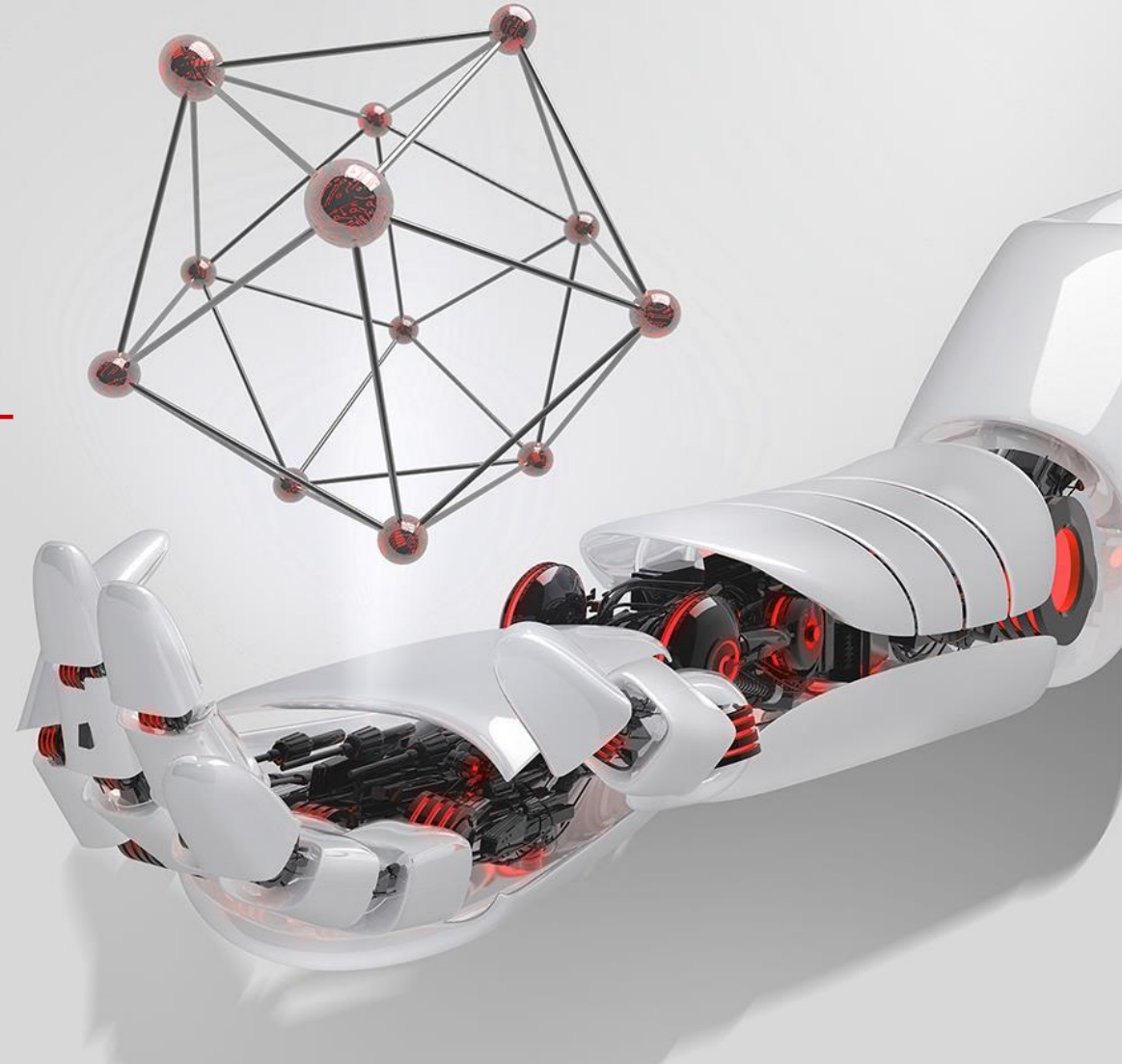


The IEEE 802.3 NEA “Ethernet for AI” Assessment

John D’Ambrosia
Chair, IEEE 802.3 NEA “Ethernet for AI” Assessment
Distinguished Engineer, IP Standards Team
Futurewei Technologies



