Overview–
IEEE 802.3 NEA
“802.3 Ethernet for AI” Assessment

John D’Ambrosia
Futurewei, U.S. Subsidiary of Huawei
12 March 2025
IEEE 802.1 Mar 2025 Plenary
Atlanta, GA, USA

“Ethernet for AI” Assessment

An assessment of "Ethernet Interconnect for AI" with an emphasis on beyond 200 Gb/s signaling.

- Activity Chair – John D’Ambrosia
- Website - https://www.ieee802.org/3/ad_hoc/E4AI/index.html
- Reflector - https://www.ieee802.org/3/ad_hoc/E4AI/reflector.html

- Jan 2025 Interim – Assessment Kickoff
  - Introduction - https://www.ieee802.org/3/ad_hoc/ngrates/public/25_01/dambrosia_nea_01_2501.pdf
- 27 Feb 2025 Electronic Meeting
  - https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0227/index.html

Introduction

- This presentation is an overview of the IEEE 802.3 NEA Ad hoc **assessment** of “802.3 Ethernet Interconnect for AI” with an emphasis on beyond 200 Gb/s signaling.
  - **This effort is NOT a call-for-interest.**
  - **It is a consensus building activity on identifying what AI interconnect needs that IEEE 802.3 could target.**

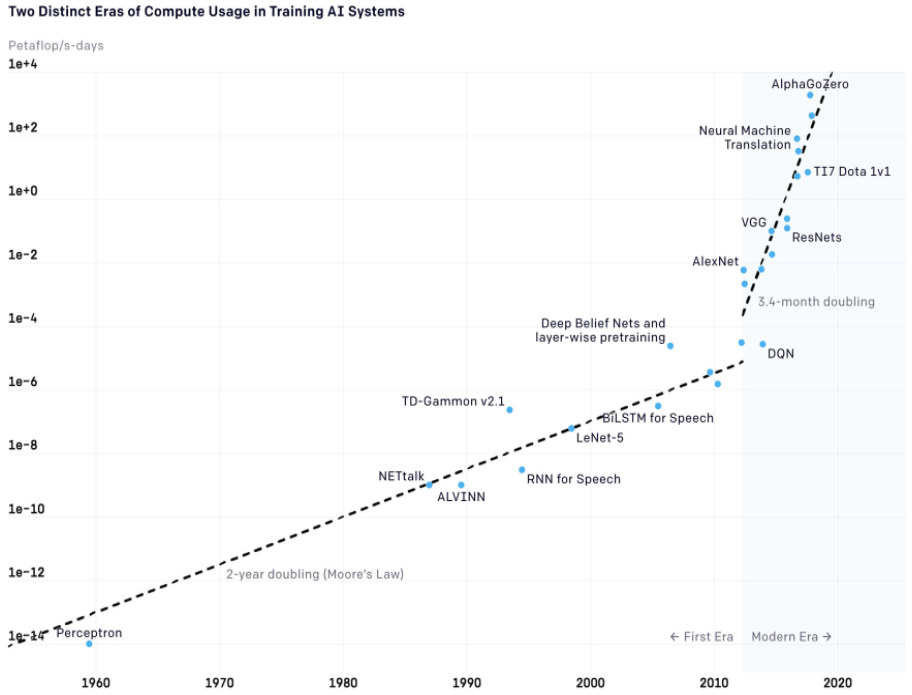
- It has been my experience:
  - Bandwidth assessments has been utilized effectively to help efforts that targeted initiation of major new efforts, i.e. new speeds of Ethernet
    - BWA1[1] > 400 GbE Study Group
    - BWA2[2] > Beyond 400 GbE Study Group
  - These efforts took time to solicit input, gather data, analyze, and build consensus
  - These assessments were of great value to the subsequent efforts.

