

Draft: 400 Gbps per lane PHYs and Interfaces

Call for Interest (CFI) Consensus Building

Kent Lusted, Synopsys
Mark Nowell, Cisco

Contributors

- John D'Ambrosia, Futurewei, US Subsidiary of Huawei
- Adam Healey, Broadcom
- Halil Cirit, Meta
- Rob Stone, Meta
- Matt Brown, Qualcomm
- Adeo Ran, Cisco
- Alan Weckel, 650 Group
- David Ofelt, HPE
- Jose Castro, Panduit
- Baron Fung, Dell'Oro Group
- Sameh Boujelbene, Dell'Oro Group

Goal for Tonight

- To present the draft CFI presentation covering the “Market Need”, “Technical Feasibility”, and “Why Now?” of 400 Gb/s per lane PHYs and Interfaces
- To build consensus towards a March 2026 IEEE 802.3 Working Group request to form a study group
- We are NOT discussing specific implementations or objectives – these are just some of the reasons that we need a study group!
 - Feedback welcome on how to improve the CFI presentation deck

Two Efforts Forming In the Pipeline

Moving Forward in IEEE 802.3

NEA (E4AI Assessment) – Consensus / Incubation - Ongoing stakeholder input

Focus of this
CFI

Effort #1

SG

Baseline
Adoption

TF Draft

Effort #2

SG

4 – 8 mos

Baseline
Adoption

TF Draft

>= 3.2 Tb Ethernet
and PHYs not in
Effort #1

* Each new project will potentially build on the previous project(s).
802.3 has the “Ethernet for AI” NEA effort which will sustain
incubation across many of these distinct Task Force projects to
encourage a fast-follow cadence.

Additional
Efforts

18

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0819/dambrosia_e4ai_01_250819.pdf

IEEE 802.3 NEA E4AI, January 2025

Objective

- Build consensus on starting a study group investigating a “400 Gbps per lane Signaling PHYs and Interfaces” project
- We do not need to:
 - Fully explore the problem
 - Debate strengths and weaknesses of solutions
 - Choose a solution
 - Create a PAR or 5 Criteria or the SG objectives
 - Create a standard
- Anyone in the meeting may vote or speak

Supporters

Overview

To keep pace with the technology demands necessary to support the projected growth in AI Infrastructure, 400 Gbps/lane data rate PHYs need to be defined in IEEE 802.3

To explain this better and propose the formation of a new Study group we plan to walk through the following material:

- Market Drivers
- Technical Feasibility
- Why Now? & Close
- Q&A (With panel)

CFI Request

With the rapid adoption of Artificial Intelligence (AI) tools and technology, there is a desire to use Ethernet in high-bandwidth, high-growth AI cluster networks, as well as traditional front-end networks for server and switch attachment. The Ethernet community must quickly respond to support the AI network time-to-market requirements and use cases. A number of technology improvements should be investigated to increase the per-lane signaling rate of electrical and optical Ethernet interconnects for the radix-optimized connectivity essential for communication between the high-performance compute devices and intermediate network switch devices. **This call for interest is to initiate a Study Group to develop a Project Authorization Request (PAR) and Criteria for Standards Development (CSD) for Physical Layer specifications supporting existing Ethernet rates using 400 Gbps-per-lane signaling for electrical interconnects as well as optical interconnects with reaches of up to 500 meters.**

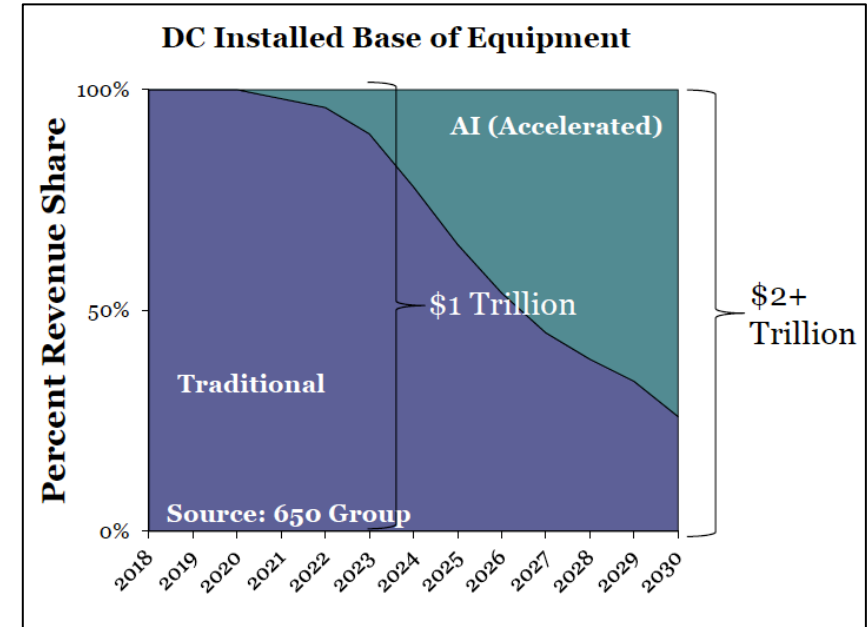
https://ieee802.org/3/cfi/request_0326_1.html

Market Need

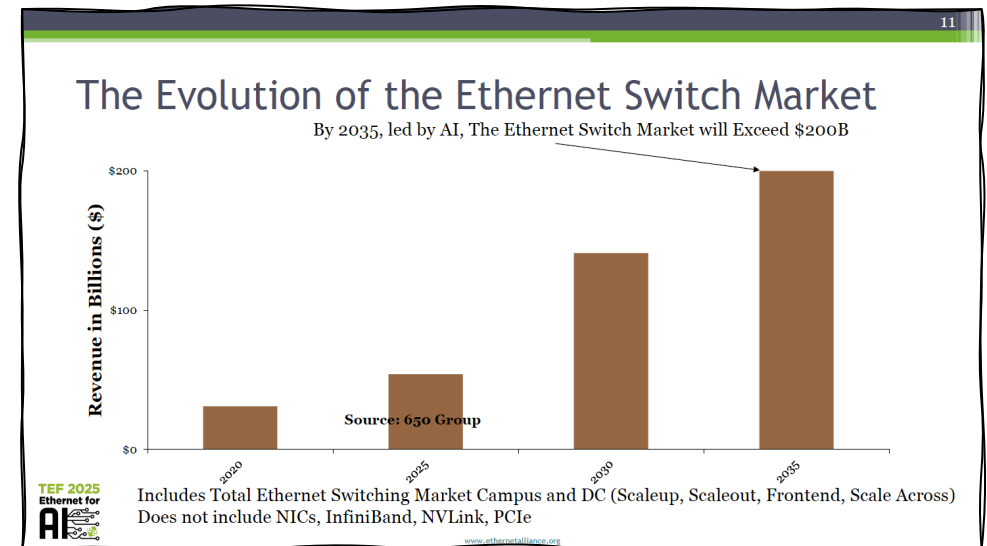
Mark Nowell

AI Explosion

- The enormous growth of AI/ML is placing unprecedented **demand on system interconnects and on scaling the bandwidth**, specifically xPU-to-xPU (via switches)
- Ethernet is uniquely positioned to be the building block of choice for AI/ML system interconnects
- The scale of the deployments requires a robust ecosystem to support the volumes.
- In networking, Ethernet wins due to its robust ecosystem and supply chain



Source: Alan Weckel, EA's TEF 2025, "Setting the stage for networking in an AI world"

