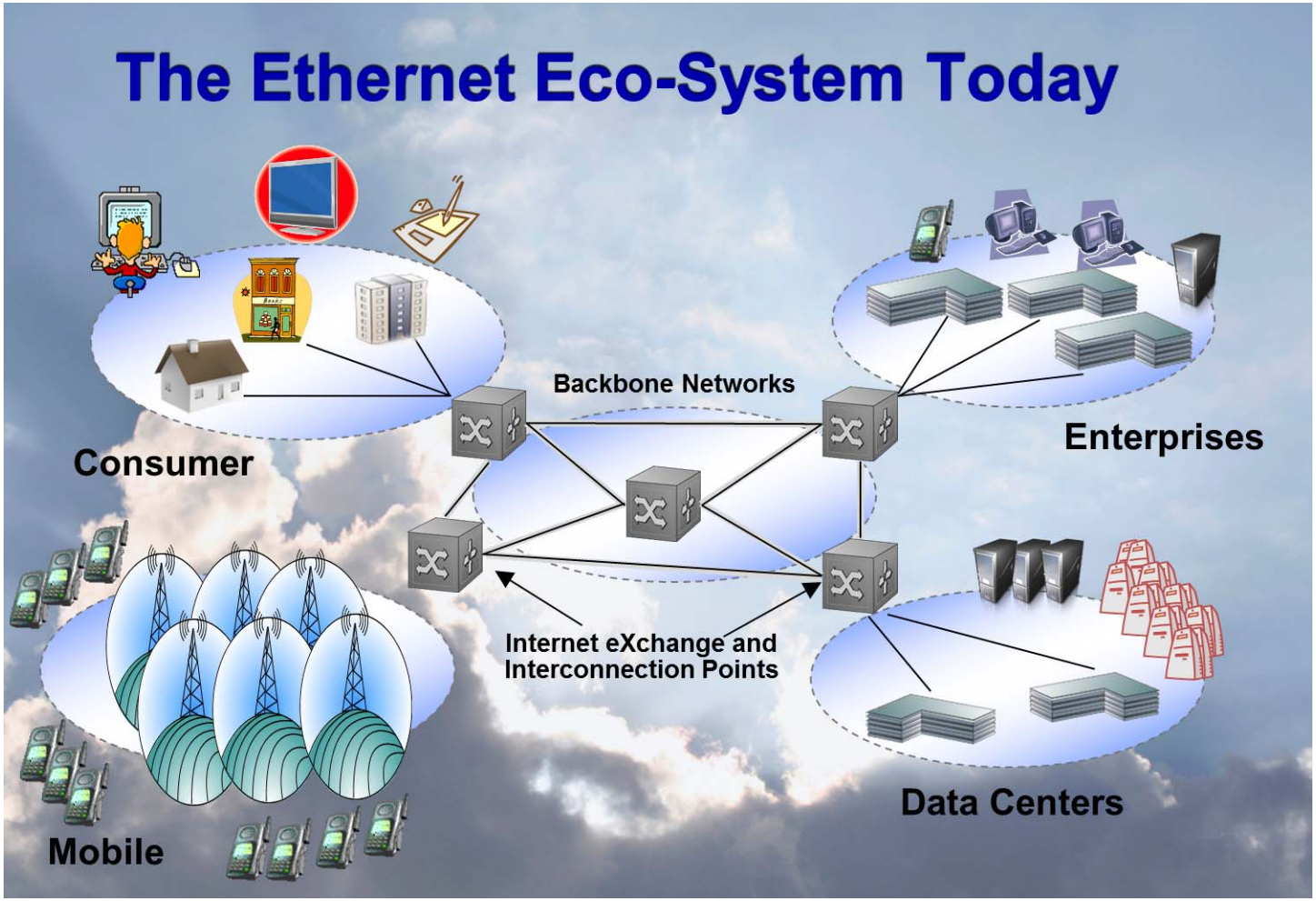
Ethernet's Rate Roadmap
SC13 Birds Of a Feather Session

Nathan Tracy, TE Connectivity, Moderator
Panelists:
- Chris Cole, Finisar
- Scott Kipp, Brocade
- Arlon Martin, Mellanox

Disclaimer

- The views WE ARE expressing in this presentation are our own personal views and should not be considered the views or positions of the Ethernet Alliance.

What does Ethernet mean to you?

