Update on Recent Developments in Industry

Lily Lyu (Huawei)
Jieyu Li (CMCC)

January Interim 2025

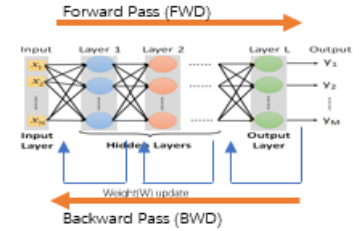
Motivation

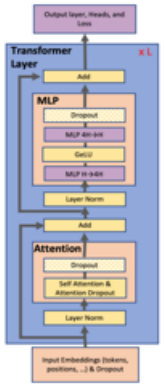
- Given the critical role of networks in AI cluster, the industry has seen rapid development.
- Industry alliances have been established to create an open ecosystem for the networks.
- This presentation aims to provide an update on the latest progress in industry.
  - The concepts of ‘scale-out’ and ‘scale-up’ are important.
  - Industry alliances have focused their efforts around ‘scale-out’ and ‘scale-up’.

Recap ‘Scale-Out’ and ‘Scale-Up’

AICN: Connecting Accelerators for AI Training

Neural Network

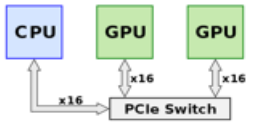




This is AI model, not the ‘network’ we are talking about. But it impacts the network development.

Scale-up

Bus technology evolves, trying to connect more GPUs in bus domain.



https://en.wikichip.org/wiki/nvidia/nvlink

Fully Connected NVLink across 256 GPUs

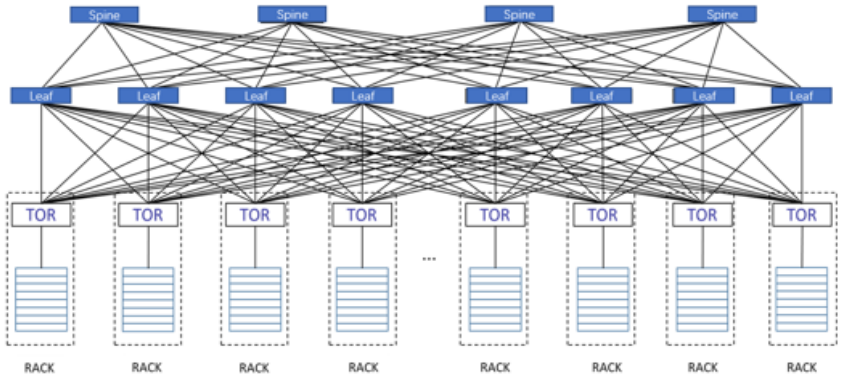
https://developer.nvidia.com/blog/announcing-nvidia-dgx-gh200-first-100-terabyte-gpu-memory-system/

- PCIe, NVLink, CXL …
  - Ultra high bandwidth: e.g NVLink5.0 is a 1.8TB/s bidirectional, direct GPU-to-GPU interconnect
- Server-scale -> rack-scale ->pod-scale
  - 10 -> 1000 GPUs

Scale-out

Network technology evolves, trying to improve performance (reliability, latency, throughput)

- Infiniband, Ethernet (RoCEv2) , …
  - High bandwidth: 800GE->1.6TGE
- Pod-scale -> across DC scale
  - Towards 10K+ GPUs



Source: https://mentor.ieee.org/802.1/dcn/24/1-24-0055-00-ICne-clarification-of-the-content-in-aicn.pdf

