Going Forward
IEEE 802.3 NEA
"Ethernet for AI" Assessment

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

11 June 2026

Supporters

- Marco Mascitto, Nokia
- Rob Stone, Meta

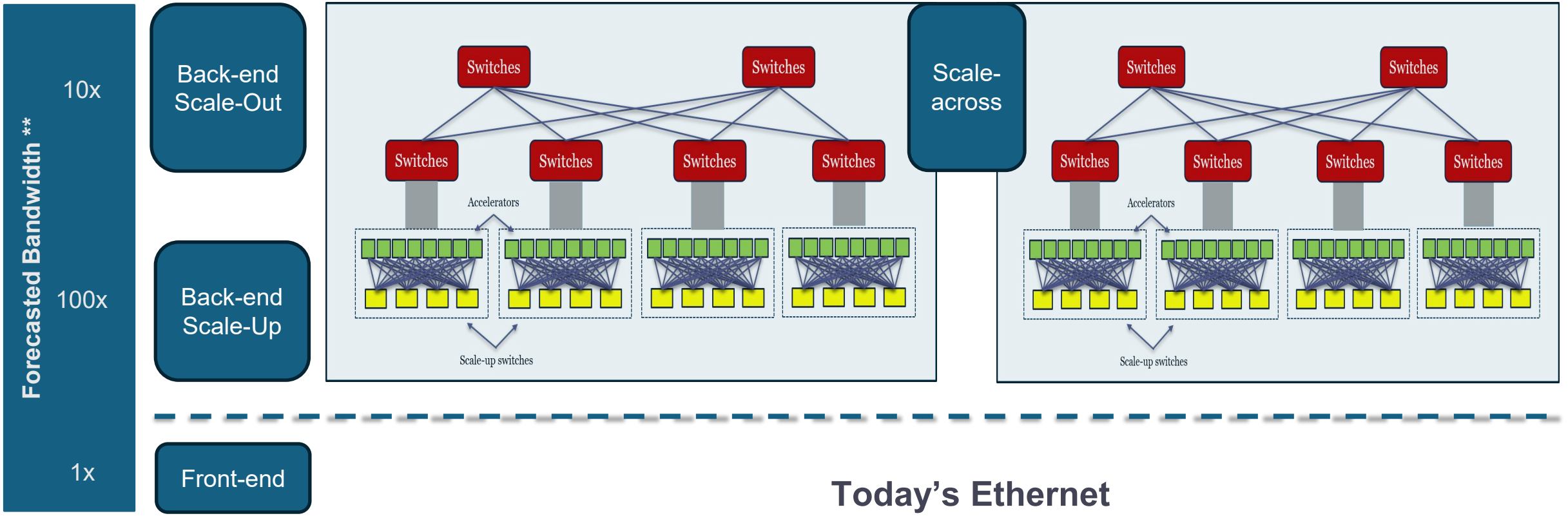
Introduction

- E4AI 2025 Plan
  - Effort #1 – Focus on approval and formation of IEEE 802.3 400 Gb/s Per Lane Study Group
  - Effort #2 - >= 3.2 Tb/s Ethernet and PHYs not in Effort #1
  - Additional efforts to be explored under E4AI Assessment
- This presentation is intended as a discussion of “Going Forward” for activities in the IEEE 802.3 NEA “Ethernet for AI Assessment”

Let’s Start with Some Basic Definitions

- For today’s presentation and the sake of subsequent conversations during the meeting, the following terms with the noted definitions will be used:
  - Lane - from one PMD to another across the transmission medium (e.g. optical fiber, optical wavelength, wire pair). A PMD may be made up of 1 to “n” lanes, where **the lane rate**, the effective data rate of operation not including any FEC overhead, across each lane is defined. The rate that the PMD supports is the number of lanes * the lane rate and is equivalent to the **MAC rate**.
  - PMD - Within 802.3, that portion of the Physical Layer responsible for interfacing to the transmission medium. The PMD is located just above the Medium Dependent Interface (MDI). A PMD supports a single **MAC rate**.
  - Port – In IEEE P802.3dj there are specifications for an MDI connector that has multiple PMD instantiations
    - “Breakout” is supported
    - The port can support a maximum “**port capacity**” (which may or may not match the MAC rate of the PMDs implemented). For example, an MDI connector defined to support a maximum of 4 PMDs, each operating at a 200 GbE MAC rate, would have a maximum port capacity of 800 Gb/s.
- For the sake of clarity, please be careful to specify what you mean.
- The presenter recognizes that subsequent standards related conversations to these definitions may be necessary at a later date.

AI Datacenter Network Hierarchy *



Volume deployment of "Scale-up" will need the development and deployment of solutions for "Scale-out" and "Scale-Across" Networks

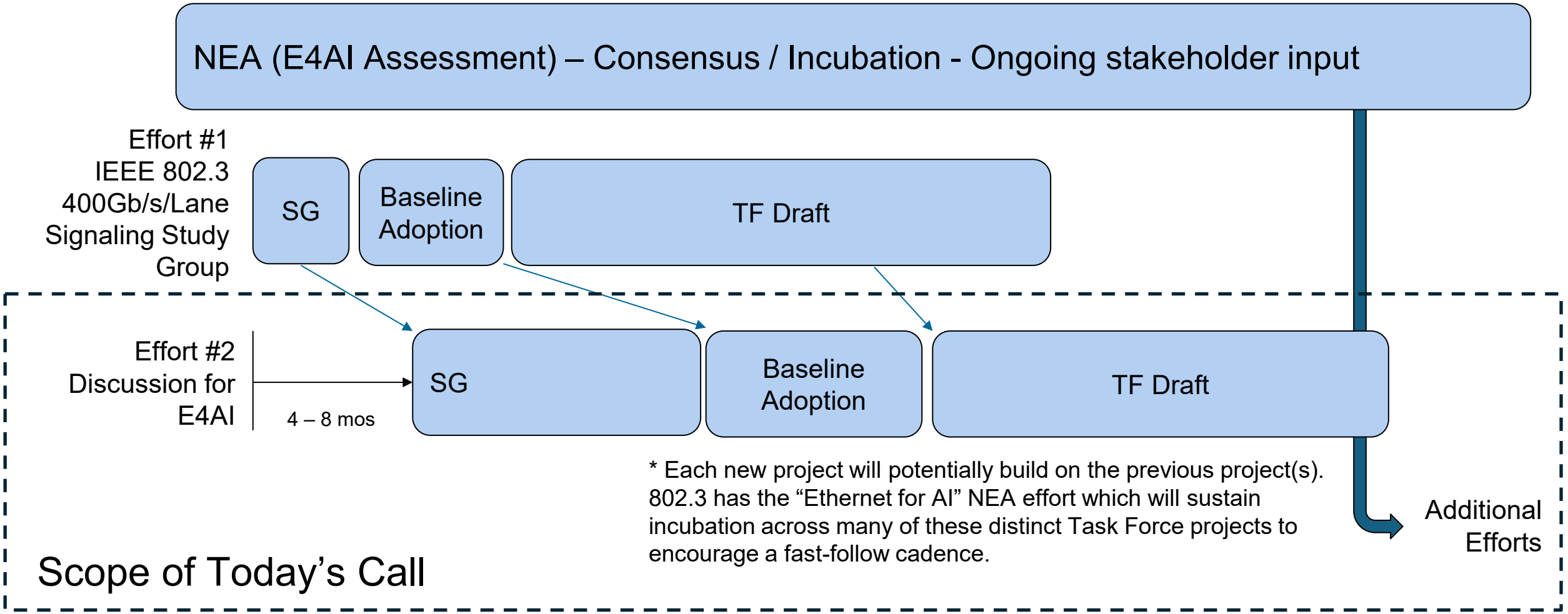
* Based on "AI Datacenters and their Diverse Network Requirements", Ram Huggahalli (Microsoft), Ethernet Alliance TEF 2024, Oct 2024.

** Alan Weckel, 650 Group, "AI Networking: What do scaleup and scaleout really mean for networking demand," 27 Mar 2025, https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/weckel_e4ai_01_250327.pdf.

