



# MultiPHY Ethernet

Drew Perkins  
[dperkins@infinera.com](mailto:dperkins@infinera.com)  
408-572-5308

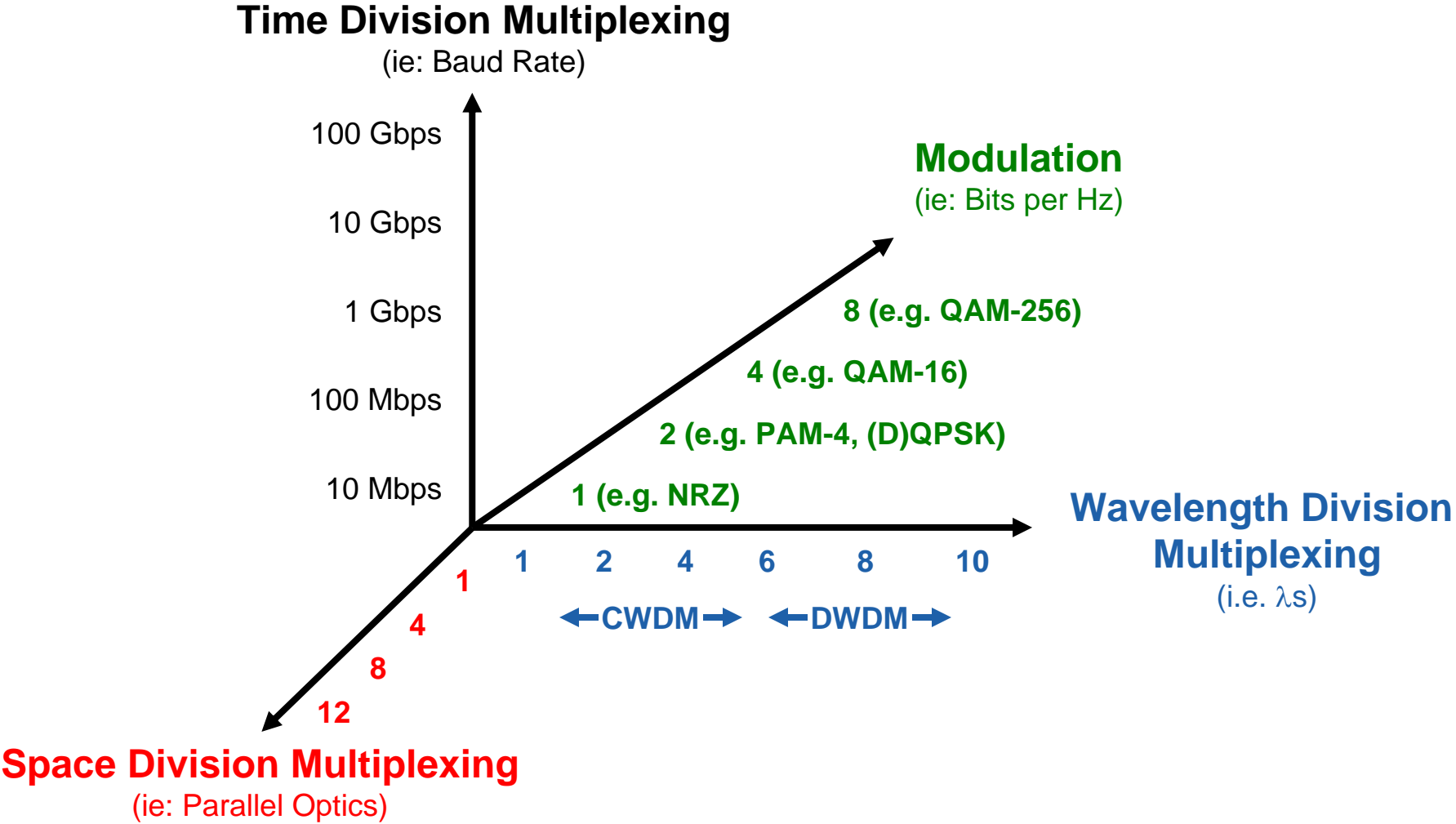
IEEE HSSG Interim  
September 20, 2006

# Agenda

---

- Technological Approaches to 100 Gb/s
- > 10 GbE Architectural Requirement
- 802.3ah Clause 61
- MultiPHY Ethernet Architecture