Regarding the Views Expressed



My Industry Involvement

- Chair, IEEE 802.3 NEA “Ethernet for AI” Assessment Activity
- Chair, IEEE P802.3dj 200 Gb/s, 400 Gb/s, 800 Gb/s, and 1.6 Tb/s Ethernet Task Force
- Vice-Chair, IEEE 802.3 New Ethernet Applications (NEA) Ad Hoc
- Distinguished Engineer, Futurewei, U.S. Subsidiary of Huawei
- Email – jdambrosia@ieee.org
- Follow me on [LinkedIn](#)

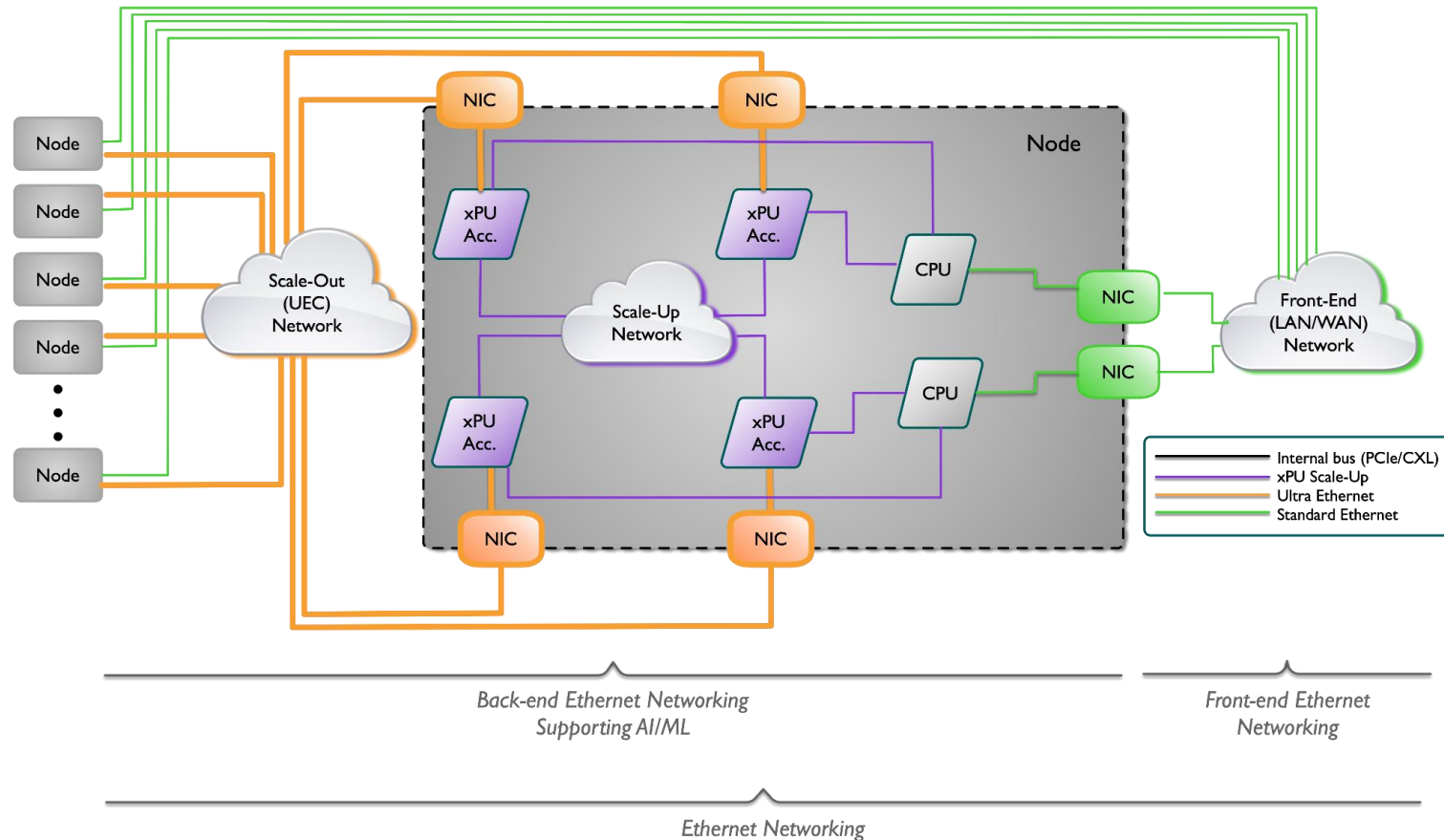


Per IEEE-SA Standards Board Bylaws

“At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position of IEEE.”