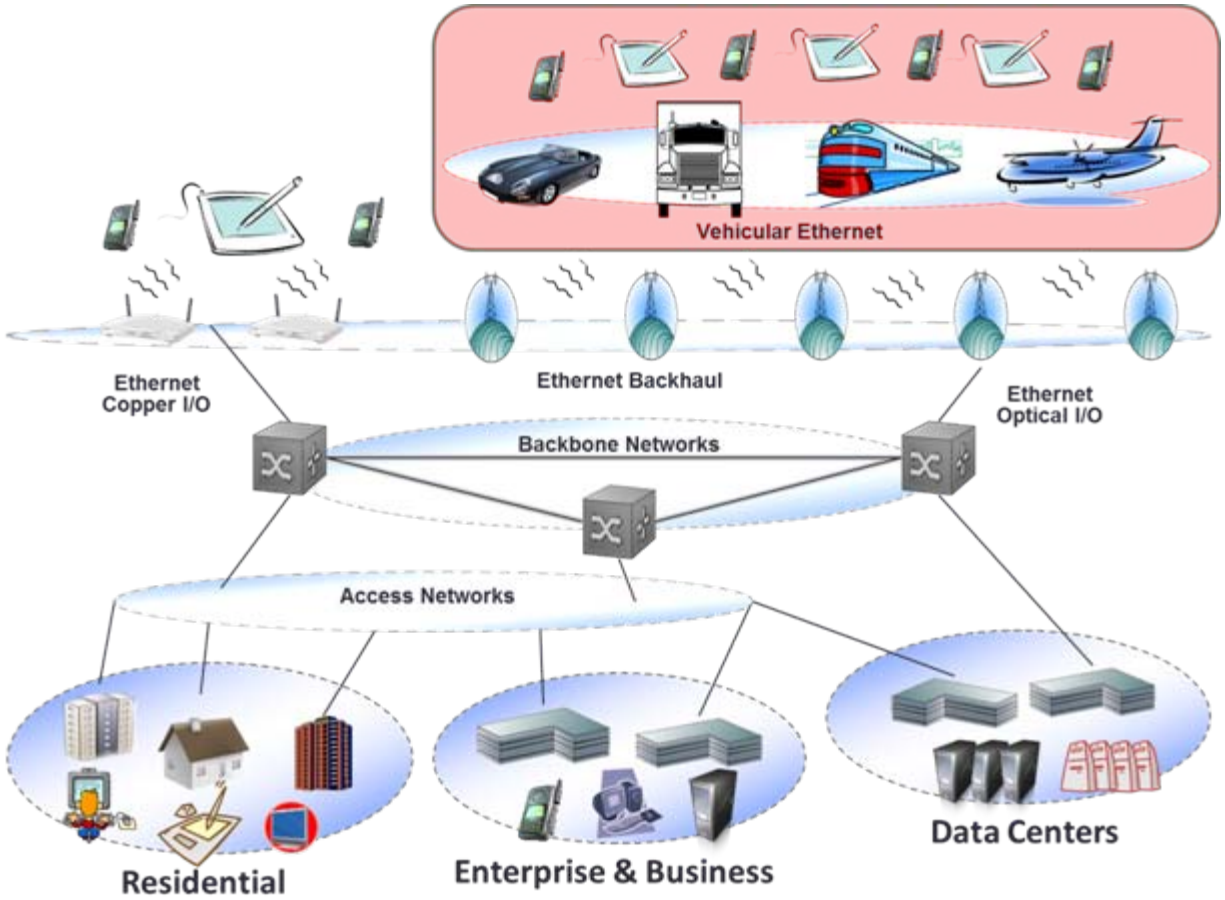


From IEEE 802.3 400G CFI

Historical perspective

- Ethernet has grown to meet market needs from 10 Mb/s to 100Mb/s to 1Gb/s to 10Gb/s
- Then 40Gb/s and 100Gb/s came along
- 400Gb/s is now an established IEEE study group
- Ethernet was originally developed to connect computers in a building….but now,
- Now Ethernet wants to be available across all media and reaches
- Ethernet is pervasive because it evolves to meet market demands, is backwards compatible, and it's large market attracts the technical "horsepower" to continue

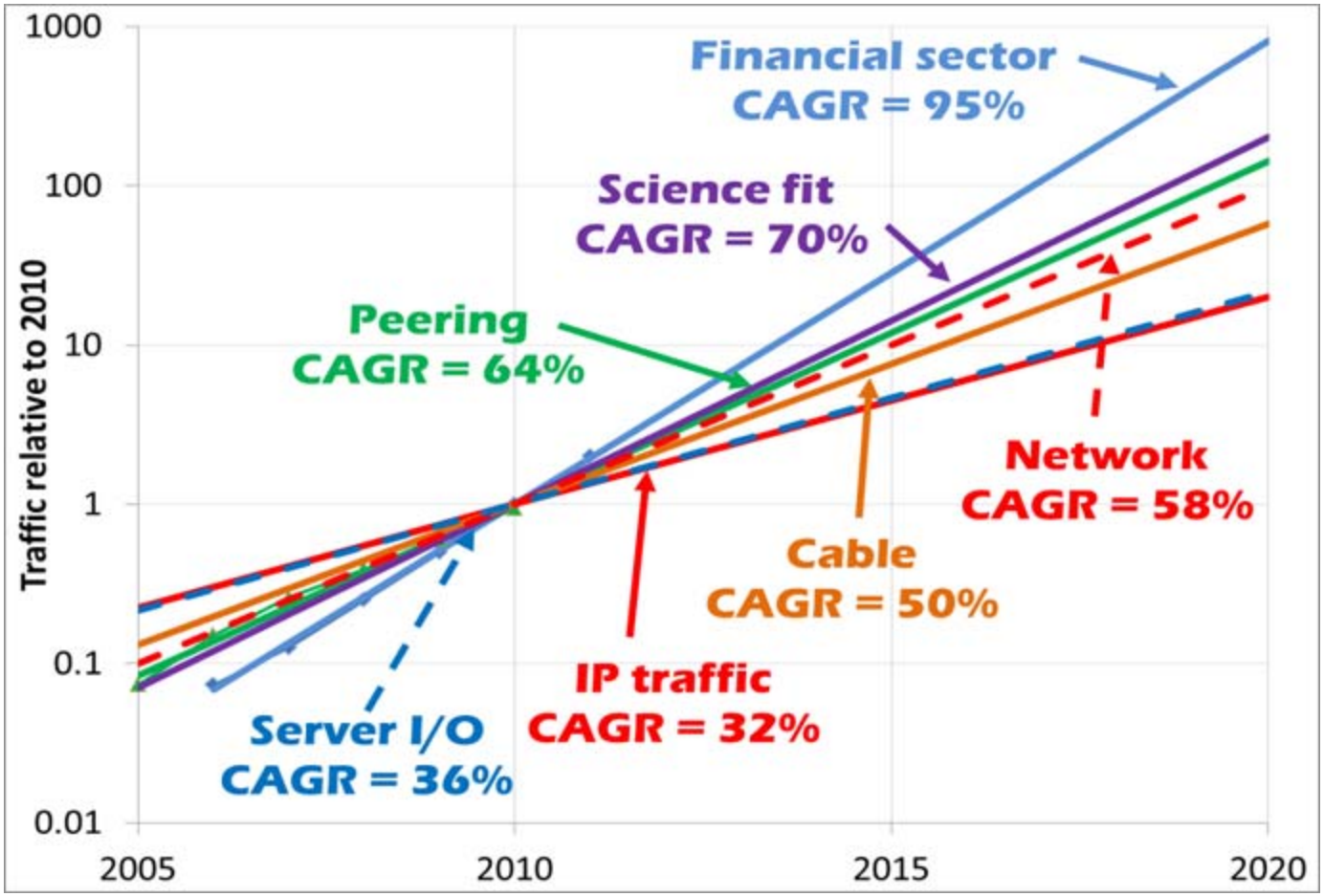
Where are we going?