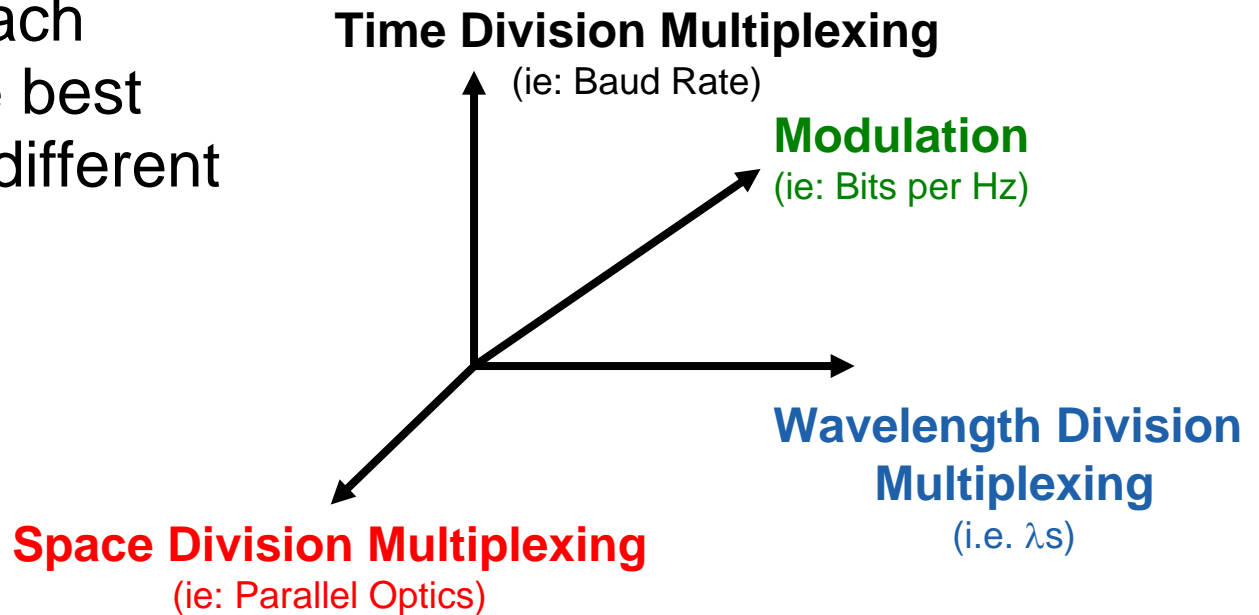
# Technological Approaches to 100 Gb/s Links



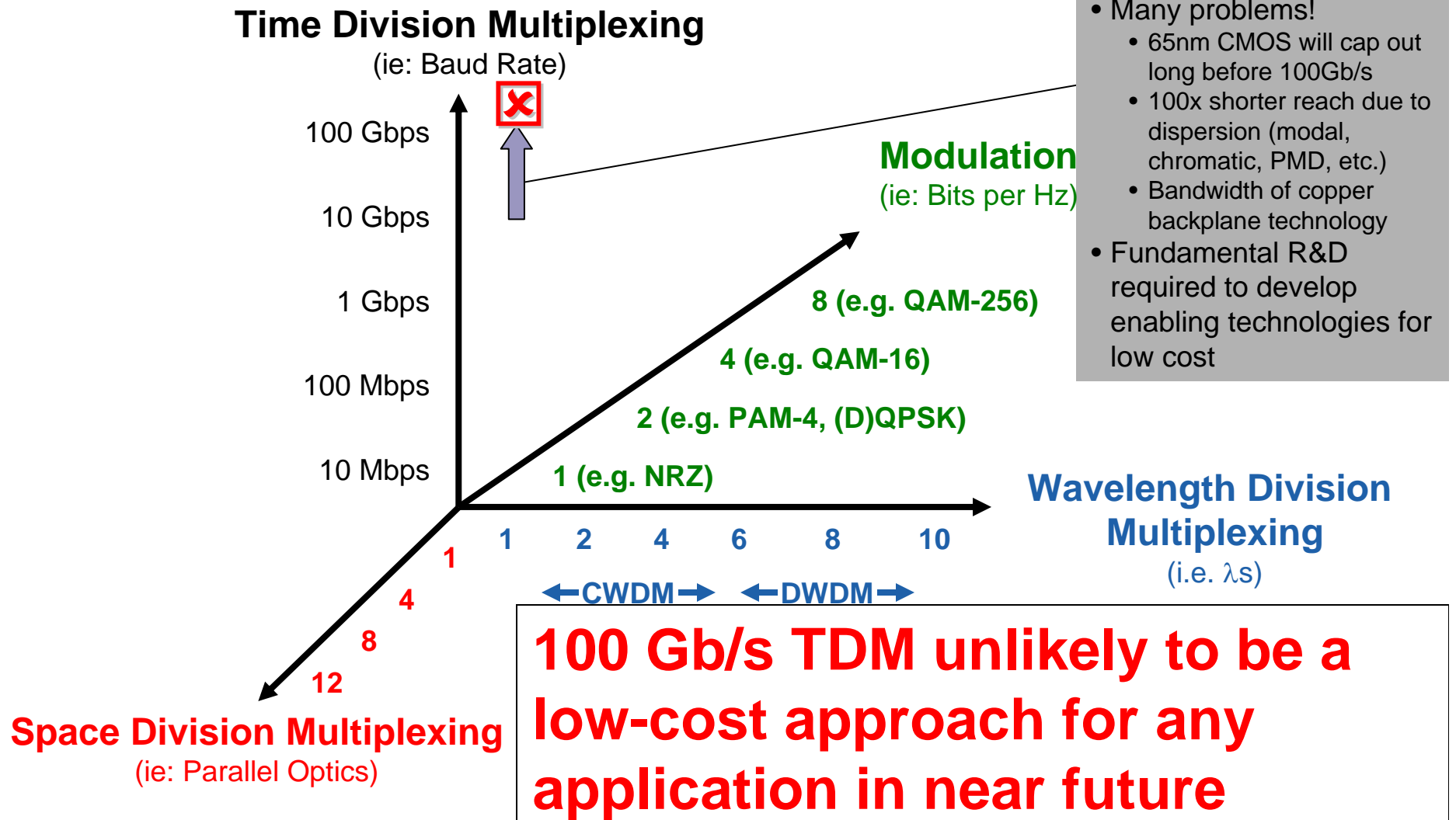
# Which Ethernet Application?

