

Enabling 400G/800G Interconnects with 56G/112G PAM-4 PHY

Rita Horner

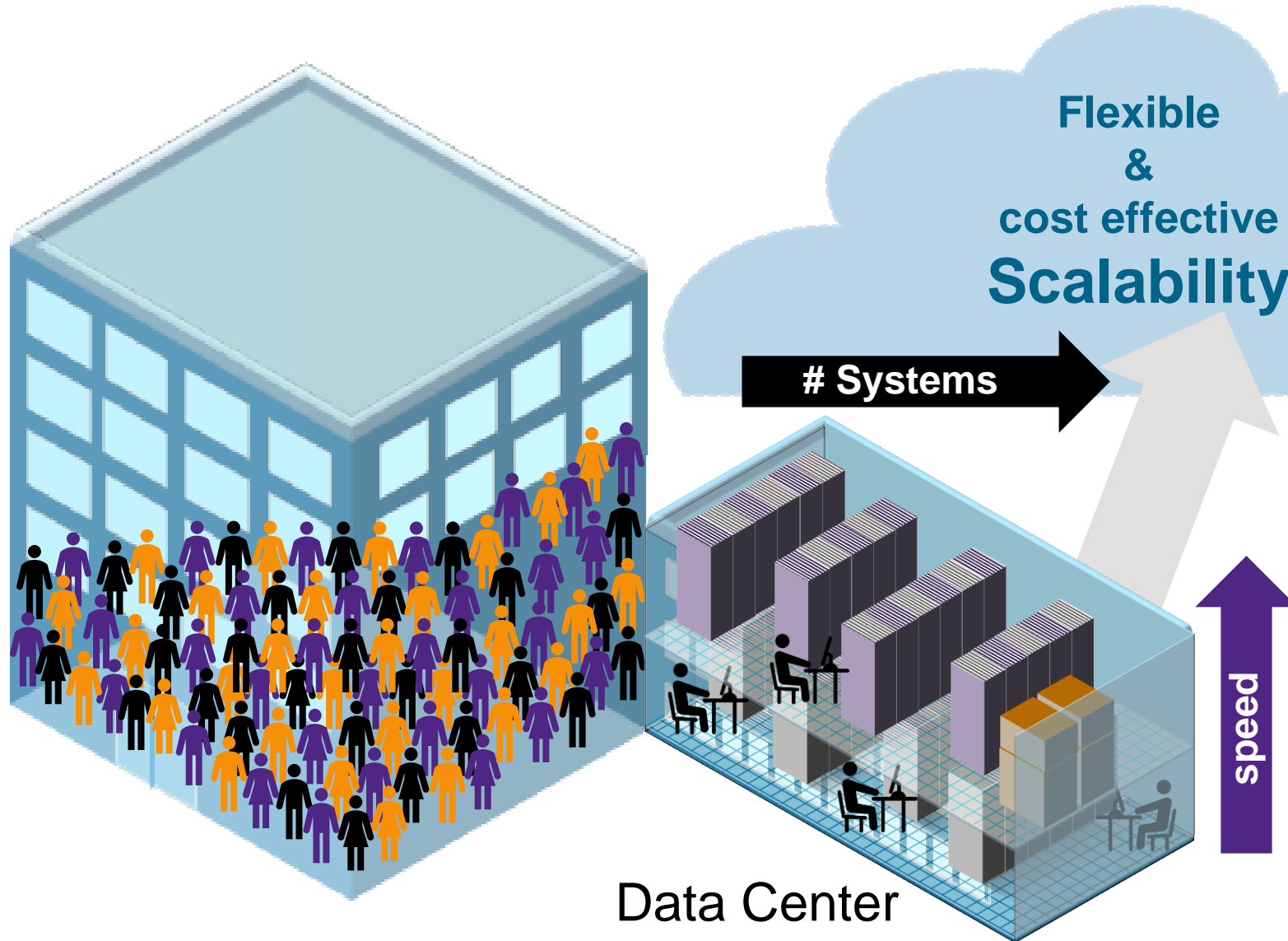
June 2019



Agenda

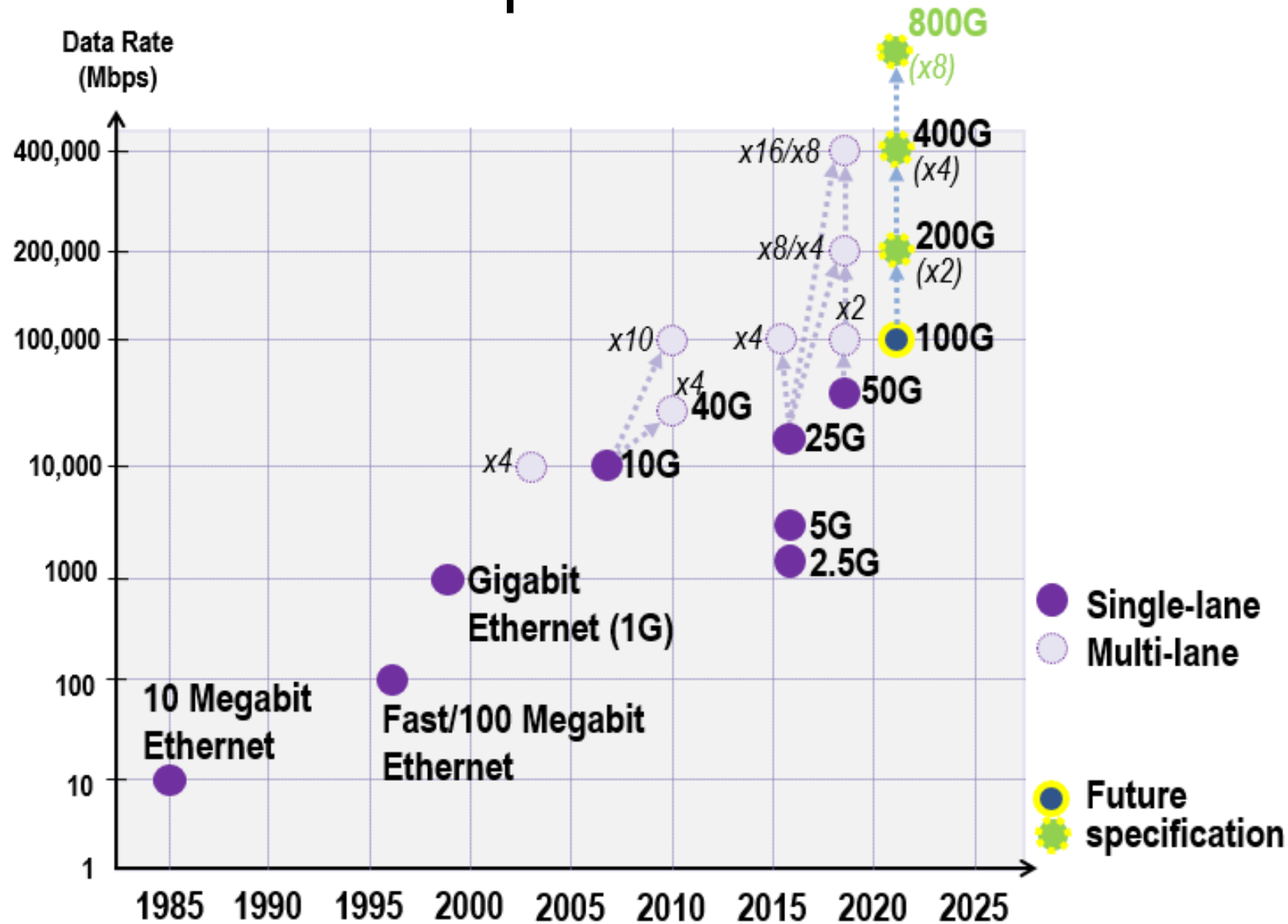
- Next-Generation Data Center Interconnects
- Multi-Level Modulation
- 56G/112G Ethernet PHYs
- Integrating High-Speed PHYs in SoCs