Related IEEE 802.3 Efforts Review

- IEEE P802.3ds 200 Gb/s per Wavelength MMF PHYs
- IEEE 802. NEA “Ethernet for AI” Assessment - An assessment of "Ethernet Interconnect for AI" with an emphasis on beyond 200 Gb/s signaling
  - IEEE P802.3dt Ethernet Metadata Services
    - the optional exchange of metadata through a physical layer service access point
  - IEEE 802.3 400 Gb/s Per Lane Study Group
    - 400 Gb/s per lane signaling for electrical interconnects as well as single-mode fiber optical interconnects with reaches up to 500 meters
    - Approved @ IEEE 802 March 2026 Plenary
  - Workshops
    - Fiber For AI (24 Feb 2026) – Multi-core, Hollow-core, Future SMF / MMF
    - Header Compression (23 Apr 2026)
    - Future MDIs (03 Dec 2026)

E4AI - Presented Plan



Ethernet Solutions in Development (based on >=200 Gb/s/lane)

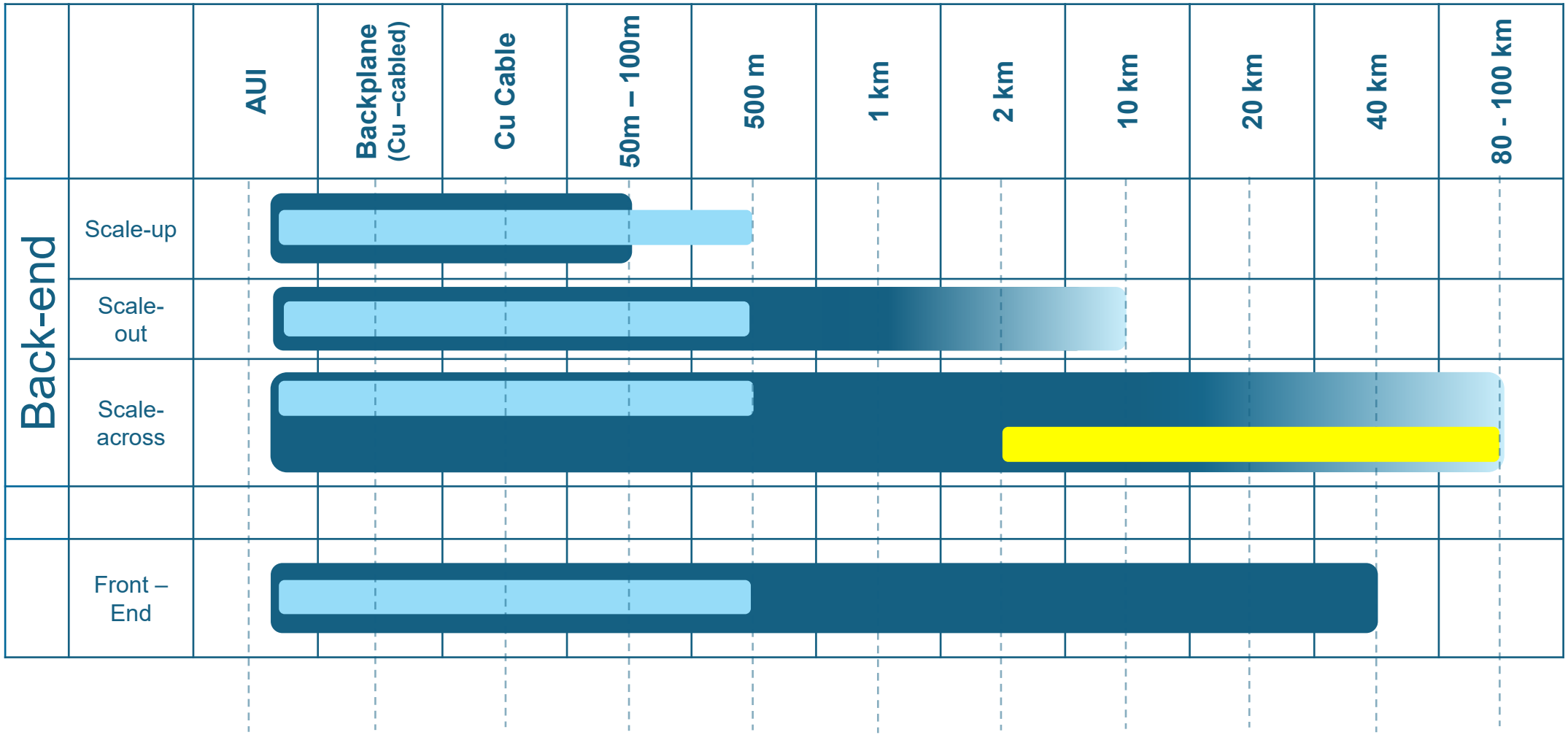
Ethernet Rate	AUI	Backplane	Cu Cable	MMF	SMF 500m	SMF 2km	SMF 10km	SMF 20km	SMF 40km
200 Gb/s	200GAUI-1 C2C C2M	200GBASE-KR1	200GBASE-CR1	1. 200GBASE-SR1-30 2. 200GBASE-SR1-50 3. 200GBASE-MR1	200GBASE-DR1	200GBASE-DR1-2			
400 Gb//s	400GAUI-2 C2C C2M	400GBASE-KR2	400GBASE-CR2	1. 400GBASE-SR2-30 2. 400GBASE-SR2-50 3. 400GBASE-MR2	400GBASE-DR2	400GBASE-DR2-2			
800 Gb/s	800GAUI-4 C2C C2M	800GBASE-KR4	800GBASE-CR4	1. 800GBASE-SR4-30 2. 800GBASE-SR4-50 3. 800GBASE-MR4	1. 800GBASE-DR4 2. 800GBASE-FR4-500	1. 800GBASE-DR4-2 2. 800GBASE-FR4	1. 800GBASE-LR1 2. 800GBASE-LR4	800GBASE-ER1-20	800GBASE-ER1
1.6 Tb/s	1.6TAUI-8 C2C C2M	1.6TBASE-KR8	1.6TBASE-CR8	1. 1.6TBASE-SR8-30 2. 1.6TBASE-SR8-50 3. 1.6TBASE-MR8	1.6TBASE-DR8	1.6TBASE-DR8-2			
					Duplex Fiber (FR)				