Source: http://www.ieee802.org/3/400GSG/public/13_11/carlson_400_01_1113.pdf

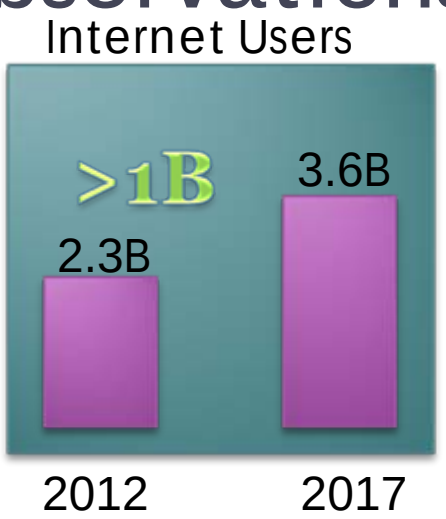
New applications contribute to huge traffic growth

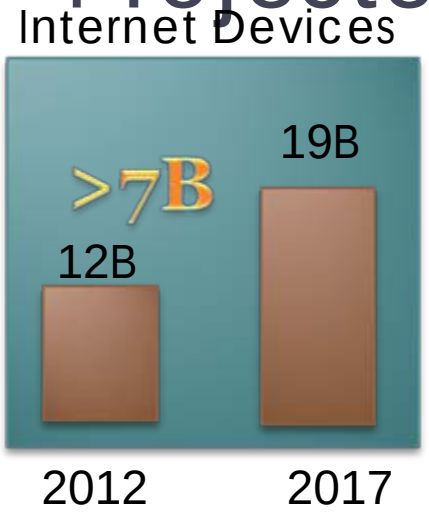
IEEE 802.3 Band Width Ad Hoc (BWA) Findings

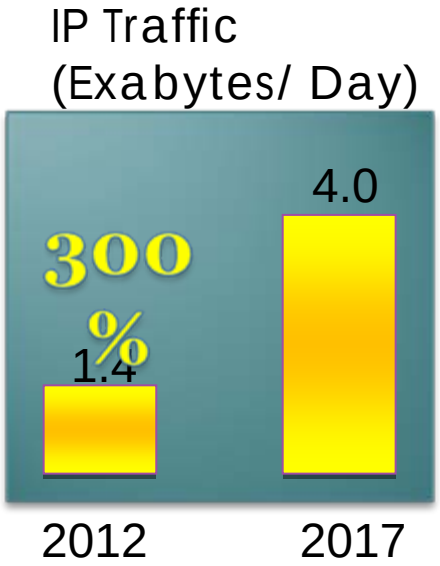


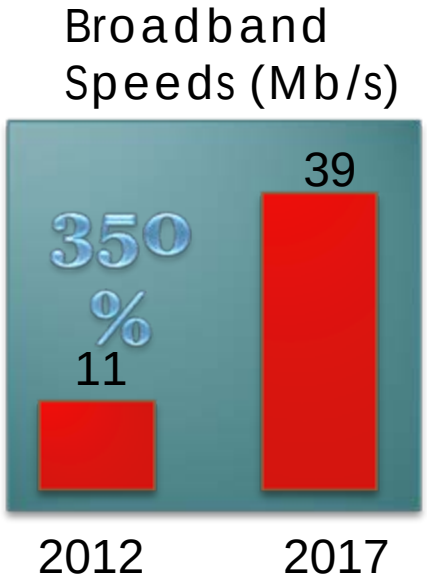
Source: http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf

Observations – Projected Growth









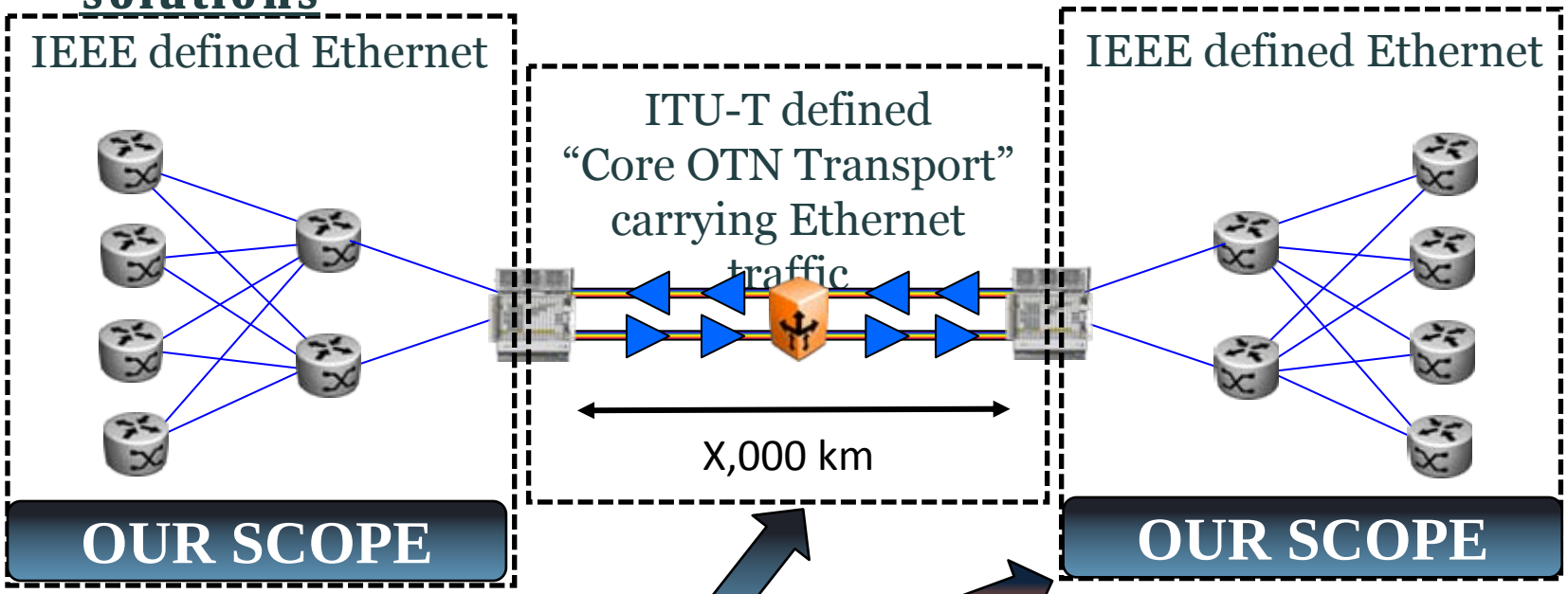
- YouTube Statistics *
  - Every minute - 72 hours of video are uploaded
  - Each month 4 billion hours of video are watched
  - *25% of global views come from mobile devices*
  - *Traffic from mobile devices tripled in 2011*