Enterprise Growth

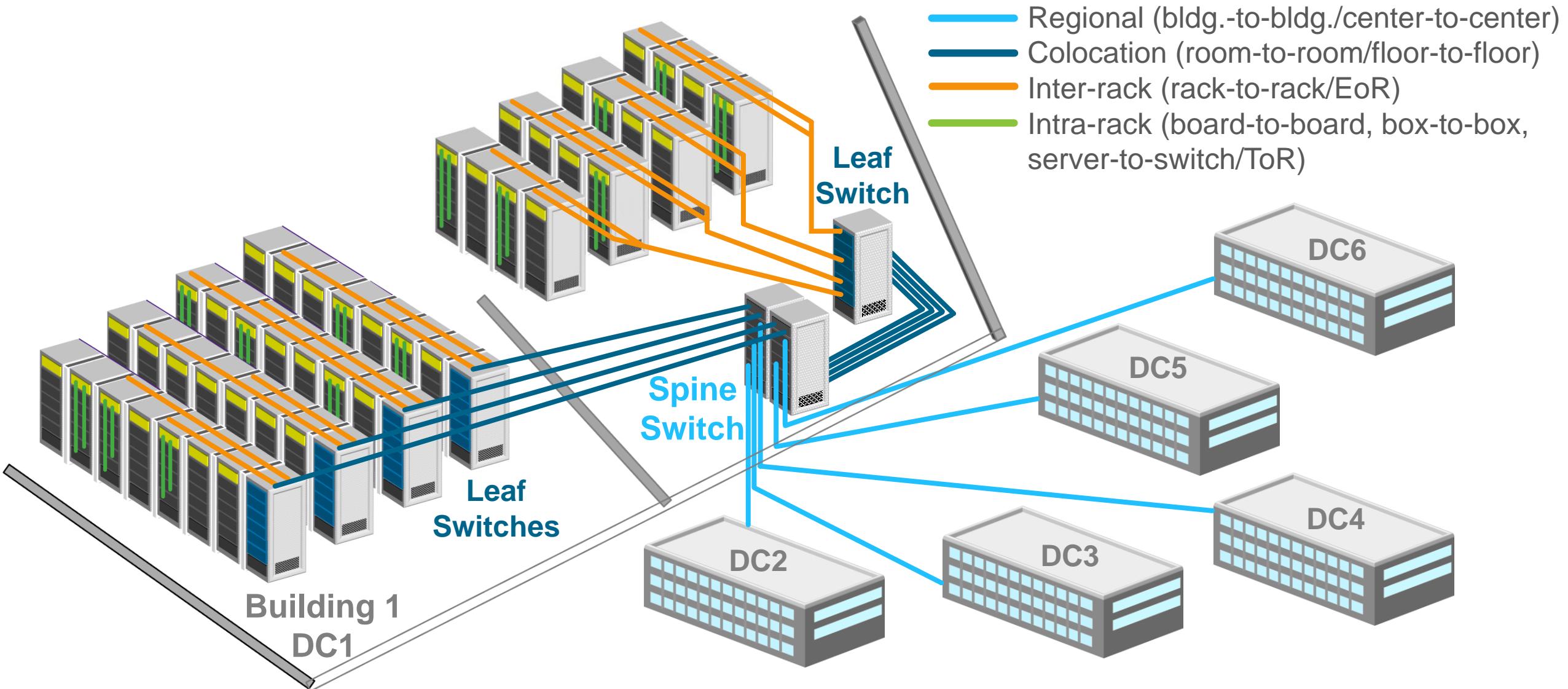


- **Horizontal Scaling** (scaling out)
 - Additional processing & storage
 - Larger housing area
 - Increased IT & cooling cost
- **Vertical Scaling** (scaling up)
 - Upgrading hardware
 - Ease of control
 - Potential availability constraints
 - More expensive equipment

