The Case for 1.6 Terabit Ethernet

IEEE 802.3 Beyond 400 Gb/s Ethernet Study Group
Electronic May 2021 Session

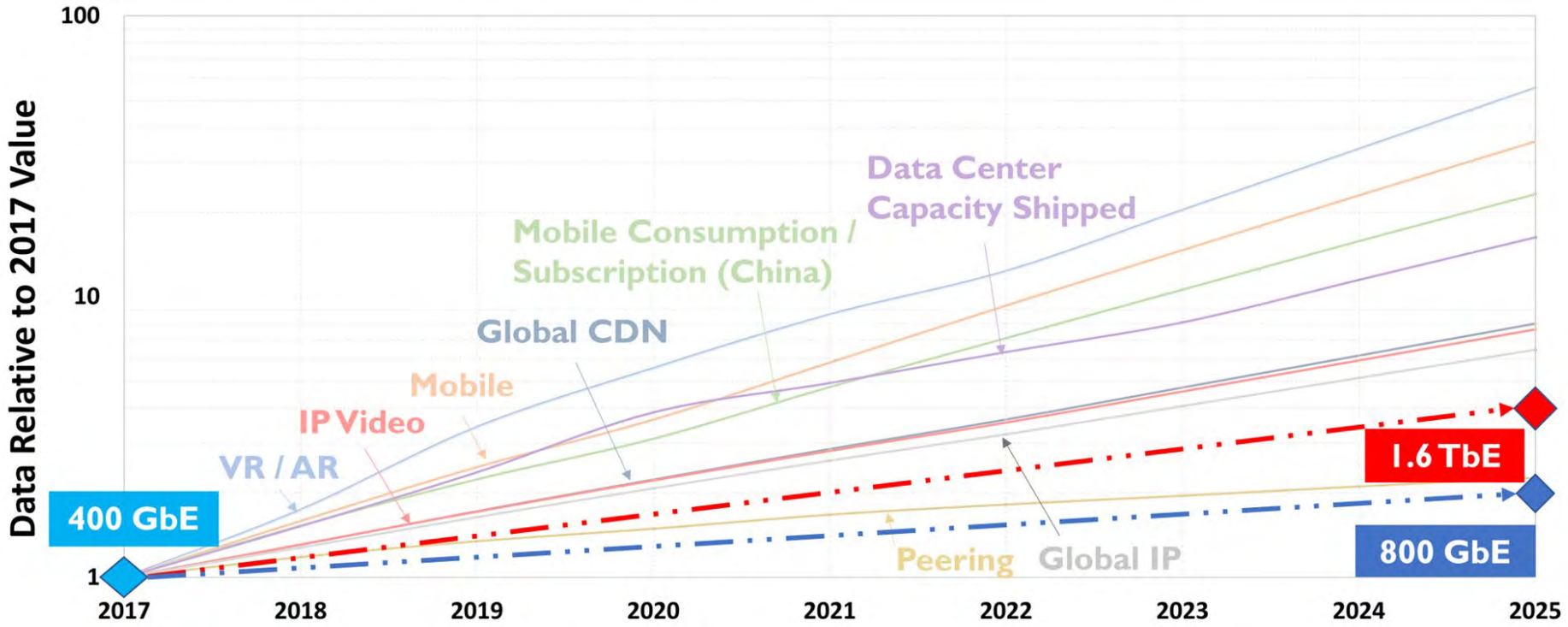
John D'Ambrosia,
Futurewei, U.S. Subsidiary of Huawei

24 May 2021 Electronic Meeting

Supporters

- Brad Booth, Microsoft
- Matt Brown, Huawei
- Cedric Lam, Google
- Sam Kocsis, Amphenol
- Kent Lusted, Intel
- Jerry Pepper, Keysight
- Scott Schube, Intel
- Jim Theodoras, HG Genuine
- Nathan Tracy, TE Connectivity
- Ed Ulrichs, Intel
- Xinyuan Wang, Huawei
- Chongin Xie, Alibaba

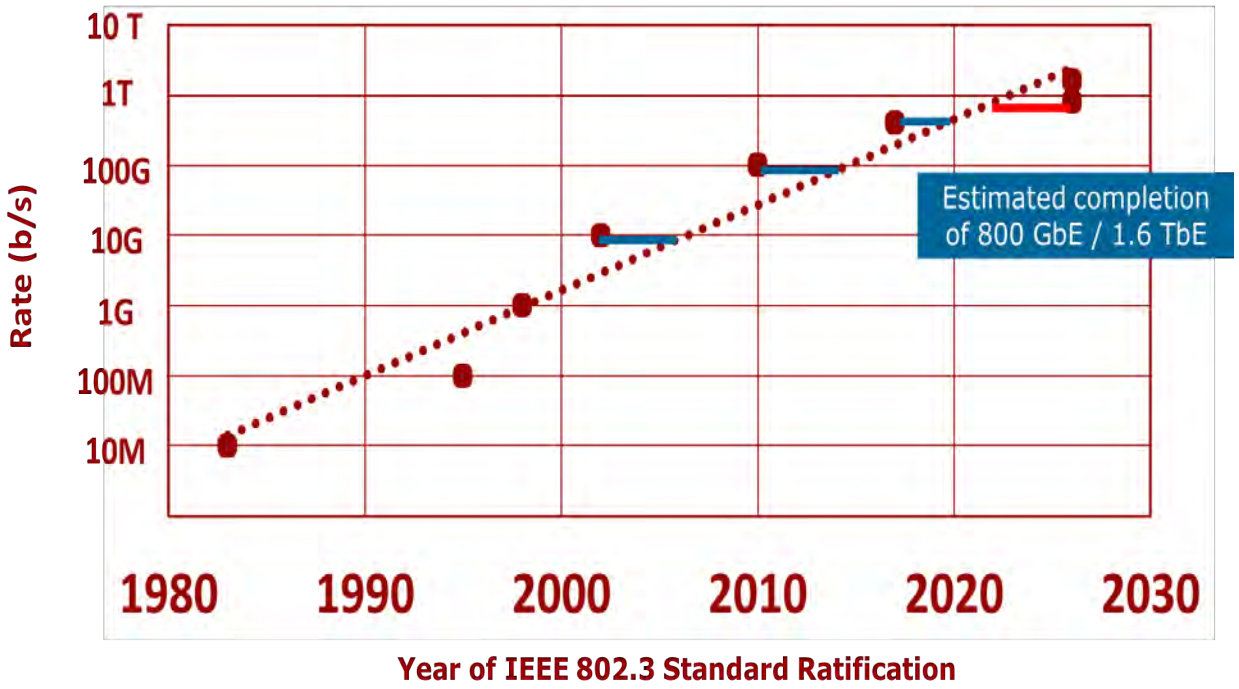
Foreword – Excerpt IEEE 802.3 Ethernet BWA