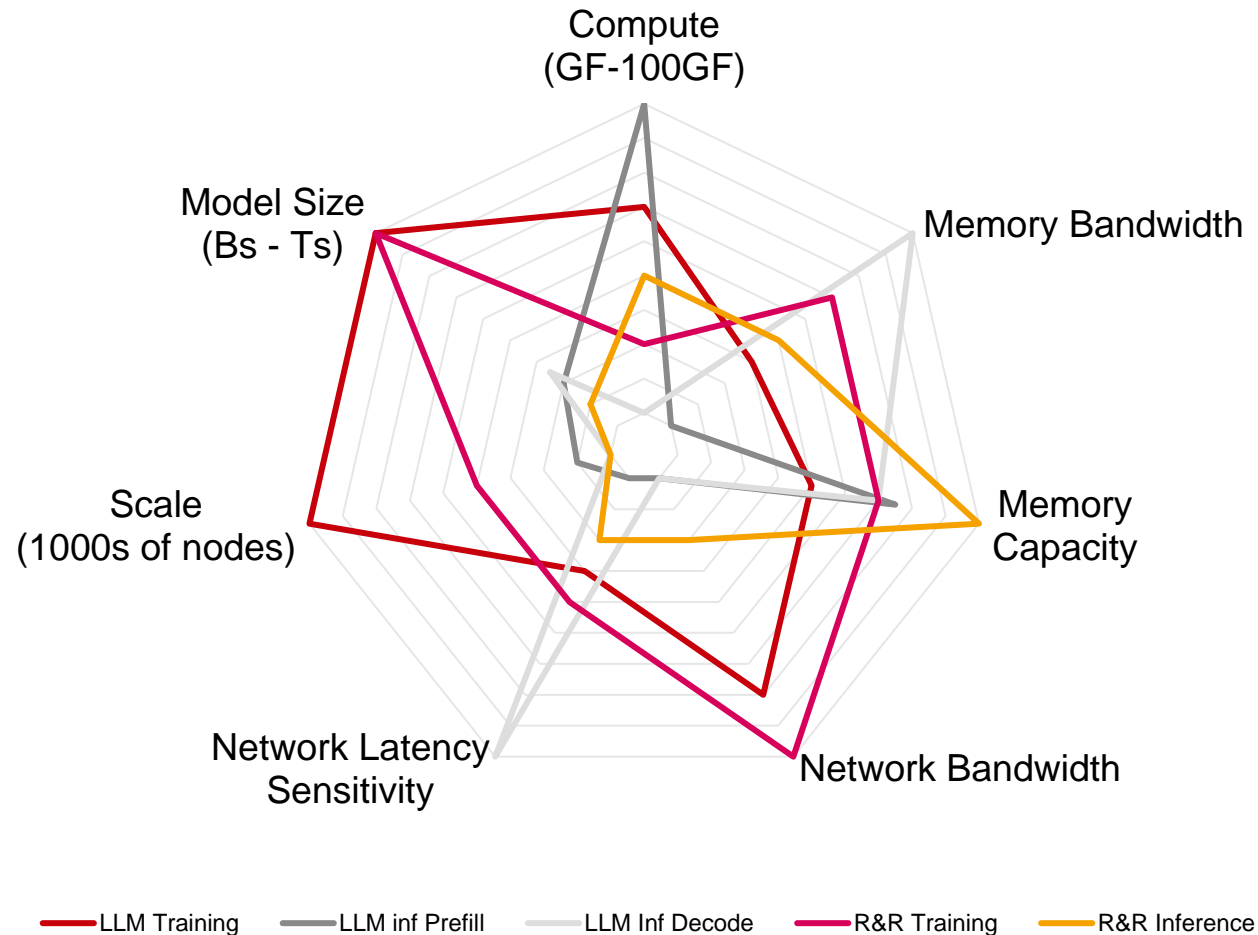
What are AI Networks?

General Purpose vs. Scale-Up versus Scale-Out (UEC) Networks



Source: Ultra Ethernet Consortium

AI Application Requirements Vary



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 are the Different Interconnect Requirements?