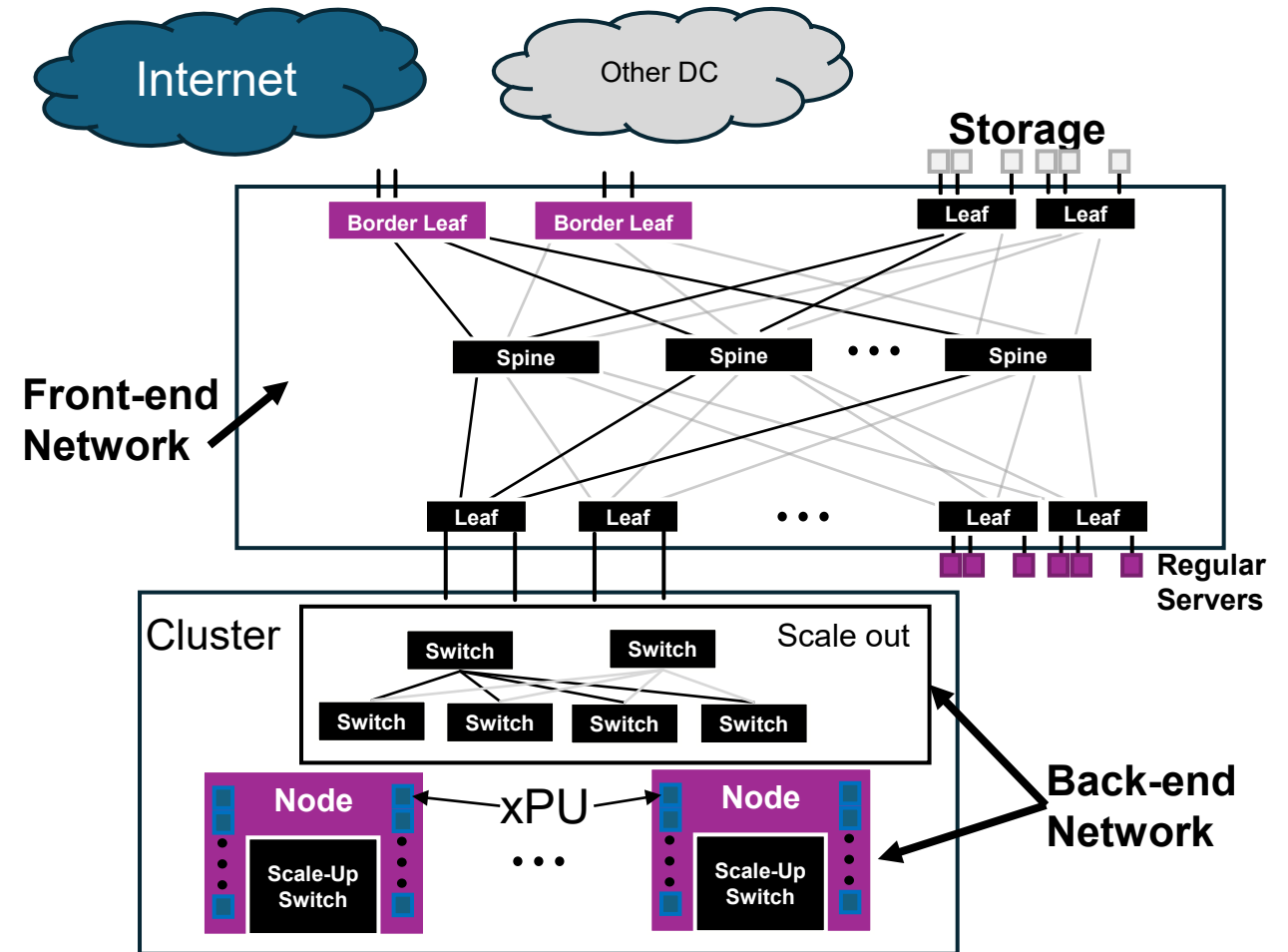


Source: Alan Weckel, EA's TEF 2025, "Setting the stage for networking in an AI world"

Network Evolution

The network is evolving into distinct networks with both common yet unique needs

- Front-end is the classic data center switching network connecting servers, storage, internet, etc.
- One of the “servers” can be a large AI cluster
- An AI cluster is a set of compute resources (CPUs, GPUs, accelerators) tied together by a “back-end” network.
 - The entire cluster can be viewed as a single server



Source: David Ofelt, HPE

Looking Closer at the Backend Network

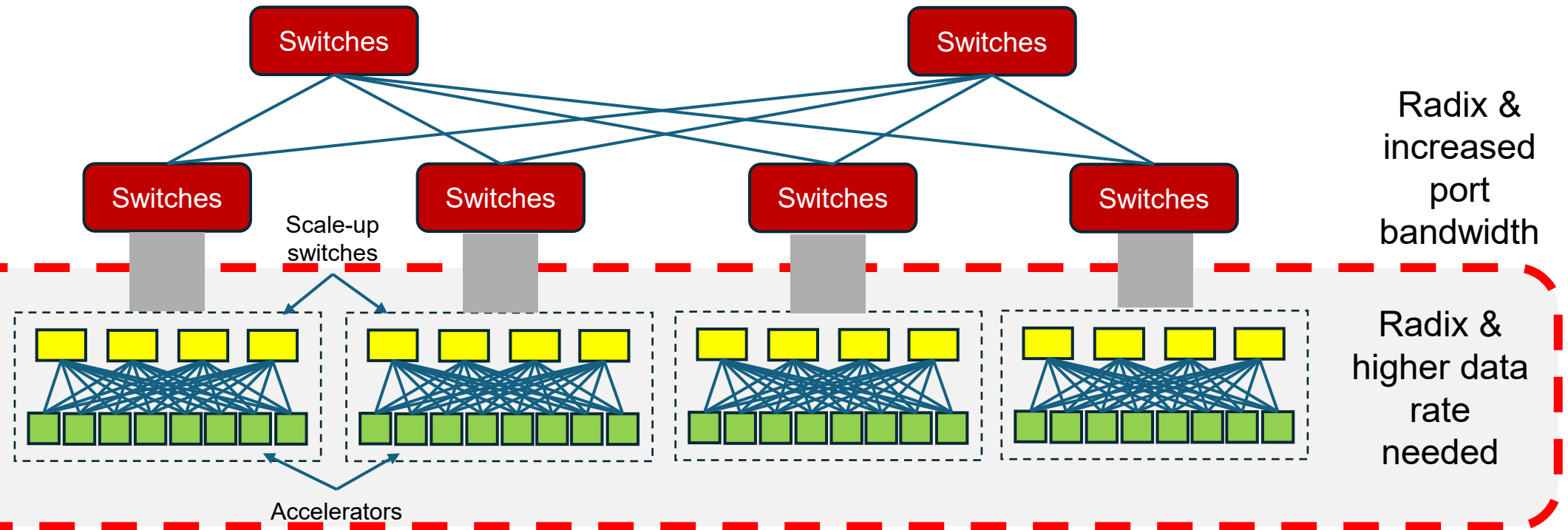
Front-end

Today's Ethernet

Fat Pipes
needed

Back-end
Scale-Out

Back-end
Scale-Up



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

IEEE 802.3 NEA E4AI, January 2025

Looking Forward

- AI back-end “scale-up” networks are driving higher signaling rates first
 - Followed by scale-out
 - Then front-end networks
- Scale-up = high radix
 - Flat network, lots of one-lane and two-lane links
- Copper excels due to low cost & power and high reliability, but the reach is limited and shrinking
- Optics have a growing footprint in scale-up (and, of course, scale-out) to meet reach and cluster size requirements
 - Power optimized solutions, such as Co-packaged optics (CPO) implementations, are increasingly important

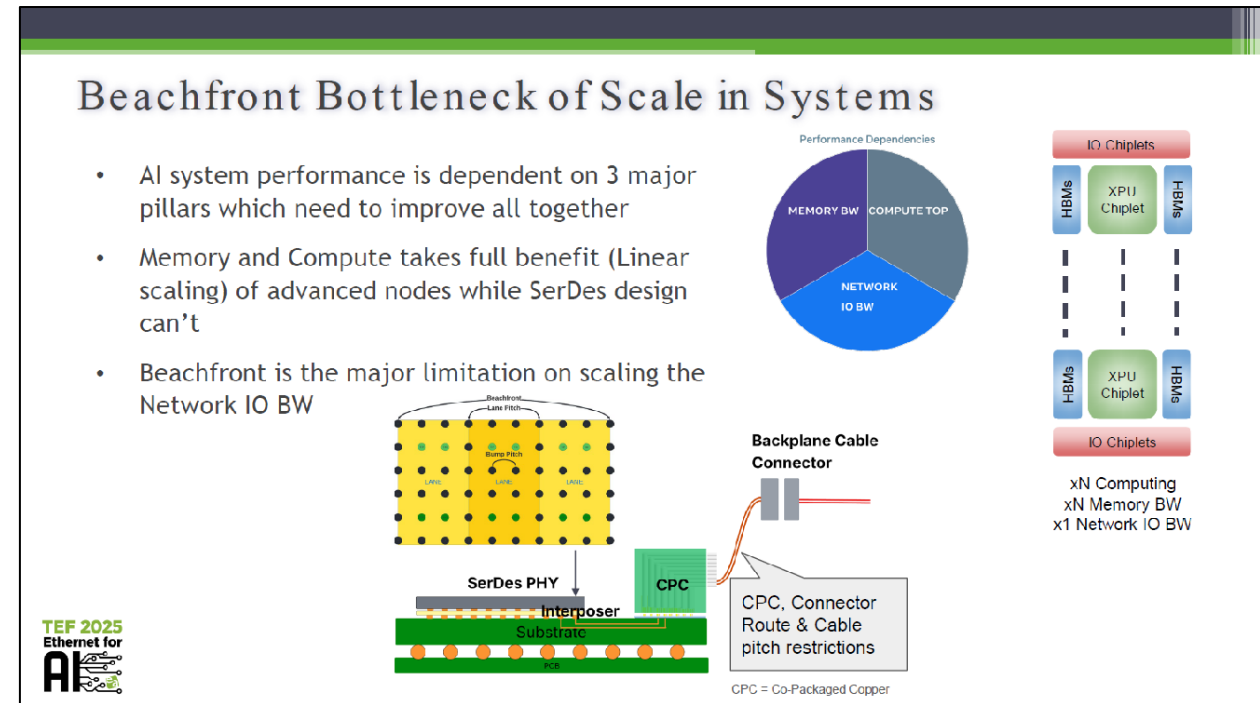
Projected relative growth of different networks (relative to front-end)

Front-end	1x
Back-end network (scale-out)	10x
Back-end network (scale-up)	100x

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/weckel_e4ai_01_250327.pdf

Why Higher Data-Rate Physical Layers?