- Ethernet is used today for many applications over different distances
  - Distances > 100m primarily use optical technologies
- Performance for each application may be best advanced using a different approach

Telecom Application Class	Translation	Reach (km)
Very Short Reach (VSR)	Intra-Room	0.1-0.3
Short Reach 1 (SR-1)	Intra-Campus	2
Short Reach 2 (SR-2)	Metro Access	10-15
Intermediate Reach (IR)	Metro Core	40
Long Reach (LR)	Regional	100
Very Long Reach (VLR)	Long-haul	N x 100

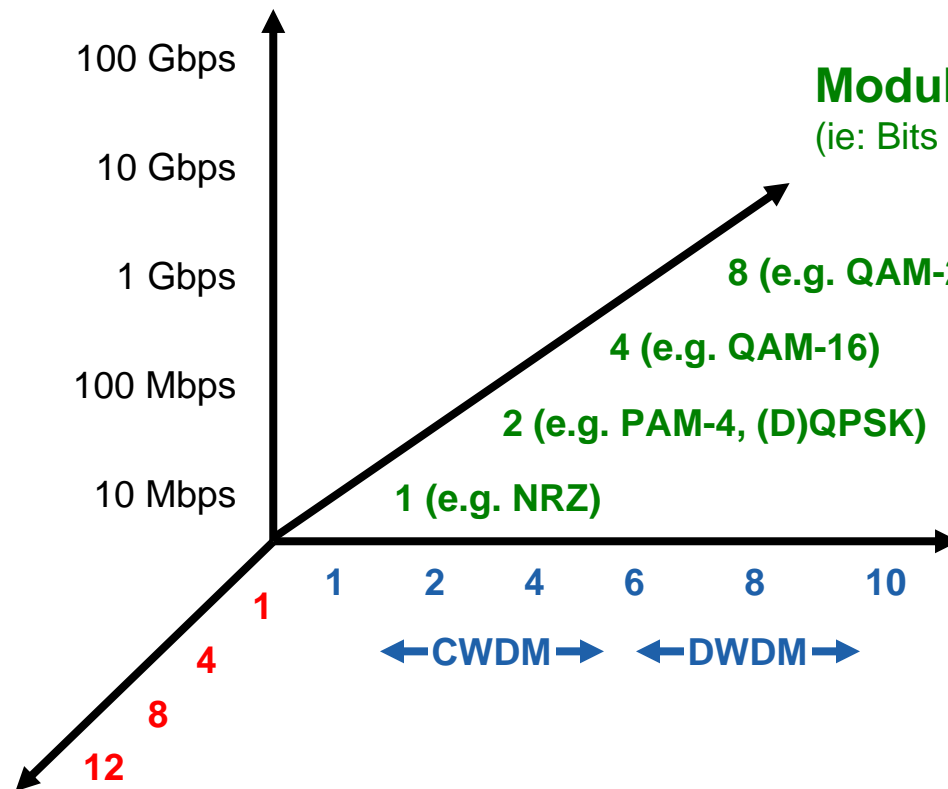


# Scaling Beyond 10 Gbps: TDM



# Scaling Beyond 10 Gbps: Modulation

## Time Division Multiplexing (ie: Baud Rate)



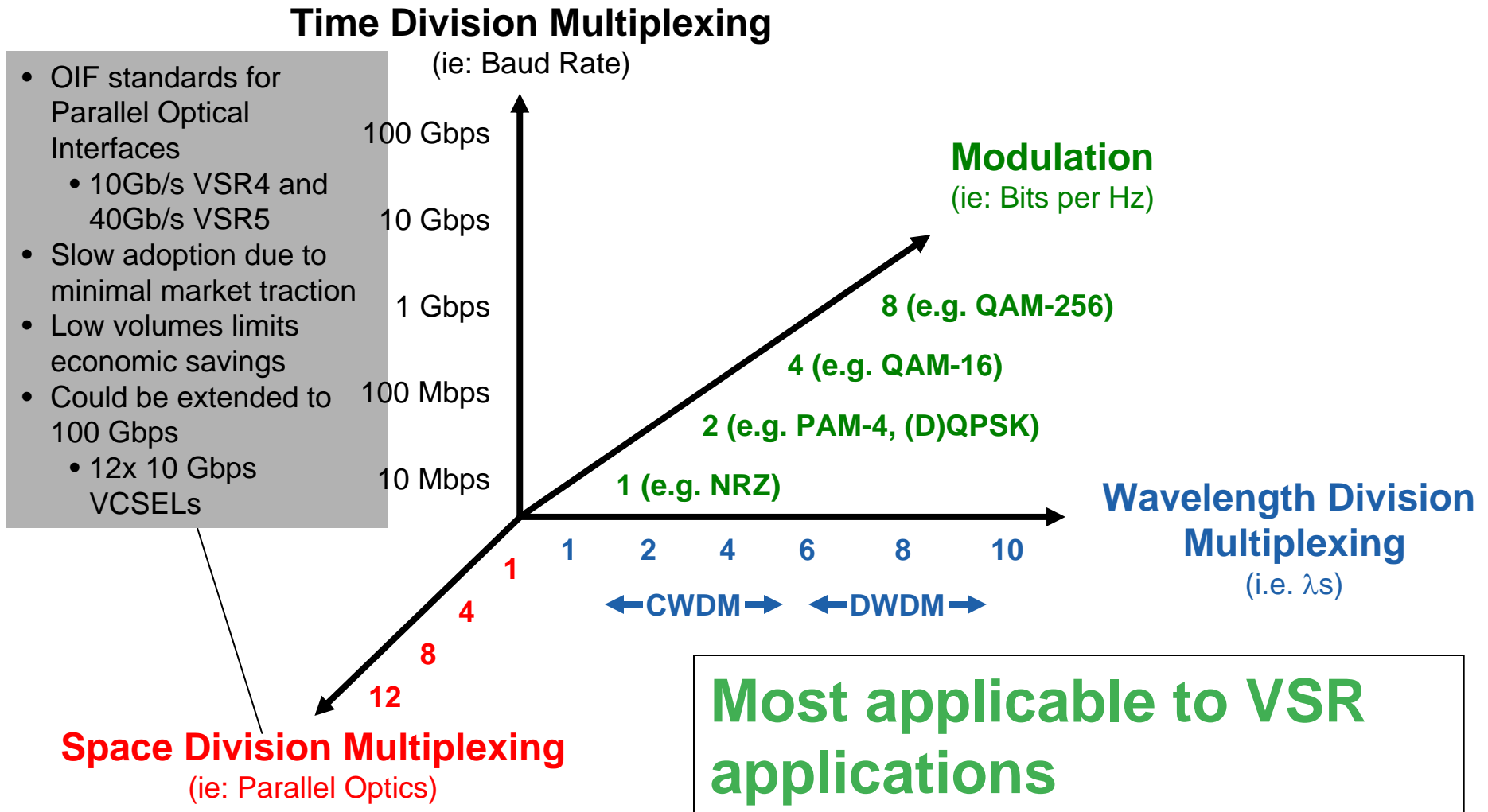
- Digital Communication theory is well-established
- Proven technology for copper technologies 1000BASE-T, DSL, Cable Modems, etc.
- Limited use with optical technology
- May be used in conjunction with other approaches

