Introduction for the Linux community

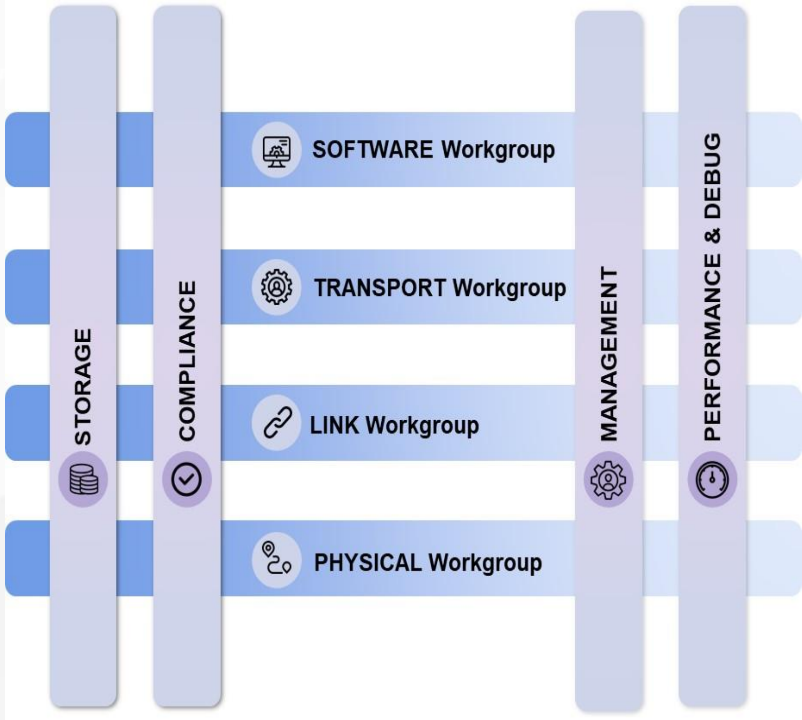
Uri Elzur, UEC Technical Advisory Committee Chair

Shrijeet Mukherjee, Enfabrica

July 15th, 2024

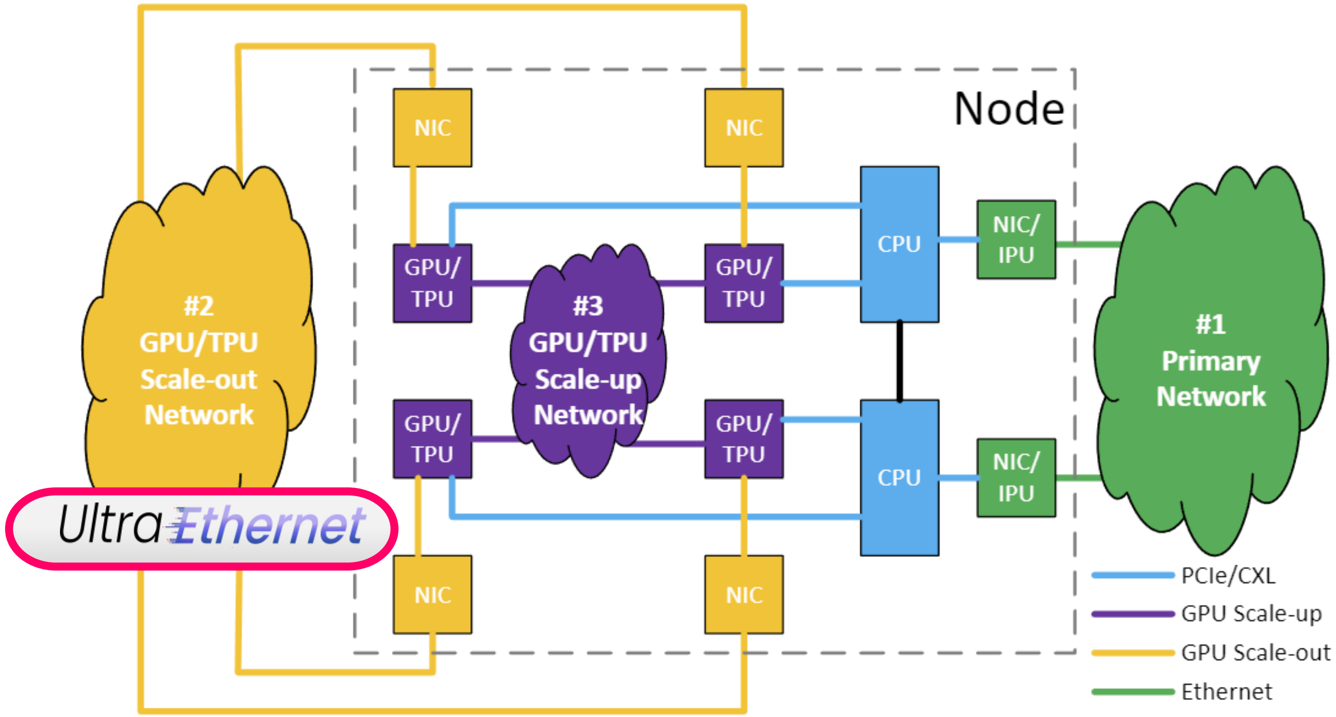
Ultra Ethernet Consortium

An Ethernet-based, open, interoperable, high performance, full-communications stack architecture to meet the growing network demands of AI & HPC at scale