- Compute devices and Switch devices are being driven to higher capacities
- Beachfront density limits the number of lanes; thus, a higher rate per lane is required
- Consistent with historical progression of Ethernet lane speeds for data center applications

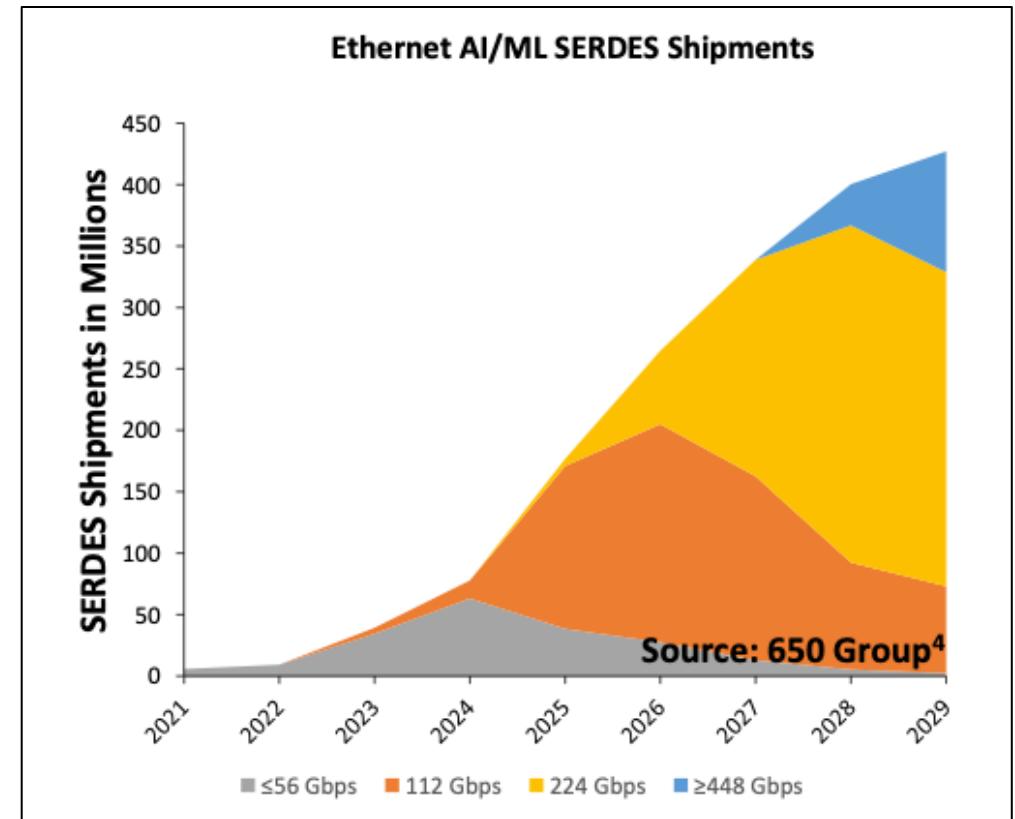


Source: Halil Cirit, EA's TEF 2025, "Transitioning to 400G SerDes: Key Drivers and System Design Implications for Future AI Workloads"

Wait for upcoming E4AI NEA session
before finalizing graphic

Market Projections for 400 G/lane

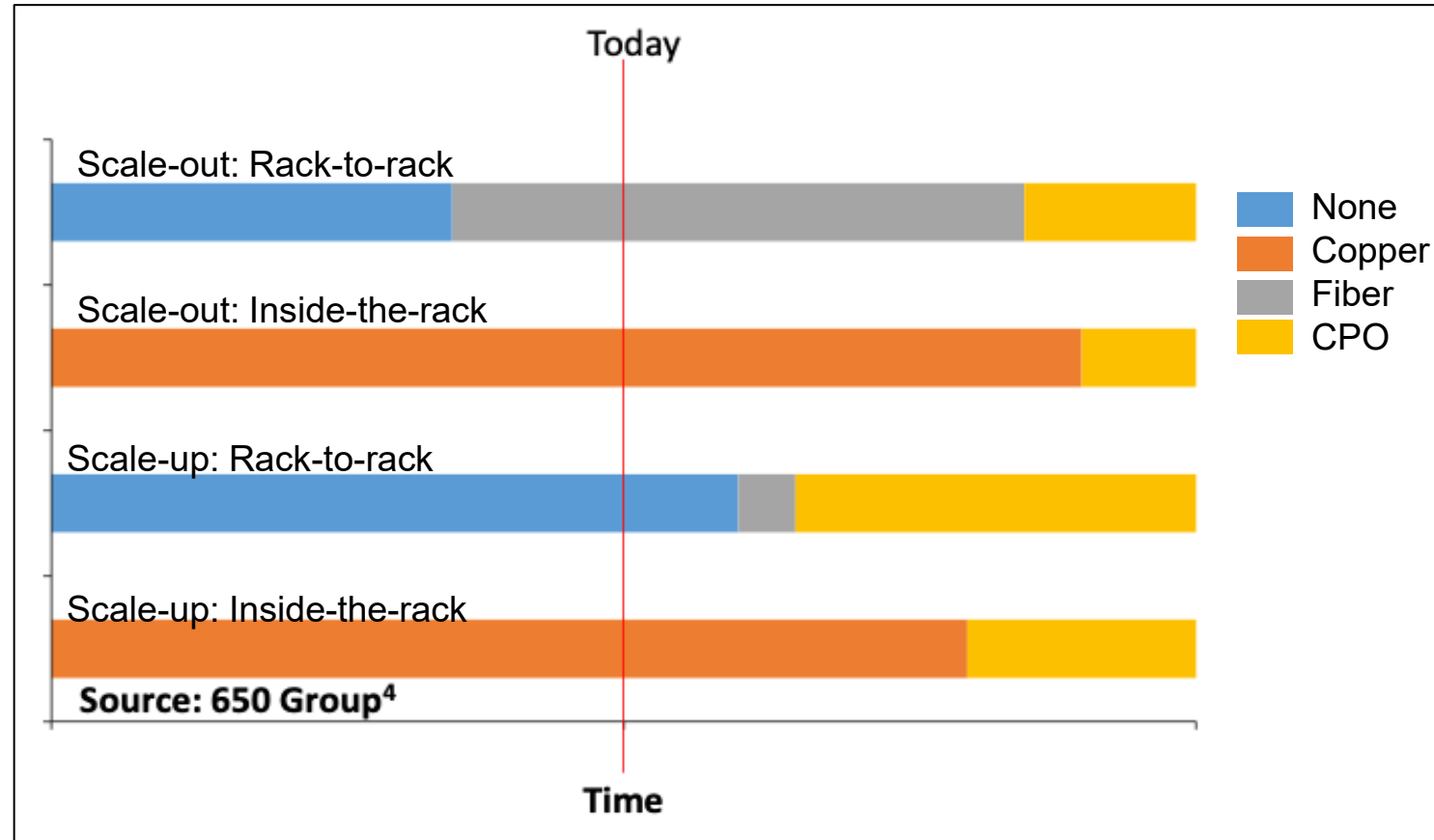
- Market analysis is predicting the market need for 400 Gb/s signaling starting in 2027
- Driven by AI networks getting bigger and constraints at the Switch/GPU, the transition to 400 G/lane technology is natural



https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/weckel_e4ai_01_250327.pdf

AI I/O Infrastructure: Both Electrical and Optics

- Current and future transitions involve Copper, pluggables, and CPO
- Copper dominates inside-the-rack connections today but shifts toward optics over time
- Rack-to-rack connections transition from pluggable to CPO
- Market requires both copper and optical PMDs @ 400G

