1.6Tb/s Data Center Optics: Is Coherent the Right Choice ?

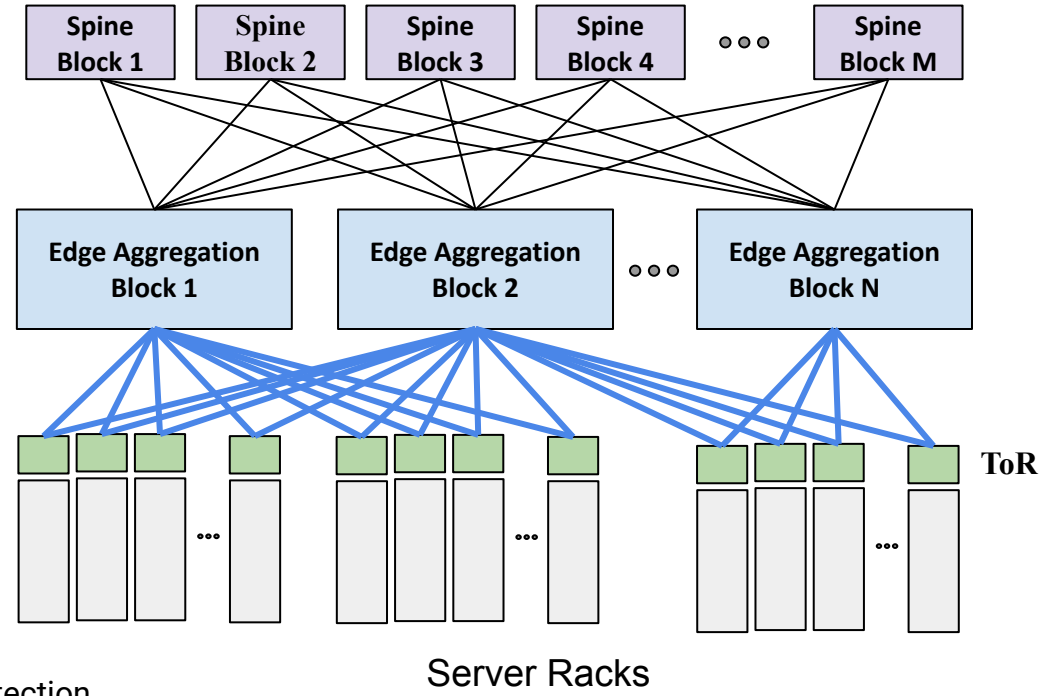
Xiang Zhou, Cedric Lam and Hong Liu

Google Systems & Services Infrastructure

OFC 2022 Panel, Tuesday, 08 March 14:00 – 16:00
What are the parallelization technologies for cost and energy efficient 1.6Tb links?

Intra-DC Cluster Interconnects (Current)

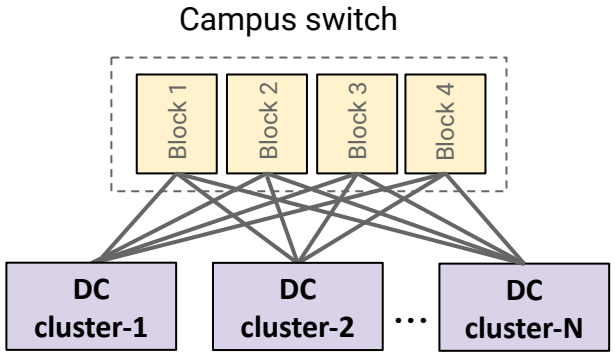
- Server to TOR
  - ~3m
  - Copper cable
- TOR to Edge Aggregation
  - ~100m SR Optics
  - IM-DD SDM
- Edge aggregation to Spine
  - ~1 km 'LR' optics
  - IM-DD CWDM