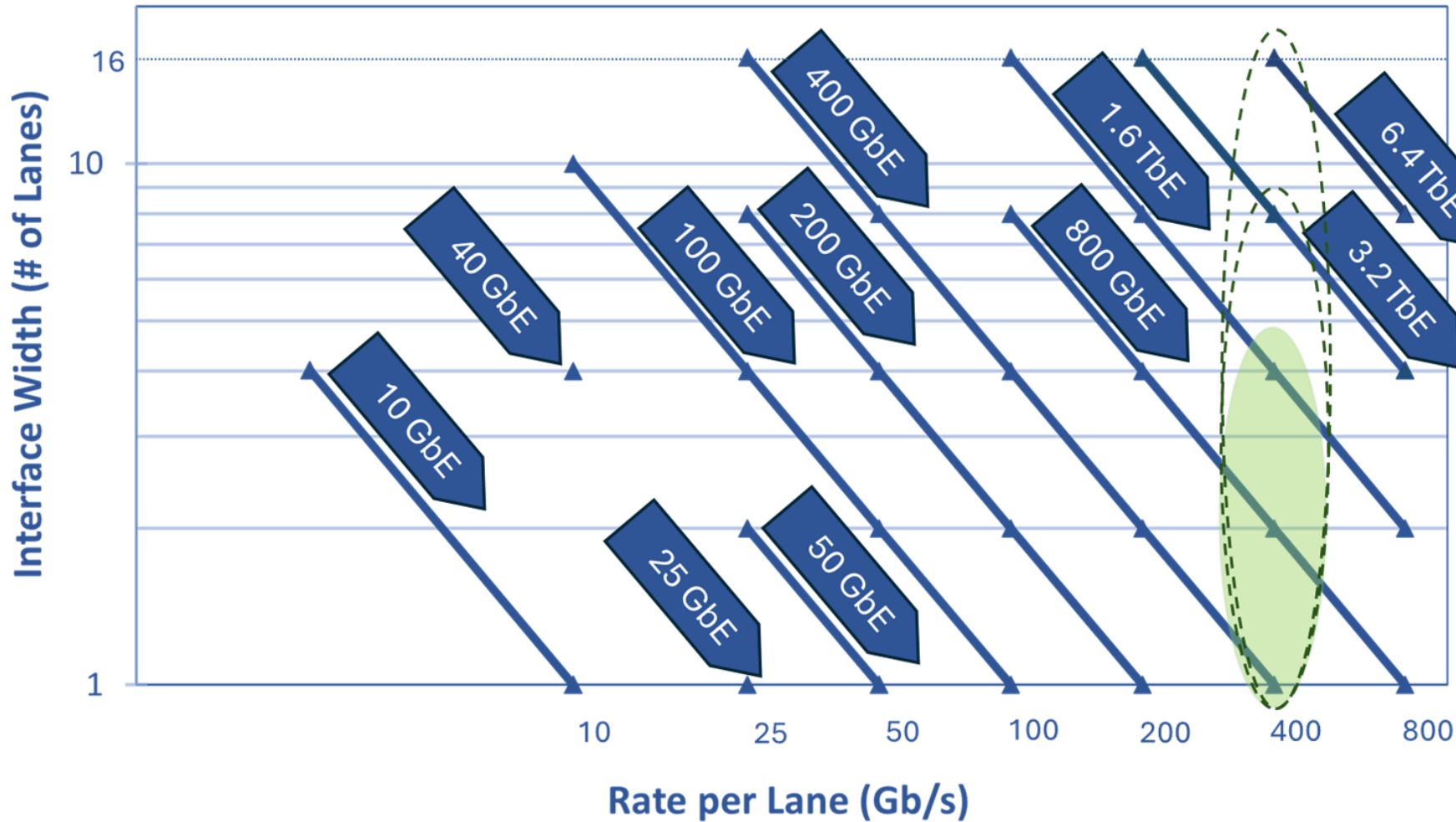
	AUI (C2C / C2M)	Backplane (+cabled backplanes)	Cu Cable	SMF IMDD (PSM)	SMF IMDD (WDM)	SMF Coherent
Front –End	X	X	?	?	?	?
Backend Scale-up	X	X	?	?	?	?
Backend Scale-out	X	X	?	?	?	?