Evolution of Ethernet Speeds

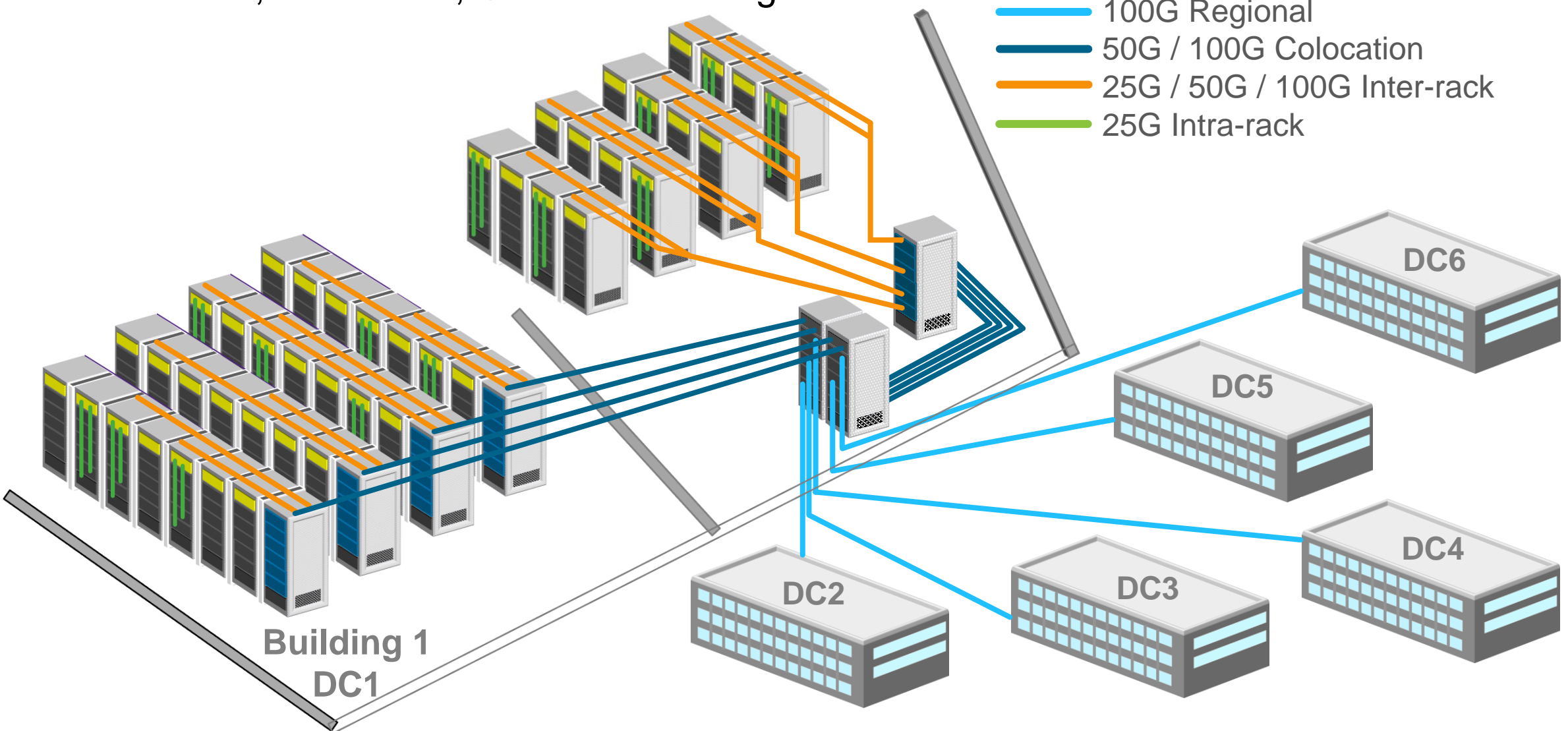


Data Center Interconnects



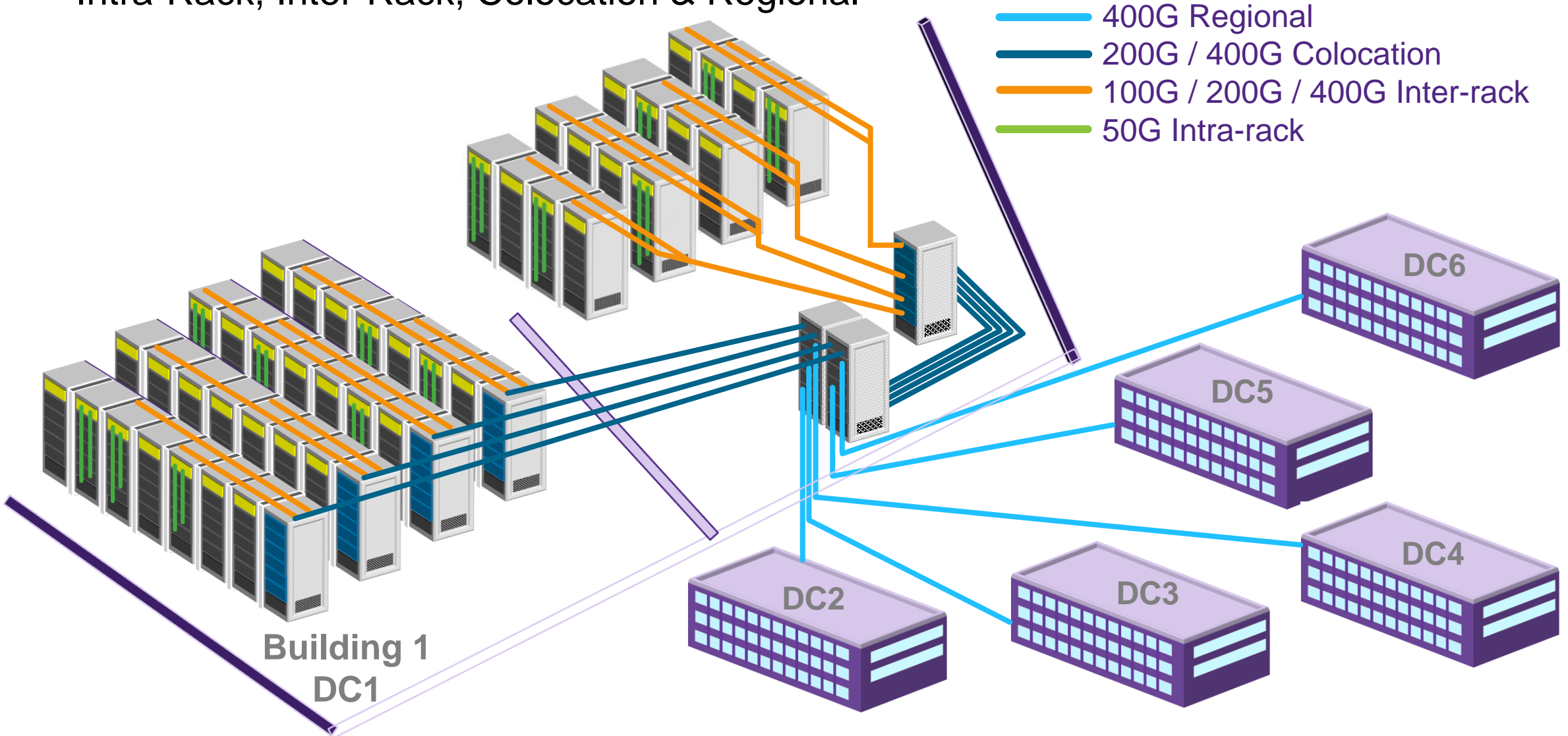
Hyperscale Data Center 100 GE Interconnects