Cisco VNI data

*Youtube Statistics, http://www.youtube.com/t/press_statistics, data obtained Feb 15, 2013.

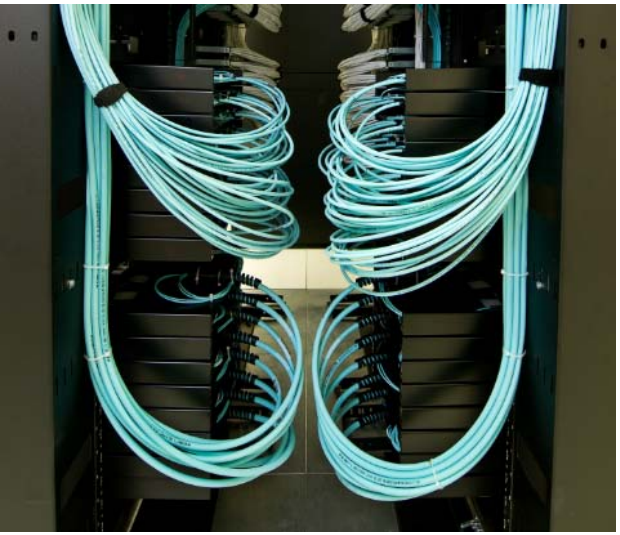
Ethernet and ITU-T

- **At highest rates Ethernet is becoming dominant traffic for client- and line-side**
  - **"Core OTN Transport" is defined by the ITU-T**
- **Interdependent problems, but not interchangeable solutions**