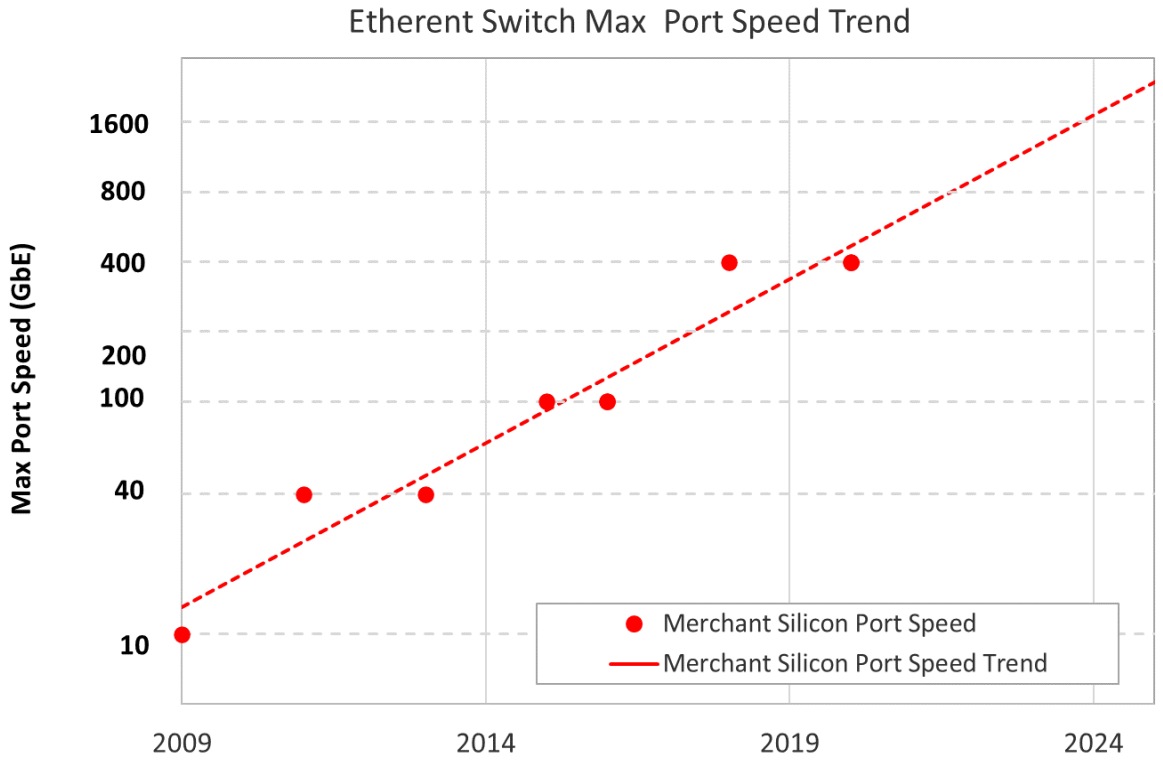


Assuming a new project to define the next rate of Ethernet begins in 2020, and takes 5 years to complete (2025), growth rate curves based on either 800GbE or 1.6TbE were also generated and compared to the submitted data. Assuming no other architectural changes in deployment, ***this overlay demonstrated a significant growth lag between 800GbE and the observed growth curves. However, the 4 × growth curve generated by a 1.6TbE solution would also lag all observed growth curves, except "Peering Traffic".*** Furthermore, all of the underlying factors that drive a bandwidth explosion, including (1) the number of users, (2) increased access rates and methods, and (3) increased services all point to continuing growth in bandwidth.

Source:https://www.ieee802.org/3/ad_hoc/bwa2/BWA2_Report.pdf

Is 800 GbE defined by IEEE 802.3 late?





Data provided by Rob Stone, Facebook

Ethernet Technology Consortium publicly released 800 GbE specification
(https://ethernettechnologyconsortium.org/wp-content/uploads/2020/03/800G-Specification_r1.0.pdf)