200 GbE – 10 / 40km solutions Duplex fiber solutions / Beyond 2km based on lower signaling lane rates

400 GbE - 10 / 40km solutions Duplex fiber solutions / Beyond 2km based on lower signaling lane rates

1.6 TbE – no WDM solutions for Duplex fiber solutions or Beyond 2km

“AI Networks” Industry Need versus Efforts

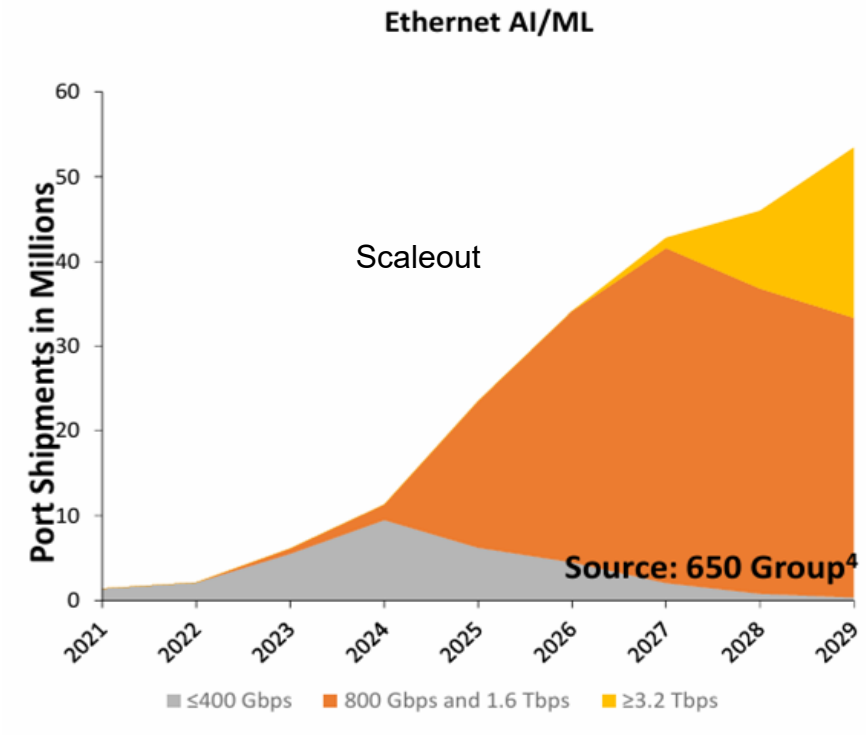


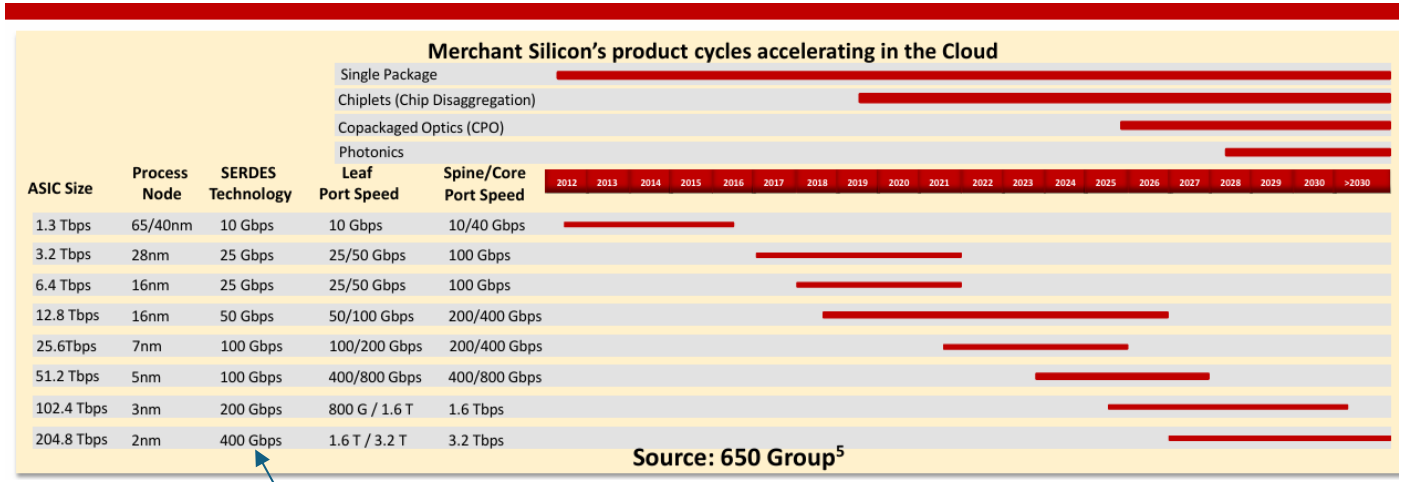
Radix “Fat-Pipe” “break-out” Point-to-point

Industry Need IEEE 802.3 400GPL SG ZR Optics – May be used as part of DWDM system

Market Update

3.2 Tbps Port Shipments [1]



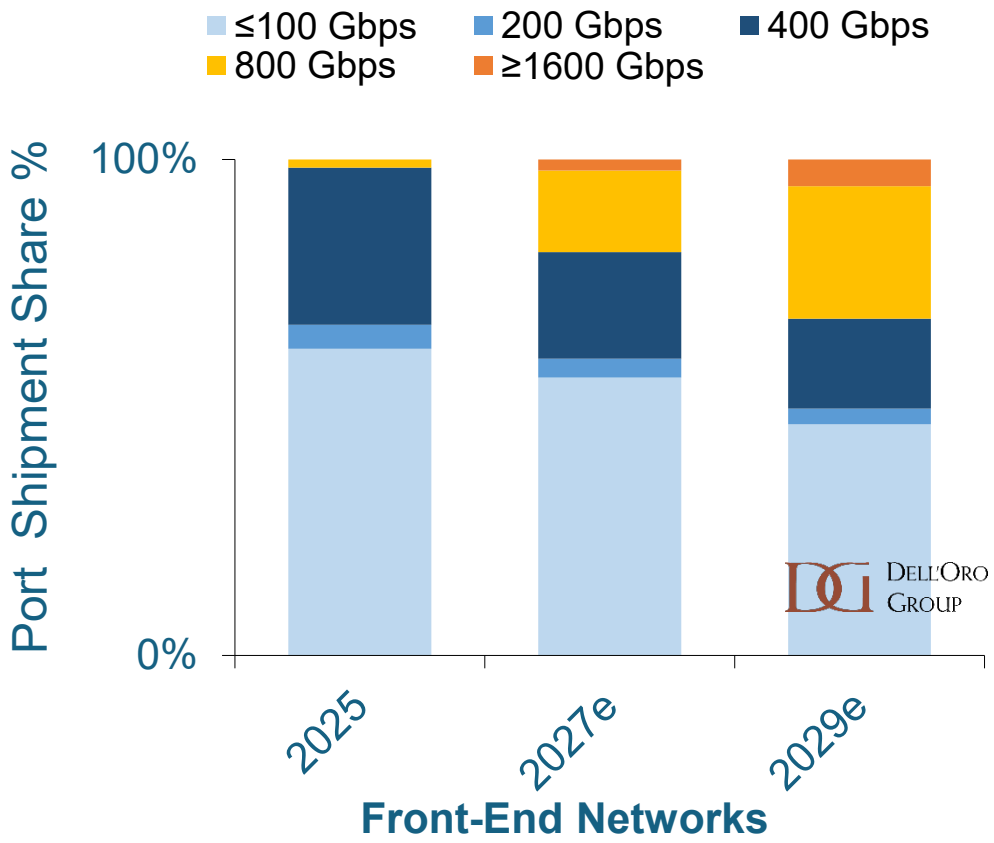


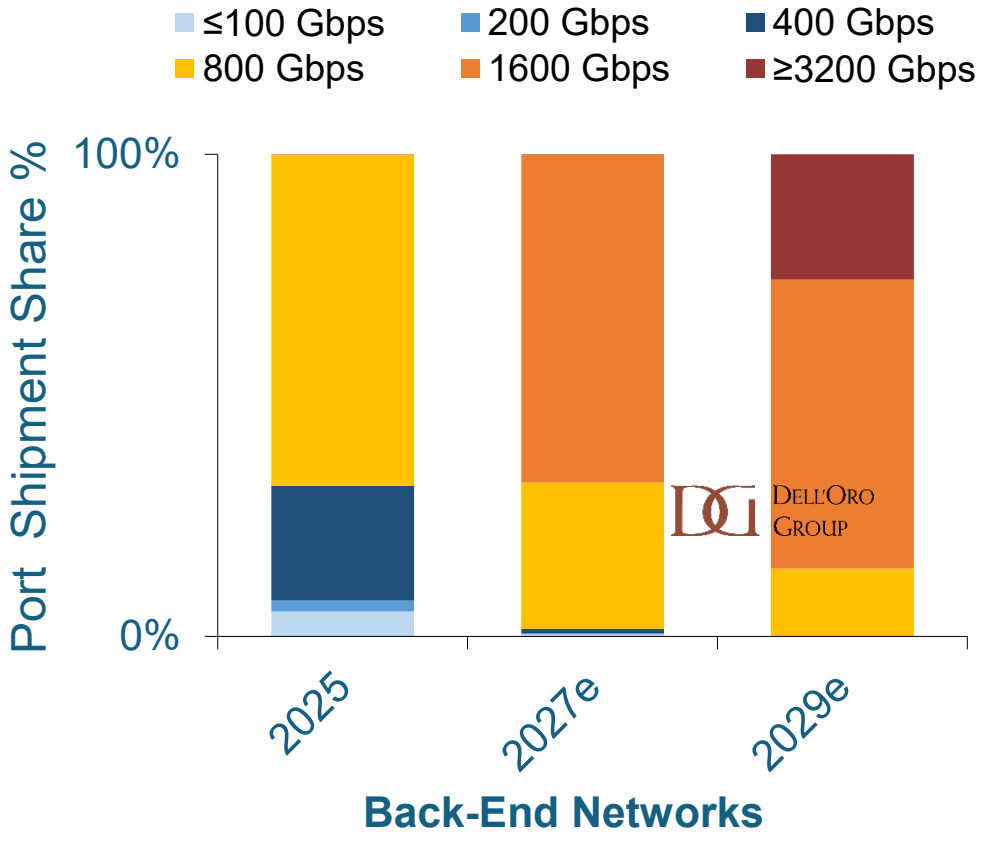
- 2027-2030: 204.8 Tbps switch ASIC size –
  - 1.6T / 3.2T Leaf Port Speed / 3.2 Tbps Spine / Core Port Speed

- A 3.2 Tb/s port can support lower rates of Ethernet (based on 400 Gb/s/lane) or 3.2 TbE – this is an implementation choice!