Commercial Product Example

Scale-out network

Nvidia

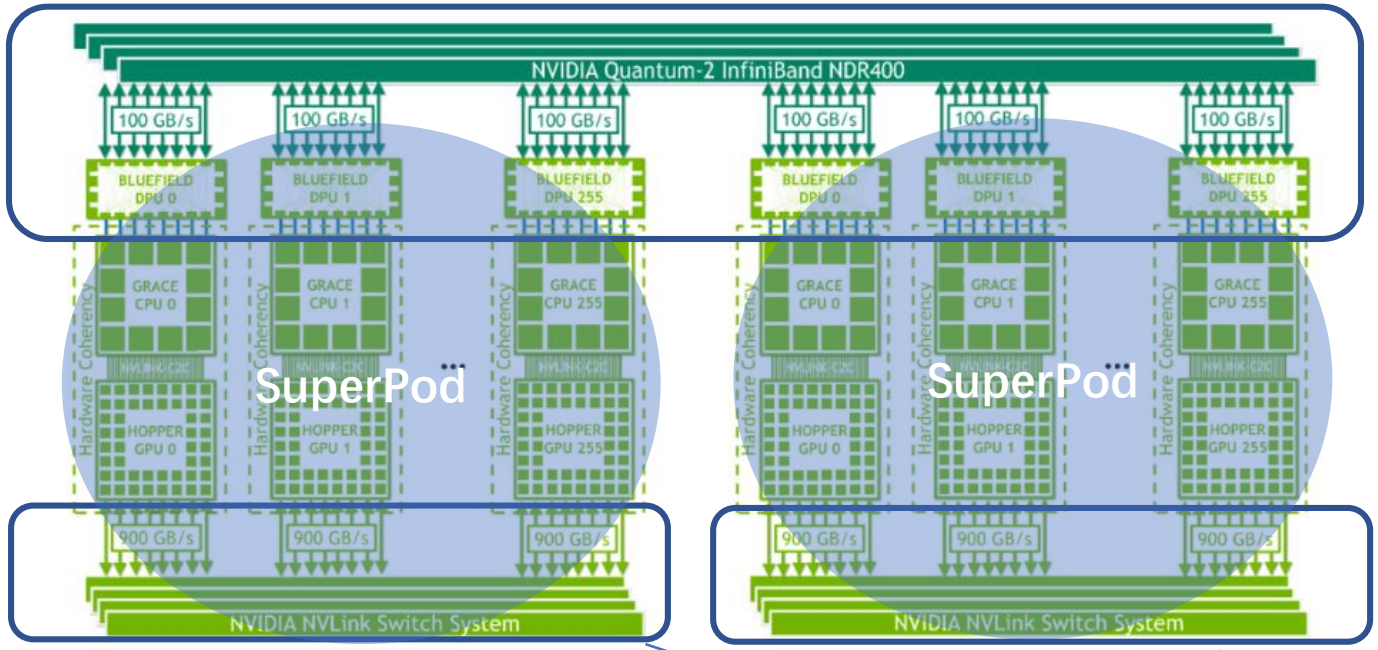


Figure 14. NVIDIA HGX Grace Hopper with NVLink Switch System for strong-scaling giant ML and HPC workloads

Source: https://resources.nvidia.com/en-us-grace-cpu/nvidia-grace-hopper-2

Meta

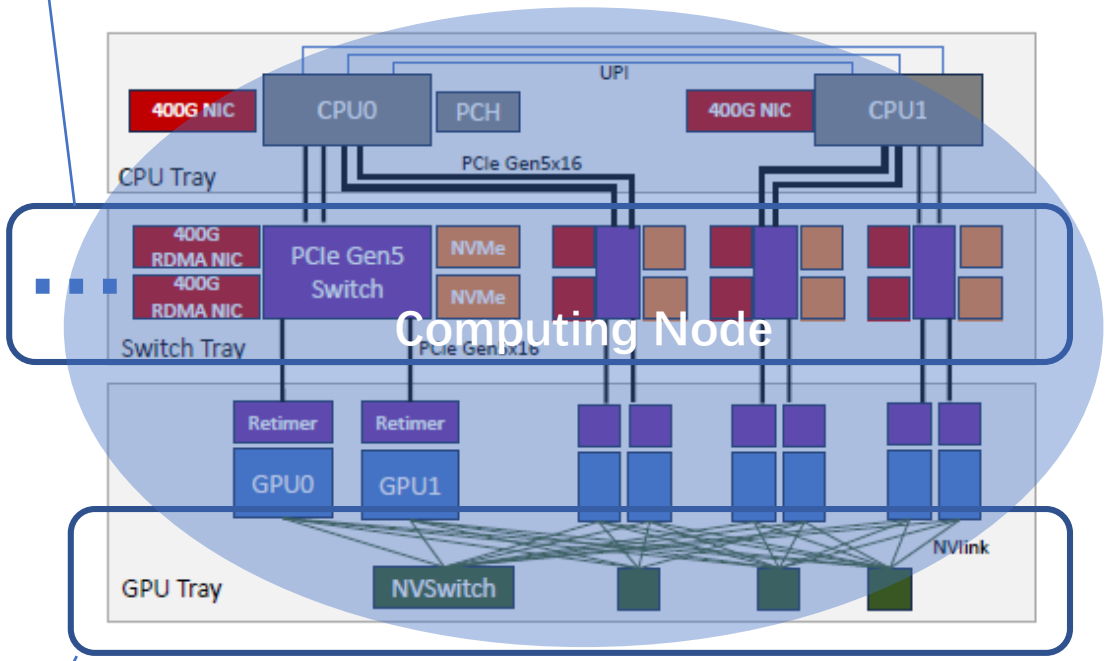


Figure 4: Grand Teton platform

Source: Sigcomm2024: RDMA over Ethernet for Distributed AI Training at Meta Scale