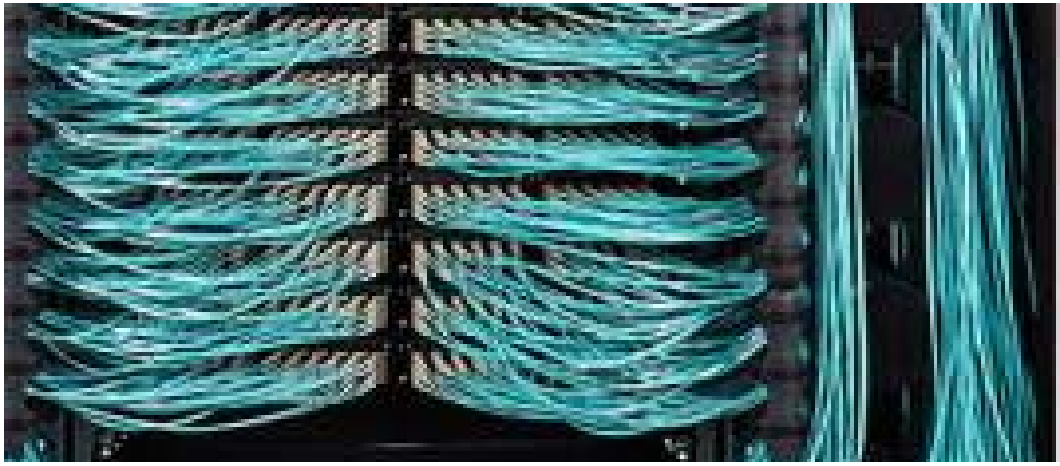


Interconnect trends in the data center

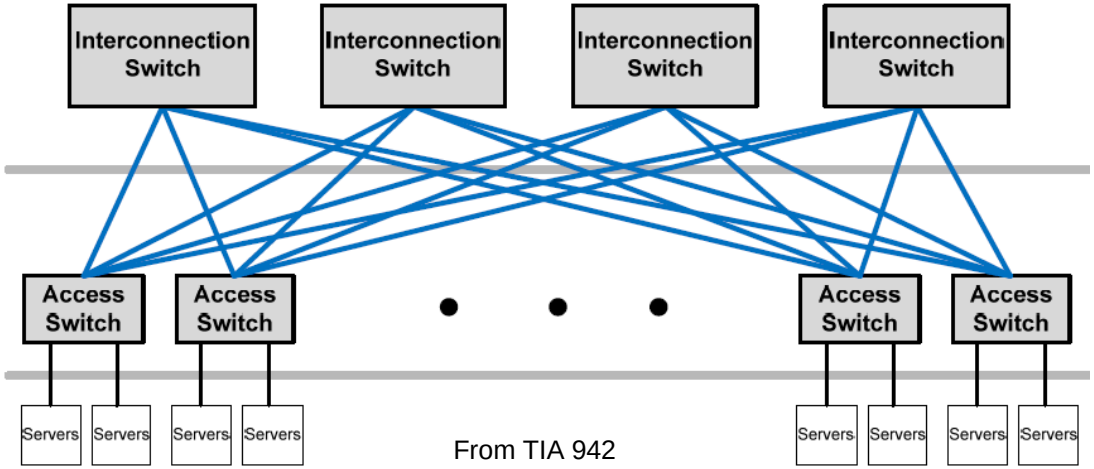
Conventional data center



ToR data center



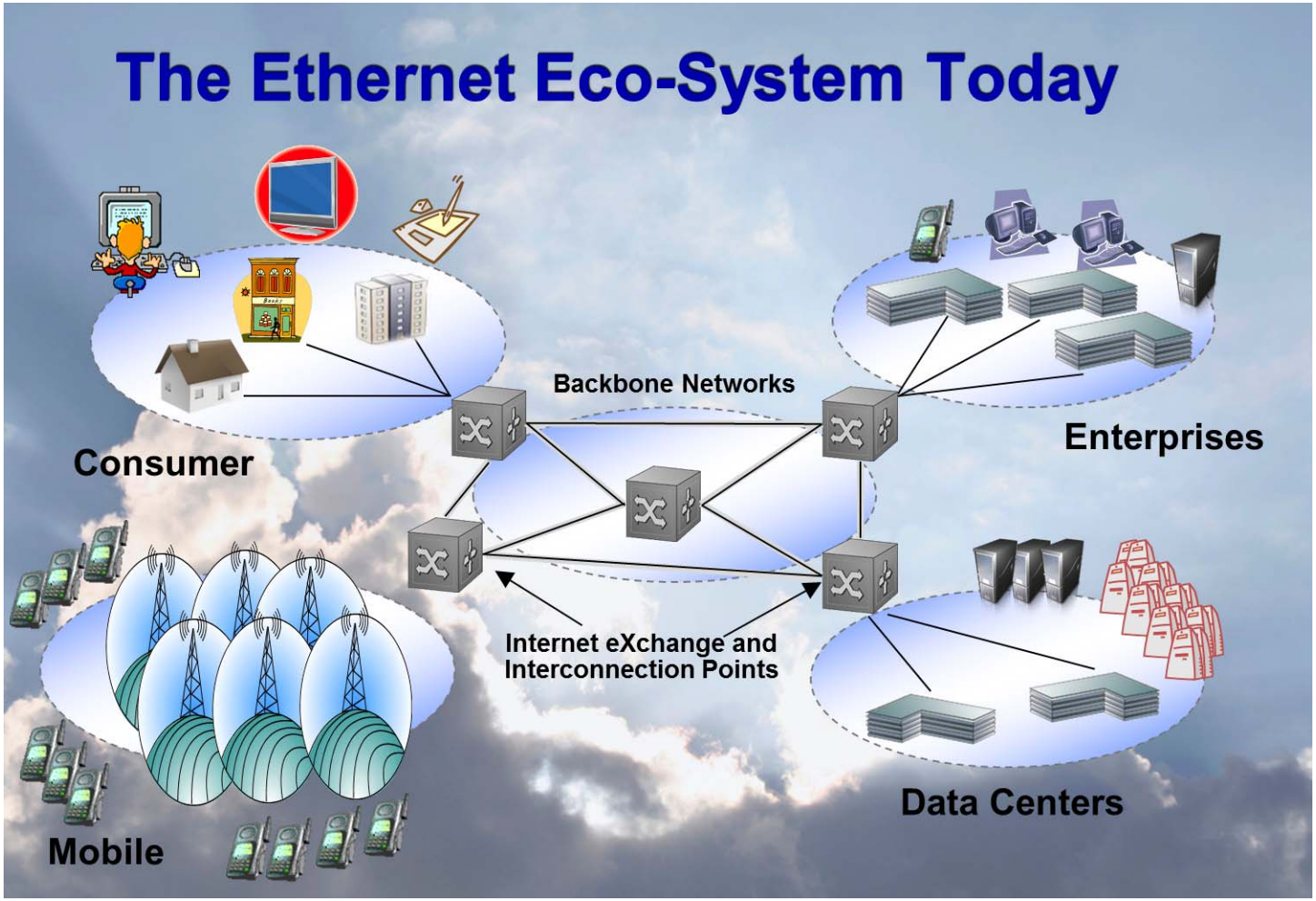
Leaf and spine data center has even higher density