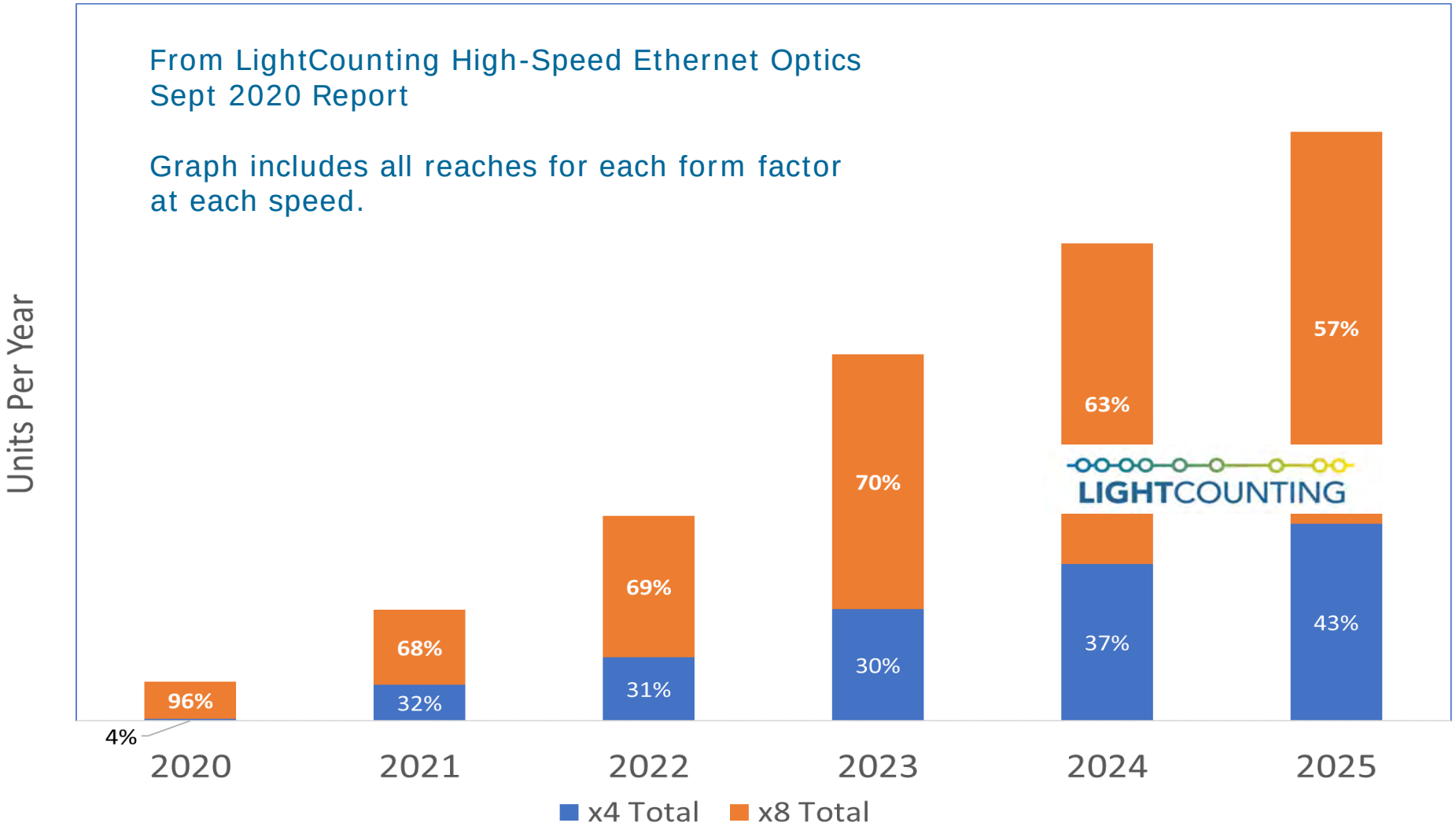
- Draft 1.0 dated March 10, 2020
- Based on 8x100G (references to various IEEE 802.3 clauses)

Industry Consortiums & Efforts - 800G related Efforts

- QSFP-DD-800G (http://www.qsfp-dd800.com/wp-content/uploads/2020/03/QSFP-DD-800-Hardware-1p0-3-6-20%20FINAL.pdf)
  - Draft 1.0 dated March 6, 2020
  - Based on 8x100G
- OSFP MSA
- 800G Pluggable MSA – 800G-PSM8 – (https://static.s123-cdn-static-d.com/uploads/2598123/normal_5f50dfed42c1e.pdf)
  - Draft 1.0 dated - August 28, 2020
  - Based on 8x100G
- COBO 8 / 16 lane solutions
- CPO

- It is recognized that an 800G port is a multi-rate port that could support 8x100GbE, 4x200GbE, 2x400GbE, or 1x800GbE

Comparison of Form Factor Adoption (200GbE / 400GbE / 800GbE)



IEEE 802.3 Ethernet Optical Landscape (based on >=50 Gb/s signaling per lane, excluding ZR)

Ethernet Rate	Signaling Rate	MMF 50m	MMF 100m	SMF 500m	SMF 2km	SMF 10km	SMF 40km
100 Gb/s	50 Gb/s		Over 2 pair				
	100 Gb/s	**Over 1 pair****	**Over 1 pair****	Over 1 λ (pair)	Over 1 λ (pair)	Over 1 λ (pair)	
200 Gb/s	50 Gb/s		Over 4 pair	Over 4 pair			
	50 Gb/s				Over 4 λ's	Over 4 λ's	Over 4 λ's
	100 Gb/s	**Over 2 pair***	**Over 2 pair***				
400 Gb/s	50 Gb/s		Over 8 pairs Over 4 pairs (4.2)				
	50 Gb/s				Over 8 λ's	Over 8 λ's	Over 8 λ's
	100 Gb/s	**Over 4 pair****	**Over 4 pair****	Over 4 pairs			
	100 Gb/s				Over 4 λ's	Over 4 λ's (6km)	
800 Gb/s*	100 Gb/s	**Over 8 pairs**	**Over 8 pairs**	**Over 8 pairs**			
	200 Gb/s			**Over 4 pairs**	**Over 4 pairs**		
	200 Gb/s				**Over 4 λ's**		
	TBD					**Over single SMF in each direction**	**Over single SMF in each direction**