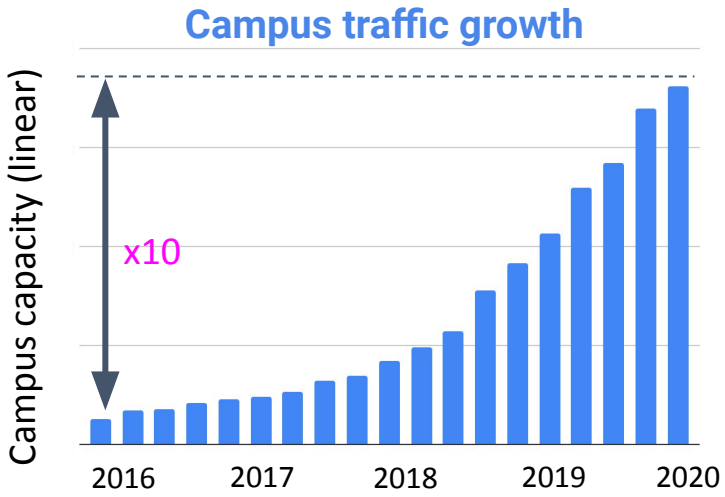


- IM-DD: intensity modulation-direct detection
- SDM: space division multiplexing
- CWDM: coarse wavelength division multiplexing

Campus Interconnects

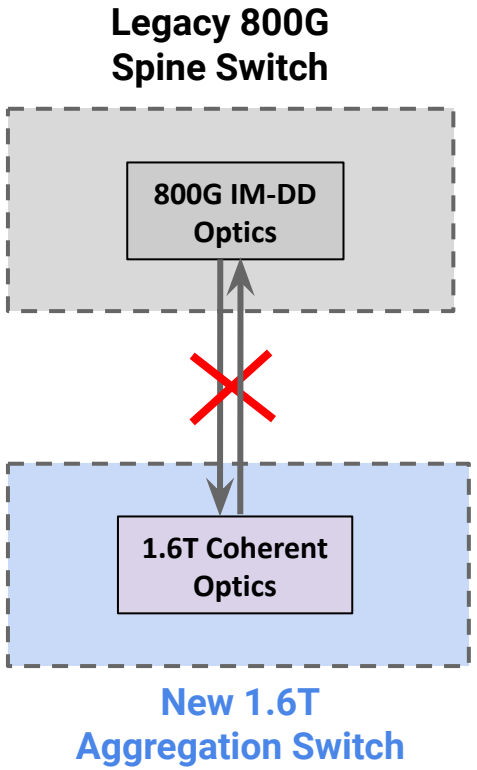
- A Fabric Interconnecting clusters distributed over multiple buildings
- 2-10km links based on campus size
  - Majority link <3km today
  - But campus size continues to grow
- Fast bandwidth growth
  - Campus capacity increased by more an a decade over 4 years (2016 to 2020)
- Cost sensitive
  - Same IM-DD CWDM ‘LR’ optics used for both Intra-DC Fabric and Campus up to 800G





1.6Tb/s DC Optics Options

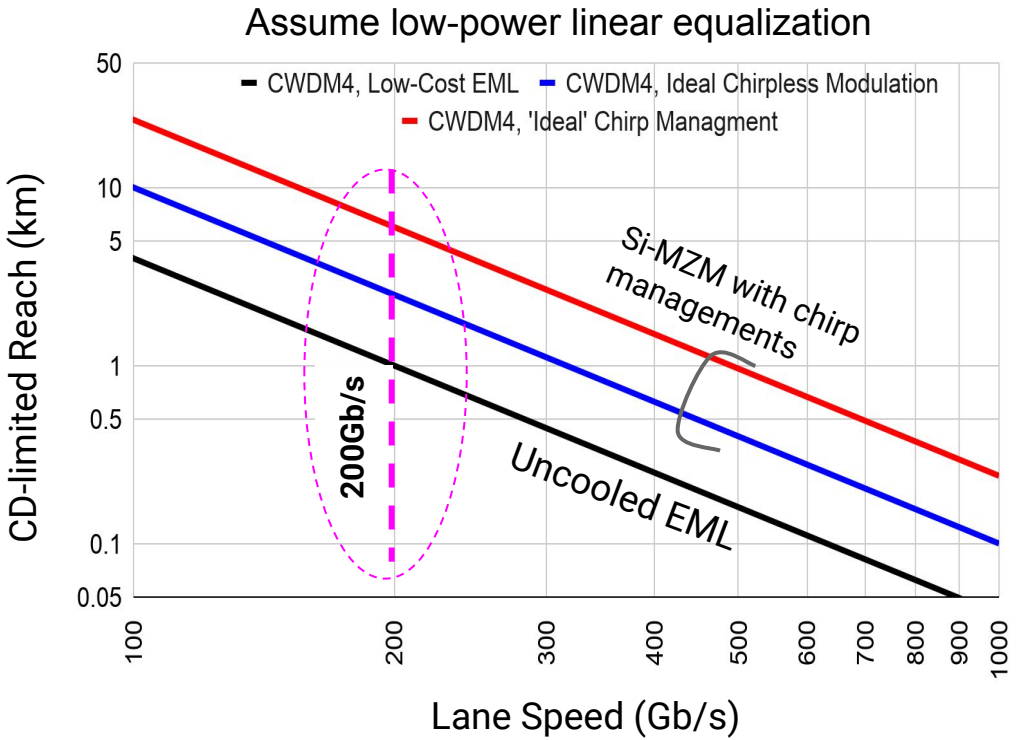
- 8x200Gb/s IM-DD (PAM4)
  - Backward compatibility
  - 200G/400G/800G fan-out granuarities
    - Important for TOR to Edge connections
  - Lowest complexity, cost and power
  - Reach limited by fiber CD and four-wave-mixing (FWM)
- 2x800Gb/s Coherent (PM-16QAM)
  - Longer reach
  - Potentially larger link budget (assume no optical amplifier)
  - Relatively low complexity and power
  - No backward optical interoperability
- 4x400G Coherent (PM-16QAM,PM-Coh-4PAM, PM-QPSK)
  - Best reach and link loss budget (assume no optical amplifier)
  - High complexity, cost and power
  - No backward optical interoperability