1. IEEE 802.3 Industry Connections Ethernet Bandwidth Assessment Ad hoc, https://www.ieee802.org/3/ad_hoc/bwa/index.html
2. IEEE 802.3 Industry Connections NEA Ad Hoc Ethernet Bandwidth Assessment, Part II, https://www.ieee802.org/3/ad_hoc/bwa2/index.html

Prior Influence of AI in IEEE 802.3

ARTIFICAL INTELLIGENCE & COMPUTE

- First Era (Before 2012)
  - Moore's Law – 2-year doubling
  - Uncommon to use GPUs for machine learning
- Modern Era (2012 and later)
  - 2012 – 2014: most results used 1-8 GPUs rated at 1-2 TFLOPS
  - 2014 – 2016: large-scale results used 10-100 GPUs rated at 5-10 TFLOPS
  - 2016 – 2017: greater algorithmic parallelism (huge batch sizes, architecture search, expert iteration), specialized hardware (TPUs), faster interconnects



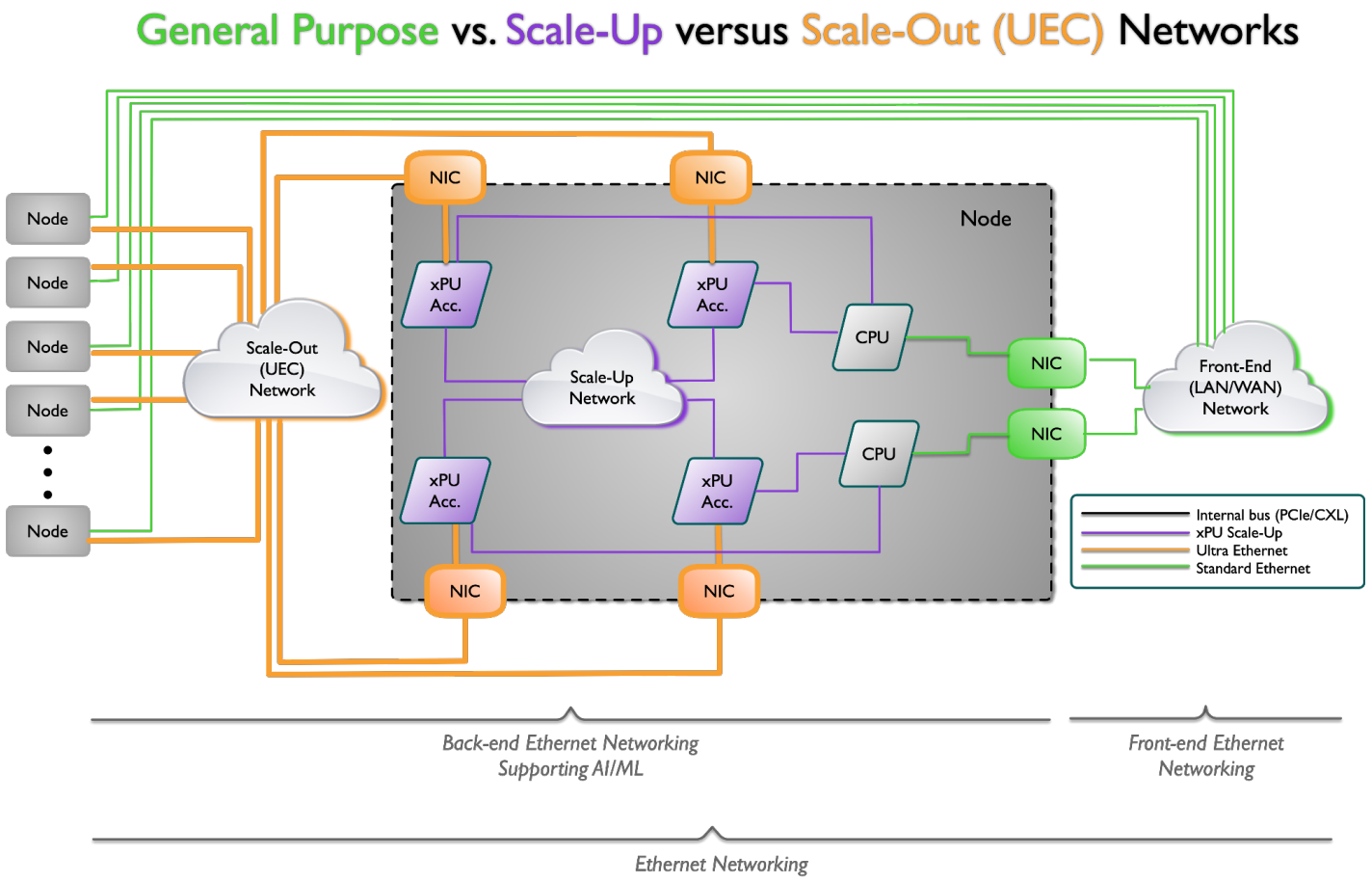
Source – OpenAI blog post 'AI and Compute' addendum 'Compute used in older headline results' posted 7th November 2019 by Girish Sastry, Jack Clark, Greg Brockman and Ilya Sutskever <https://openai.com/blog/ai-and-compute/>.