Intra-Rack, Inter-Rack, Colocation & Regional



Hyperscale Data Center 400 GE Interconnects

Intra-Rack, Inter-Rack, Colocation & Regional

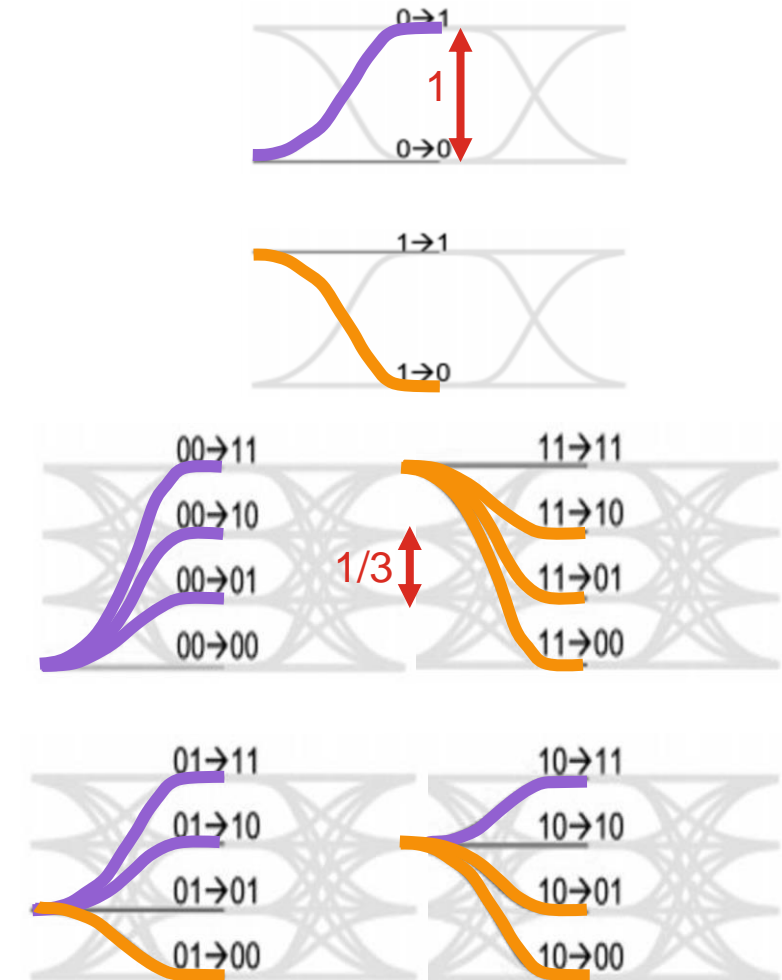
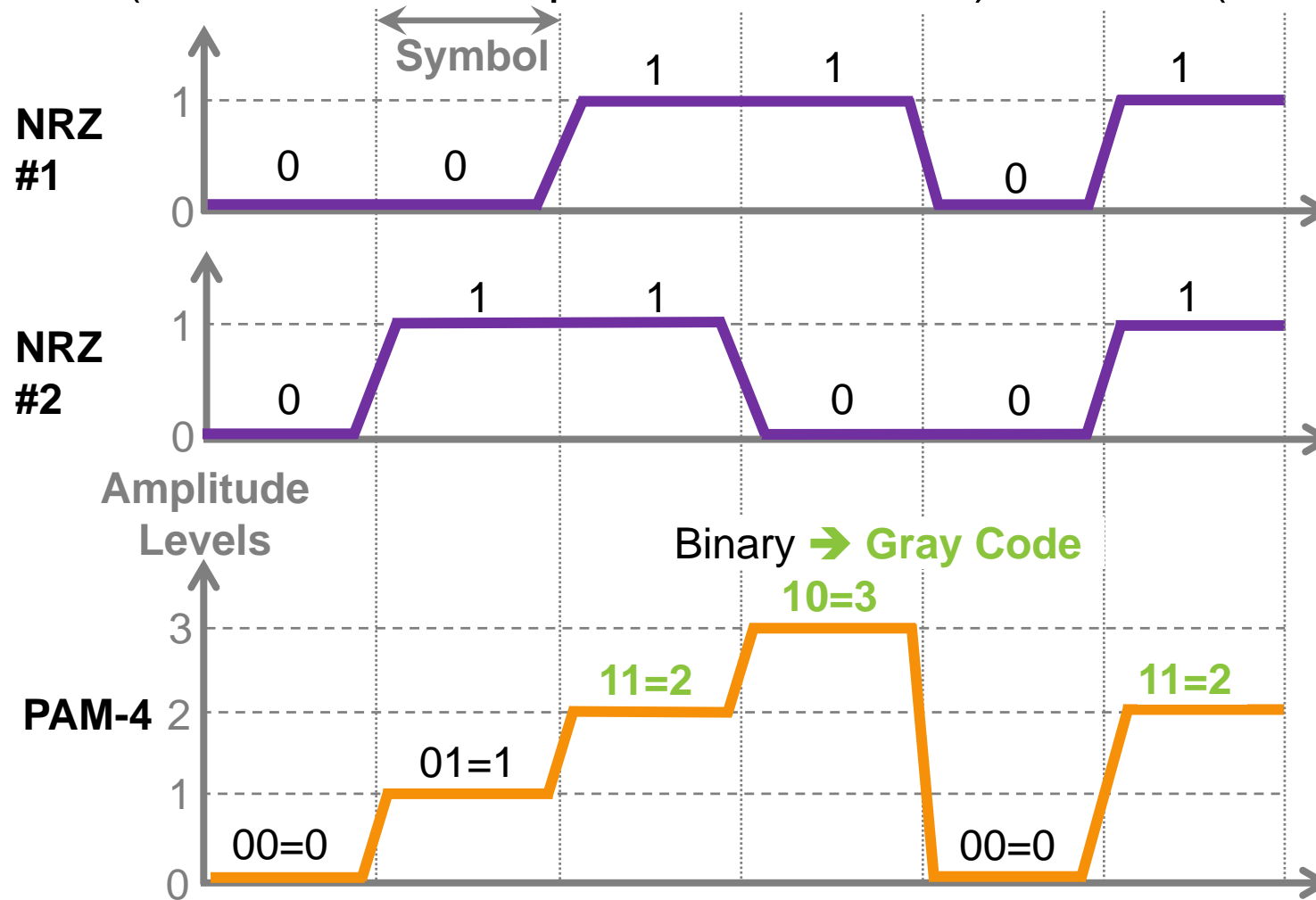


Agenda

- Next-Generation Data Center Interconnects
- **Multi-Level Modulation**
- 56G/112G Ethernet PHYs
- Integrating High-Speed PHYs in SoCs

Multi-Level Signaling vs. Binary Modulation

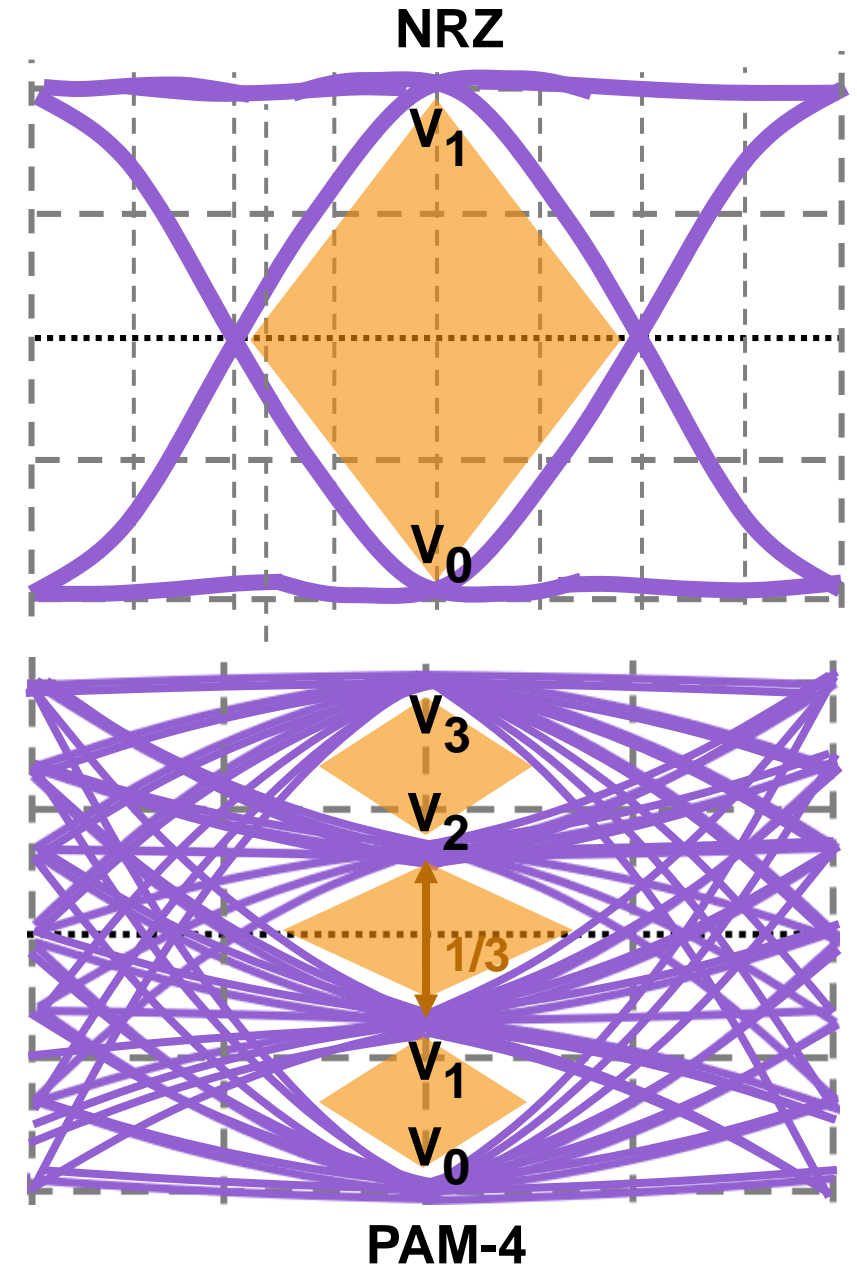
PAM-4 (4-Level Pulse Amplitude Modulation) vs. NRZ (Non Return to Zero)