8x200G IM-DD PAM4: Fiber CD-Limited Reach

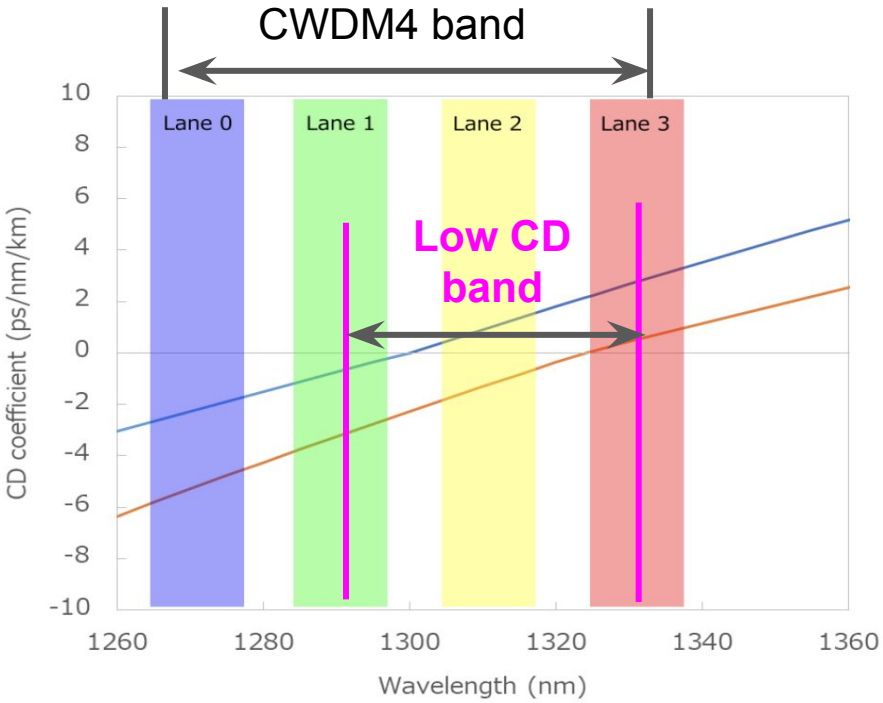
Assume CWDM4 optical bandwidth (1264.5nm to 1337.5nm)