29 Oct 2020 IEEE 802.3 NEA Ad hoc - Beyond 400 Gb/s Ethernet CFI Consensus Presentation Page 14

- AI was cited as a factor influencing bandwidth growth in the "Beyond 400GbE" CFI

Source: https://www.ieee802.org/3/ad_hoc/ngrates/public/calls/20_1029/CFI_Beyond400GbE_Rev7_201029.pdf

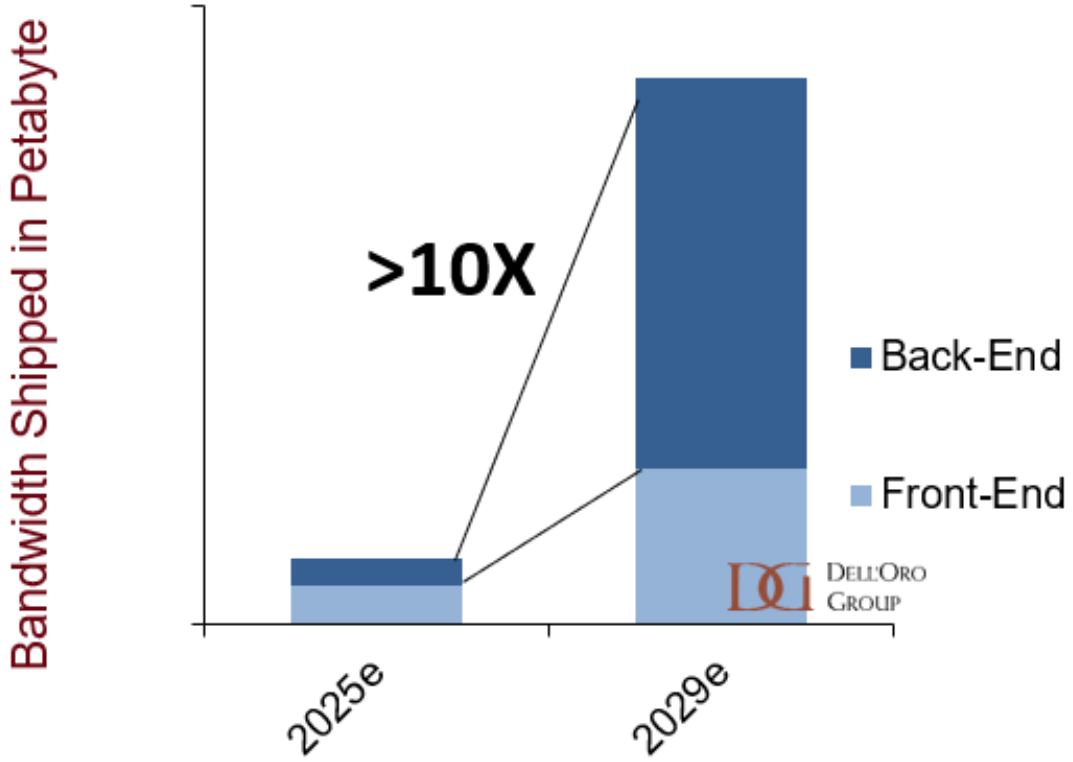
What are AI Networks?