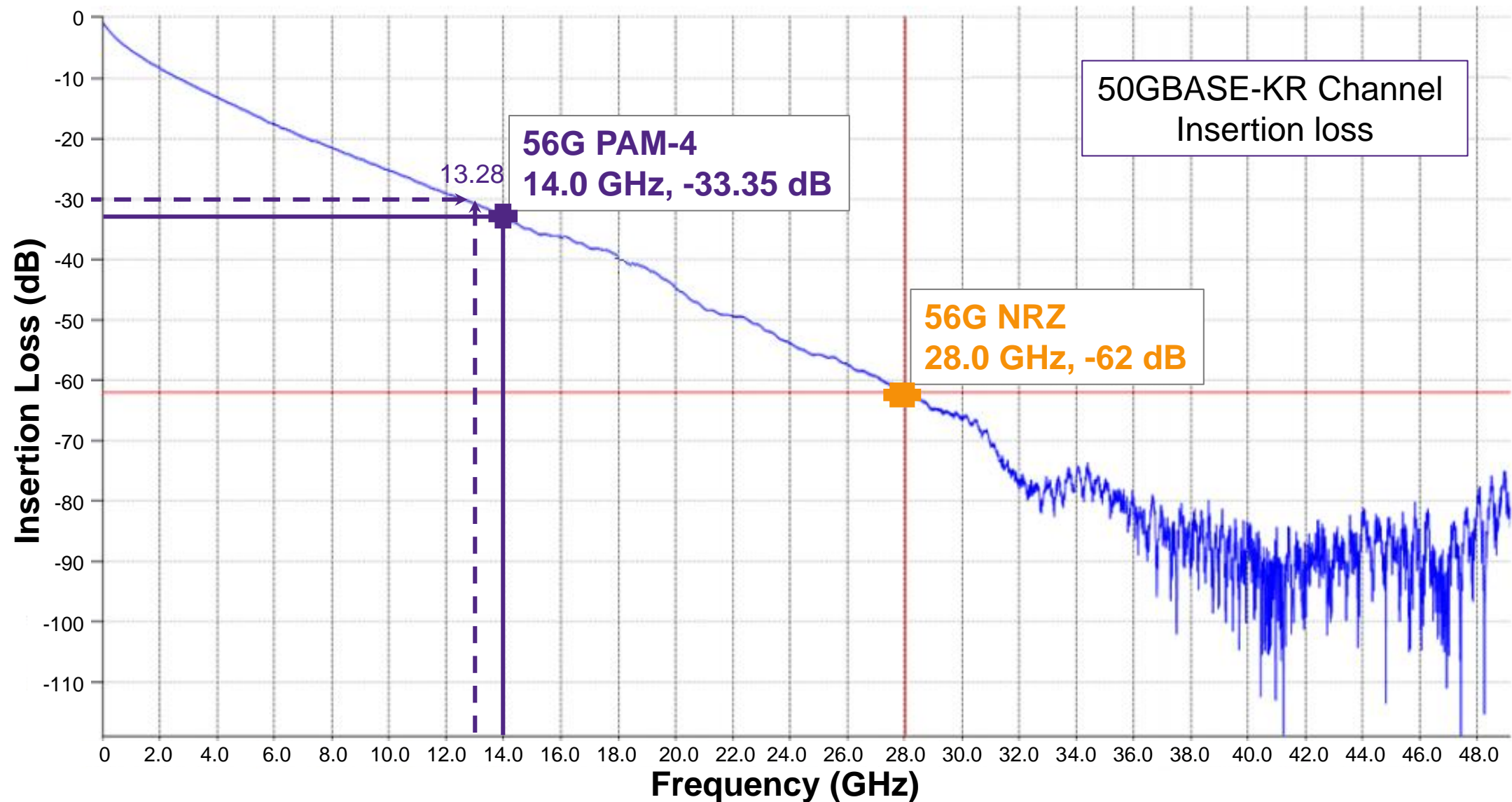
Channel Impairment Impact

PAM-4 vs. NRZ

- Additional voltage levels reduces the eye height by a factor of 3 in PAM-4
 - Signal-to-noise ratio (SNR) degrades
- Tx output eye width \rightarrow 1/2 to 2/3 of NRZ
 - Middle eye is the most symmetrical
 - Top and bottom eyes do not match the middle eye
- Impairments impact each of the three eyes differently
- Nonlinearity can significantly impact bit error rate
- Crosstalk and reflection have greater signal degradation impact




Shift From NRZ To PAM-4



Data Center Interconnects @ 100 GE → 400 GE

Channel Length, Media, Loss, Width ~~(x4)~~ → (x8)

As speeds go up, technology requirements migrate down, and channels shrink



Connection Area	Medium Type	100 GE (25G / 50G / 100G)	400 GE (50G / 100G / 200G / 400G)
Regions (Data center-to-data center)	Optical	40 km SMF	→ 802.3ct, 80 km SMF 10 km SMF → 802.3cn, 40 km SMF
Colocation (Building-to-building)		10 km SMF	2 km SMF
Within the building (Room-to-room)		100 m MMF	500 m SMF → 802.3cm, 100 m MMF
Intra-rack (Rack-to-rack)		30 m AOC, MMF	30 m AOC, MMF
Inter-rack (within the rack)	Copper -- cable	5 m DAC, Twin-axial	3 m DAC, Twin-axial
Intra-box (server/switches)	Copper PCB/mezzanine/backplanes	35 dB @12.89 GHz	30 dB @13.28125 GHz

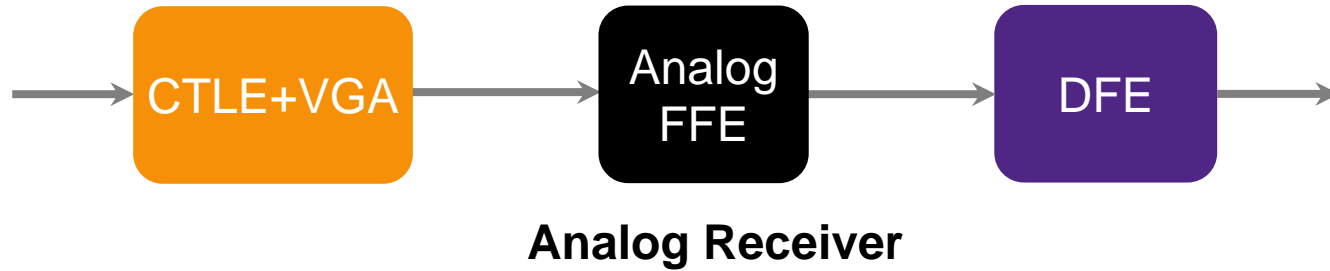
Single Mode Fiber (SMF), Multi Mode Fiber (MMF), Active Optical Cable (AOC), Direct Attached Cable (DAC)

Agenda

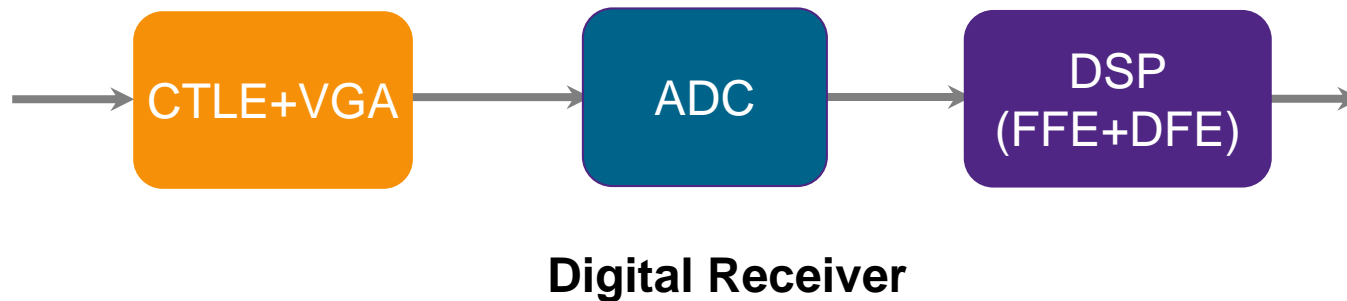
- Next-Generation Data Center Interconnects
- Multi-Level Modulation
- **56G/112G Ethernet PHYs**
- Integrating High-Speed PHYs in SoCs