## Wavelength Division Multiplexing (i.e. $\lambda$ s)

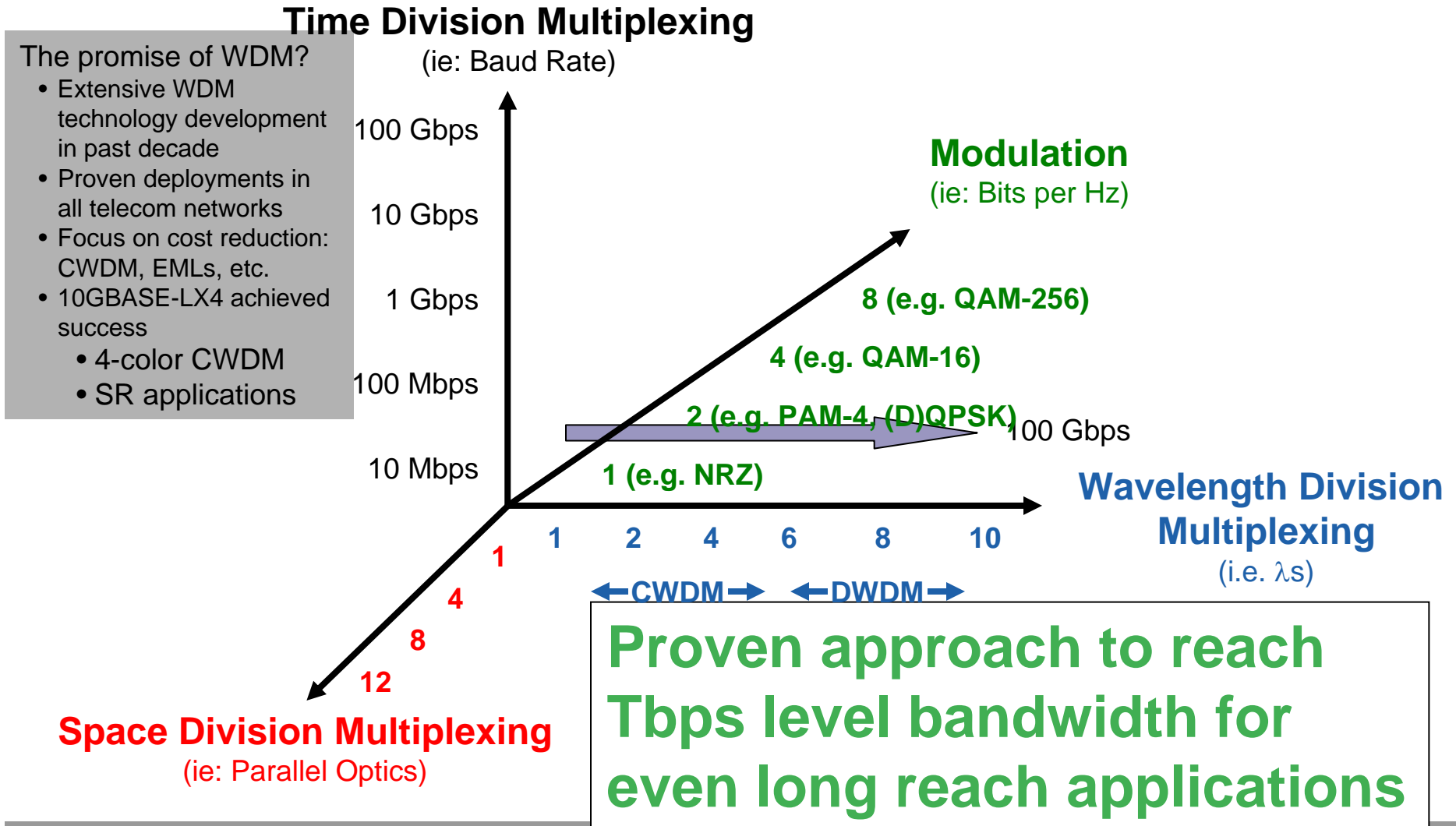
## Space Division Multiplexing (ie: Parallel Optics)

**Never applied to a high-volume optical standard**  
**Difficult for most applications of interest**

# Scaling Beyond 10 Gbps: SDM



# Scaling Beyond 10 Gbps: WDM





# Requirements

---

- Ability to bond multiple channels, whether SDM or WDM
- Support up to  $2^5$  to  $2^6$  channels for several years of exponential growth
- Support increasing individual channel speed over time
  - Define protocol once to scale over time
- Distribute data at level lower than “conversation” level
  - Bit, byte, word, fragment, packet level

# Additional Highly Desirable Qualities

---

- Support existing 10G PHYs
  - 10GBASE-R, 10GBASE-W, 10GBASE-X
  - 10GBASE-L, 10GBASE-S, 10GBASE-LRM, etc.
- Rapid recovery from individual channel failures
- Hitless addition and removal of channels
- Support LAG as well at higher layer for additional scalability