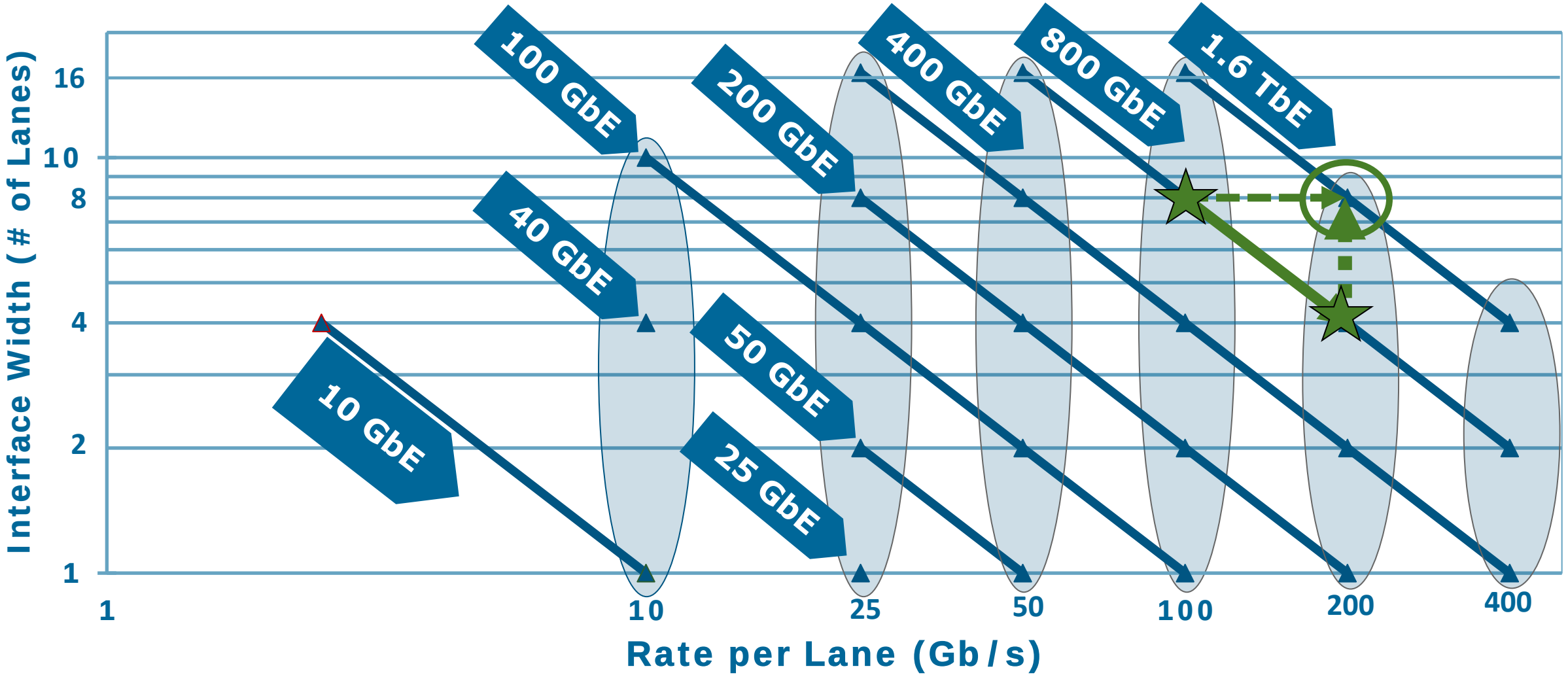
*- Objectives for B400G Study Group

** - Objectives for IEEE P802.3db

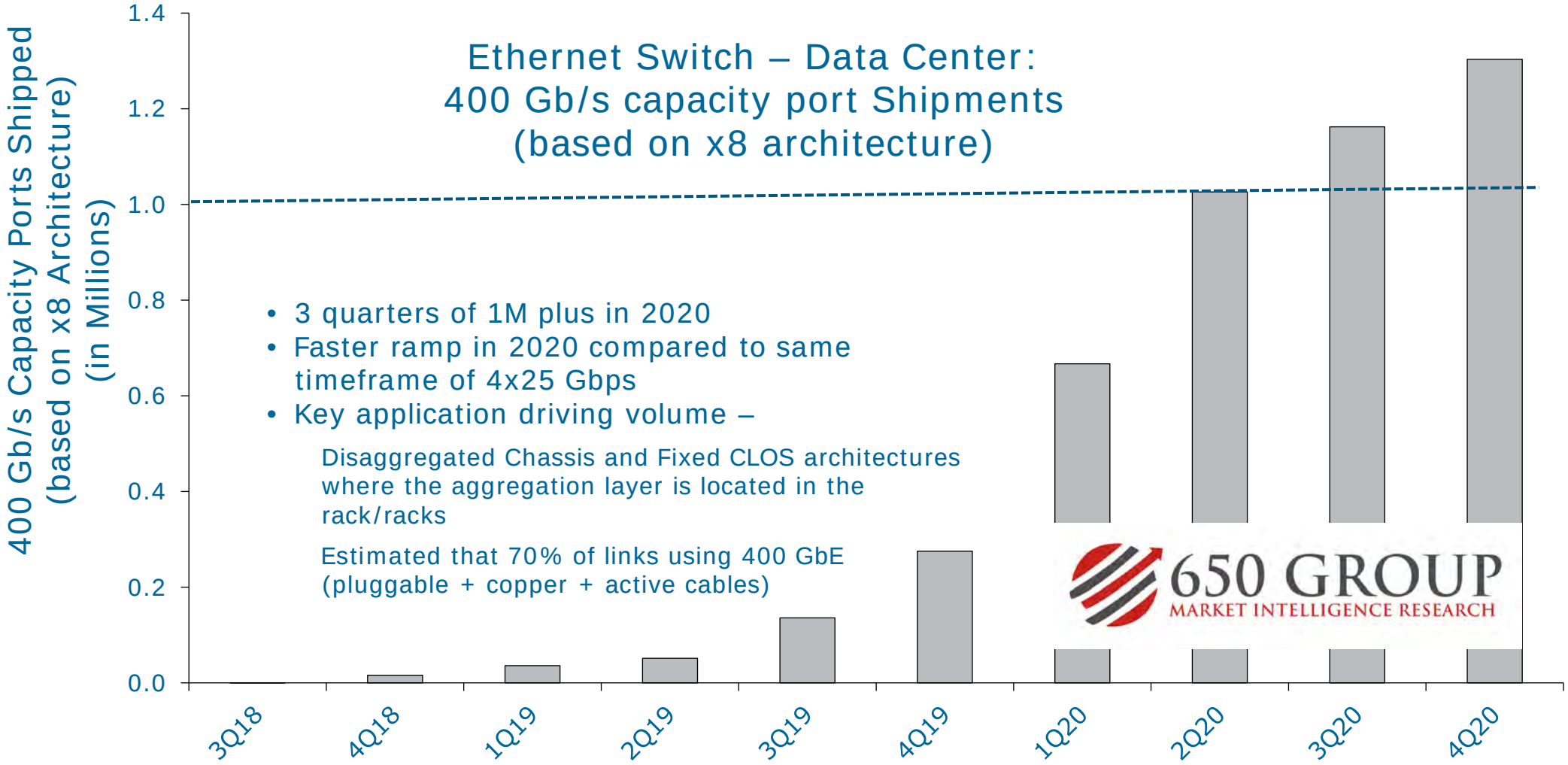
Summary of Trends

- Industry efforts at developing 800 GbE and 800 Gb/s capacity solutions have been underway prior to March 2020.
- Optical form factors based on x8 will be >50% for 200/400/800
- IEEE 802.3 existing PHYs
  - Signaling technologies are leveraged across multiple ethernet rates
  - MMF: use of 1 / 2 / 4 / 8 pairs
  - PSM: use of 1 / 4 / 8 pairs
  - SMF: use of 1 / 4 / 8 λ's
- IEEE 802.3 Beyond 400 Gb/s Ethernet
  - Optional 800 Gb/s attachment unit interfaces (AUIs) for chip-to-module and chip-to-chip applications
    - 8 lanes (assumed 8 x 100 Gb/s)
    - 4 lanes (assumed 4 x 200 Gb/s)
  - Physical Layer Objectives targeting 800 Gb/s, <= 2km
    - 8 lanes (assumed 8 x 100 Gb/s)
    - 4 lanes (assumed 4 x 200 Gb/s)