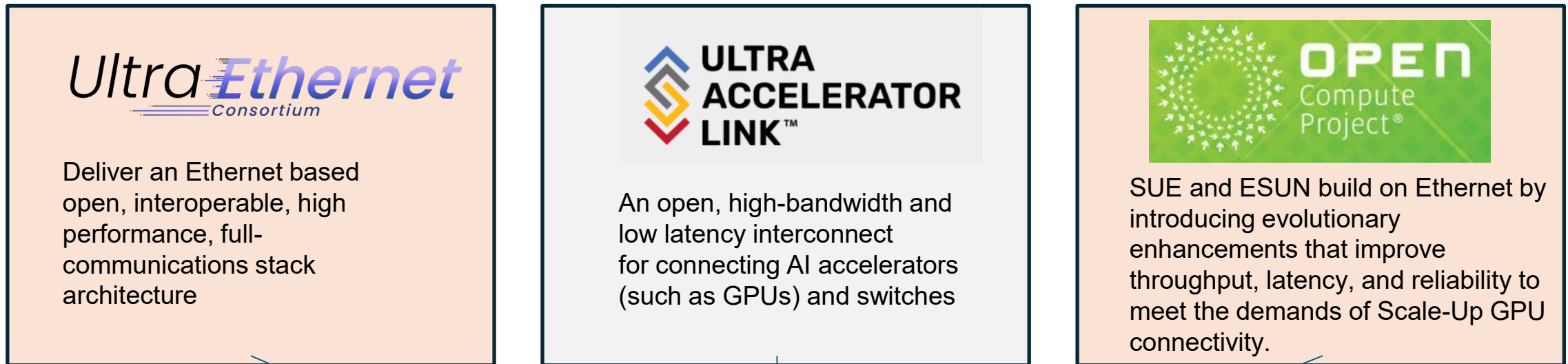


Source: https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/weckel_e4ai_01_250327.pdf

Requirements for BOTH electrical and optical PMDs

Many Consortia are Relying on Ethernet

To support the breadth of AI applications, multiple organizations exist:



All are looking for IEEE 802.3 to define the Ethernet Physical Layer Specifications

Summary: “Market need”

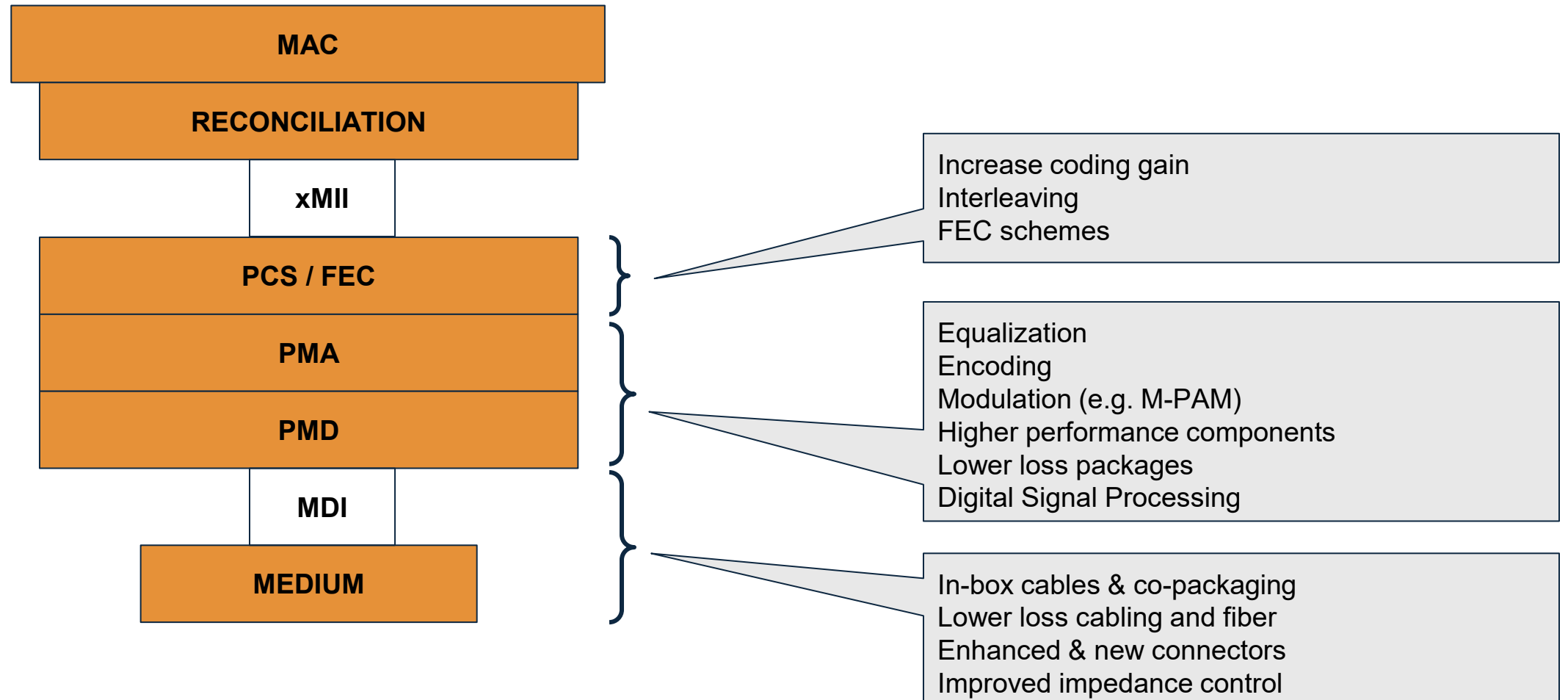
- AI scale-up networks are the fastest growing segment and require 400 Gb/s lane rates
- AI networks needs rapid development of the next physical layer data rate
- Ecosystem dynamics are all pointing to Ethernet (and IEEE 802.3 specifically) to provide the industry with these key building blocks
- Initial market is for high radix solutions supporting both electrical and optical interfaces

Technical Feasibility

We have done this before...

1G → 10G → 25G → 50G → 100G → 200G

Opening the Toolbox Again



SERDES Capability is the Enabler

SERDES is a key enabler for every interface: optical (FPP, NPO, CPO), cable, connector, backplane, retimers, etc.

In the past, we found new means to increase the per-lane signaling rate, including modulation, encoding, FEC

Electrical interface evolution

Data rate per lane, Gb/s	10	25	50	100	200	400
Modulation	PAM-2	PAM-2	PAM-4	PAM-4	PAM-4	?
Nominal cable reach, m	7	5	3	2	1~2 ¹	?
Technology added	DFE	RS FEC	Stronger RS FEC	Floating-tap DFE	MLSD	Inner FEC?
Year ²	2007	2014	2018	2022	2026 (est.)	?

- New tools pulled from the toolbox with each generation
- Used to address challenges presented by doubling the data rate
- What if the next set of tools includes an inner error correcting code?



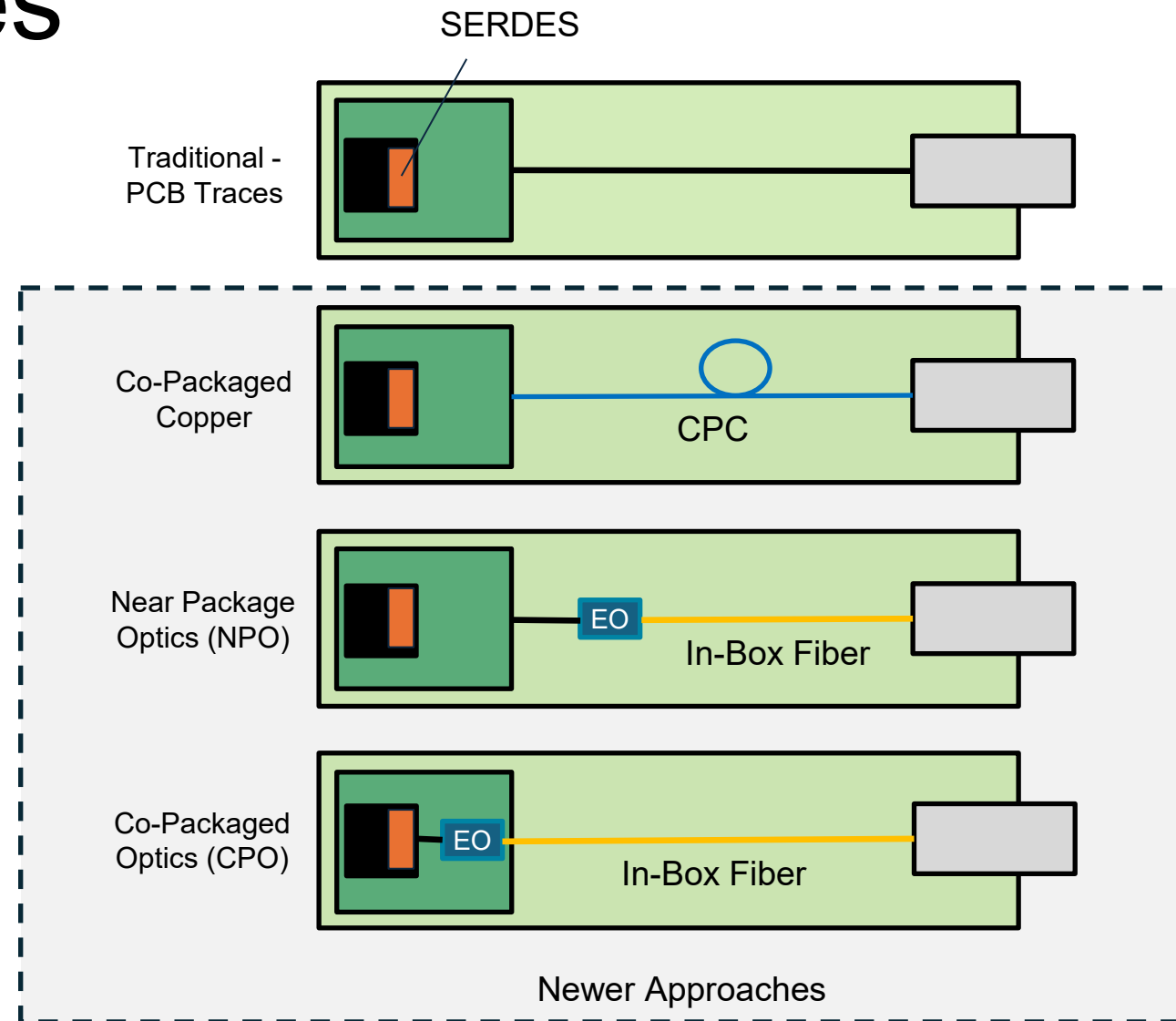
¹ Nominal cable reach is a function of host class
² Year of publication of the IEEE 802.3 Ethernet standard.