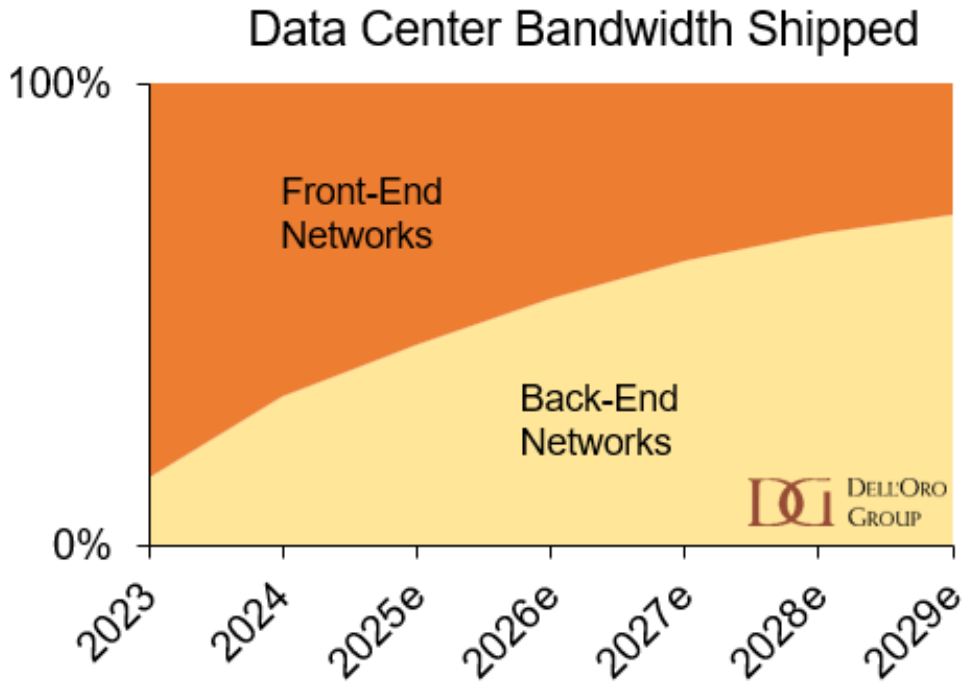


Source: Ultra Ethernet Consortium

- The author is aware that there are different representations of different implementations of AI Networks.
- The key takeaway is there are three types of networks for AI:
  - Front-end / traditional Ethernet
  - Back-end networks
    - Scale-up
    - Scale-out

Bandwidth Requirements in AI Networks





- Source: Dell'Oro AI Networks for AI Workloads Report
- Includes both Ethernet and InfiniBand
- Excludes Scale-up