The Relationship Between Ethernet & Signaling Rates



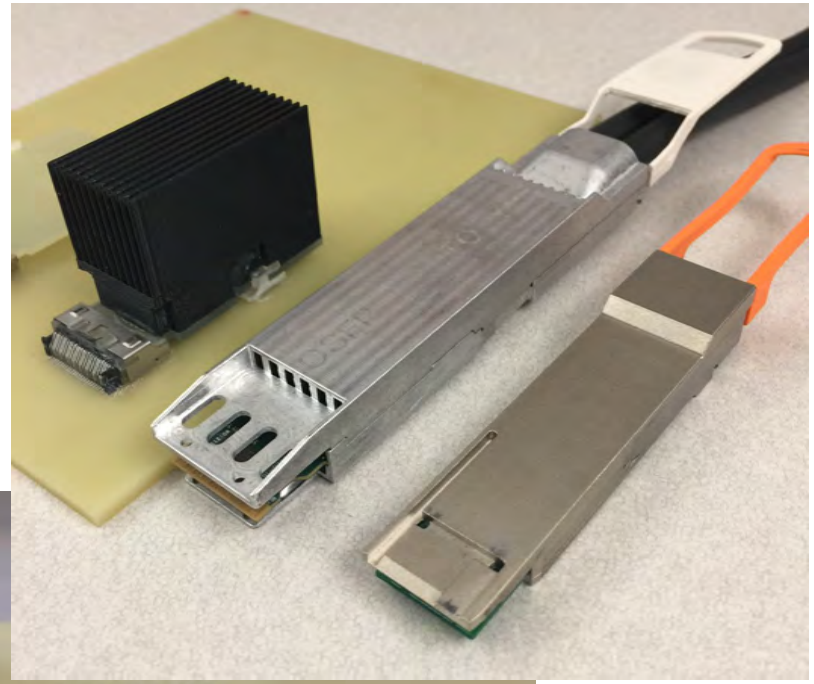
Growing Acceptance of x8 Solutions

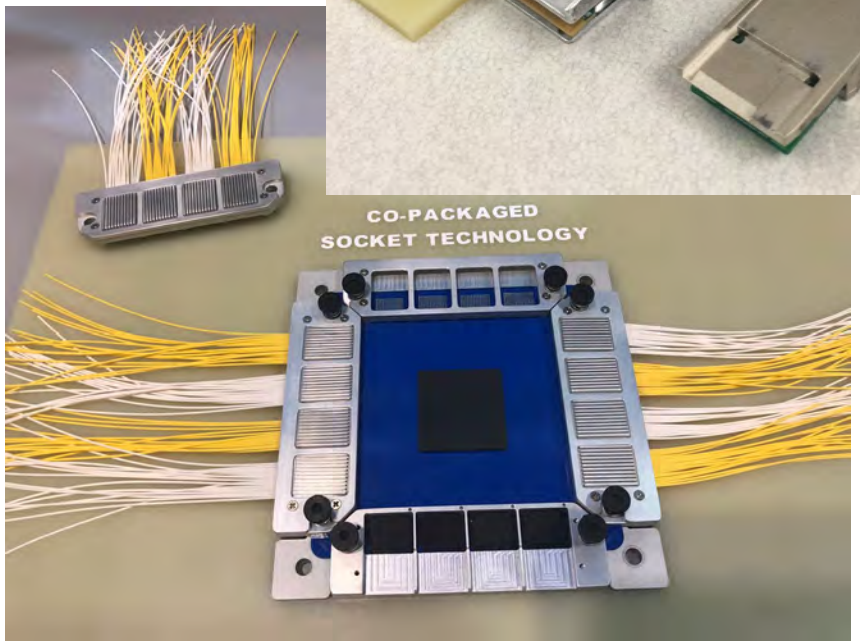


Used with permission, Alan Weckel, 650 Group

The New Paradigm

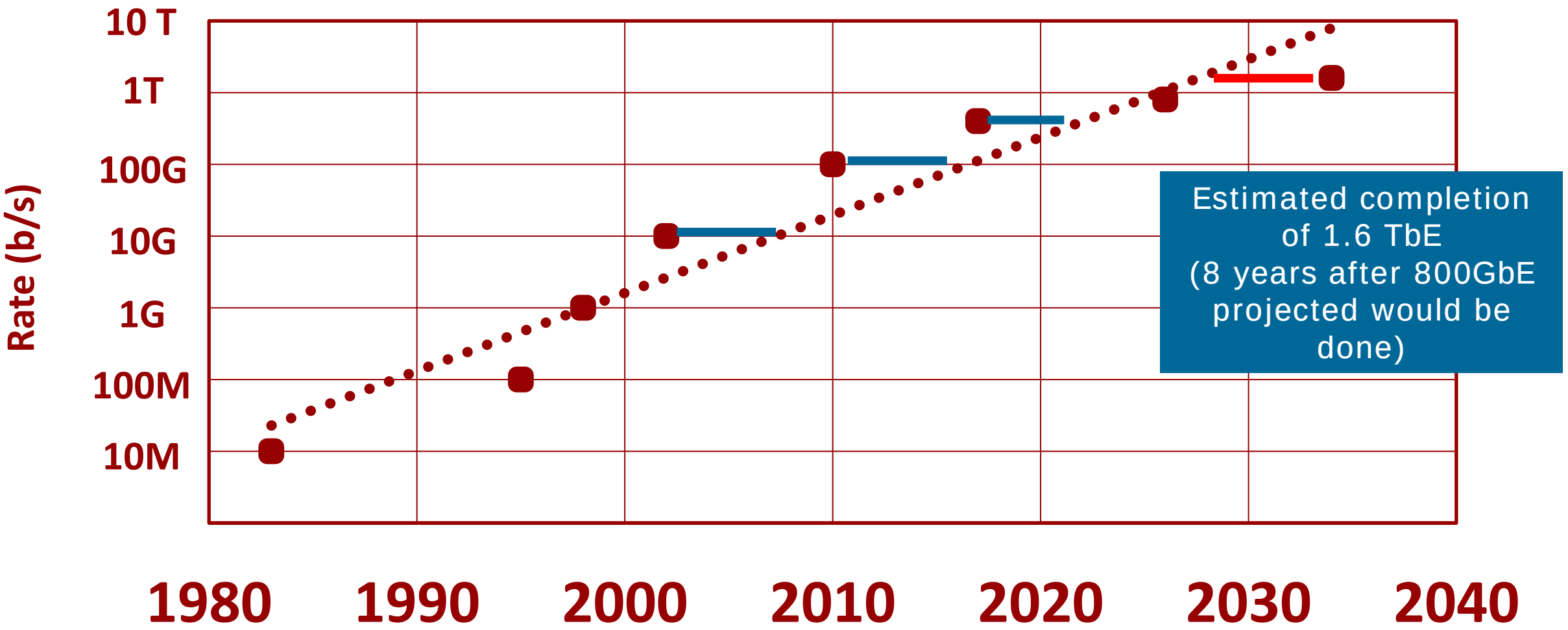
- There has been a shift to multi-rate ports: x1, x2, x4, x8
- Ethernet PHYs have been developed based on x1, x2, x4, and x8 lane rates
- Market acceptance has been highlighted.
- This project
  - is continuing this trend with 8 x 100 Gb/s physical layer specifications
  - should continue that trend based on 200 Gb/s signaling