---

*How can we utilize SDM and WDM  
without incurring LAG's problems?*

# IEEE 802.3ae (10 GbE) Architecture

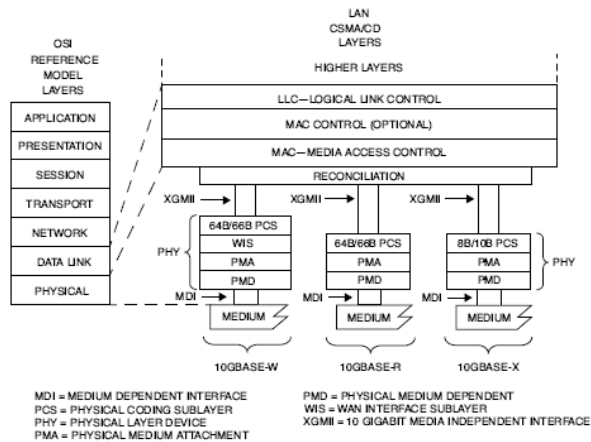


Figure 44-1— Architectural positioning of 10 Gigabit Ethernet

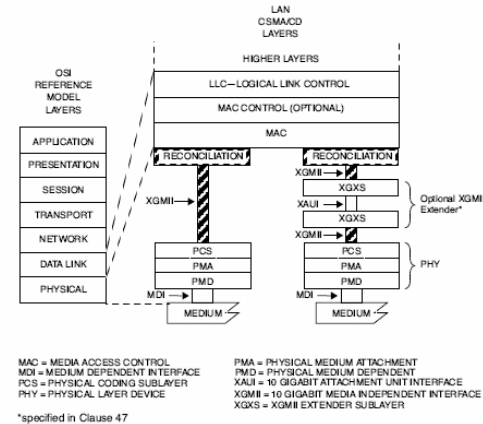
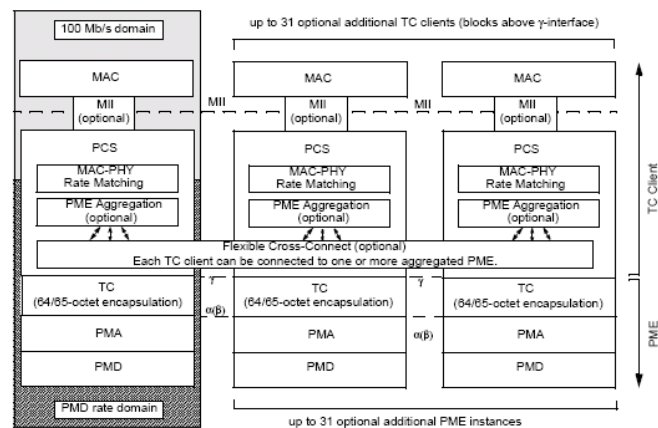
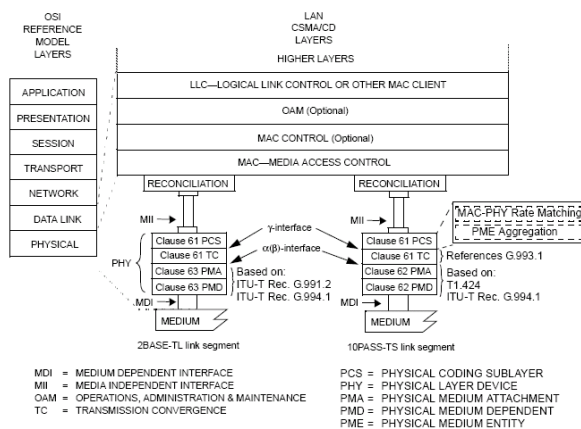


Figure 46-1—XGMII relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and the IEEE 802.3 CSMA/CD LAN model

- MAC is rate independent
- XGMII is 10G-specific
  - Could define a new CGMII for 100 Gb/s
- Straightforward to reutilize existing 10G PHYs, one per link
  - LAN PHY and/or WAN PHY can be reused
- How do we bond a number of 10G links together?

# IEEE 802.3ah (EFM) Clause 61 Architecture



- Ethernet in the First Mile (EFM) encountered the same problem at Nx 10 Mb/s
- Defined a new Physical Medium Entity (PME) Aggregation layer to inverse mux data over multiple PHYs
- Specifies a protocol virtually identical to Multilink PPP