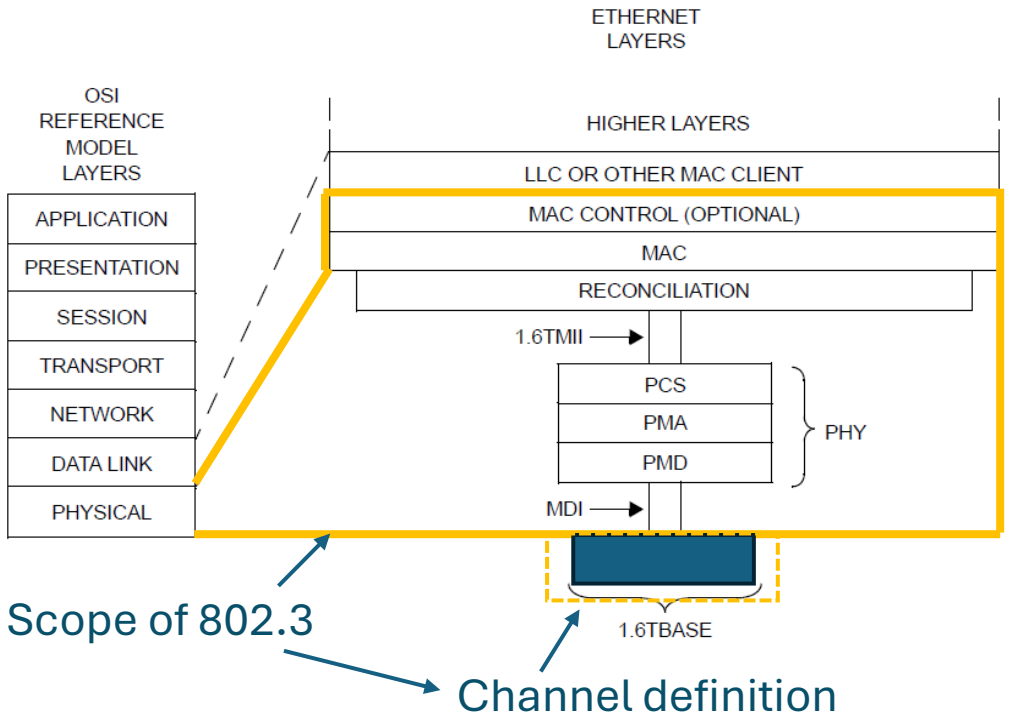
Recent Relevant Liaisons to IEEE 802.3 WG

- Nov 2024 – UA Link
  (https://www.ieee802.org/3/minutes/nov24/incoming/UALink%20liaison%20to%20IEEE802.3%2010-2024%20v4.pdf)
  - "…an industry body focusing on exploring and specifying optimizations for networks based on the 802.3 physical layer and supporting Artificial Intelligence (AI) and Machine Learning (ML) **scale-up networks** and workloads."

- July 2024 – Ethernet Alliance TEF
  (https://www.ieee802.org/3/minutes/jul24/incoming/Liaison_EA_to_802d3_240718_final.pdf)
  - "…there are calls to begin exploration of 400Gb/s per lane electrical and optical signaling to support the AI networks of the future…."

- Sept 2023 – Ultra Ethernet Consortium
  (https://www.ieee802.org/3/minutes/sep23/incoming/UEC%20liaison%20to%20802d3%20-%20Signed_Redacted.pdf)
  - "We wish to inform the IEEE 802.3 Working Group of the formation of the Ultra Ethernet Consortium (UEC), an industry body focusing on exploring and specifying optimizations for Ethernet-based networks supporting Artificial Intelligence (AI) and Machine Learning (ML) workloads."
  - Note –white paper / URL referenced in liaison changed –
    - https://ultraethernet.org/wp-content/uploads/sites/20/2023/10/23.07.12-UEC-1.0-Overview-FINAL-WITH-LOGO.pdf
  - Focus for D1.0 – **Scale-out Networks**