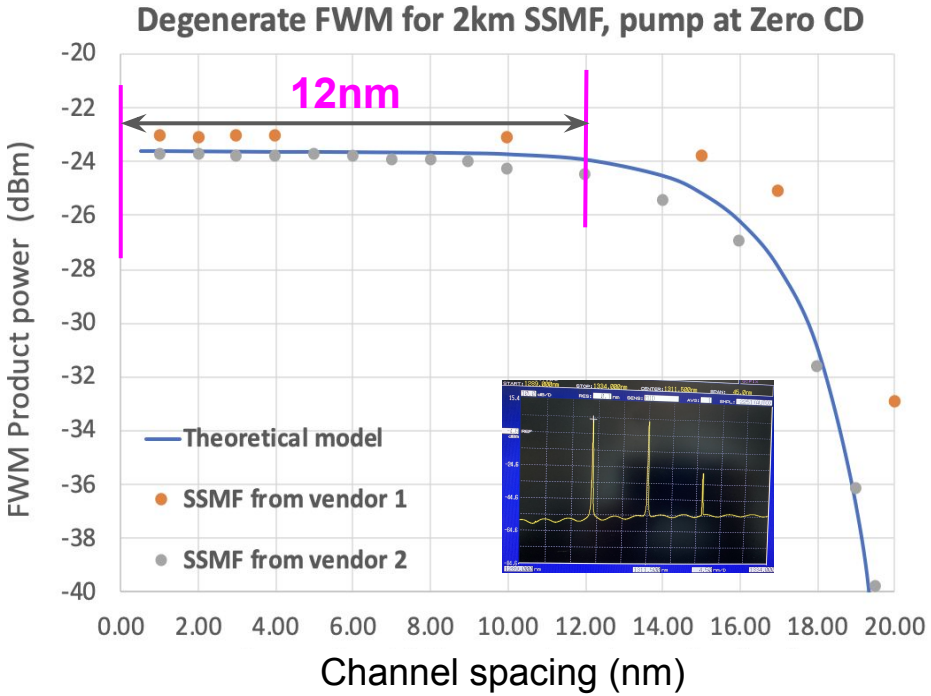
- CD Penalty ∝ (baud rate)^2
- Fiber CD puts an upper limit on the supported reach
- 10nm-spaced WDM8 (1.6T) or 20nm-spaced CWDM4 (800G)
  - ~1km with uncooled EML
  - ~3km with ideal MZM
- 8x200G IM-DD PAM4 can support Intra-DC reach, but not enough to support Campus reach