1. Alan Weckel, 650 Group, "AI Networking: What do scaleup and scaleout really mean for networking demand," 27 Mar 2025, https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/weckel_e4ai_01_250327.pdf.

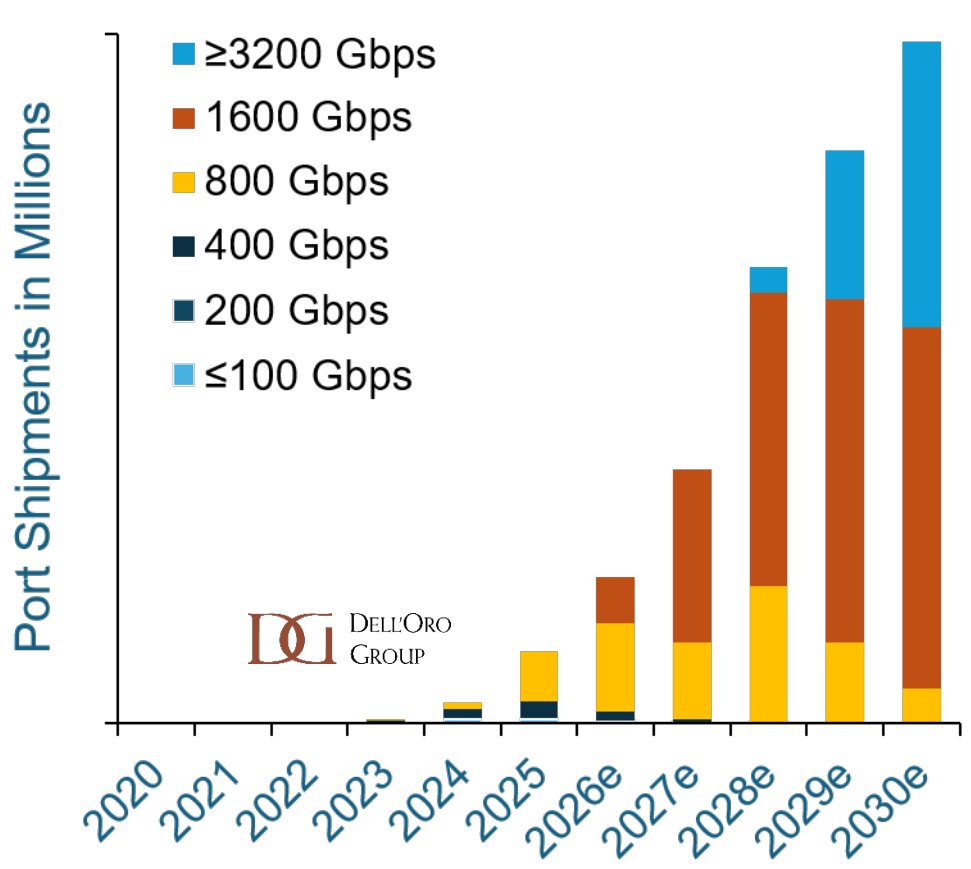
Ethernet Switch Port Speed Transition (100 Gbps+)



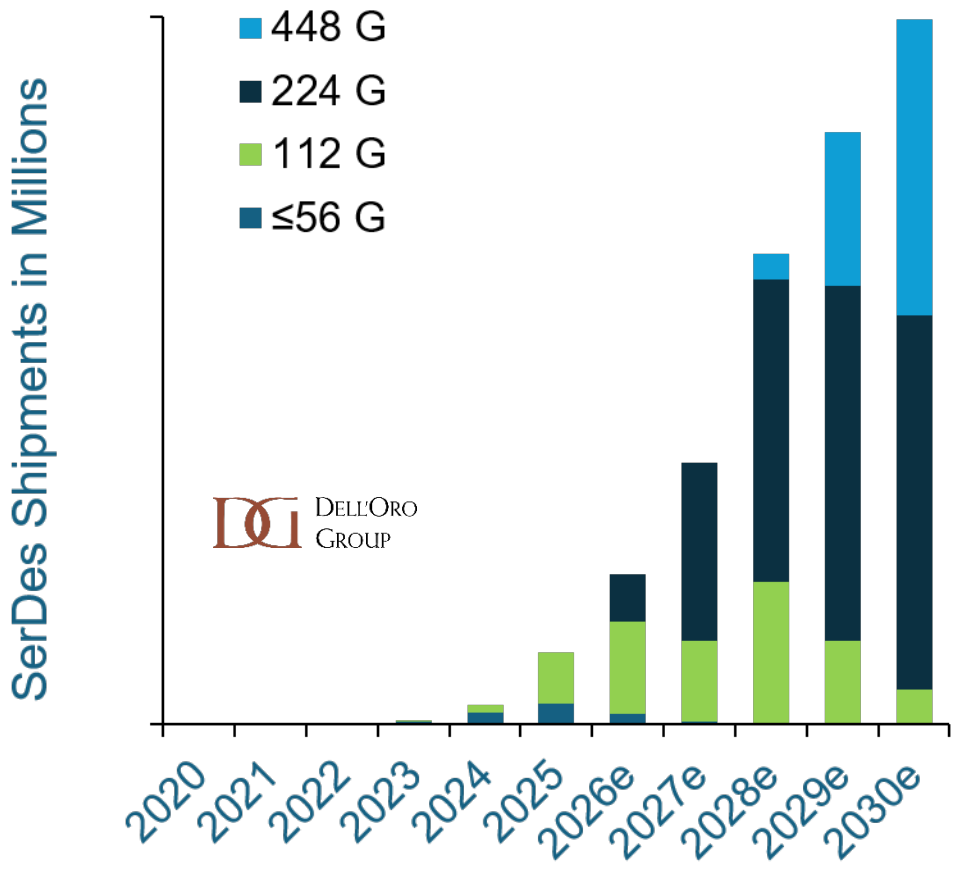


- Source: Sameh Boujelbene, Dell'Oro Data Center Switch—AI Back-end Networks, January 2026
- Back-end includes scale-out and scale-up

Ethernet Switch Shipments in AI Back -End Networks (100 Gbps+)