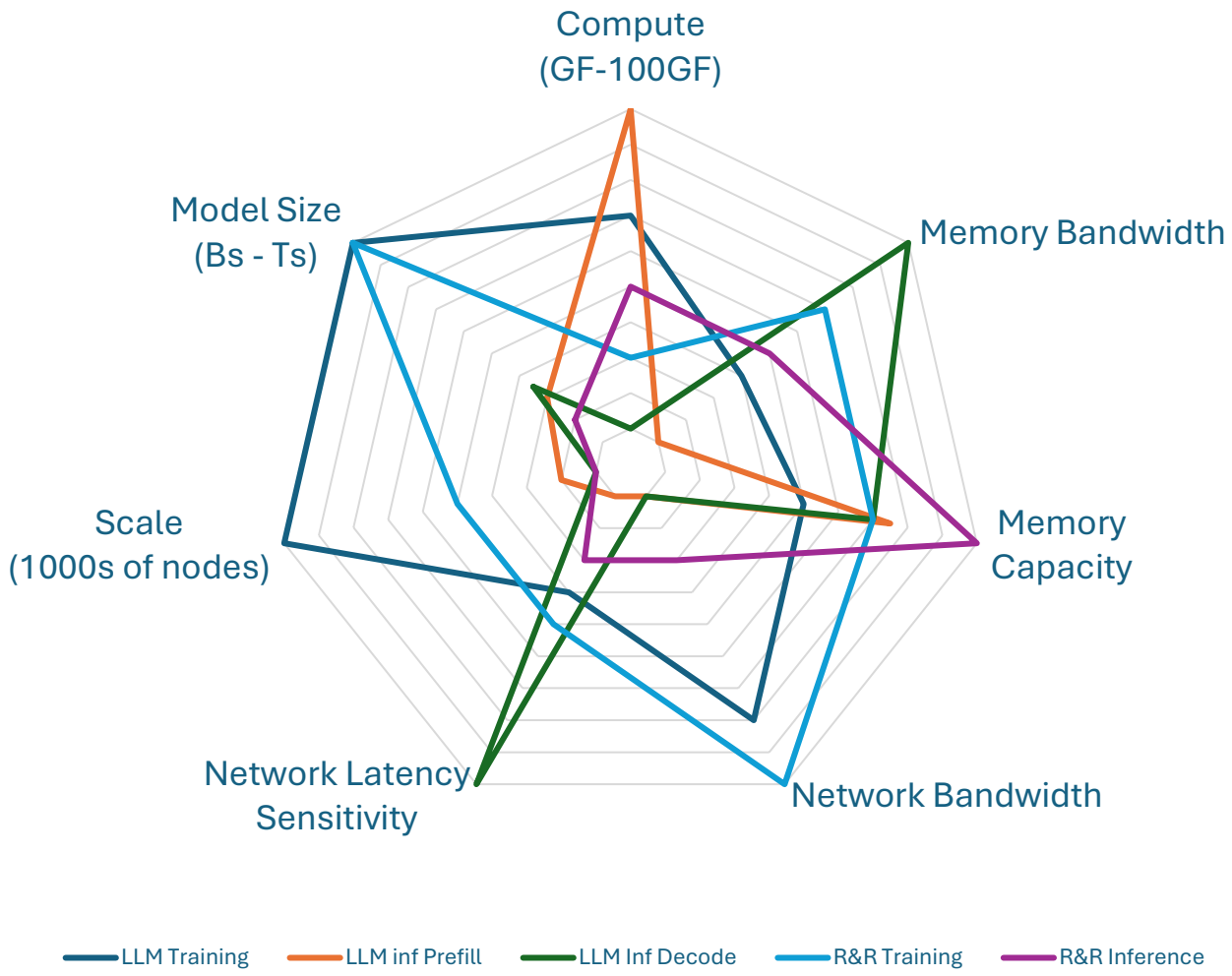
Important –
Scope of 802.3



Industry Events

- Ethernet Alliance – “Ethernet in the Age of AI” TEF
  - Reference - https://ethernetalliance.org/tef-2024-ethernet-in-the-age-of-ai/
  - Event Proceedings – https://ethernetalliance.org/tef-2024-ethernet-in-the-age-of-ai-presentations-form/
  - Keynotes
    - Ram Huggahalli (Microsoft) - “AI-Centric Datacenters and their Diverse Network Requirements”
    - Moray McLaren (Google) - “The Future of Networking for AI in Hyperscale data centers”
    - Nicolaas Viljoen (Meta) - “Ethernet-The foundation of AI @ Meta”
  - Presentations focused on 400 Gb/s electrical / optical signaling
    - Halil Cirit (Meta) – “System Overview and Exploring Alternative FEC Techniques for 400G Performance Enhancement”
- SFF / SNIA – “400G AI Workshop”
  - Focus on 400G copper channel requirements and designs for AI applications
  - Reference - https://www.eventcreate.com/e/aiworkshopbysniasff