From TIA 942

Photos from TE Connectivity

Ethernet will continue to grow

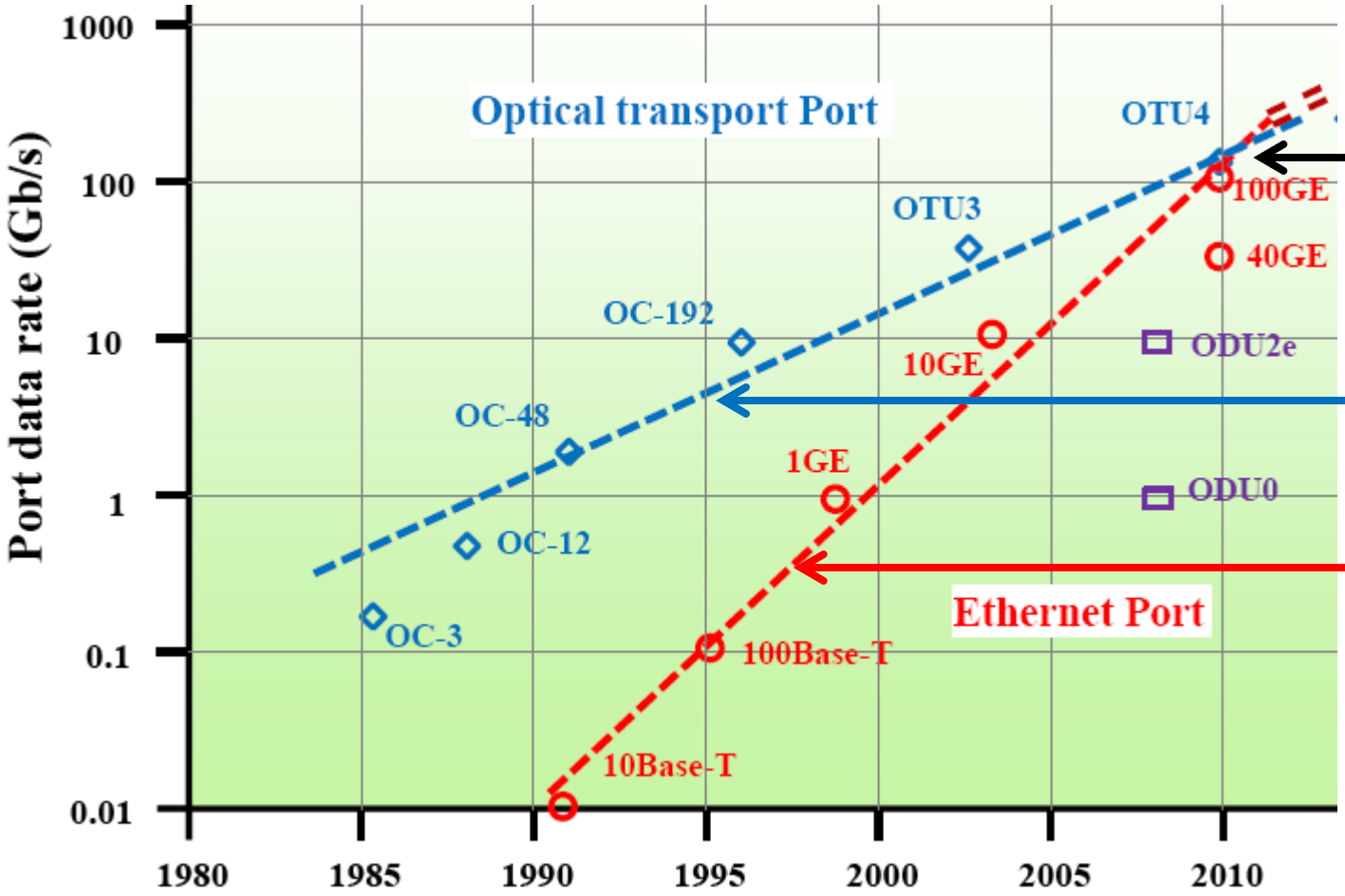


Things to think about

- What is the market need? Speed, reach
- What is the market capability? Technical, economic, etc
- Timing?
- Specific needs of HPC applications?
- When?
- How do we pay for it? Volume? At what price?

High Speed Technology
Chris Cole, Finisar

Historical Rate Increases



M. Vissers, “OTN Evolution”, OIF Q4’13 Meeting, Nov. 1, 2013

Future Rates

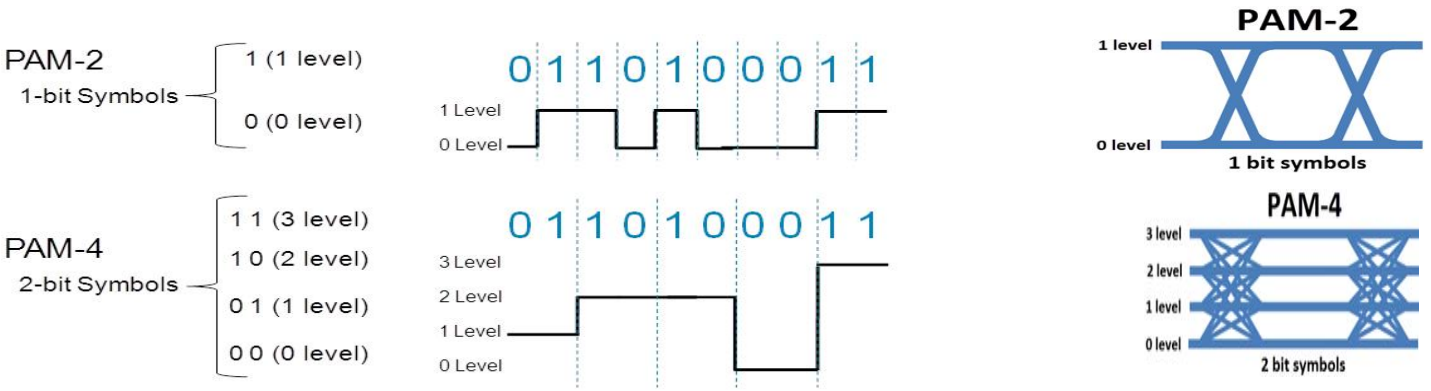
- 10x Ethernet rate increases are unlikely
- 4x OT & Ethernet rate increases are likely
- 400Gb/s: Next Rate after 100Gb/s
  - 400GbE (in IEEE 802.3 Study Group)
  - OTU5 (w/ 200Gb/s & 300Gb/s in ITU-T SG15)
- 1.6Tb/s: likely rate after 400Gb/s
  - 1Tb/s is only 2.5x increase from 400Gb/s
  - 4x is min. increase that justifies huge R&D
  - Different than increase to 100GbE which was defined before 40GbE

Bit Rate Parameters

- Symbol rate (Baud)
- Number of channels (fibers or wavelengths)
- Number of bits/symbol (modulation order)
- 10G & below optics are 1 channel NRZ
- 40G & 100G (25GB) optics are 4 channel NRZ
- 40GB NRZ (Serial) and 100G 2x50GB NRZ are future low cost alternatives
- 400G options: 25GB * 16 * 1(NRZ)
  50GB * 8 * 1 (NRZ)
  100GB * 4 * 2(bits/symbol)