DFE = decision feedback equalizer
FEC = forward error correction
MLSD = maximum-likelihood sequence detector
RS = Reed-Solomon

Source: Adam Healey, EA's TEF 2025, "Benefits and limitations of inner error-correcting codes for 400 Gb/s per lane electrical links"

System Architectures

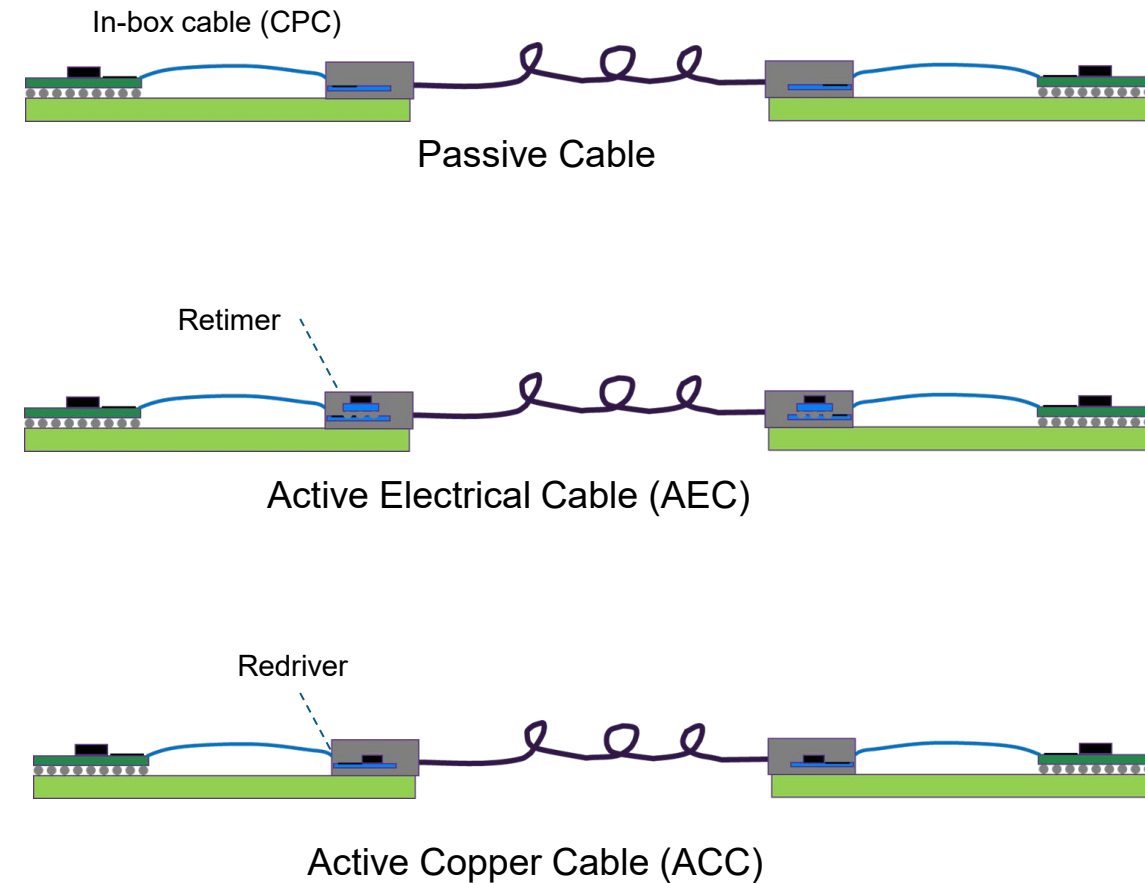
- Broader mix of system architectures
 - PCB traces to front panel pluggable (FPP)
 - Co-packaged copper to FPP
 - Near-Package Optics (NPO)
 - Co-Packaged Optics (CPO)
- A variety of approaches depending on system requirements & limitations
- All need SERDES



Copper Interconnect

- Numerous solutions for copper can be used to address the needs of the market
 - Passive:
 - Well known and well-established methodologies
 - AEC:
 - Comparable to a pluggable optical transceiver
 - ACC:
 - Not well specified across industry

Type	IC	Reach	Power
Passive	None		
Active Electrical Cable (AEC)	Retimer		
Active Copper Cable (ACC)	Redriver		



Optics and Fiber

Optics provides extended reach vs. copper, but with more overall power

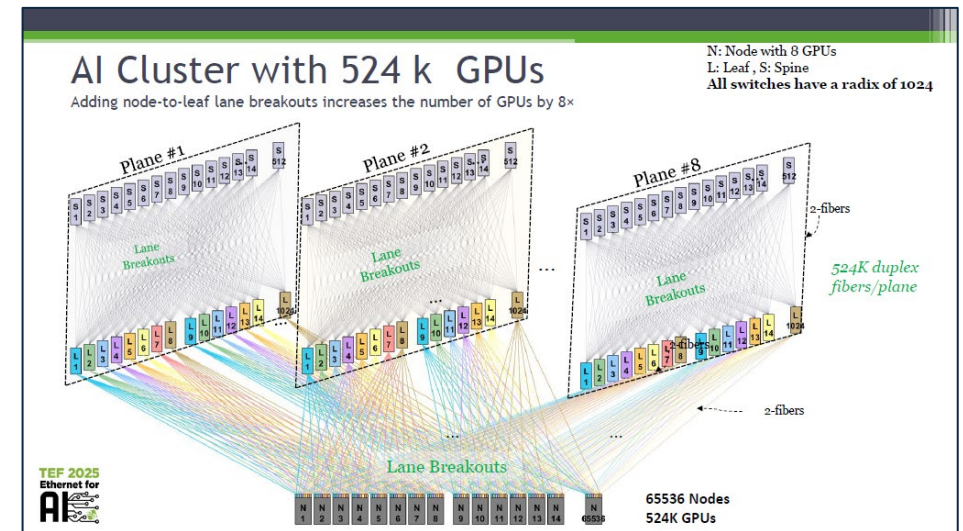
Single-lane optics and parallel single-mode provide the best support for high-radix scale-up (and some scale-out) networks

- High bandwidth per fiber enables dense connectivity
- Low dispersion from single-wavelength operation
- Potential lower power through shared laser resources

Minimal additional FEC beyond host FEC

Aligned with industry directions (breakout, pluggables, CPO)