AI Application Requirements



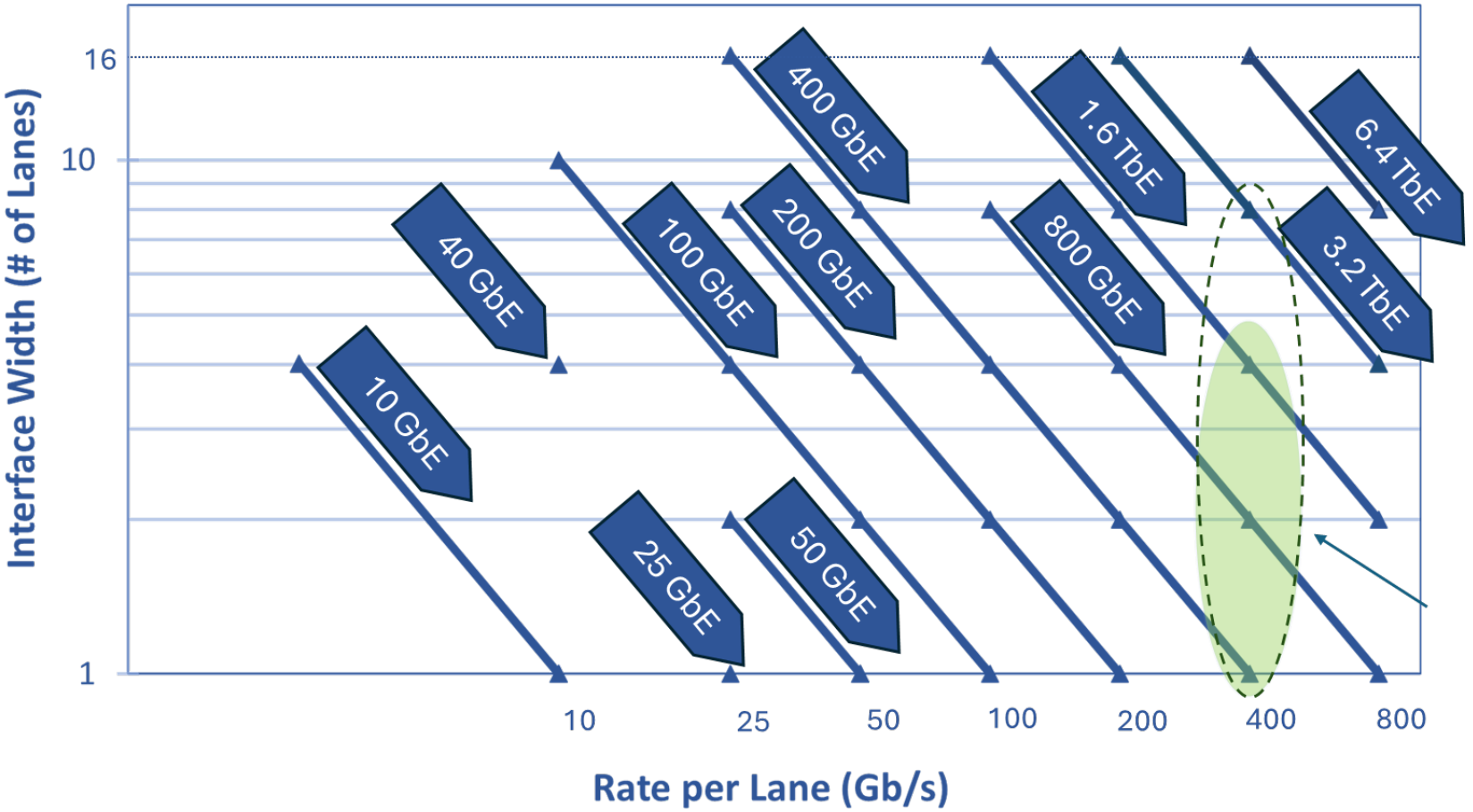
Ethernet Priorities (2 – 5 years)

- Resilience
- Reach
- Beachfront (I/O BW Limitation)
- Power

Latency can be traded off for resilience, reach and power

Source: Ethernet Alliance TEF 2024 Keynote – Ethernet – The foundation of AI @ Meta, Nic Viljoen

What lane rate / Ethernet rates are being targeted?