Scale-up network

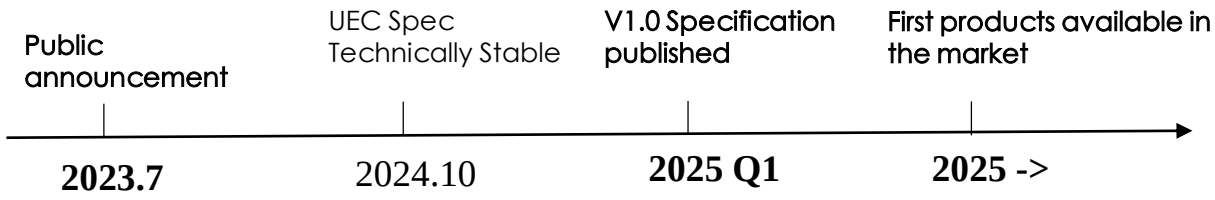
Scale-out Network Technologies Converging on Ethernet Stack

Ultra Ethernet Consortium(UEC) is formed by industry giants

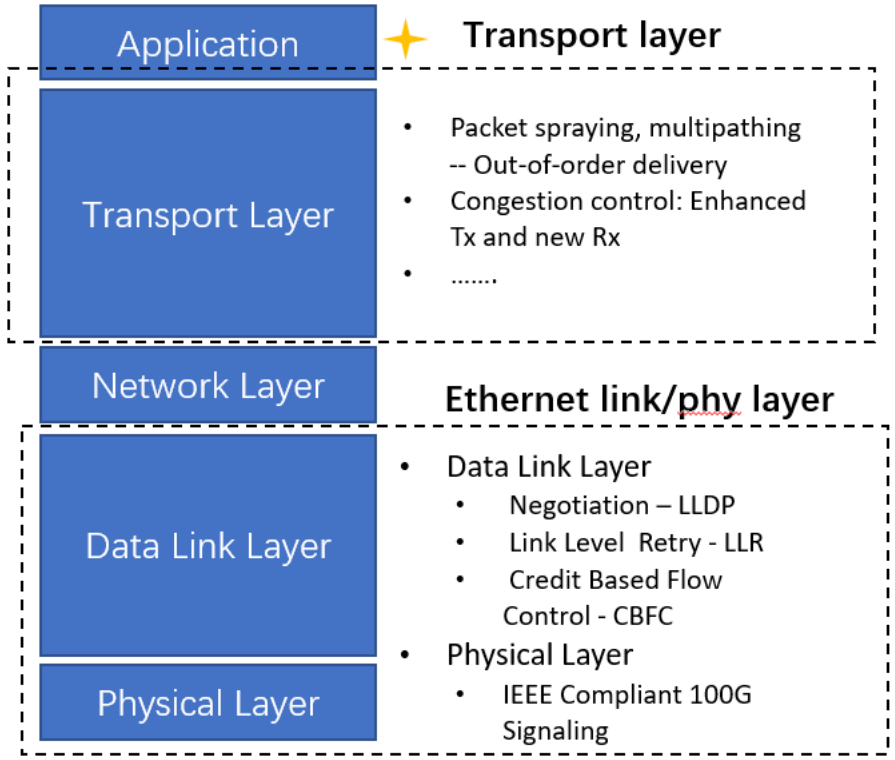
Background:

- Steering member include AMD, Broadcom, Arista, Cisco, Eviden, HPE, Intel, Meta, Microsoft and Oracle.
- Aim to build an Ethernet-based, open, interoperable, high performance, full-communications stack architecture to meet the growing network demands of AI & HPC at scale
- Spec 1.0 targets the scale-out network
- Website: https://ultraethernet.org/

Timelines:



Key technologies:



Source: 2024 OCP Global summit

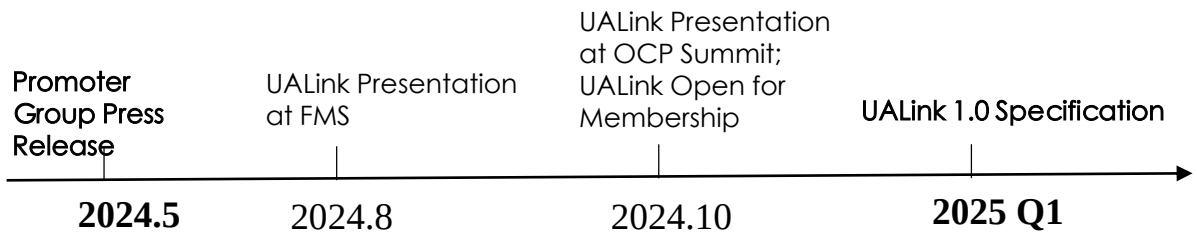
Scale-up Network Technologies Under Discussion