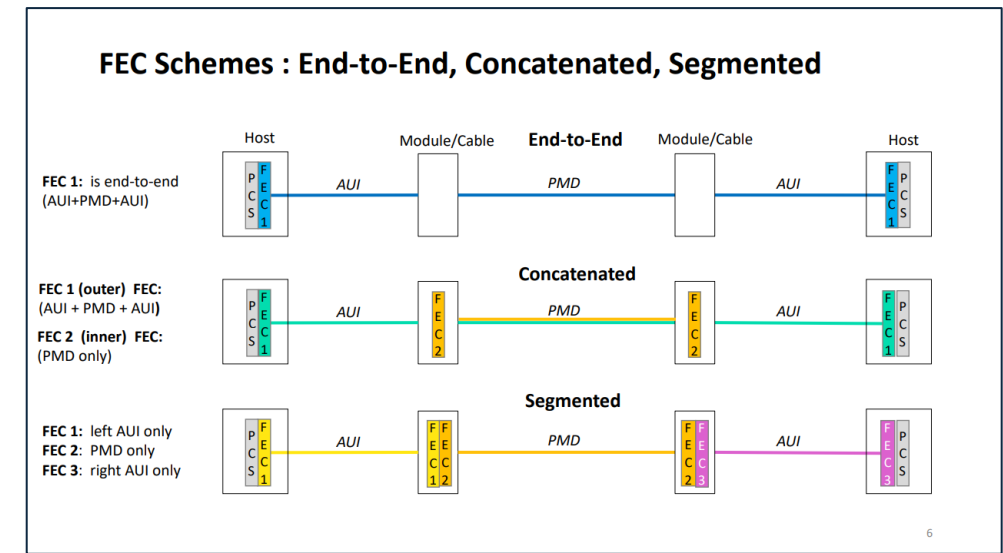
Robust, mature ecosystem of components and technologies



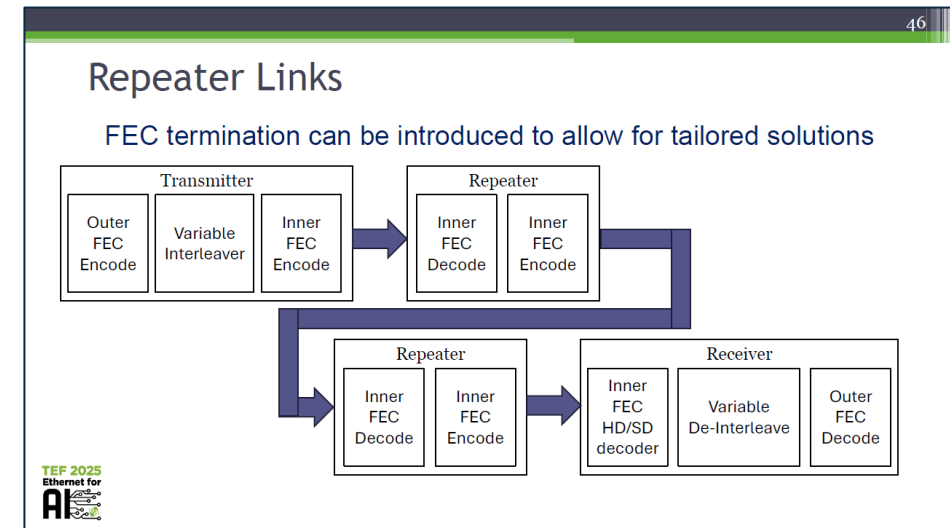
Source: Jose M Castro, EA's TEF 2025, "Enabling Massive Scale-Out AI Networks with Ethernet and Optical Lane Breakouts"

FEC Schemes

- Multiple FEC schemes are available for 400 G/lane
- Each type has distinct latency, coding-gain, and implementation-complexity tradeoffs that require detailed study
- The diversity of options enables selecting the FEC scheme and associated architecture that best aligns with system-level performance, power, and interoperability requirements



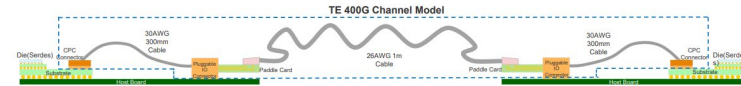
https://www.ieee802.org/3/df/public/22_02/shrikhande_3df_01a_220203.pdf



Source: Tony Chan Carusone, EA's TEF 2025, "Channel-Aware Modulation and FEC Selection for 400G+ Ethernet"

Work is Already Underway

Direct Attach Copper Cable Channel (tracy_eai_03_250430)



CPC Connector

- Includes connection to substrate and internal cable termination.
- Does not include substrate footprint and breakout.

Pluggable IO Connector

- Includes internal cable termination.

Paddle Card

- Includes external cable termination and paddle card routing and bulk cable

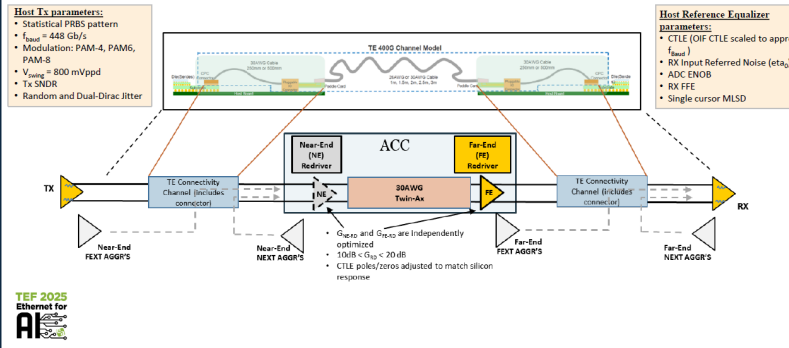
Channel File Name: tracy_eai_03_250430.

Models are preliminary and under development.

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0430/tracy_eai_01_250430.pdf

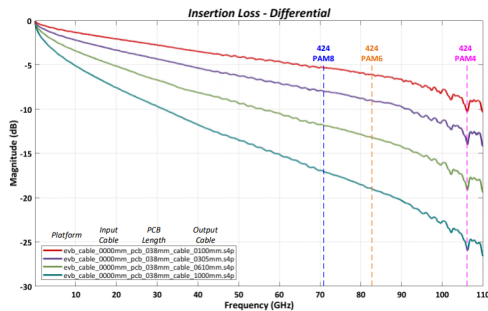
- Cable, Backplane and C2M, C2C channels
- Active Cables
- Connectors

ACC Detailed System Block Diagram



Source: Ashika Pandankeril Shaji, EA's TEF 2025, "448G/lane Interconnects for AI: Channel Feasibility, Reach Extension, and Early Measurement Results"

EVB Measured Insertion Loss



- 4.2dB to 14dB 212Gbps Channels
- No notch filter response to 110GHz
- 1dB ILD BW ~ 95GHz
- 15dB RL BW ~ 95GHz
- P/N Skew < 0.5ps
- 1.65dB/25mm PCB Striplines @53GHz
- Independently measured coaxial 2-ports provided as low skew 4 ports.

Contents in : josephson_e4ai_02_250430.zip

IEEE 802.3 New Ethernet Applications Ad Hoc

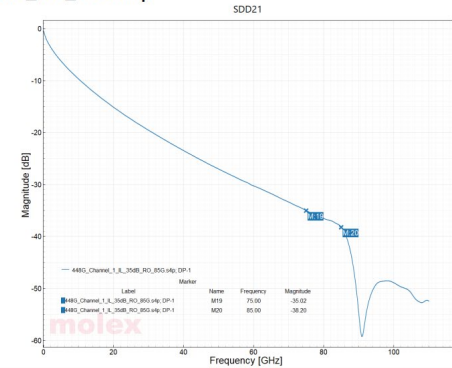
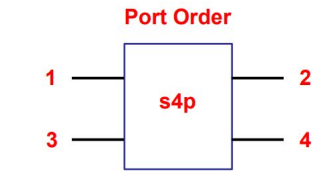
3

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0430/josephson_e4ai_01a_250430.pdf

"TEF 2025 presentations will be publicly available in early February 2026. Images in this presentation are used with permission from the Ethernet Alliance and the respective authors."

IL Channels

Filename: 448G_Channel_1_IL_35dB_RO_85G.s4p

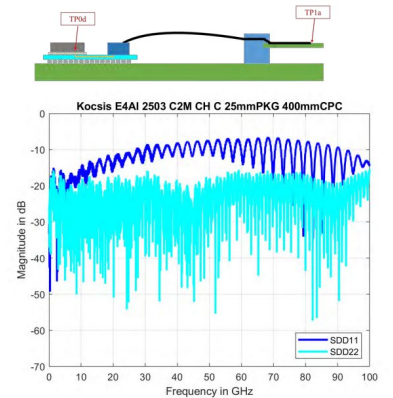
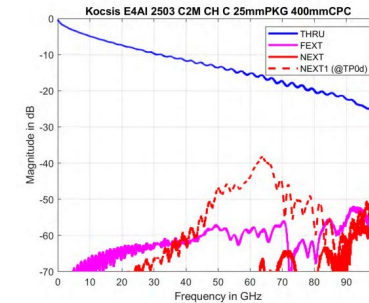


molex

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0430/shah_eai_01a_250430.pdf

Channel "C"

"New Pluggable, New Paradigm"

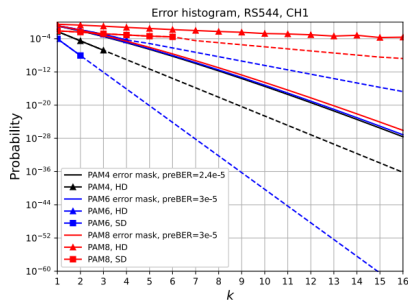


https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0327/kocsis_eai_01_250327.pdf

Underway... (2/3)

Block Error Ratio Evaluation for CH1

- We plotted the error statistics for CH1 using block error ratio method using RS(544,514).