Major Types of PAM-4 Designs

Analog vs. Digital Receiver Architecture



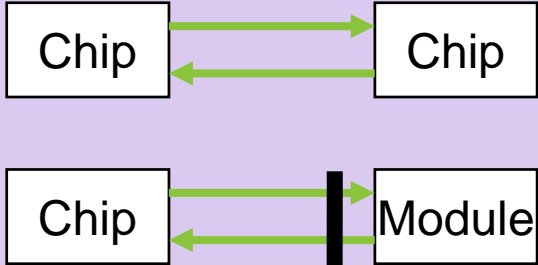


- Similar to traditional NRZ receiver
- PVT variation concern
- Technology scalability
- DFE taps (area + power)
- Ideal for short channels
- Power advantage



- Complex architecture
- Flexibility in process scaling
- More capable signal processing techniques → better SNR
- Ideal for broad range of channels
- Power disadvantage

IEEE802.3 Specifications for Single & Multiple Channel

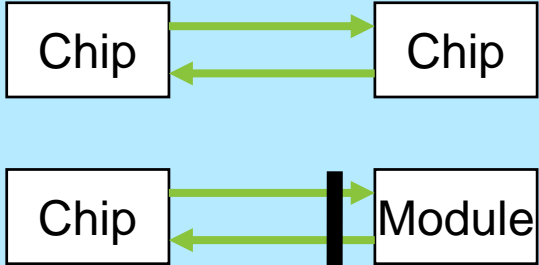


802.3bs (200G & 400G)/802.3-2018 & 802.3cd (50G, 100G, 200G) @ 53.125 Gbps

Specifications	Description	
50GAUI-1 100GAUI-2 200GAUI-4 400GAUI-8	Chip-to-Chip & Chip-to-Module	
50GBASE-CR 100GBASE-CR2 200GBASE-CR4	1-, 2- or 4-lane shielded balanced Twinaxial copper cabling (CR), 2 m & 3 m	
50GBASE-KR 100GBASE-KR2 200GBASE-KR4	1-, 2- or 4-lane Electrical backplane (KR)	

Single & aggregated link rates of up to 400 Gbps

IEEE802.3 Specifications for Single & Multiple Channel

802.3ck (100G, 200G & 400G) @ 106.25 Gbps (1st Draft Targeted for July 2019)

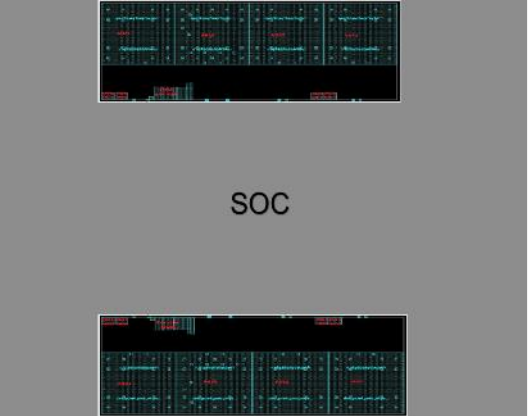
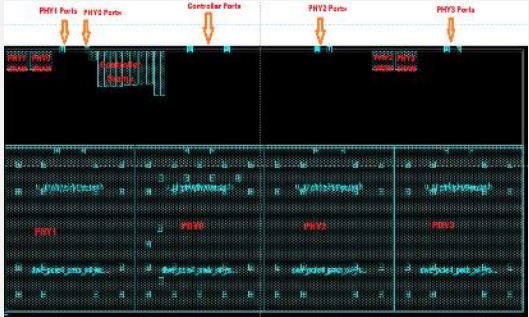
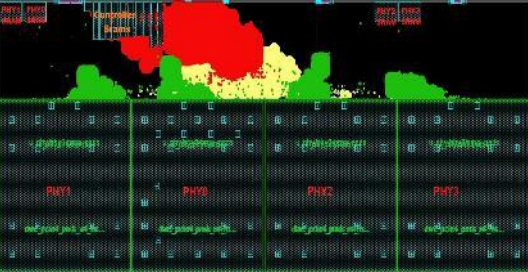
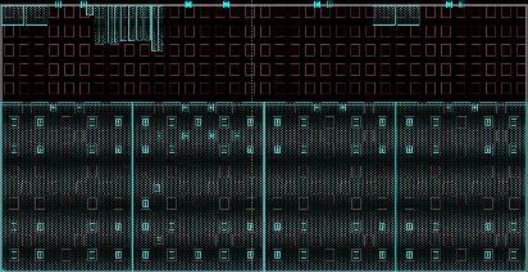
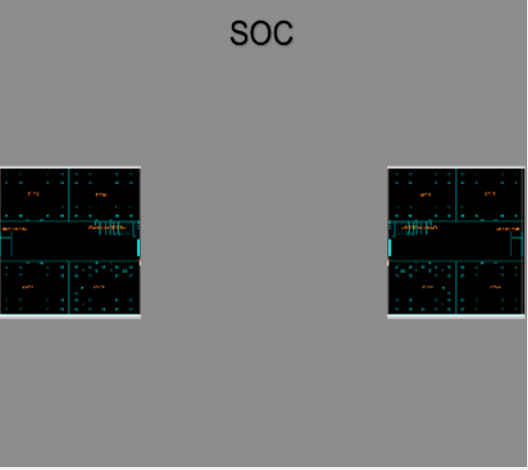
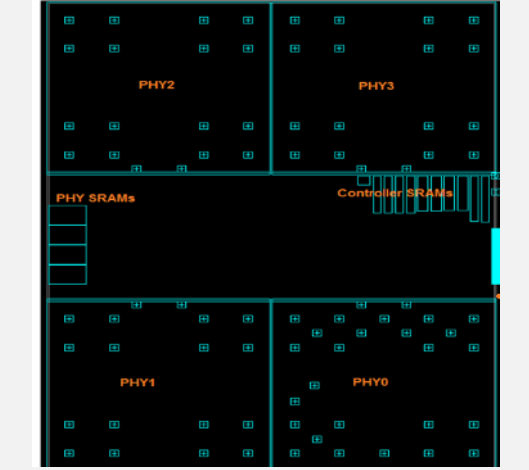
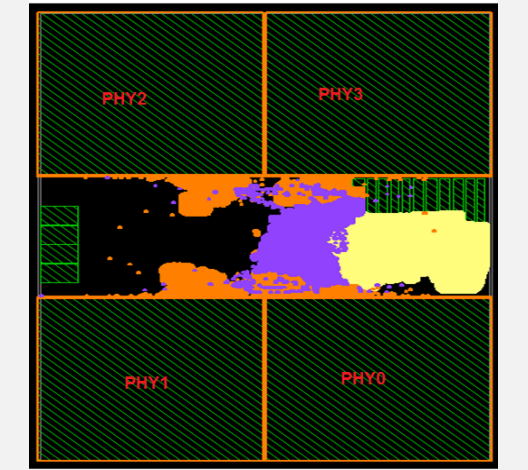
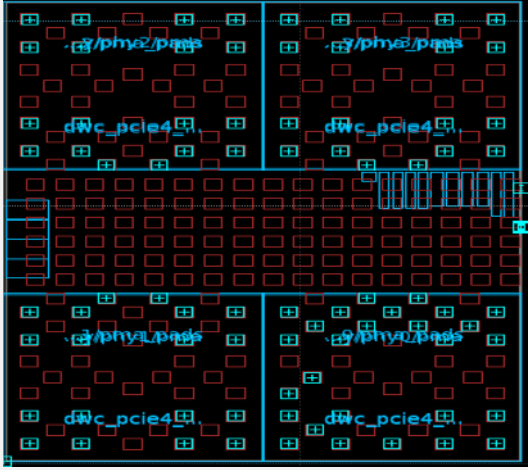
Specifications	Description	
100GAUI-1 200GAUI-2 400GAUI-4	Chip-to-Chip (-S vs. -L?) & Chip-to-Module	
100GBASE-CR1 200GBASE-CR2 400GBASE-CR4	1-, 2- or 4-lane shielded balanced Twinaxial copper cabling (CR), 1.5 m & 2 m	
100GBASE-KR1 200GBASE-KR2 400GBASE-KR4	1-, 2- or 4-lane Electrical backplane (KR)	