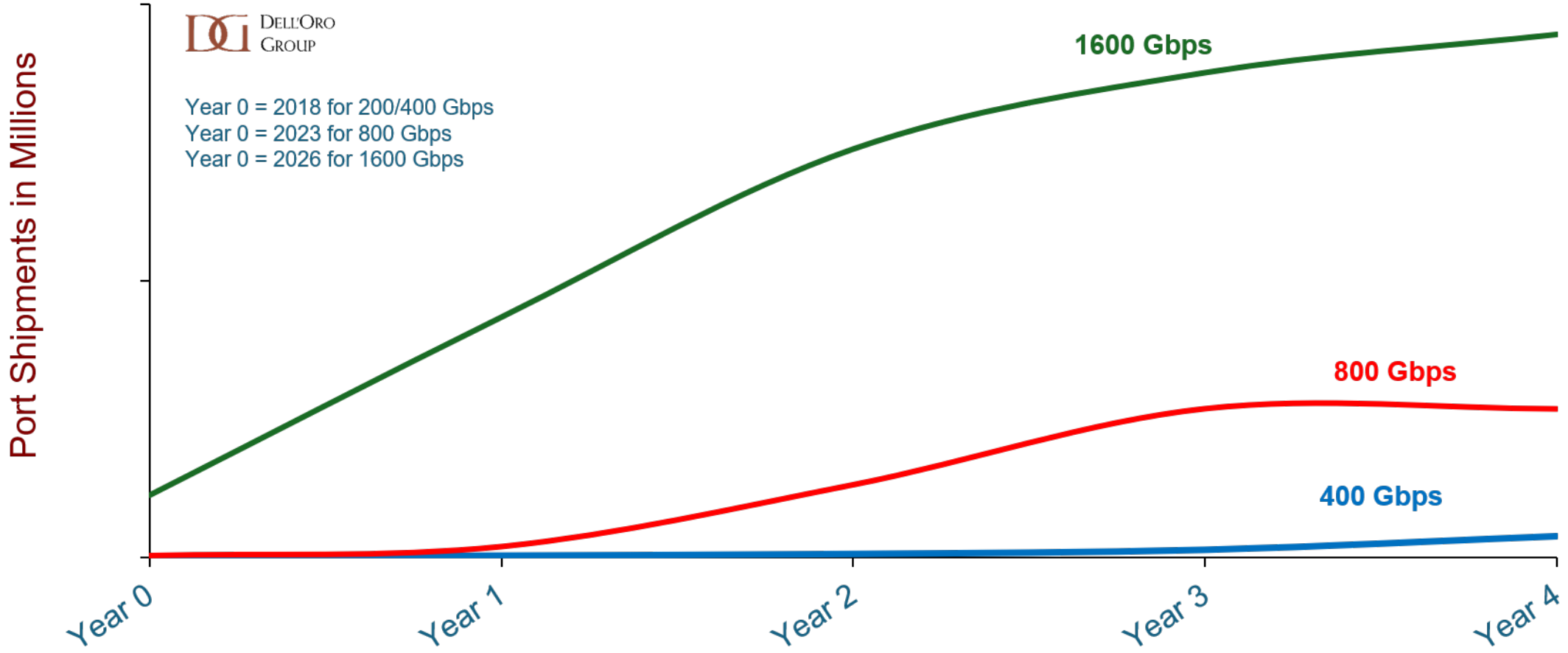
Switch Port Shipments



SerDes Shipments

- Source: Sameh Boujelbene, Dell'Oro Data Center Switch—AI Back-end Networks, January 2026
- Back-end includes scale-up and scale-out

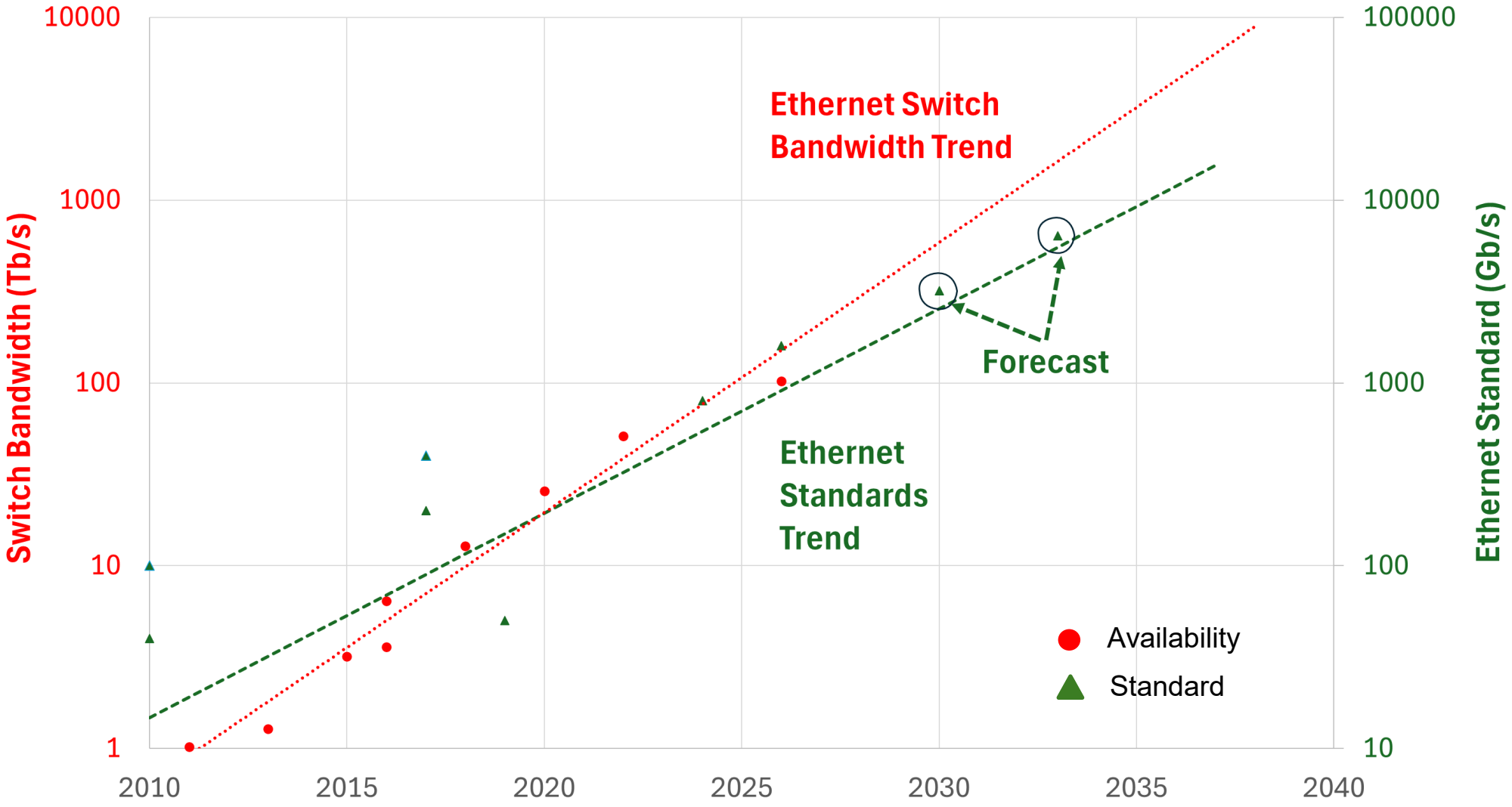
400/800/1600 Gbps Ethernet Switch Port Shipment Ramp



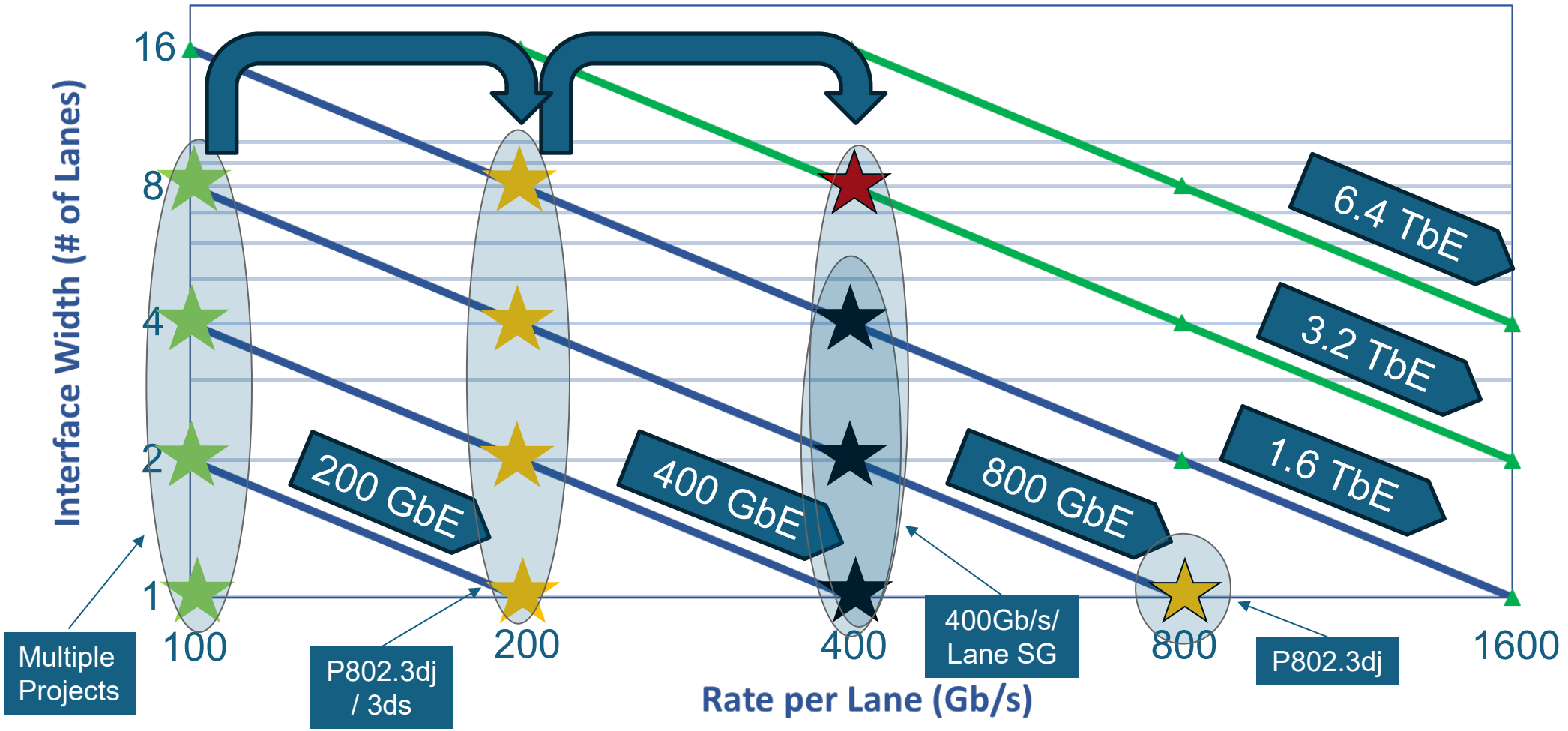
- Source: Sameh Boujelbene, Dell'Oro Data Center Switch—AI Back-end Networks, January 2026
- includes front-end, back-end (including scale up and scale out) but back-end is the majority