- PAM4 performance is excellent, passing the error mask with broad margin.
- PAM6 using RS(544,514) only cannot pass the error mask.
 - PAM6 using Inner FEC has best performance.
- PAM8 using RS(544,514) with or without Inner FEC cannot meet KER requirement.

IEEE 802.3 NEA "E4AI" ad hoc - Oct. 23, 2025

8

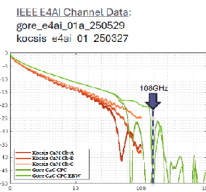
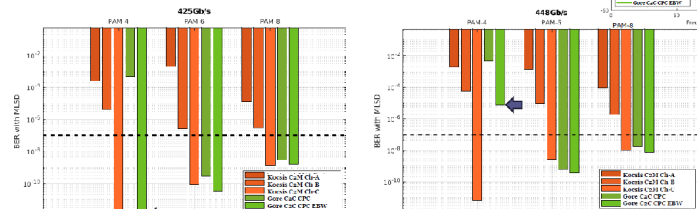
https://www.ieee802.org/3/ad_hoc/E4AI/public/25_1023/he_e4ai_01_251023.pdf

- FEC
- Package
- Encoding & Modulation
- COM analysis

Data here

Pre-FEC BER Results

- CPC-based C2C (0.5m) and C2M (0.4m) channels
- With 7% additional overhead, a pronounced pre-FEC BER increase is observed for some channels



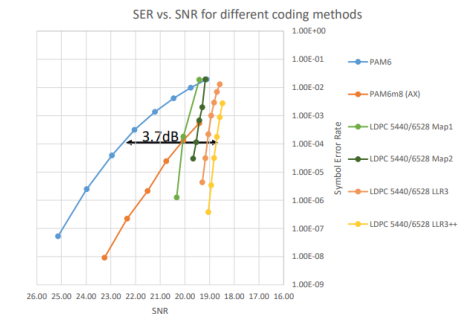
Source: Tony Chan Carusone, EA's TEF 2025, "Channel-Aware Modulation and FEC Selection for 400G+ Ethernet"

"TEF 2025 presentations will be publicly available in early February 2026. Images in this presentation are used with permission from the Ethernet Alliance and the respective authors."

PAM6m8 compared to PAM6

- Plot SER vs. SNR for

- PAM6
- PAM6m8 AX CC8
 - 2.4 dB coding gain vs. PAM6
- LDPC Simple LLR Map
 - 2.4 dB coding gain vs. PAM6
- LDPC Improved LLR Map
 - 3.0 dB coding gain
- Updated LLR Maps
 - 3D Map gives 3.4 dB coding gain
 - 3D Map + Patching gives 3.7dB coding gain



Note: Coding gains taken at SER=1e-4

17

IEEE 802.3 New Ethernet Applications Ad Hoc Ethernet for AI
October 23, 2025

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_1023/graumann_e4ai_01_251023.pdf

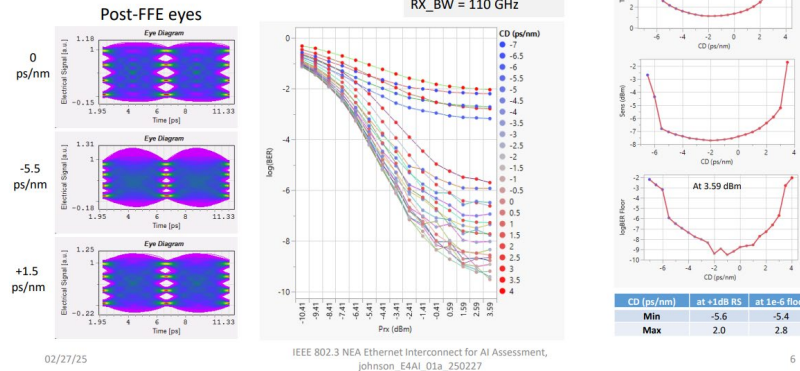


Underway... (3/3)

- Optics Analysis
- Test & Measurement
- Demos

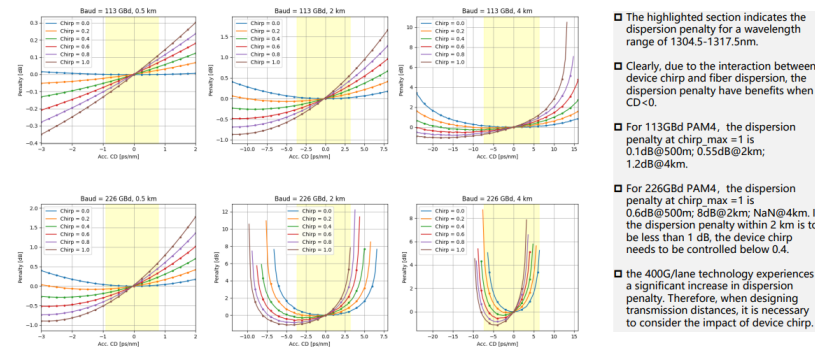
226 GBaud PAM-4

TX_BW = 110 GHz
TX_RIN = -150 dB/Hz
TX_Jrms = 0.035 UI
RX_BW = 110 GHz



https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0227/johnson_e4ai_01a_250227.pdf

The penalty of chromatic dispersion with different device chirp



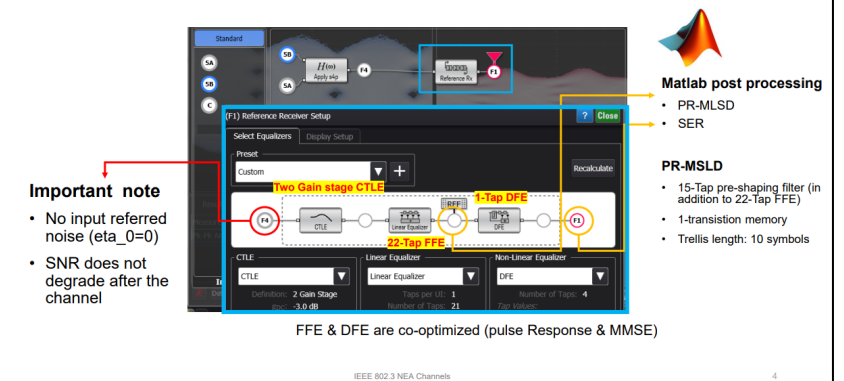
- The highlighted section indicates the dispersion penalty for a wavelength range of 1304.5-1317.5nm.
- Clearly, due to the interaction between device chirp and fiber dispersion, the dispersion penalty have benefits when $CD < 0$.
- For 113GBd PAM4, the dispersion penalty at chirp_max = 1 is 0.1dB@500m; 0.55dB@2km; 1.2dB@4km.
- For 226GBd PAM4, the dispersion penalty at chirp_max = 1 is 0.6dB@500m; 8dB@2km; NaN@4km. If the dispersion penalty within 2 km is to be less than 1 dB, the device chirp needs to be controlled below 0.4.
- the 400G/lane technology experiences a significant increase in dispersion penalty. Therefore, when designing transmission distances, it is necessary to consider the impact of device chirp.

https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0819/fan_e4ai_01_250819.pdf

IEEE 802.3 NEA E4AI, January 2025

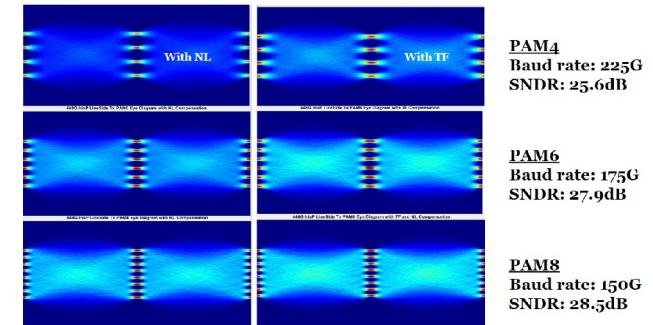
"TEF 2025 presentations will be publicly available in early February 2026. Images in this presentation are used with permission from the Ethernet Alliance and the respective authors."