“Inside the box”

“Box-to-box”

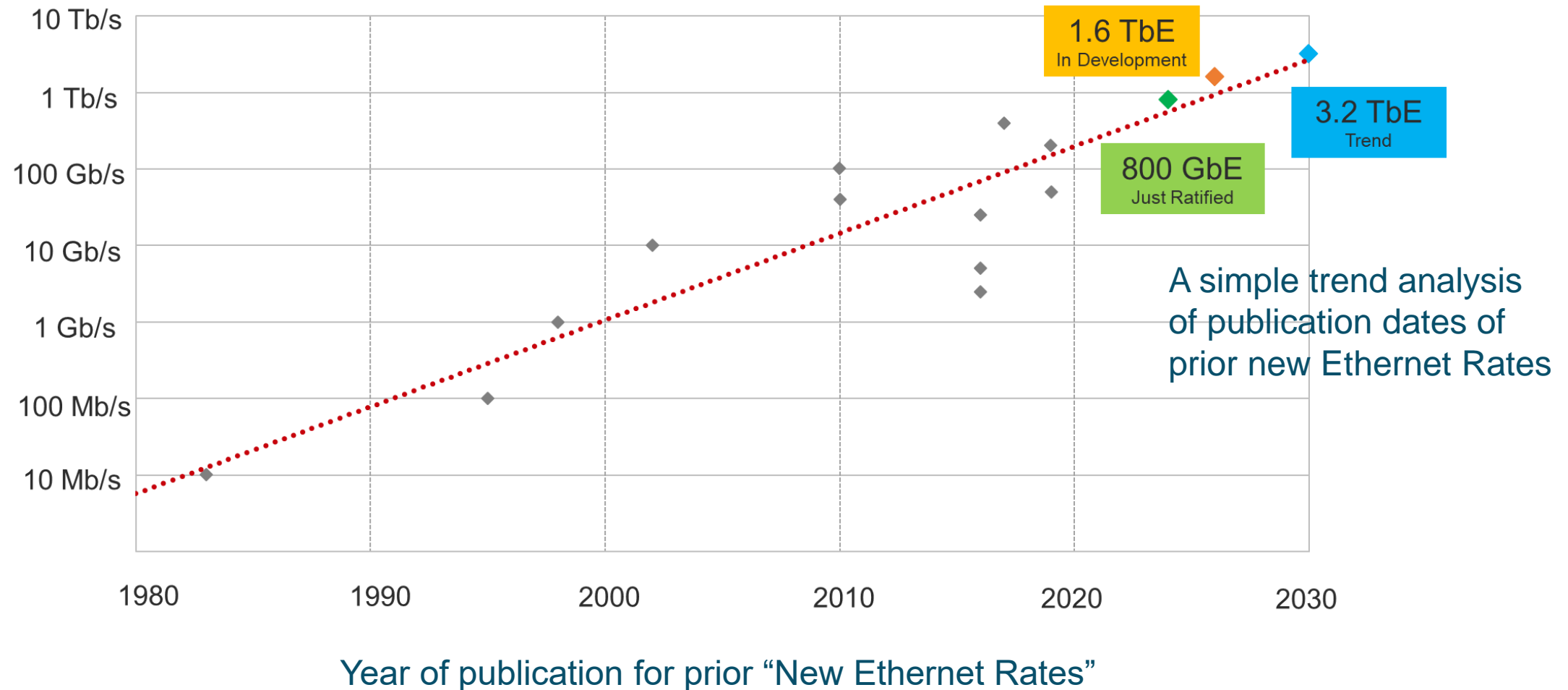
- Front-end networks leverage IEEE 802.3 Ethernet. Additional needs?
- Assumed that all networks will need “inside the box interconnect”
- What are the box-to-box interconnect reach requirements for each network?
- Priority for development?
- Identify latency performance requirements for each network

What 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)

Trend of IEEE 802.3 New “Higher Speed” Introduction



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

IEEE 802.3 “Ethernet for AI” Assessment

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

- Website - https://www.ieee802.org/3/ad_hoc/E4AI/index.html
- Reflector - https://www.ieee802.org/3/ad_hoc/E4AI/reflector.html