Ethernet Rate
vs
Port Capacity

Ethernet Trends – Standards / Switch Bandwidth



Lane Rate versus Ethernet Rate



Ethernet Lanes / PMDs / Capacities

Table 180A–1—Number of PMDs supported by each MDI connector type

MDI connector type	200GBASE-DR1, 200GBASE-DR1-2	400GBASE-DR2, 400GBASE-DR2-2	800GBASE-DR4, 800GBASE-DR4-2	1.6TBASE-DR8, 1.6TBASE-DR8-2
Single-fiber connector	1	—	—	—
Single-row 12-fiber	1, 2, or 4	1 or 2	1	—
Single-row 16-fiber	1, 2, 4, or 8	1, 2, or 4	1 or 2	1

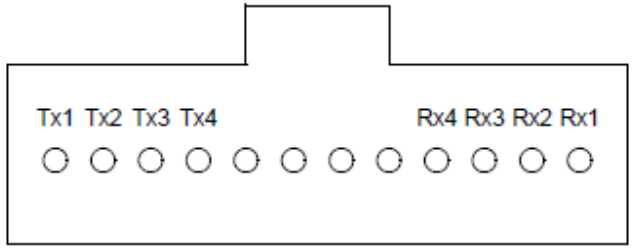


Figure 180A–2—Optical lane assignments for a single-row 12-position connector

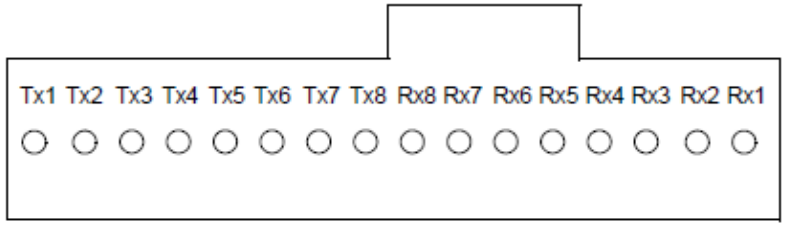
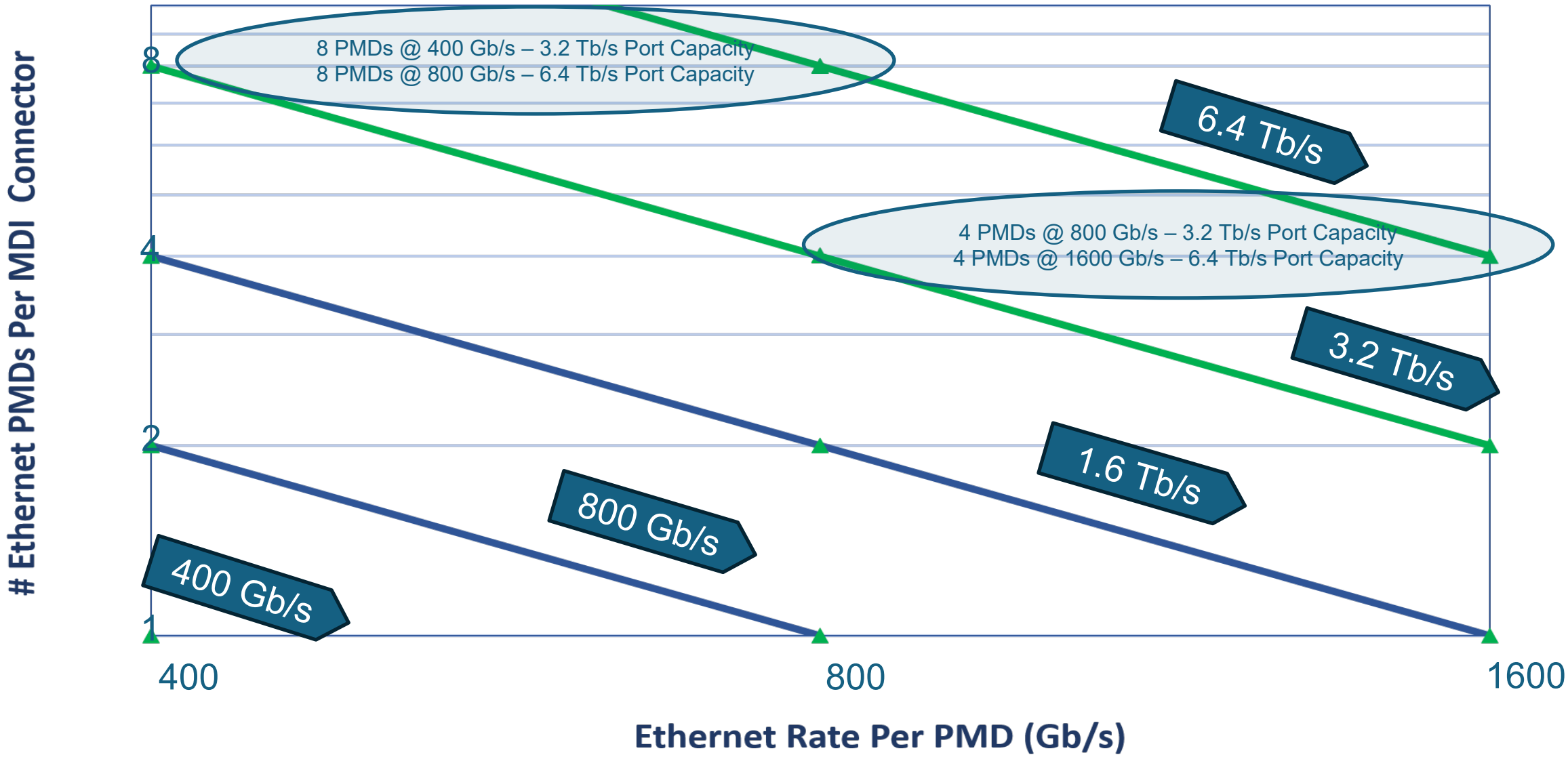