425Gb/ Electrical – Equalizer Investigation



https://www.ieee802.org/3/ad_hoc/E4AI/public/25_0529/calvin_e4ai_01a_250529.pdf

448Gb/s in PAM4, PAM6 and PAM8



Improvement of SNR does not justify the increase in cardinality

Source: Naim Ben-Hamida, EA's TEF 2025, "Ethernet for AI: 448G technology for next gen networking"

Summary: Technical Feasibility

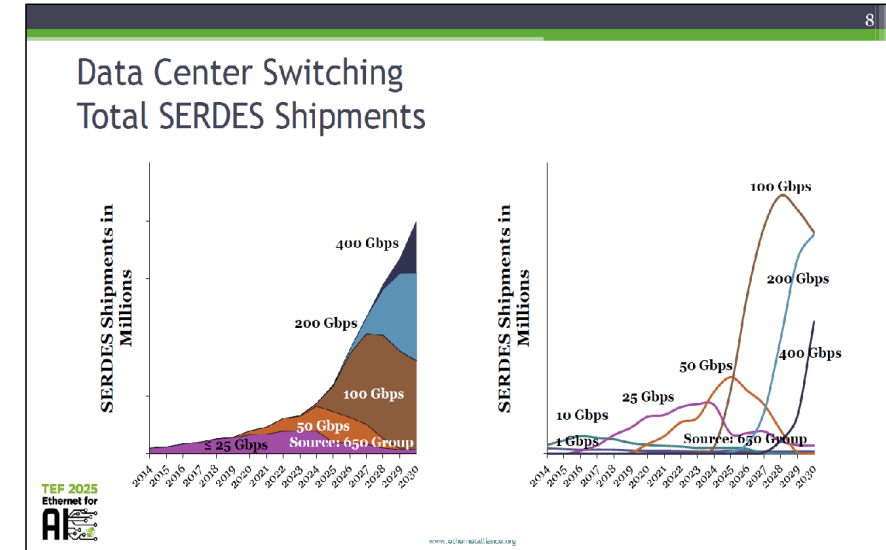
- Technical feasibility is no longer the question — industry is already investing across SERDES, channels, connectors, modulation, FEC, packaging, etc.
- IEEE 802.3 must define the interoperable standard
- The work is already underway across industry
 - IEEE 802.3 E4AI, EA TEF, SNIA, UEC, UALink, OCP ESUN, OIF ([OIF 448G Framework Document](#))

Why now?

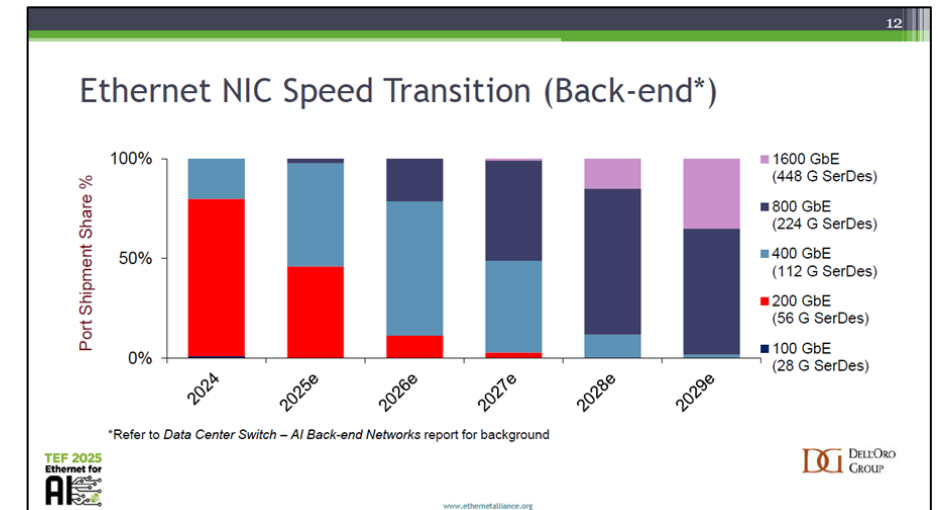
Kent Lusted

Growth & Bandwidth Needs

- AI scale-up networks need 400G/lane for xPU-XPU and flattest network
 - Scale-out and front-end networks will leverage this work
- XPU and Switch designs need higher speed interconnect to remain feasible
- Copper interfaces specs are needed to define SERDES capabilities and for low power
- Optics interface specs are needed to enable reach beyond copper at lowest power and cost



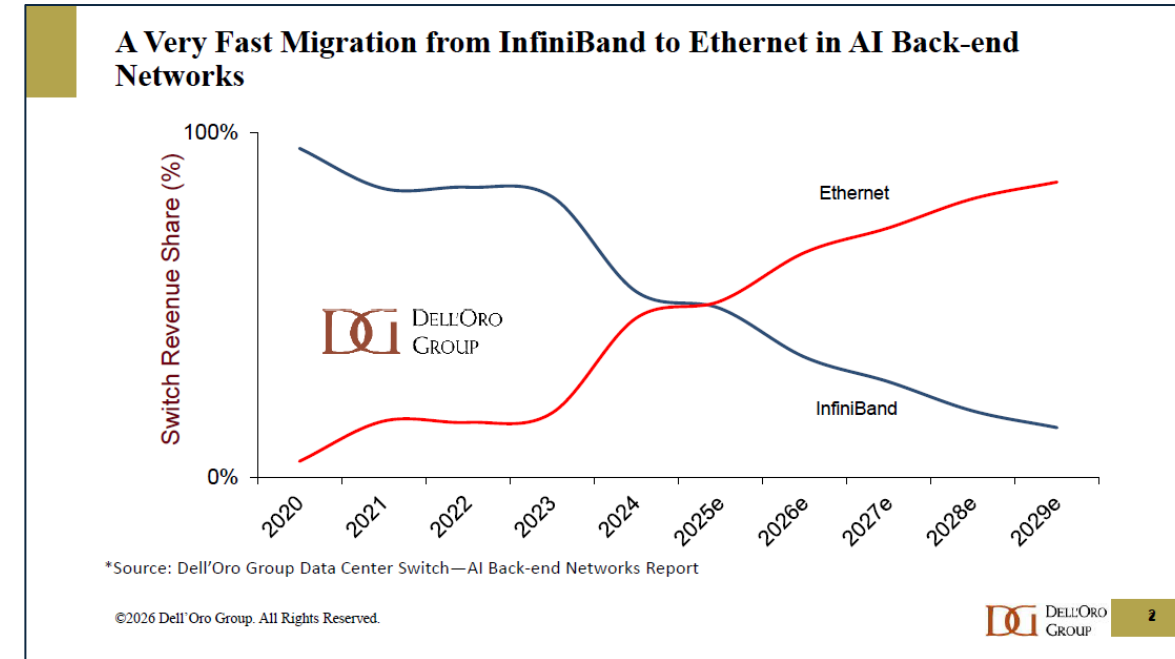
Source: Alan Weckel, EA's TEF 2025, "Setting the stage for networking in an AI world"



Source: Baron Fung, EA's TEF 2025, "XPUs and the Future of AI Connectivity"

Ethernet is the Choice

- Ethernet standards are the physical layer building blocks for AI networks
- Industry partners want to leverage Ethernet physical layer specs for broadly interoperable AI applications (scale-up and scale-out)



Source: Sameh Boujelbene, Dell'Oro Group

“Ethernet is winning the race in AI back-end networks and was able to eclipse InfiniBand in 2025.”, Sameh Boujelbene, Dell'Oro Group

Why Now?

There is no current IEEE 802.3 standard effort for 400 Gbps/lane, yet hardware investment is already underway

- 400 Gbps/lane signaling is essential to support next-generation AI network architectures, including high-radix system designs and practical chip-package integration.

The industry increasingly depends on common, shared technology foundations that span multiple applications, making standardized Ethernet PHYs even more valuable.

- Timely standardization is key to preventing fragmentation and ensuring long-term multi-vendor interoperability

Ethernet's widespread adoption provides immediate leverage of existing ecosystems, delivering significant economies of scale for new signaling technologies

Continuing to extend Ethernet's long-proven model of delivering scalable, interoperable physical-layer solutions

Summary

- AI infrastructure demands capacity growth
- Ethernet physical layer is the foundation for emerging networks (UALink, UEC, ESUN, etc)
- Scale-up is the fastest growing and largest segment, followed by scale-out and front end
- 400 Gbps/lane is essential to maintain practical switch and XPU designs
 - Evolution of known techniques plus advancing technology provide paths forward
- Meeting time-to-market needs requires a focused Ethernet PHY project now
 - Additional PHY developments can follow in future efforts
- Let's form a Study Group in March 2026!!

Supporters

- Your name here

Thank You!

Straw Polls

Should a study group be formed to develop a PAR, CSD responses, and objectives for “400 Gbps per lane Signaling PHYs and Interfaces”?

If formed, will you participate in this Study Group?

Unique affiliations for those who indicated they’d participate (post processed)