# IEEE 802.3ah Clause 61 Fragmentation

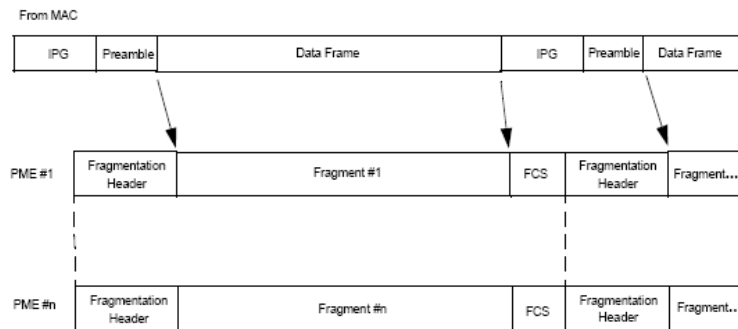


Figure 61-9—Data frame fragmentation

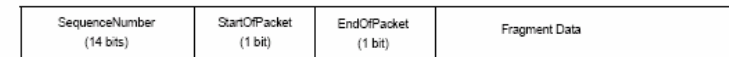
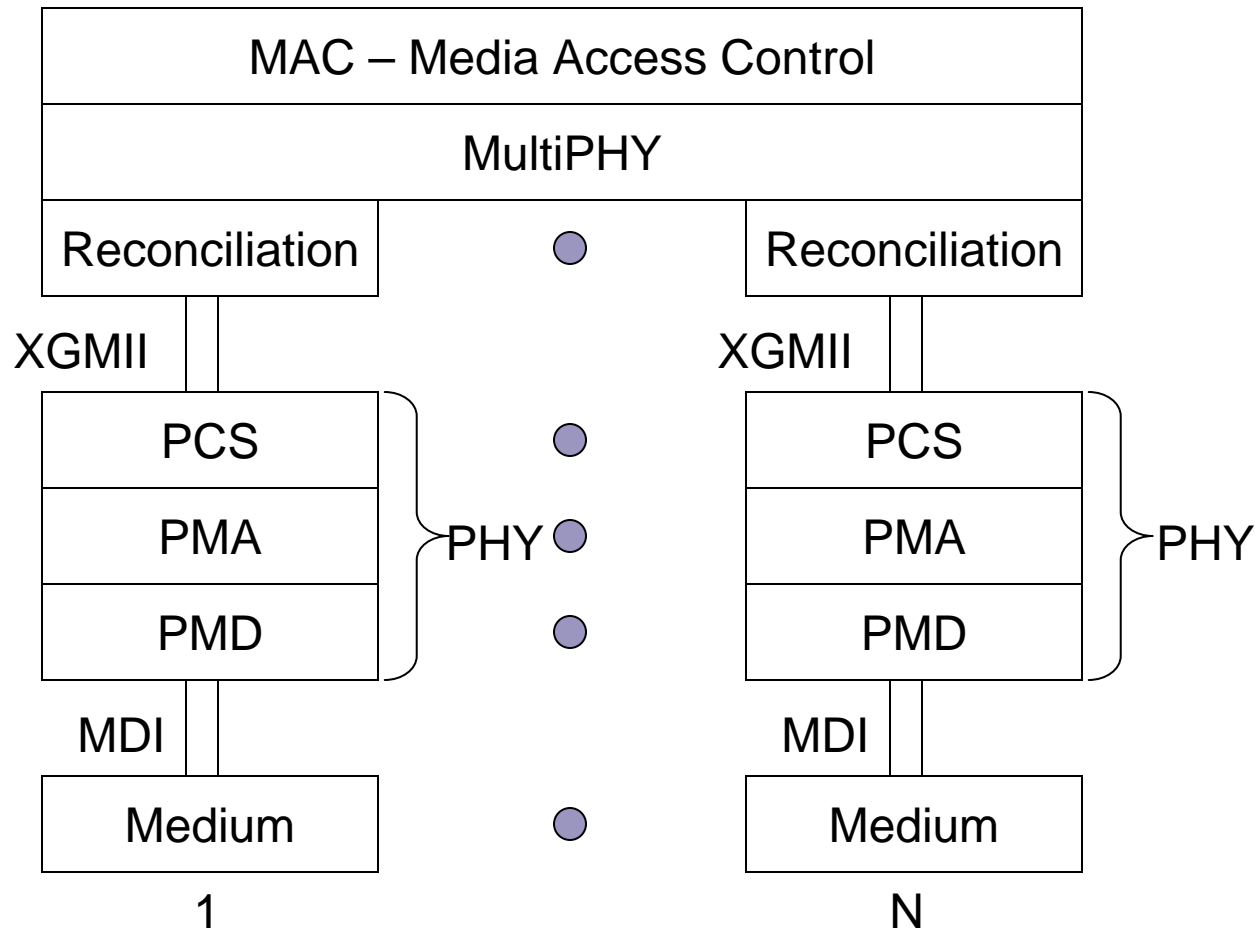