Figure 180A–3—Optical lane assignments for a single-row 16-position connector

Table and figures from IEEE P802.3dj D3.1

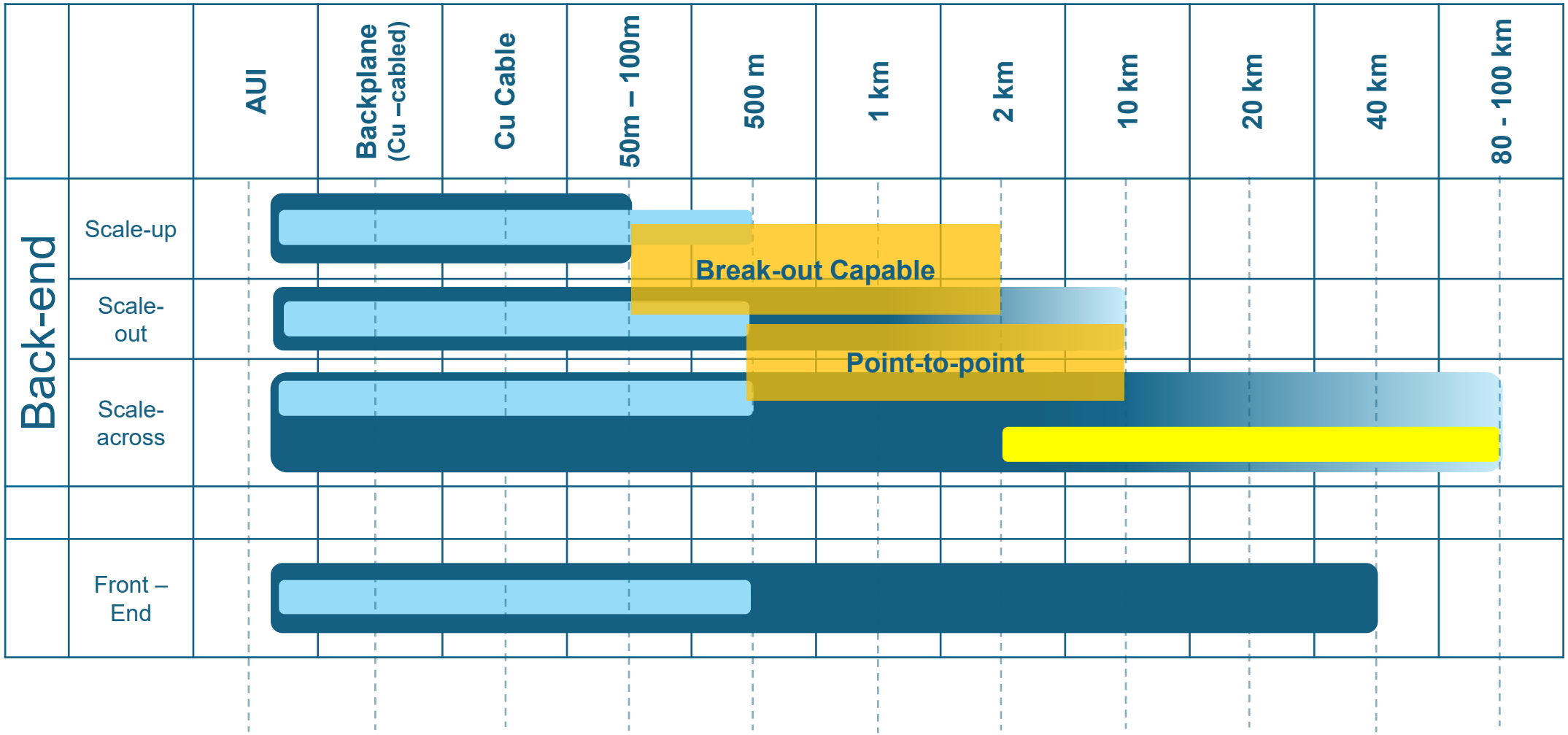
- An MDI connector can support multiple # of PMDs and different PMD types
  - Note - # in PMD suffix indicates # of lanes
- Single-row 12 fiber MDI connector type supports up to 800 Gb/s capacity, based on combinations:
  - 200GBASE-DR1 / DR1-2
  - 400GBASE-DR2 / DR2-2
  - 800GBASE-DR4 / DR4-2
- Single-row 16 fiber MDI connector type supports up to 1.6 Tb/s capacity, based on combinations:
  - 200GBASE-DR1 / DR1-2
  - 400GBASE-DR2 / DR2-2
  - 800GBASE-DR4 / DR4-2
  - 1.6TBASE-DR8 / DR8-2
- An MDI connector type could be specified for multiple PMDs to achieve a higher capacity without supporting the maximum capacity as a new MAC rate
- Could leverage FR / LR1 / ER1 / ER1-20 solutions per fiber, as well as DR solutions

PMD Rate vs Port Capacity per MDI Connector



Fiber Issues

Industry Need versus Efforts



Radix
"break-out" Point-to-point
"Fat-Pipe"

Industry Need
IEEE 802.3 400GPL SG
ZR Optics – May be used as part of DWDM system

New Fiber Technologies & Interconnect

- Currently, there are no industry specifications addressing multi-core or hollow-core fiber
- There is no industry “channel” model addressing multi-core or hollow-core fiber
- There is growing industry interest observed in using both fiber types for high-speed optical links, and they could be suggested for future IEEE 802.3 projects defining Ethernet optical PMDs
- Joint ITU-T / IEEE Workshop July 2026 – update on ITU-T efforts
- Future efforts pending @ this time.

Optical Cabling Challenges– Rack Level

Evolution of fiber density in data centers

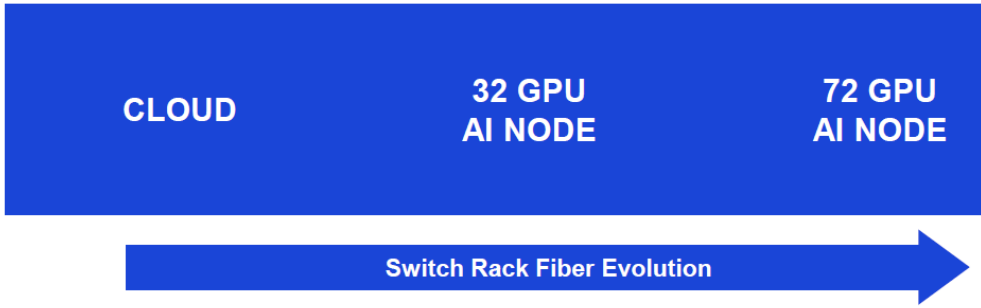
Cloud: Single-digit Enterprise growth pre-Gen AI

32 GPU Node: 4x more fiber than Cloud

72 GPU Node: 4x more fiber than 32 GPU node (16x more than cloud)

Continued scale-up & scale-out of AI networks creates density, installation, & sustainability problems that MCF solutions can address

Corning

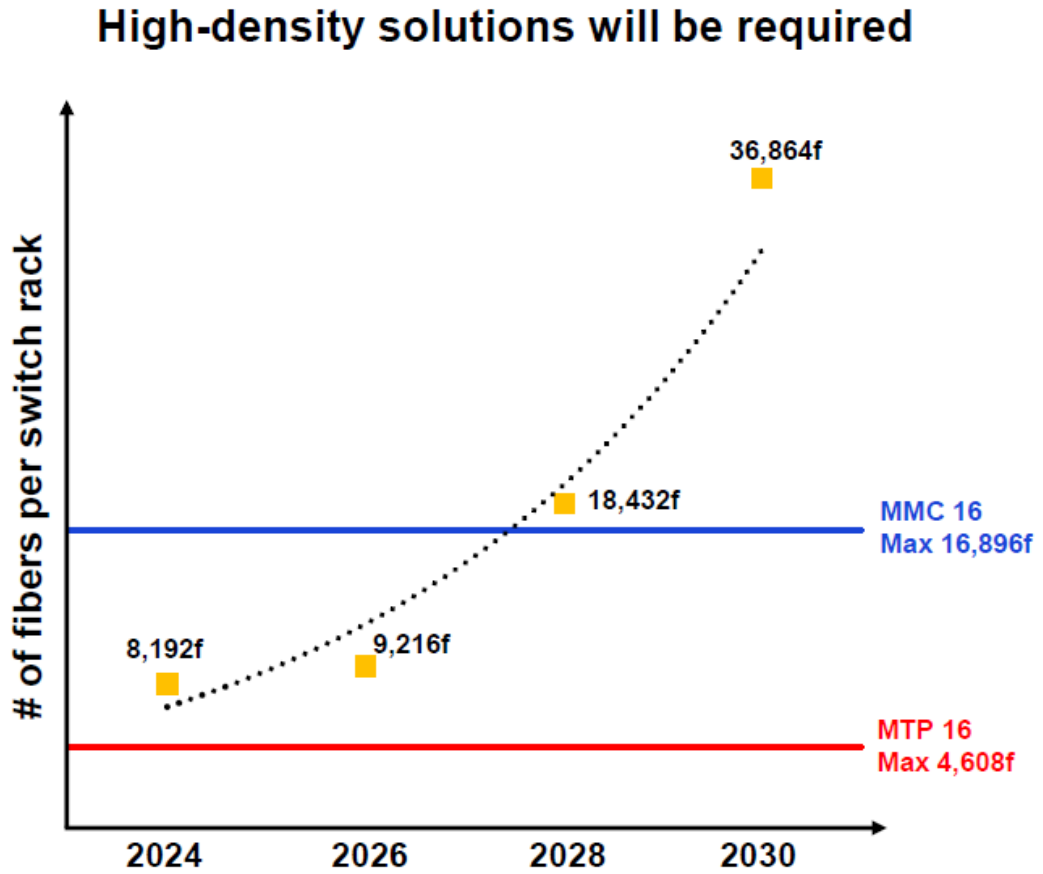


Used with permission from Vince Ferretti, Corning:
https://www.ieee802.org/3/ad_hoc/E4AI/public/26_0224/mi_e4ai_01a_260224.pdf