Ultra Accelerator Link Consortium (UALink) Leveraging Ethernet PHY

Background:

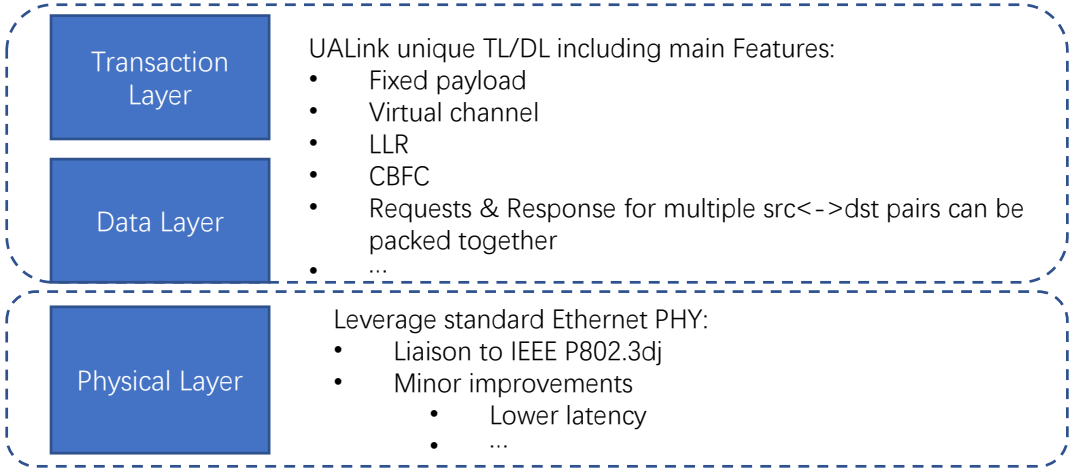
- Promoter members include AMD, AWS, AsteraLabs, Cisco, Google, HPE, Intel, Meta, Microsoft, Synopsis, Alibaba, Apple.
- Aim to develop interconnect technical specifications that facilitate direct load, store, and atomic operations between AI Accelerators (<1K endpoints), supporting AI/ML scale-up networks and workloads
- Website: https://www.ualinkconsortium.org/

Timelines:



Source: 2024 OCP Global summit

Key technologies:



Source: SC24

Scale-Out + Scale-Up = Interconnection Between Accelerators

The Scale-Out and Scale-Up networks constitute Accelerator-to-Accelerator Interconnection, but there is no clear boundary between the two.

UEC point of view：

- Scale-Out Network
  - Scale: Cluster -10k nodes and ↗
  - Distance: <100m; RTT <10 uS + ; BW ~100GB/S
  - Network semantics (DMA and packetized I/O)
- Scale-Up Network
  - Scale: Within a node; small scale e.g., 256 XPU?
  - Distance: ~1m ; RTT ~1 uS +; BW ~1200 GB/S↗
  - Memory and Network semantics

UALink point of view：

- Memory shared across Accelerators
  - Memory semantic (Direct load, store, and atomic operations between accelerators)
- The ability to make accelerators in a single Rack Pod or several Rack Pods act like a one giant accelerator to complete the task
- Complementary with scale out approaches such as UEC