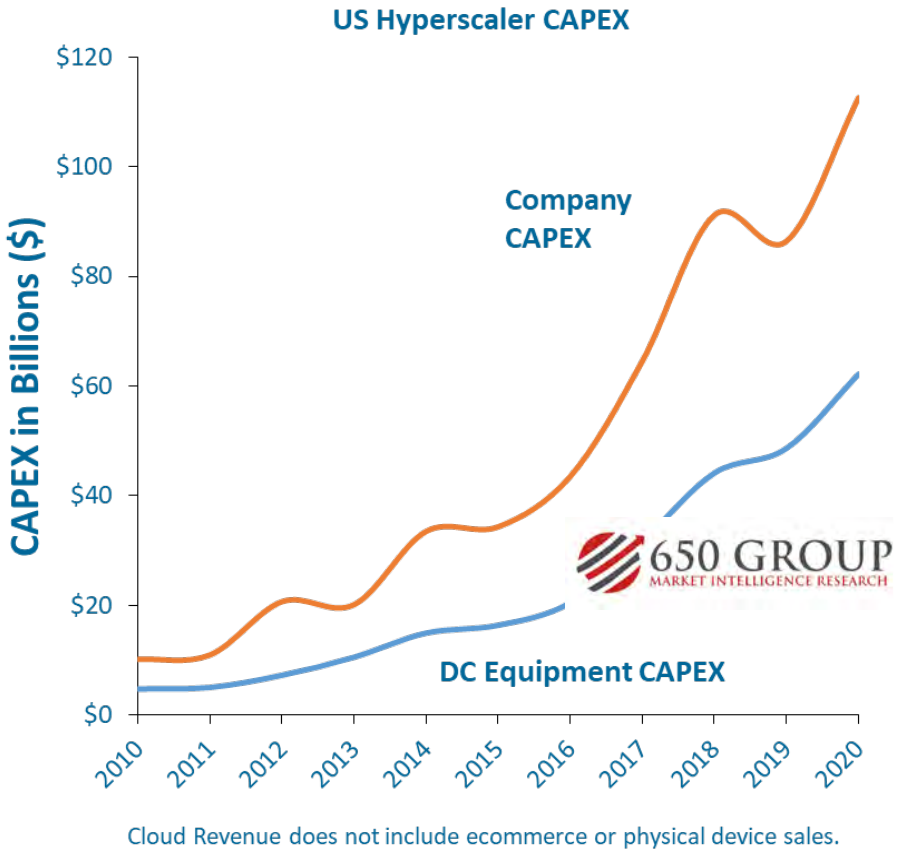


Used with permission, Nathan Tracy, TE Connectivity

Developing 800 GbE / 1.6 TbE in separate “Next Speed” Projects



Looking at the Cloud

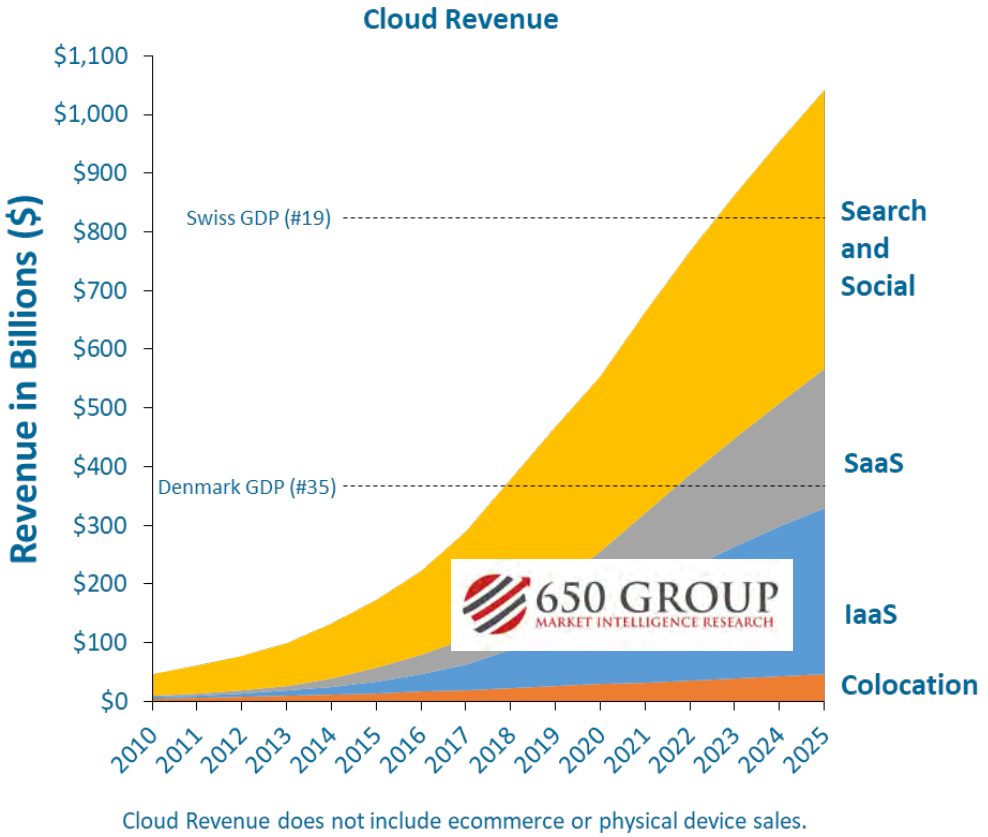


Cloud Revenue does not include ecommerce or physical device sales.