Single & aggregated link rates of up to 400G → 800 Gbps

Agenda

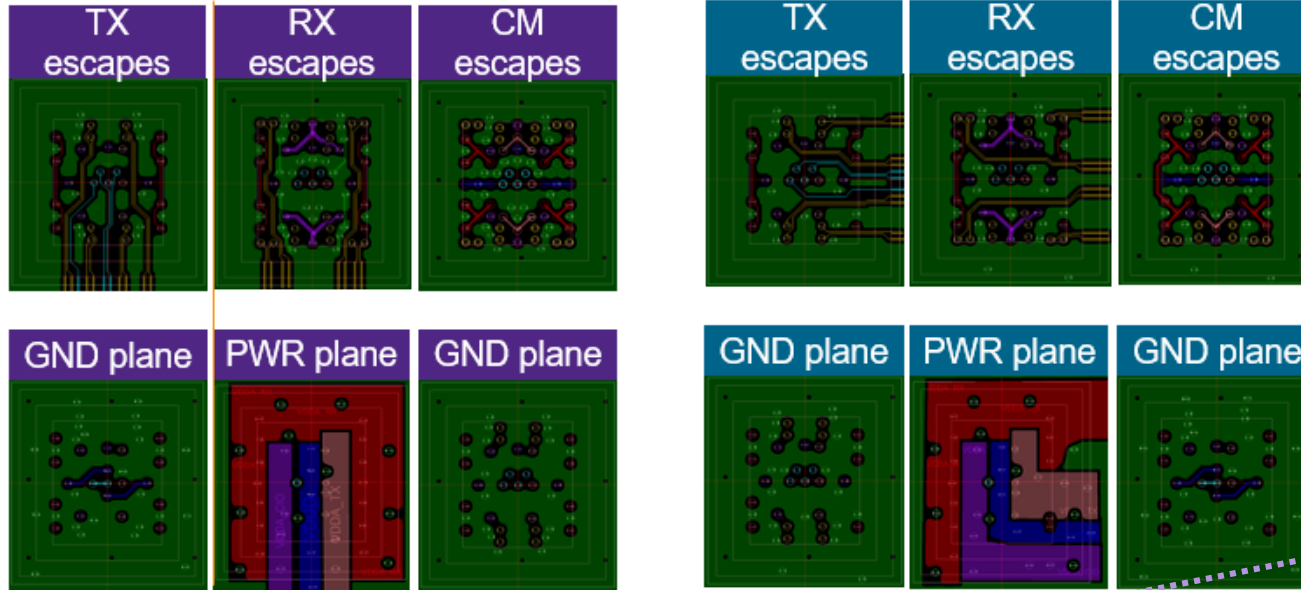
- Next-Generation Data Center Interconnects
- Multi-Level Modulation
- 56G/112G Ethernet PHYs
- Integrating High-Speed PHYs in SoCs

PHY Integration

	Placement	Floor Plan	Hierarchy	Bumps
North-South	 <p>SOC</p>	 <p>PHY1 Ports PHY2 Ports Controller Ports PHY3 Ports</p>	 <p>PHY1 PHY2 PHY3</p>	
East-West	 <p>SOC</p>	 <p>PHY2 PHY3 PHY SRAMs Controller SRAMs PHY1 PHY0</p>	 <p>PHY2 PHY3 PHY1 PHY0</p>	 <p>PHY2 PHY3 PHY1 PHY0</p>

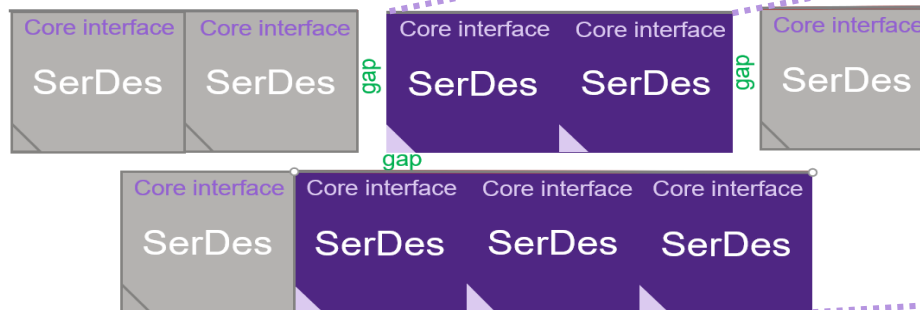
Package Study

Escape Route & Substrate Routing

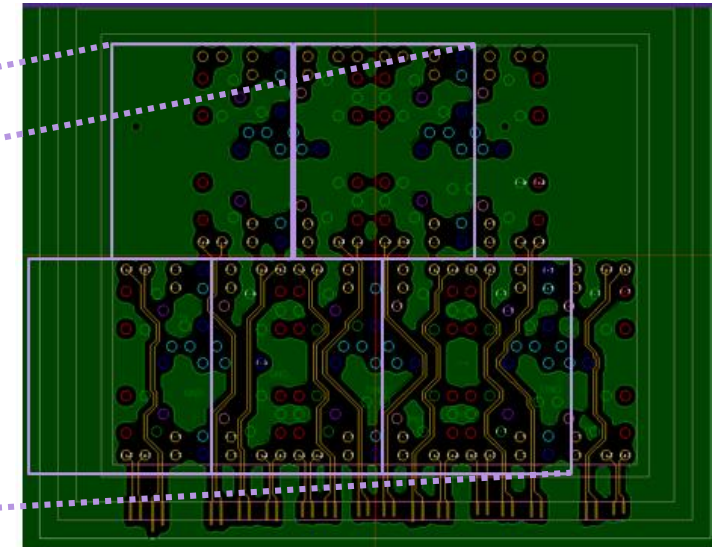
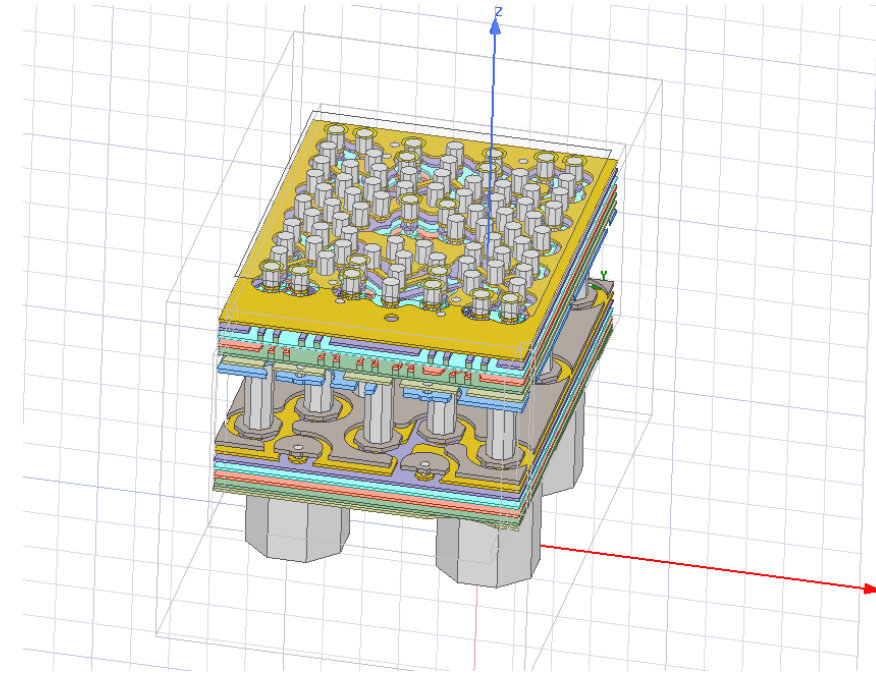