CD: chromatic dispersion

How About Narrower Channel Spacing to Lower CD Penalty

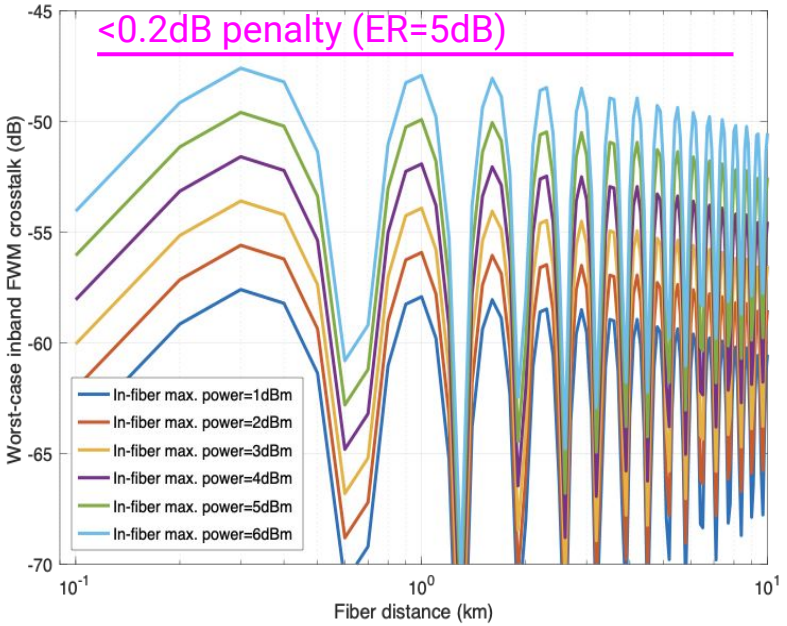




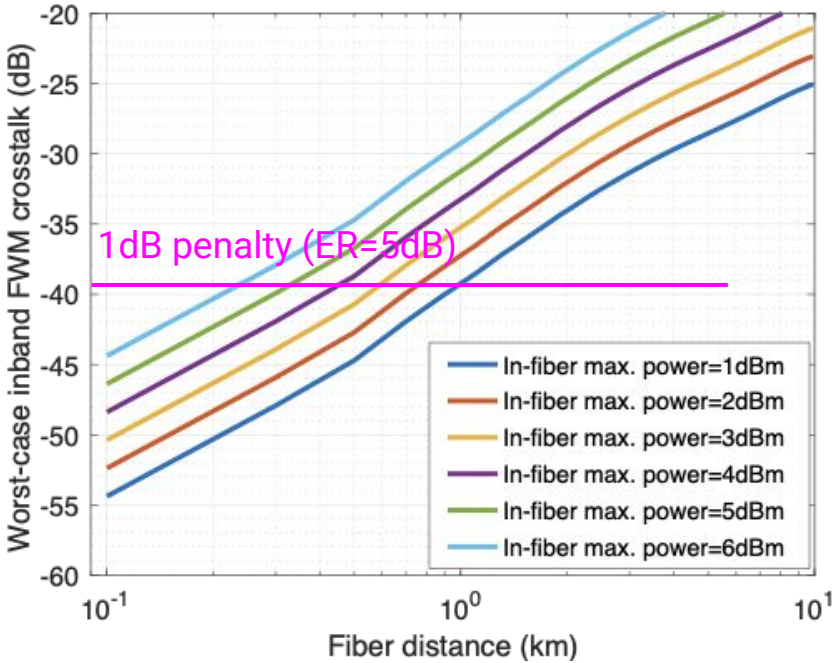
- FWM efficiency quickly increases as channel spacing reduces from 20nm
- Near perfect FWM phase matching observed when channel spacing <12nm @2km SSMF