- DC spend increasingly important
- Increased spend on Networking
- New networking TAMs
  - AI
  - DCI

Invest

To address

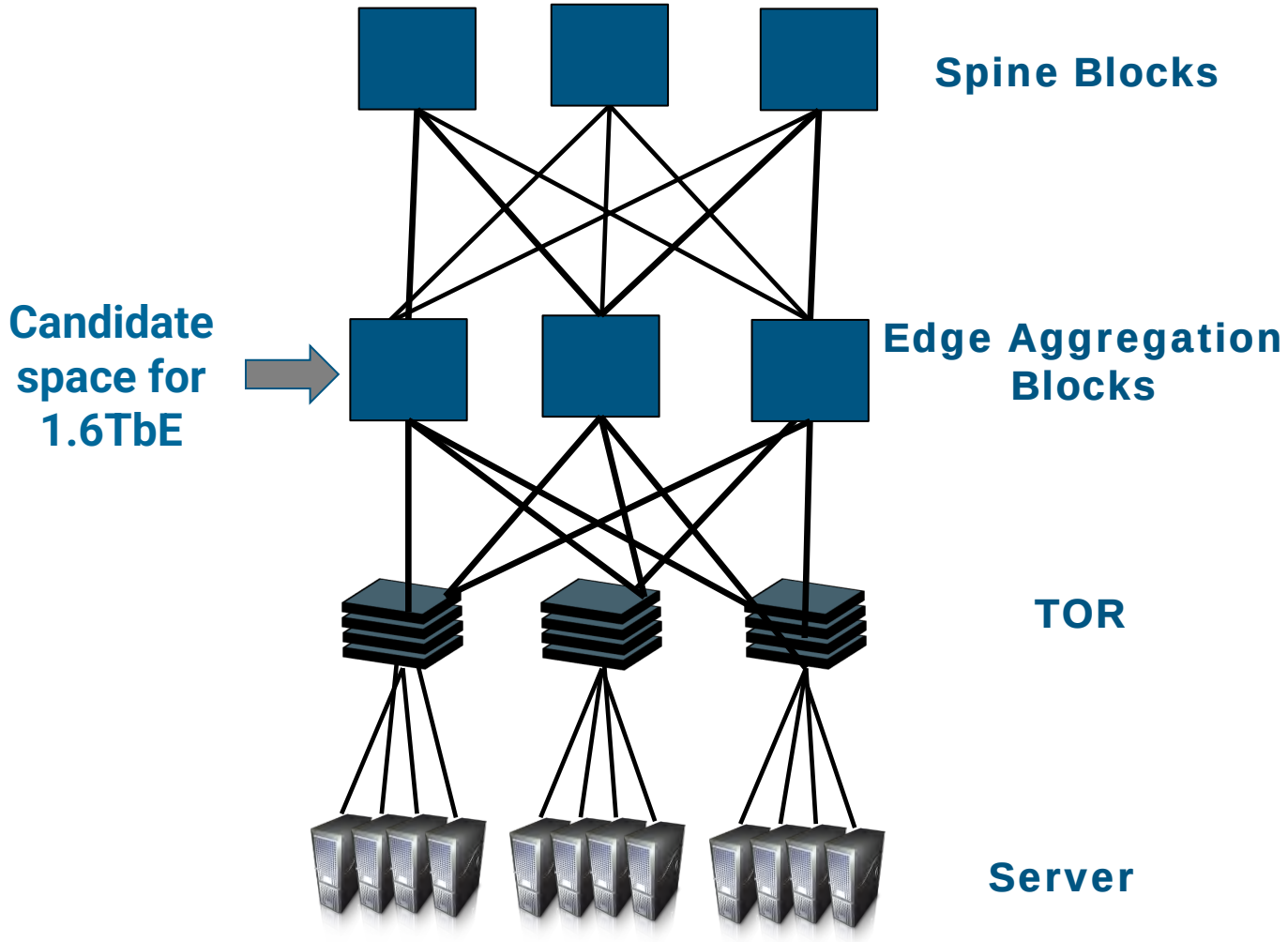


Cloud Revenue does not include ecommerce or physical device sales.

- Cloud Revenue has been doubling in size ever 2-3 years and will double again by 2025
- IaaS, the consumer of highest speed ports will outpace the market
- Server Utilization continues to improve

Used with permission, Alan Weckel, 650 Group

Where would 1.6 TbE make sense in hyperscale datacenters?