- What is approach to target lane rate –
  - Supports prior Ethernet rates?
  - Supports interconnect capacity?
- What Ethernet rates need to be supported?
  - Radix (and existing Ethernet rates) – 1, 2, 4 lanes (400/800/1600G)
  - Fat pipe (new Ethernet rate) - - 8 lanes (3.2T)

IEEE P802.3dj Objectives

Ethernet Rate	Signaling Rate	AUI	Backplane	Cu Cable	SMF 500m	SMF 2km	SMF 10km	SMF 20km	SMF 40km
200 Gb/s	200 Gb/s	200GAUI-1 C2C C2M	200GBASE-KR1	200GBASE-CR1	200GBASE-DR1	200GBASE-DR1-2			
400 Gb/s	200 Gb/s	400GAUI-2 C2C C2M	400GBASE-KR2	400GBASE-CR2	400GBASE-DR2	400GBASE-DR2-2			
800 Gb/s	200 Gb/s	800GAUI-4 C2C C2M	800GBASE-KR4	800GBASE-CR4	1.800GBASE-DR4 2.800GBASE-FR4-500	1. 800GBASE-DR4-2 2. 800GBASE-FR4	800GBASE-LR4		
	800 Gb/s						800GBASE-LR1	800GBASE-ER1-20	800GBASE-ER1
1.6 Tb/s	100 Gb/s	1.6TAUI-16 C2C C2M							
	200 Gb/s	1.6TAUI-8 C2C C2M	1.6TBASE-KR8	1.6TBASE-CR8	1.6TBASE-DR8	1.6TBASE-DR8-2			

What are the Target Interconnects?

	AUI (C2C / C2M)	Backplane (+cabled backplanes)	Cu Cable (CR)	SMF 500m (PSM / WDM)	SMF 2km (PSM / WDM)	SMF 10km (WDM/COH)	SMF 20km (COH)	SMF 40km (COH)
Front –End	X	X	X	X	X	X	X	X
Backend Scale-up	?	?	?	?	?	?	?	?
Backend Scale-out	?	?	?	?	?	?	?	?

- Front-end networks leverage IEEE 802.3 Ethernet.
  - Assumption- IEEE P802.3dj objectives are starting point
  - Additional needs?
- Backend network requirements? Similar to Front-End?
- Input needed

Identify latency performance requirements for each network

Another Perspective

	AUI (C2C / C2M)	Backplane (+cabled backplanes)	Cu Cable (CR)	SMF 500m (PSM / WDM)	SMF 2km (PSM / WDM)	SMF 10km (WDM/COH)	SMF 20km (COH)	SMF 40km (COH)
Front –End	X	X	?	?	?	?	?	?
Backend Scale-up	X	X	?	?	?	?	?	?
Backend Scale-out	X	X	?	?	?	?	?	?

“Inside the box interconnect”

“Box-to-box interconnect”

- Front-end networks leverage IEEE 802.3 Ethernet. Additional needs?
- Assumed that all networks will need “inside the box interconnect”
- What are the box-to-box interconnect requirements for each network?
- Phased Approach?

- Identify latency performance requirements for each network

Questions the Assessment will Explore

- What are the interconnect requirements for the different AI networks?
- What are the performance requirements of these interconnects?
- What are the priorities for the development of these interconnects?
- What tradeoffs can be made between latency and resilience / reach / power?

- This is a dynamic scenario where all of the answers appear to be interdependent

Liaison Relationships Request

These organizations are either exploring or defining technologies related to AI applications / networks or related signaling and interconnects:

- Open Compute Project (OCP) Foundation
- Ultra Ethernet Consortium (UEC)
- SNIA / SFF
- Ethernet Alliance
- Optical Interworking Forum (OIF)
- UALink Consortium

These organizations have access to significant technical competence and expertise in the application of Ethernet technology and specifications in AI networks. This knowledge could be beneficial to the IEEE 802.3 NEA “Ethernet for AI” Assessment.

THANKS