8x200G IM-DD PAM4: Fiber FWM Impacts

4x200G 20nm-spaced CWDM4

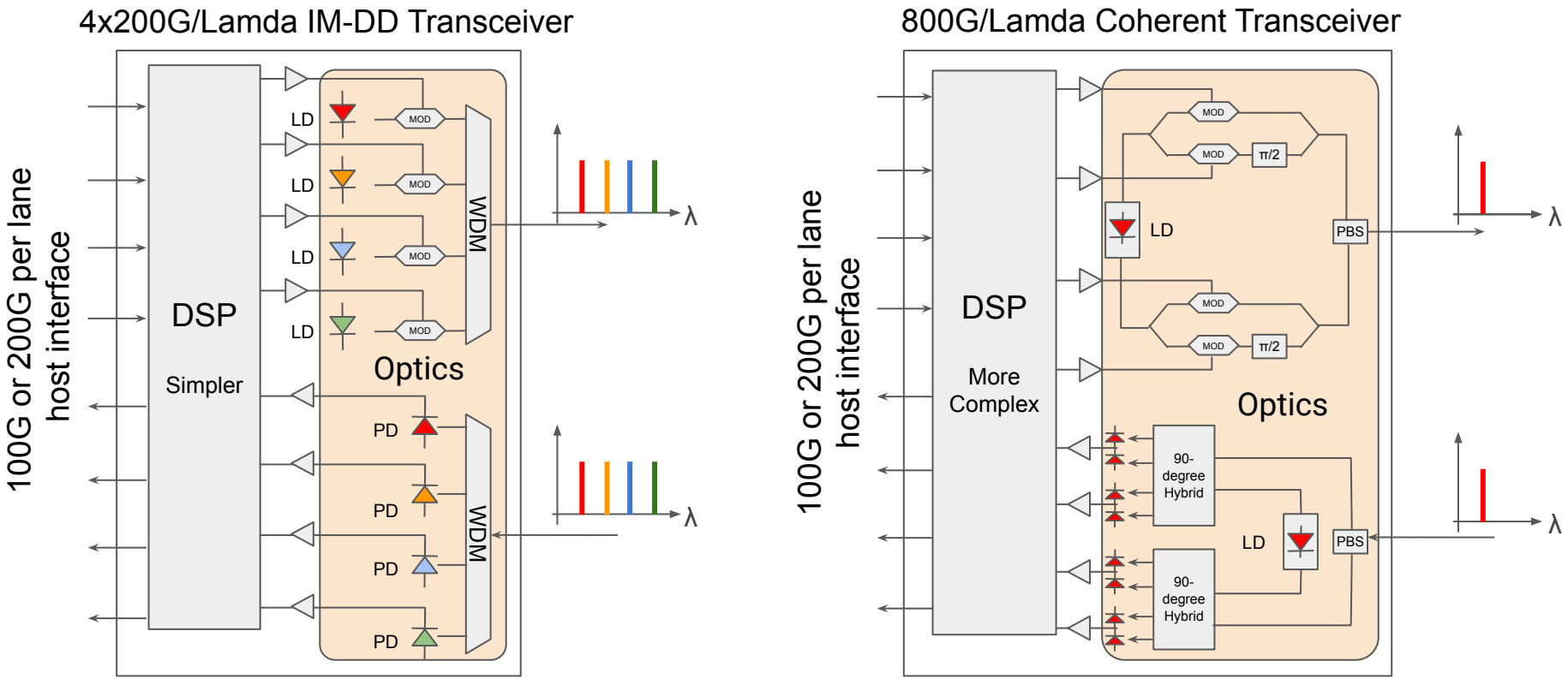


8x200G LAN-WDM8 (800GHz ch spacing)



- FWM crosstalk depends on Ch spacing, Tx power and fiber length
- Supported reach at 1dB FWM crosstalk penalty
  - >10km @ 6dBm Tx power for 20nm-CWDM4
  - <1km @1dBm Tx power for 800GHz-spaced LAN-WDM8

200G/Lane IM-DD vs 200G/Dim Coherent: Complexity



- **Overall similar complexity, although coherent require more sophisticated control circuits**
- **Power and cost of coherent will be higher**

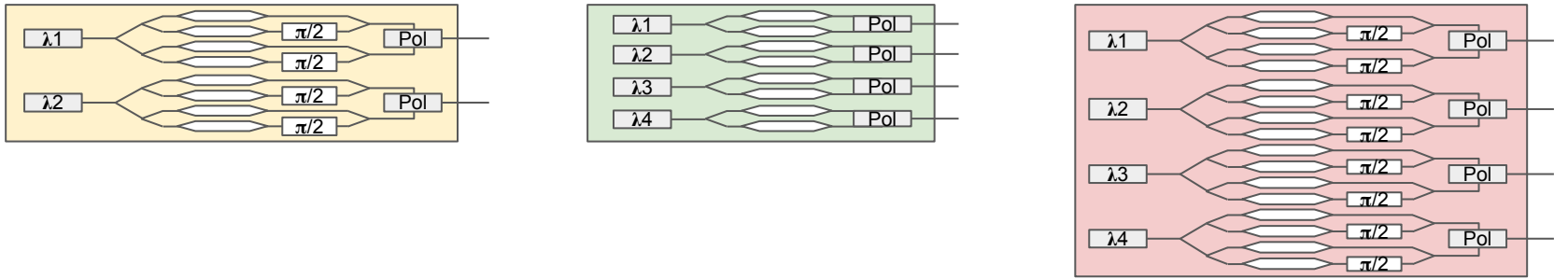
2x800G Coherent vs 4x400G Coherent: Complexity