Higher Order Modulation (HOM)

- At 1.6Tb/s and above HOM will be a must
- HOM is under industry investigation for 100Gb/s and 400Gb/s Ethernet applications



- Advanced HOM is used for 100Gb/s OTU4 :
  - Dual Polarization QPSK (Coherent)

400Gb/s Applications

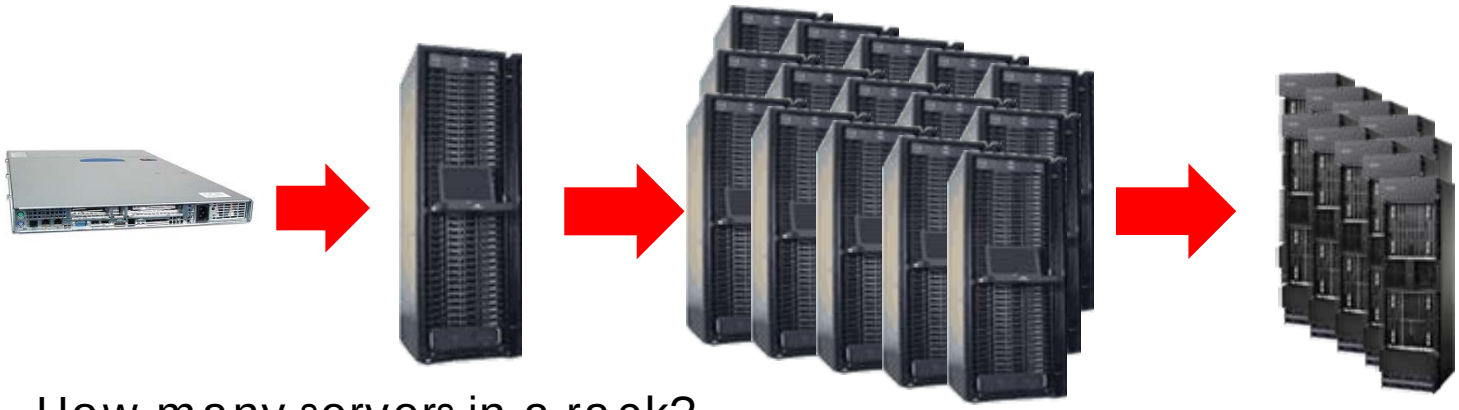
Scott Kipp, Brocade

Intra-POD Bandwidth

Each server producing 10-80 Gb/s

Rack of 40 servers producing 0.4 -3.2 Tb/s

Each cluster of 25 racks (1,000 servers) producing 10-80 Tb/s



Cluster Switches /Routers

How many servers in a rack?

I/O per server (Gb/s)	10	20	40	80
Servers / rack	40	40	40	40
Bandwidth / rack (Gb/s)	400	800	1600	3200

How much bandwidth from Top-of-Rack switch to Cluster switch?

100G? 200G? 400G? 800G?

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf

Inter-POD Bandwidth

Cluster Switches /Routers

How much bandwidth from Cluster to Core?

Data Center 1	Containers Containers	
Cluster 2	Cluster 3	Cluster 4
Cluster 1	CORE	Cluster 5

Is 100 meters long enough within a Cluster?

Is 500 meters long enough within a data center?

How much bandwidth to WAN?