Increasing cable volume and complex connections are driving increased costs, longer installation times, and delaying AI compute

Step change in infrastructure requirements

AI datacenters require more connectivity → larger facilities that require more cable trays, steel, concrete, and labor



Corning

Used with permission from Vince Ferretti, Corning:
https://www.ieee802.org/3/ad_hoc/E4AI/public/26_0224/mi_e4ai_01a_260224.pdf

“Inside the World's Largest AI Supercluster xAI Colossus”

- YouTube Video (Dated Oct 28, 2024)
  - Some Optical Cabling “Moments”
    - Timestamp1 2:12
    - Timestamp2 9:40

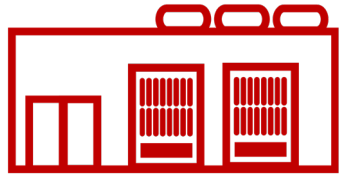
Realities of Inter-Building Fiber

Image and input provided by Rob Stone, Meta

- Meta inter-building fiber pathway
  - ~4M fibers.
  - Pathway construction is becoming increasingly challenging to scale to support inter-building demand

DC Evolution:
Traditional Compute to AI Campus (2026 / 2027)

	Traditional Data Center	AI Campus (1M XPUs)
Modeling & Data Provided by Alan Weckel, 650 Group		
Number of Buildings	1	4 - 12
Servers per Building	100K	100K – 200K
CPUs per Server	2	2 - 4
GPUs per Server	0	8
NIC Ports	100K	1.1M – 1.4M
Management Ports	100K - 120K	1.2M – 2.5M
DC Switch Ports	300K - 400K	3.4M – 6M *
Cables / Fibers **	300K – 1.2M	6M – 24M



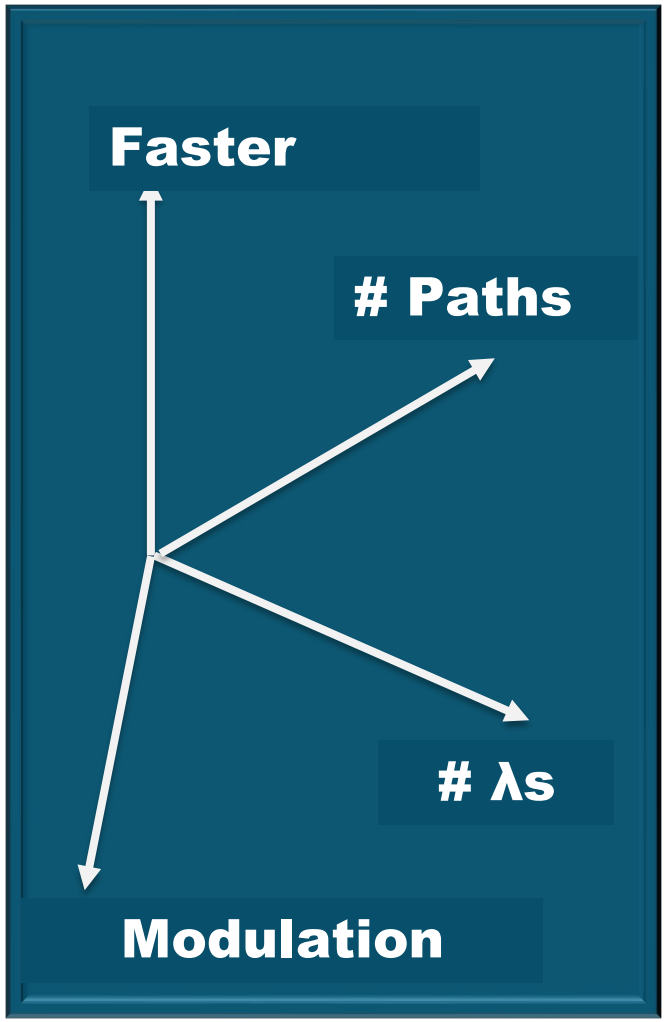
Traditional
Data Center

AI Campus
(1M XPUs)

Notes –
* - DC Switch Ports includes Front-end + Scale-up + Scale-out + Scale-across switch ports.
** - Cables / Fibers includes ToR links (mix copper / fiber) + Links to Aggregation / Core throughout Structured cable plant. Cabling inside the rack not included

Moving Forward

Leveraging All Approaches



- IEEE 802.3 has leveraged all of these approaches
  - Faster Speed
  - Modulation
    - PAM4
    - DP-DQPSK
    - DP-16QAM
  - Parallel Fiber
    - DRx
  - # λs
    - WDM Optics
    - DWDM Optics
  - Bi-directional Optics ?
- IEEE 802.3 must be open to “out-of-the-box” approaches

Areas for Further Exploration

- Fiber congestion throughout AI Networks observed to be a serious issue and should be considered going forward.
  - Bi-directional Optics?
- 1.6 TbE
  - Duplex Fiber (WDM) Optics – 500 m - 2 km
  - 2 km > Reach > 10km: IMDD? Coherent? Both?
  - Reach > 10km: Leverage coherent signaling (Not DWDM)
- 3.2 Tb/s
  - Define new MAC rate
  - Define 3.2 Tb/s Capacity solutions (breakout)
    - Leverage future IEEE P802.3dv technologies
- 6.4 Tb/s
  - ~~New MAC Rate?~~
  - **Define 6.4Tb/s Port Capacity Solutions –**
    - Define Multi MDI Connector to support breakout based on 800 Gb/s or 1.6 Tb/s WDM technologies
    - Define Multi MDI Connector to support breakout based on 800 Gb/s or 1.6 Tb/s coherent technologies

Potential Ethernet Efforts Going Forward

Ethernet Rate	AUI	Backplane	Cu Cable	MMF	SMF 500m	SMF 2km	SMF 10km	SMF 20km	SMF 40km
200 Gb/s	200GAUI-1 C2C C2M	200GBASE-KR1	200GBASE-CR1	1. 200GBASE-SR1-30 2. 200GBASE-SR1-50 3. 200GBASE-MR1	200GBASE-DR1	200GBASE-DR1-2			
400 Gb/s	400GAUI-2 C2C C2M	400GBASE-KR2	400GBASE-CR2	1. 400GBASE-SR2-30 2. 400GBASE-SR2-50 3. 400GBASE-MR2	400GBASE-DR2	400GBASE-DR2-2			
800 Gb/s	800GAUI-4 C2C C2M	800GBASE-KR4	800GBASE-CR4	1. 800GBASE-SR4-30 2. 800GBASE-SR4-50 3. 800GBASE-MR4	1. 800GBASE-DR4 2. 800GBASE-FR4-500	1. 800GBASE-DR4-2 2. 800GBASE-FR4	1. 800GBASE-LR1 2. 800GBASE-LR4	800GBASE-ER1-20	800GBASE-ER1
1.6 Tb/s	1.6TAUI-8 C2C C2M	1.6TBASE-KR8	1.6TBASE-CR8	1. 1.6TBASE-SR8-30 2. 1.6TBASE-SR8-50 3. 1.6TBASE-MR8	~~1.6TBASE-DR8~~	1.6TBASE-DR8-2	FR Optics	FR Optics	FR Optics
3.2 Tb/s	3.2TAUI-8 C2C C2M	3.2TBASE-KR8	3.2TBASE-CR8		3.2TBASE-DR8	3.2TBASE-DR8-2			

Currently not in development

Leverage signaling technologies developed by IEEE P802.3dv

Next Steps

- Industry fiber / connector issues need to be addressed for IEEE 802.3 to provide solutions
- Multiple areas of interest for “Effort 2” have been identified for follow-up. Contributions needed!
  - IEEE 802.3 NEA E4AI to meet August xx
    - Details to be announced
  - Areas for further presentations
    - Technical presentations to related to “Areas of Further Exploration”
    - Prioritization
- Potential for next CFI
  - Jan 2027 (possibility pending approval of 802.3 rules change)
  - Mar 2027
- The IEEE 802.3 Standard and amendments is a documentation of what has been accepted in the past, and should not necessarily restrict us in the future, i.e. – think outside the box!