2x800G PM-16QAM

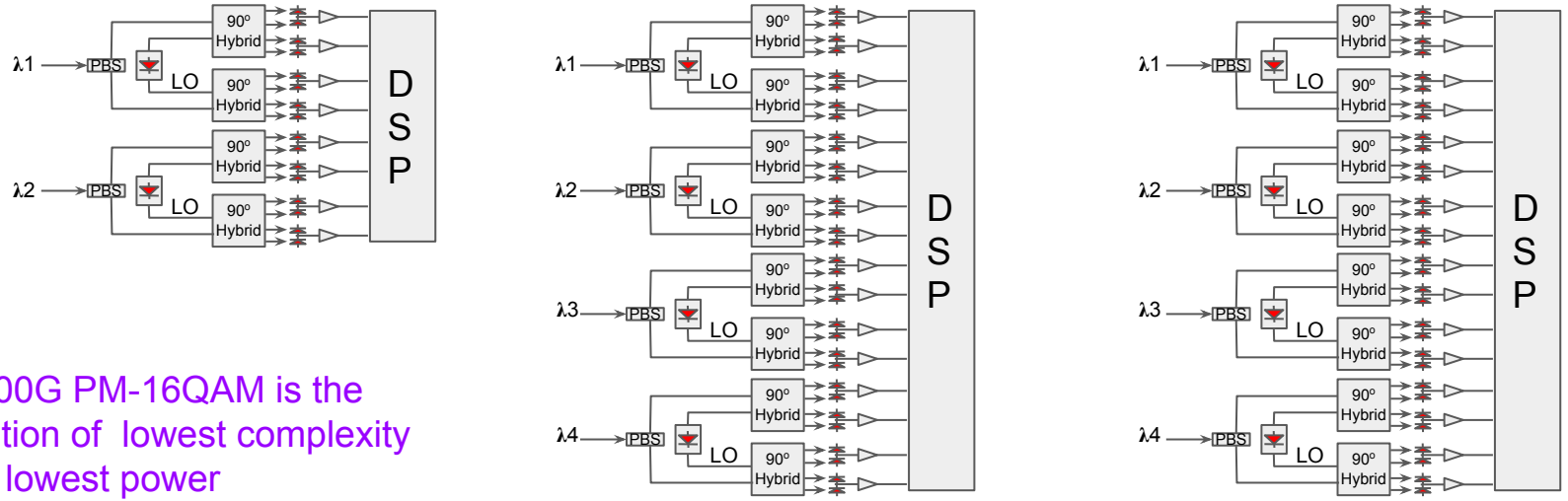
4x400G PM-Coh-4PAM

4x400G PM-QPSK or PM-16QAM

Tx



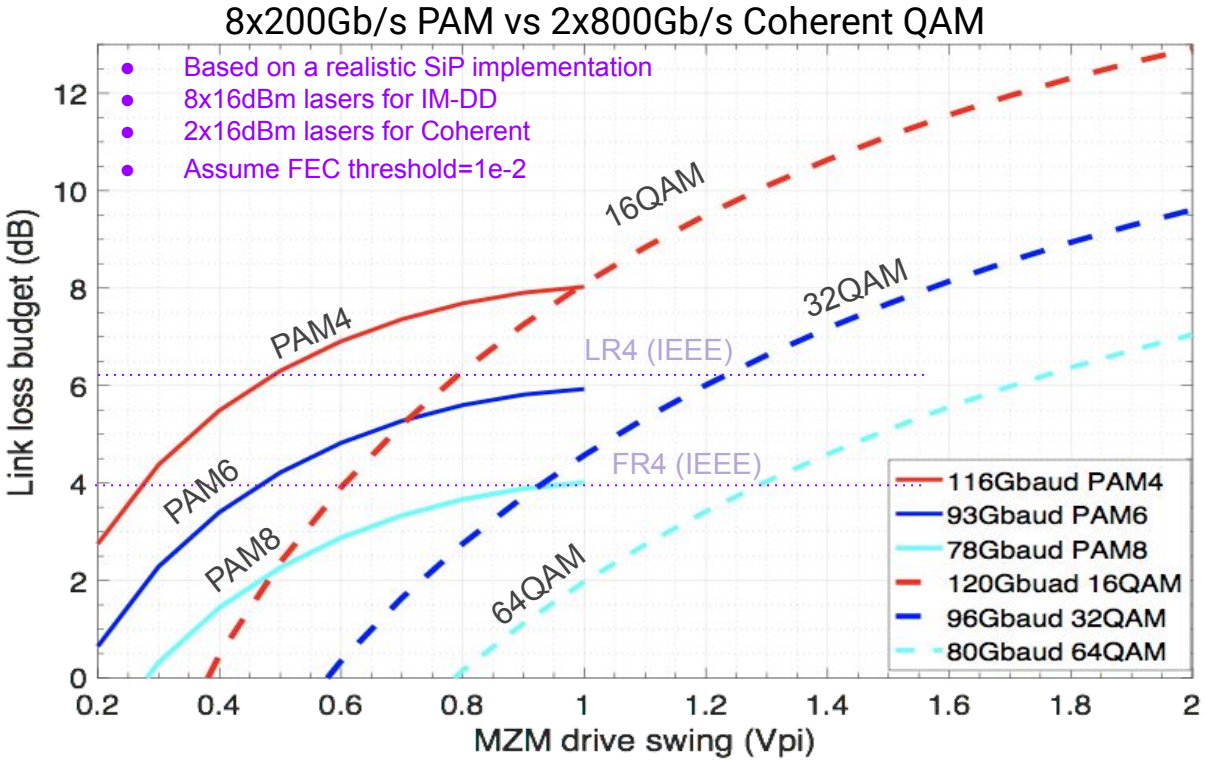
Rx



- 2x800G PM-16QAM is the solution of lowest complexity and lowest power

8x200G IM-DD vs 2x800G Coherent: Link Budget

Assume No Optical Amplifier



More detail: X. Zhou et al, JLT VOL. 38, NO. 2, pp.475-484

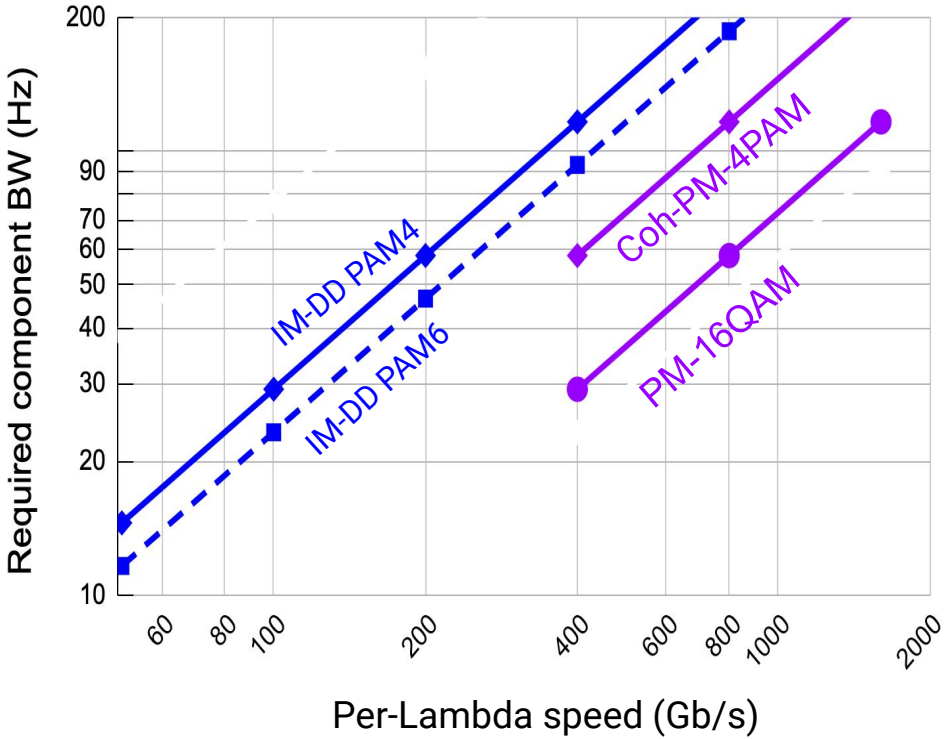
- Under identical per laser power (16dBm), IM-DD PAMn can achieve more link budget under moderate MZM drive swing
- Coherent Optics can achieve up to 4.5dB higher link budget with full 2Vpi drive (at higher power consumption)
- 2x800G PM-16QAM can close campus link (~6dB) with moderate MZM drive swing (~0.8 Vpi)

1.6Tb/s DC Optics: Google's Perspective

- 8x200G IM-DD PAM4 still the best option for <1km Intra-DC reach
  - Adequate reach
  - Lowest cost and power
  - Backward compatible
  - Support both 1.6Tb/s 10nm-WDM8 and 200G/400G breakout use cases

- Coherent Lite needed to support 1-10km campus reach
  - 8x200G IM-DD reach limited by fiber CD and FWM effects
  - 10km-optimized 2x800G PM-16QAM could be viable solution for campus
    - Component bandwidth requirements similar to 8x200G IM-DD
    - Transceiver only slightly more complex than 8x200G IM-DD

How about 3.2Tb/s Intra-DC Optics (<1km)



- 8x400G IM-DD still preferable
  - Backward compatible
  - Likely lowest cost and power
  - But high component BW requirements (>90GHz) could be a challenge
- 4x800G PM-16QAM could be a viable option if component bandwidth cannot continue to scale