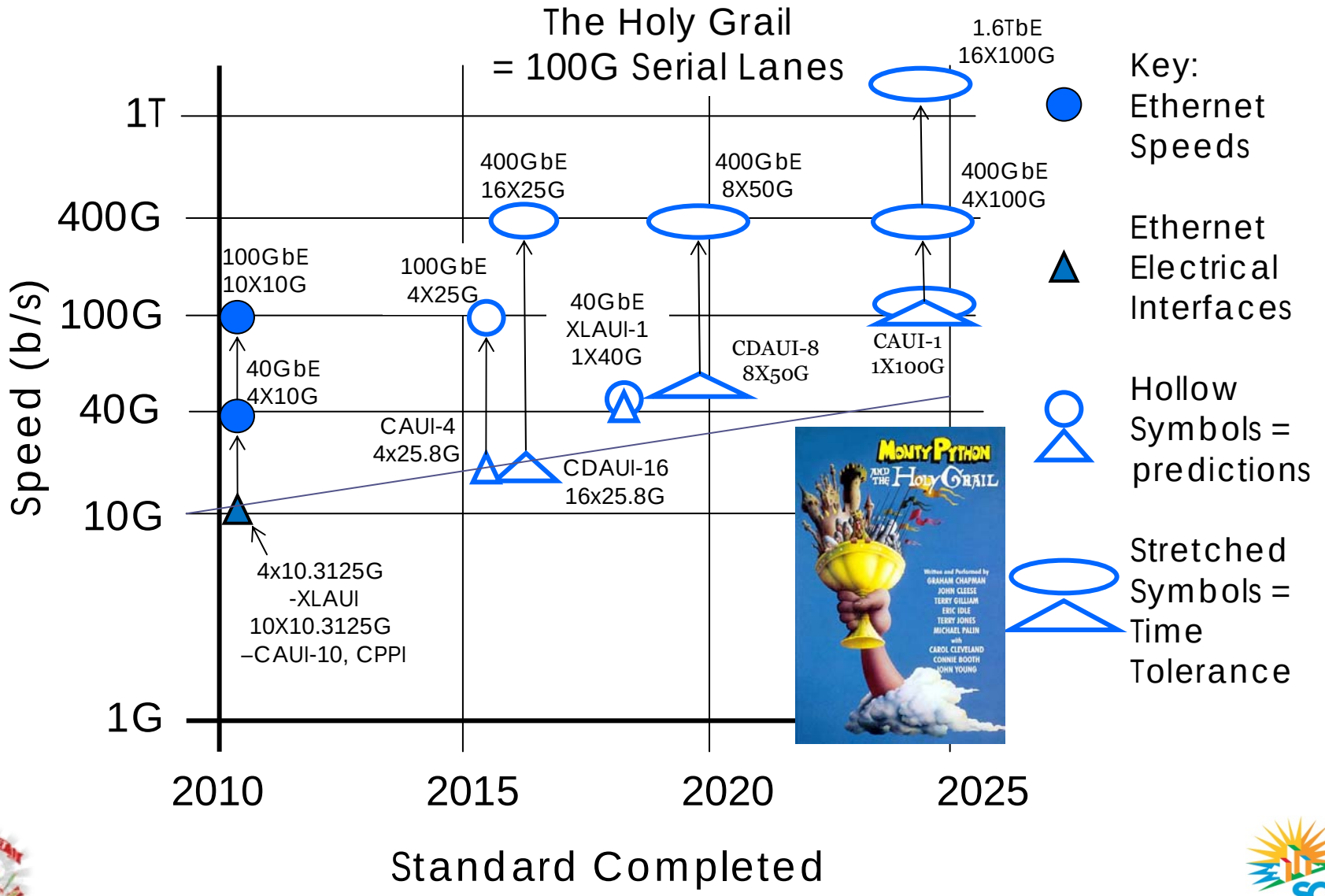
Core Switches /Routers

Source: http://www.ieee802.org/3/ad_hoc/bwa/public/may11/kipp_01_0511.pdf

IEEE 802.3 400Gb/s Study Group Objectives

- IEEE 802.3's 400Gb/s Study Group has begin the task of determine the objectives for a future task force
- So far….
  - At least 100m over MMF
  - At least 500m over SMF
  - At least 2km over SMF
  - At least 10km over SMF

Ethernet Speeds 2010-2025



Economic Considerations
Arlon Martin, Mellanox

Economic Factors for 400GbE
How do we make money?

State of the Industry

Average Net Income Across the Supply Chain (in Millions of Dollars)

Companies	2008	2009	2010	2011	2012
Content Providers	$609	$882	$1,083	$1,271	$1,426
Service Providers	$4,628	$7,298	$8,868	$6,312	$5,862
System Vendors	$802	$797	$1,267	$905	$687
IC Vendors IC	$417	$316	$1,059	$1,080	$773
OC and Module Vendors	$19	($8)	($1)	($13)	($20)

Note: Numbers shown in red indicate a loss.

Average Net Margin Across the Supply Chain

Companies	2008	2009	2010	2011	2012
Content Providers	12%	16%	16%	14%	12%
Service Providers	6%	10%	12%	8%	8%
System Vendors	5%	6%	9%	6%	5%
IC Vendors IC	10%	9%	22%	20%	15%
OC and Module Vendors	2%	(4%)	(2%)	(6%)	(5%)

Note: Numbers shown in red indicate a loss.

Source: Published information of public companies

Slide reused courtesy of Dale Murray, Lightcounting

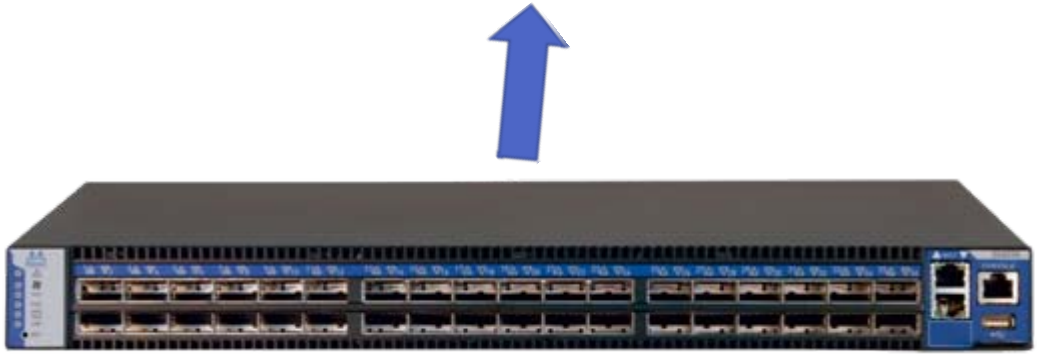
Economic Considerations for 400GbE

- A 400GbE standard could fragment an already small 100GbE market
  - 40GbE is expanding fast, but 100GbE adoption is slow
  - Anticipation of 400GbE could drain market share from 100GbE
    - Market hangs on to 40GbE until 400GbE available
- 400GbE could be extremely costly investment
  - Advanced modulation
  - WDM
  - Parallel
    - Every approach has challenges
- 400GbE might make margins even worse

For 40GbE, the volume package is QSFP; enabling TOR/Leaf/Spine switches of 1.5-2 Tb/s

- Max 1RU density is 36 QSFP ports
- Enables switches up to 2 Tb/s
- When 100G QSFPs becomes available
  - Will enable 4 Tb/s switches
  - Will be volume driver for 100GbE
  - Lower cost/bit on an end-to-end basis
  - Volumes improve margins
- 400GbE *in the right form factor*
  - Will enable new generations of switches
  - Can be an economic winner

Many Terabit Switch



SX6036 – 36 QSFP Ports Switch

SX1016 – 1.3 Tb/s: 64-Port 10GbE Switch

10G SFP+ and 40/56G QSFP pluggable transceivers provide excellent front panel density!

Need to view economics from the data center perspective

- $/bit calculated using end to end metrics
  - Pluggables, switches, fabric, and power
- 40GbE is ramping fast, because the end to end metrics are favorable
- Putting 100GbE into QSFP packages enables higher capacity switches
  - Improves the economics and volumes for 100GbE
- 400G solutions can further improve the end to end $/bit