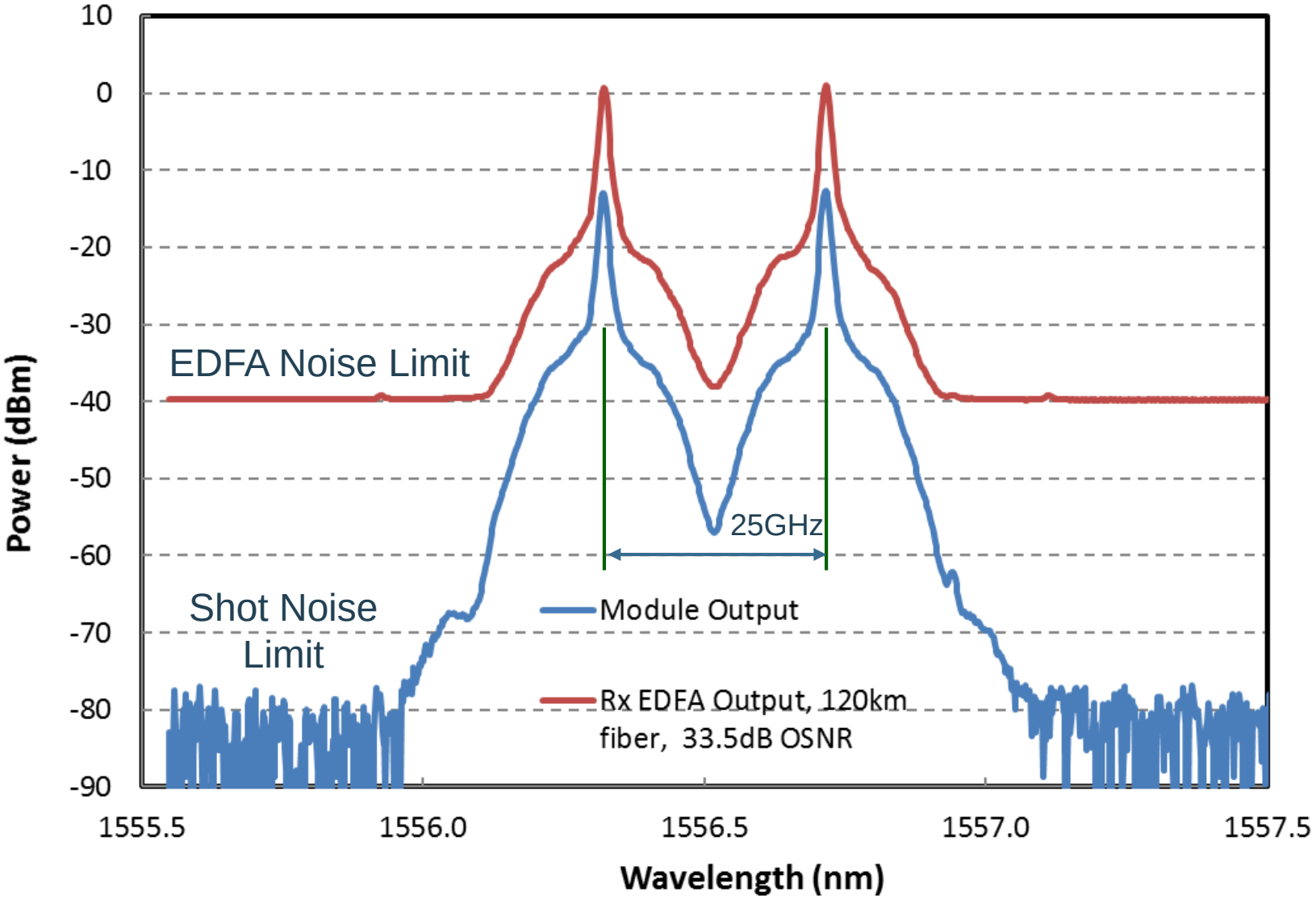


- Polarization Beam Splitter
- Insertion Loss and Crosstalk