Figure 61-10—Fragment format

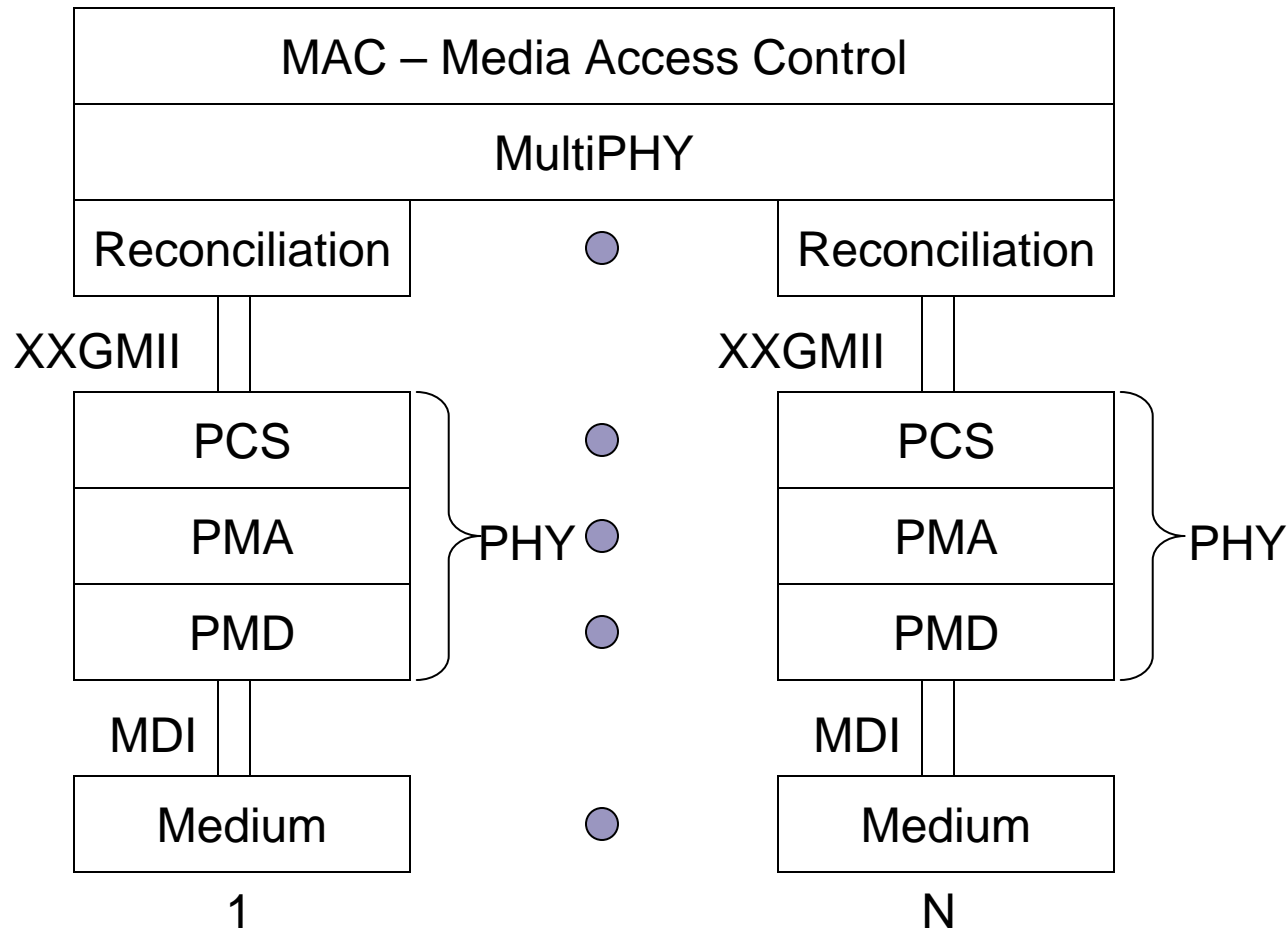
- Allows both link bandwidth *and latency* to be scaled linearly with each link added
- Supports up to 32 PMEs (320 GbE)
- Can scale link 1 PHY at a time
- Fault tolerant: individual links can fail without taking entire group down

# MultiPHY Ethernet – Nx 10G



- Aka *Aggregation at the Physical Layer (APL)*

# MultiPHY Ethernet – Nx 20G



- Aka *Aggregation at the Physical Layer (APL)*



# MultiPHY Benefits

---

- Preserves all existing 10G PHYs
- Scales into future as individual channel speeds increase
- Orthogonal to and lower level than LAG
- Support up to  $2^5$  to  $2^6$  channels with incremental hitless growth

# Summary

---

- Several off-the-shelf technologies allow parallel 10 Gb/s links *today*
- Off-the-shelf protocol solutions available to bond parallel channels into higher speed virtual link
- The time is *NOW* to standardize a physical link aggregation technology to take us to 100 GbE and beyond!



Thanks!

Drew Perkins  
[dperkins@infinera.com](mailto:dperkins@infinera.com)  
408-572-5308