N/S Package Escape

E/W Package Escape



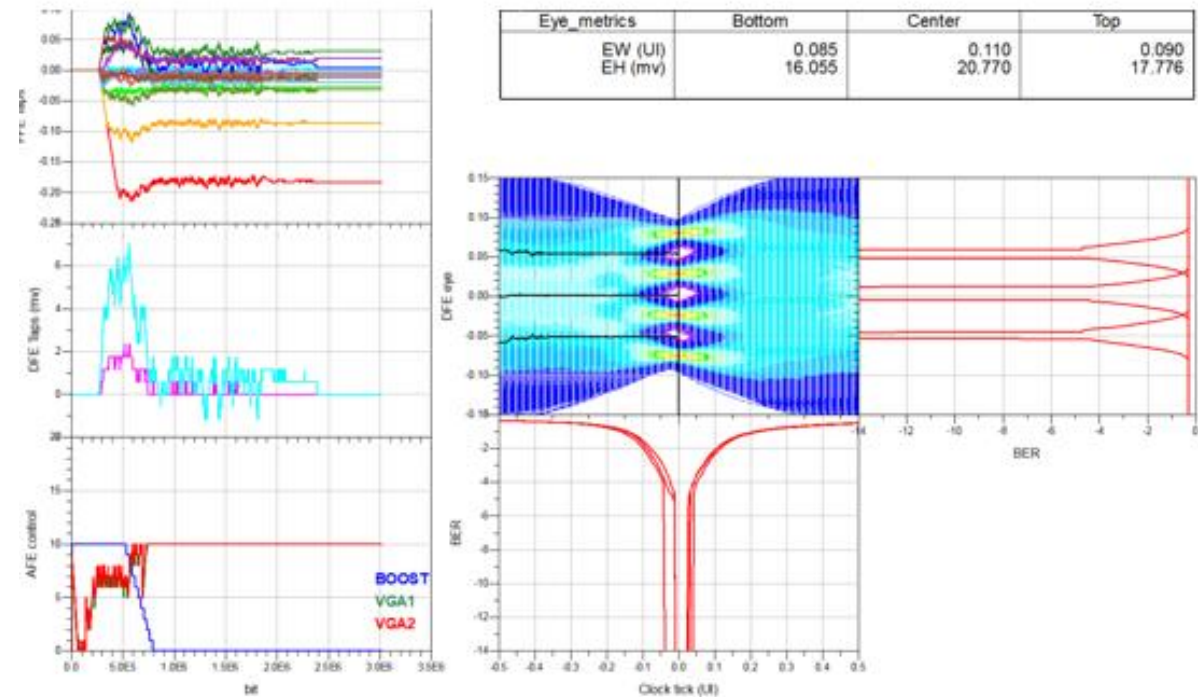
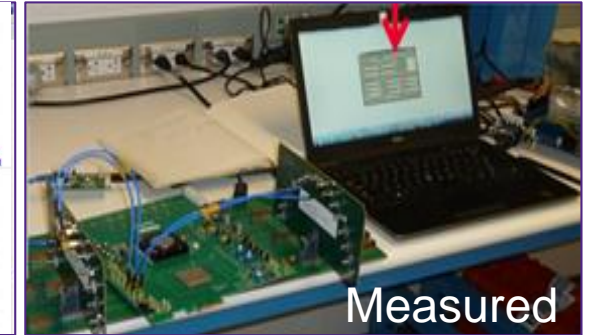
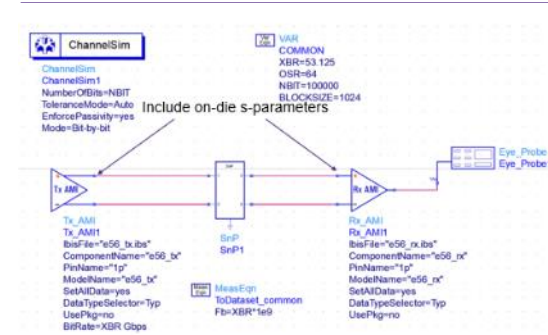
die edge



Silicon Correlated Models for Accurate SI Analysis

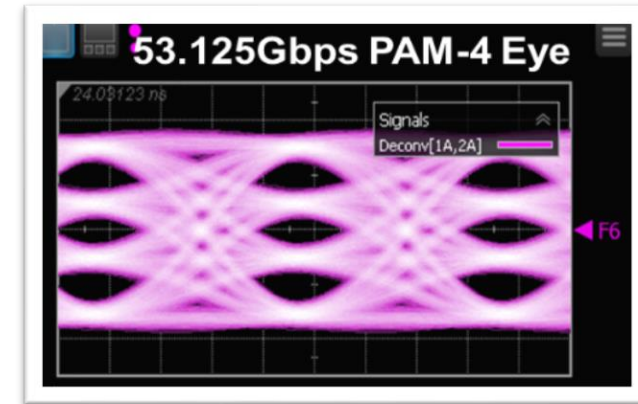
IBIS-AMI Models

- Model elements to match architecture & implementation (including adaptation algorithms)
- Further tune the model with information from hardware characterization data
- Correlated with silicon characterization for accurate analysis
- Verify models with different AMI& SPICE simulators



DesignWare 56G PAM-4 Ethernet PHY IP for Next-Generation 400G & 800G Hyperscale Data Centers

- Available in advanced process technologies (16-nm to 7-nm FinFet)
- Supports OIF & IEEE 802.3 Ethernet based channel types
 - Backplane, copper cables (DAC) and optical links
 - Backward compatible with NRZ data rates
- Delivers high performance, >35 dB @ 14 GHz PAM-4 Nyquist
- Digital (AFE + ADC + DSP) based receiver with power-performance knobs
- Robust performance over Voltage & Temperature (VT) corners
https://www.synopsys.com/dw/dwc_56g_ethernet_phy
- Scalable architecture to 112 Gbps
https://www.synopsys.com/dw/dwc_112g_ethernet_phy



Thank You