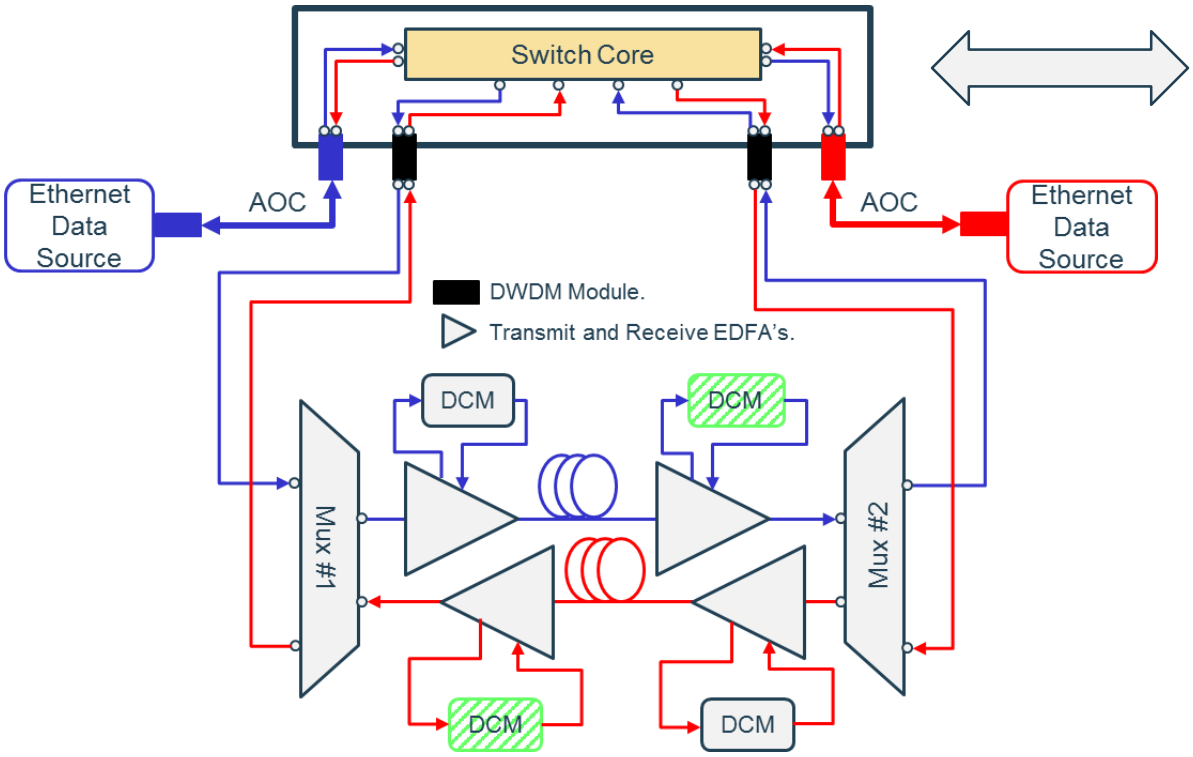
Commercial 100GHz Multiplexer: Flat Top Gaussian