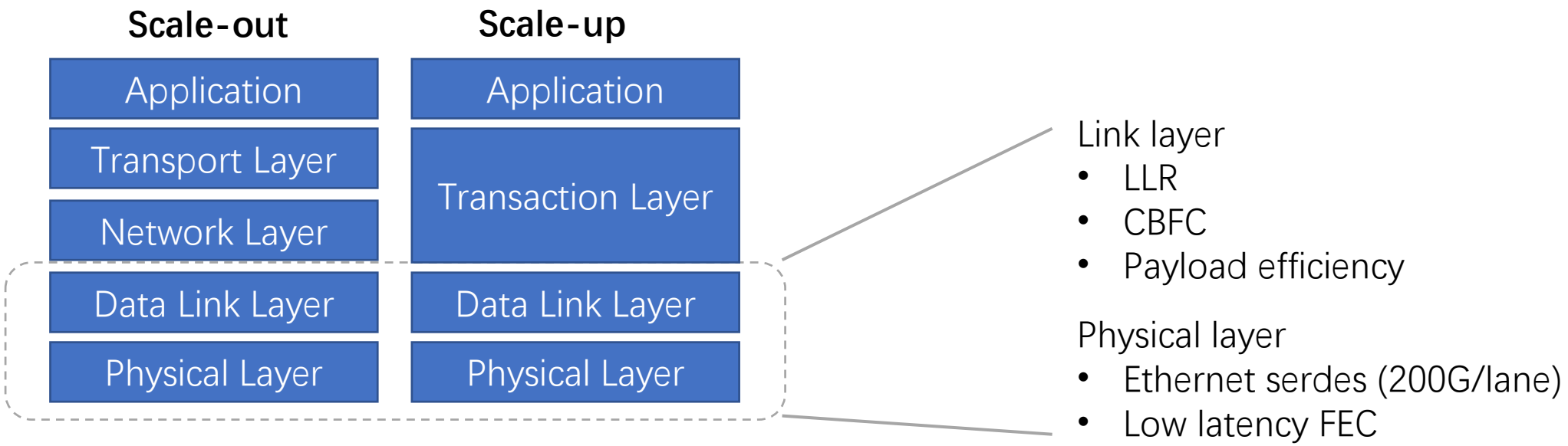
Source: 2024 OCP Global summit

My point of view：

- Supporting memory semantic (e.g. load/store operation) is the key characteristic of scale-up.
  - Both scale-up and scale-out require low latency/high bandwidth network. Scale-up network has more demanding requirements in order to support memory semantic.
- Defining boundary between scale-out and scale-up networks is a comprehensive issue， impacted by factors such as network technology, accelerator capability, model architecture etc.
  - Open discussion in industry

Scale-Out More Focus on Network LB & CC; Both Scale-Out and Scale-Up Enhance PHY/LL with Similar Strategies

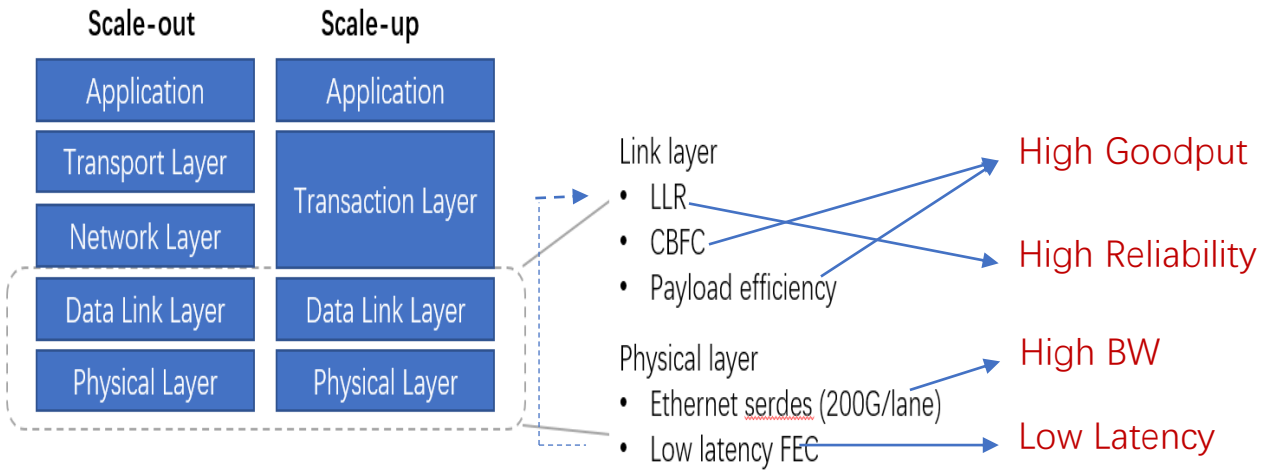
- Scale-out: load balancing →packet spray
- PHY/LL: similar strategies in the technical approaches to enhance the performance of underlying links, such as LLR,CBFC etc.



Discussion: IEEE 802's Role in AI Network Development

IEEE802 covers Ethernet physical layer and link layer.

- ✓ IEEE802.1 used to standardize DCB technologies. PFC is widely used and QCN has enlightened other congestion control mechanisms (e.g. DCQCN).
  - CBFC was discussed in ieee802.1
  - LLR was in ieee802
  - LB was analyzed in NENDICA report
- ✓ IEEE802.3 is working on Beyond 400G project which is recognized as high bandwidth opportunity for AI networks.
  - There was a contribution about low latency PHY in NEA



IEEE802 could start/re-start investigations into the PHY/link technologies toward performance matrix required for AI network.

Thanks!