To Spine

- Duplex SMF

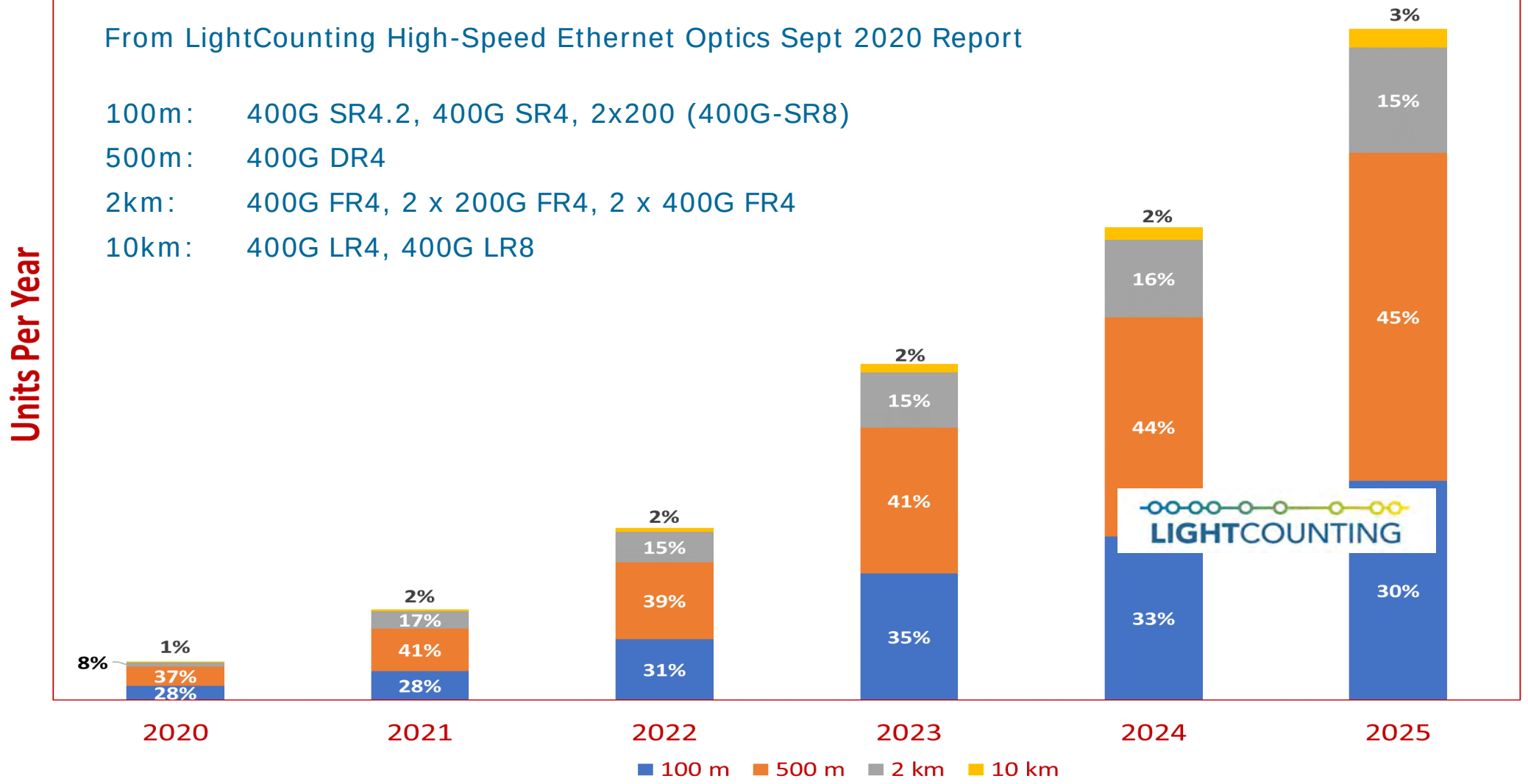
Typical PHYs within Edge Aggregation Blocks

- Depends on the link length.
- Connections within the same rack copper links
- PSM

Forecast – Transceiver Modules Targeting 200 GbE or 400 GbE

From LightCounting High-Speed Ethernet Optics Sept 2020 Report

100m: 400G SR4.2, 400G SR4, 2x200 (400G-SR8)
500m: 400G DR4
2km: 400G FR4, 2 x 200G FR4, 2 x 400G FR4
10km: 400G LR4, 400G LR8



Comments on 1.6 Tb/s Already

- Per nowell_b400g_01_210118 - ”Lack of 1.6 TbE doesn’t preclude 1.6T pluggable modules (e.g. 2x 800 GbE)”
- Per chopra_b400g_01_210208
  - forecasted 51.2T switch silicon BW in 2022 (based on doubling ≈ every two years)
  - Define 1.6TE MAC
    - Only if there is a cost effective PMD solution
    - Over 212G to enable 102.4T with radix 64 (64x1.6TE)
- My observations
  - Assume doubling then slows to every four years, 102.4T switches by 2026 (estimated end of future TF)
  - B400G SG has adopted
    - Objectives to define (assuming 100 Gb/s signaling):
      - Optional 800 Gb/s AUIs over 8 lanes for chip-to-module and chip-to-chip applications
      - over 8 pairs of SMF with lengths up to at least 500 m
    - Objectives to define (assuming 200 Gb/s signaling):
      - Optional 800 Gb/s AUIs over 4 lanes for chip-to-module and chip-to-chip applications
      - over 4 pairs of SMF with lengths up to at least 500 m
      - over 4 pairs of SMF with lengths up to at least 2 km
  - Assumed - 1.6 Tb/s Port with module supporting 2 800GbE PHYs will be developed.
  - Provides market with solutions for multiple applications
    - 8 ports of 200 GbE, 4 Ports of 400 GbE, 2 Ports of 800 GbE, 1 Port of 1.6 TbE
  - Simply put – this is obvious – if IEEE 802.3 doesn’t address this, it will be done outside of IEEE – and technical decisions outside this body could influence future 802.3 technical decisions
  - 1.6 Tb/s interfaces being touted already - https://blogs.cisco.com/sp/ciscociscosilicononeg100announcement

Summary

- 800 GbE is late
  - Industry has stepped in via ETC 800 GbE specification based on IEEE 802.3 work
- Industry adopting x8 solutions supporting existing multiple Ethernet rates
- The B400G Study Group has adopted 800 Gb/s objectives targeting the following
  - Optional 800 Gb/s attachment unit interfaces for chip-to-module and chip-to-chip applications
    - 8 lanes (assumed 8 x 100 Gb/s)
    - 4 lanes (assumed 4 x 200 Gb/s)
  - Physical Layer Specifications
    - Over 8 fibers of SMF for 500m (8x 100 Gb/s)
    - Over 4 fibers of SMF for 500m (4 x 200 Gb/s)
    - Over 4 Fibers of SMF for 2km (4 x 200 Gb/s)
- 1.6 TbE needs to be considered now so we won't be late with it
  - Focus the physical layer specification work –
    - Leverage x8 solutions
    - Leverage specifications based on 200 Gb/s signaling being developed for 800 GbE
- Everything is in place to be leveraged to develop 1.6 TbE, AUIs, and PHYs for target application space

Recommendation

- **Adopt the following objectives**
  - Support a MAC data rate of 1.6 Tb/s
  - Define a physical layer specification that supports 1.6 Tb/s operation:
    - over 8 pairs of SMF with lengths up to at least 500 m
    - over 8 pairs of SMF with lengths up to at least 2 km