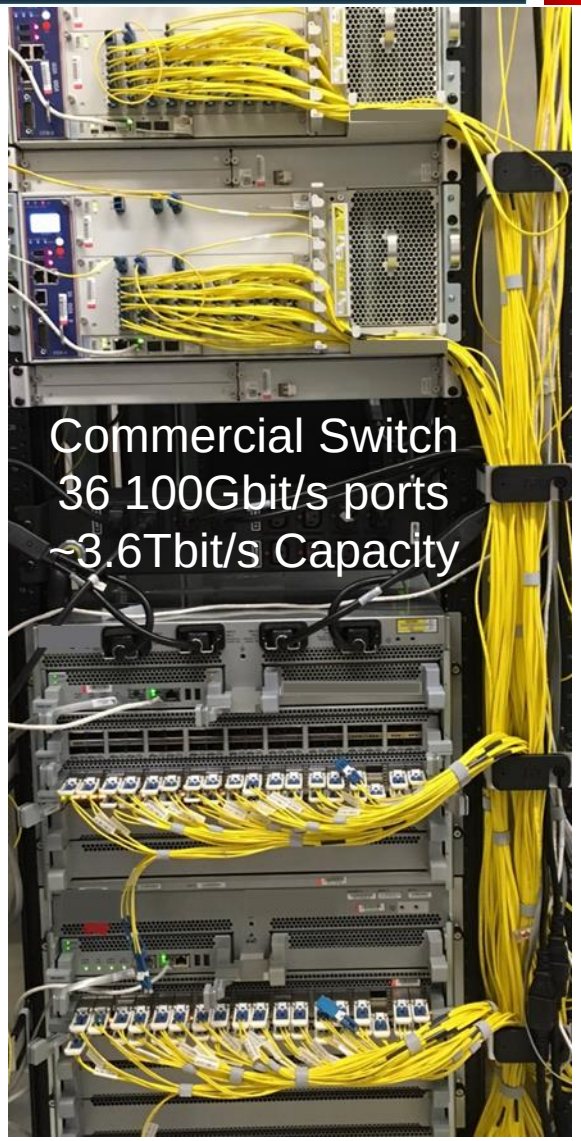


Module Dual λ Optical Output: 28.125GBaud Data



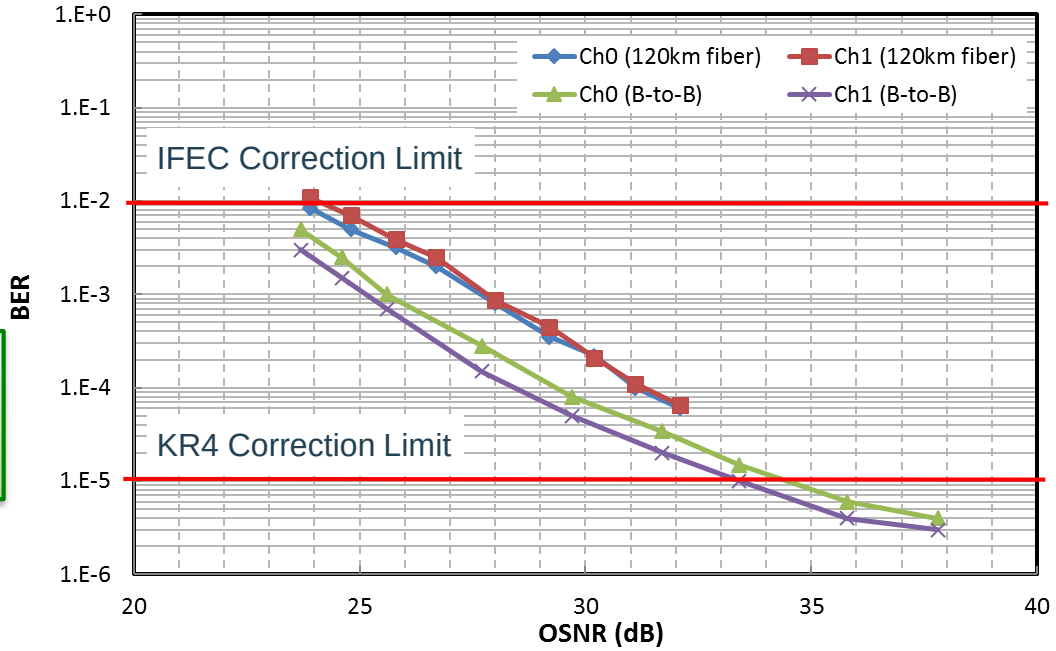
Two Way Traffic Using a Commercial Switch



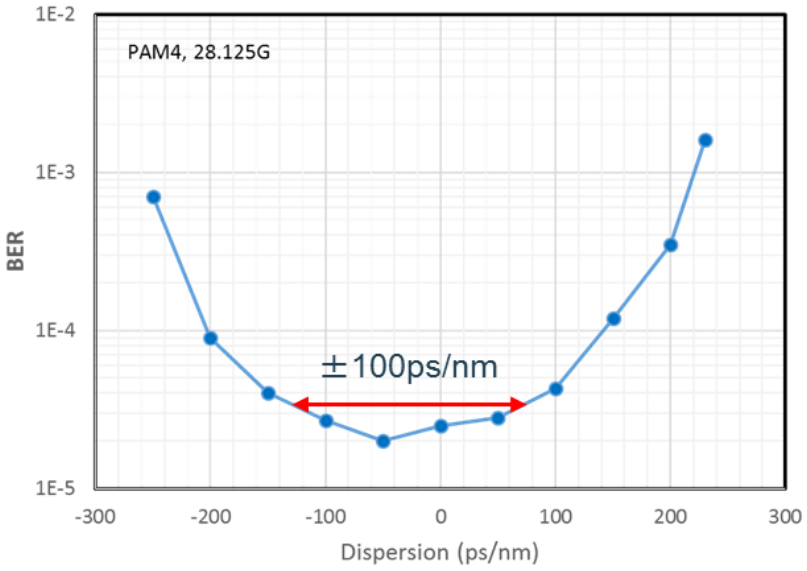


- Line system ➔ Transmit, Receive EDFAs.
- Dispersion Compensation at Mid-Stage of EDFA's.

120km BER vs. OSNR Performance

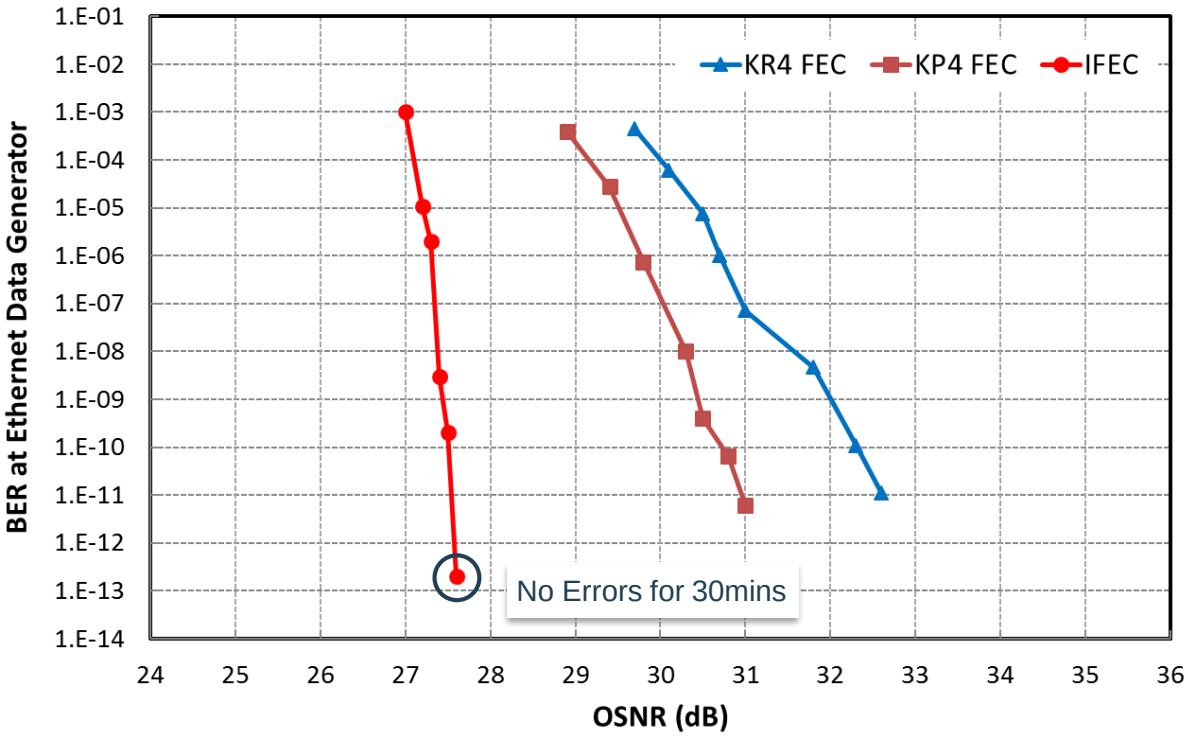


- Dispersion Tolerance ±100ps/nm
- ~1dB Penalty



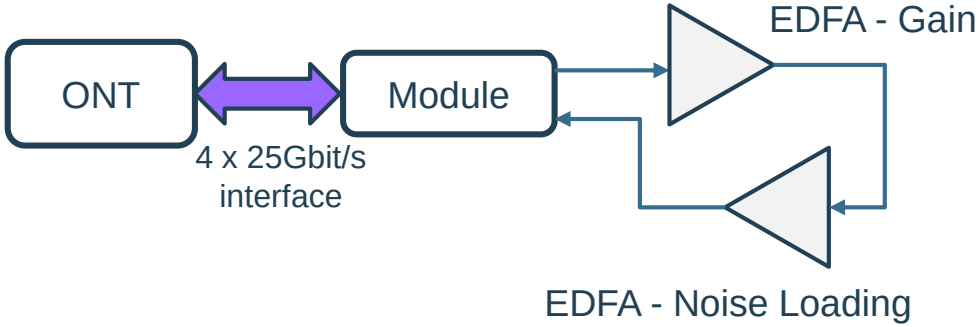
- 120km SMF Fiber ➔ 26.2dB Loss
- Baud Rate = 28.125GBaud

PAM 4 Link: FEC Performance



FEC Code	Baud Rate (Gbaud)
IFEC	28.125
KP4	26.5625
KR4	25.78125

- KP4 and KR4 are Reed-Solomon FEC codes.
- IFEC is a Iterative Multi-Layer code.



- BER vs. OSNR for 100GbE.
- Error free overnight at OSNR = 28dB.

Thank You!