>80 member companies
>800 active members

AI/HPC Networks of Interest: Basic Characteristics



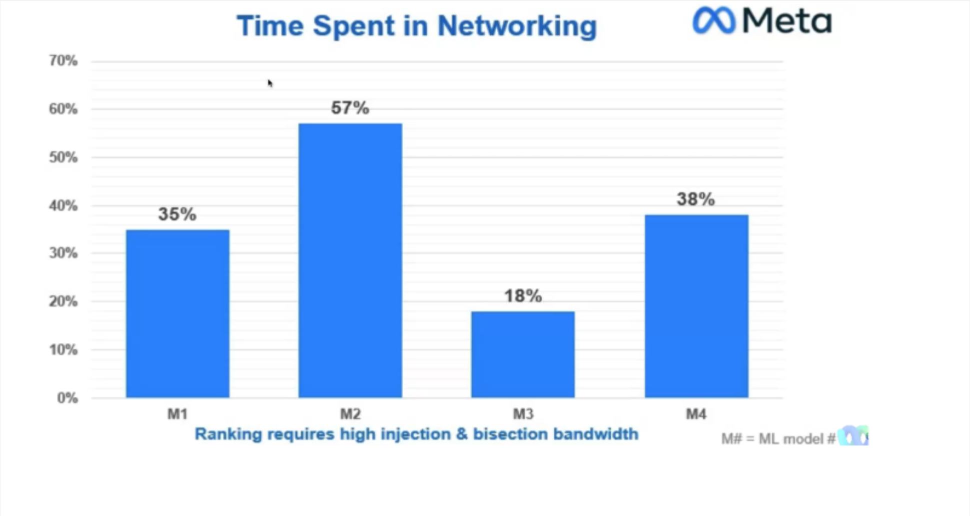
85% utilization

35-40% utilization

1. **Primary DC network**
   - Used by all 3 deployment models
   - Main network for some HPC At Scale
   - Very large scale: up to 100K-1M Endpoints
   - Distance: >150m ; RTT ~100 uS +; BW/GPU **~10GB/S**
   - Storage attached e.g., over RoCE RDMA
   - Network semantics
2. **GPU/TPU Scale-Out Network**
   - DL/Inference Cluster -10k nodes and ↗
   - Distance: <100m ; RTT <10 uS + ; BW **~100GB/S**
   - Main network for some HPC At Scale
   - Network semantics
3. **GPU/TPU Scale-Up Network**
   - Within a node; small scale e.g., 256 XPU?
   - Distance: ~1m ; RTT ~1 uS +; BW **~1200 GB/S↗**
   - Direct connect and/or switched
   - Memory and Network semantics

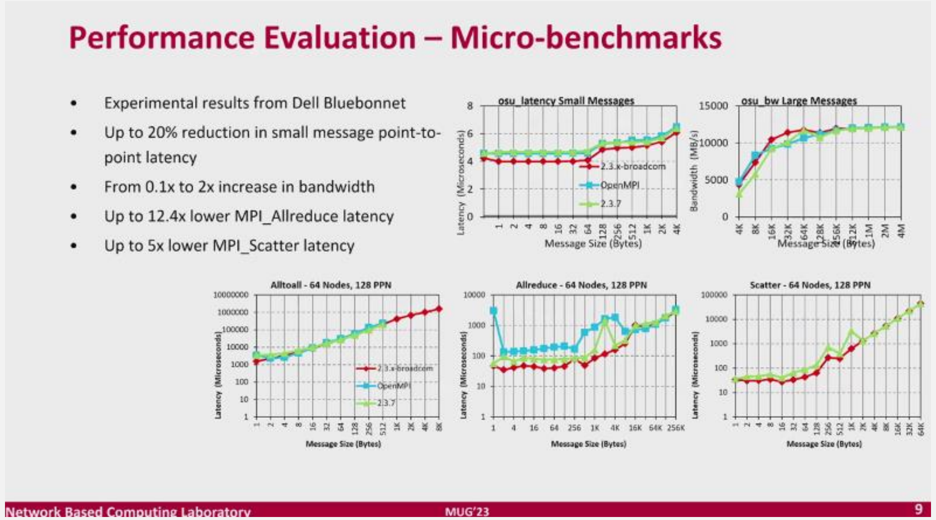
The Network Directly Influences Workload's Performance !

AI *GPU Scale-out*



- Framework coordinated – systolic
- High Bandwidth
- Large messages
- In Network Compute – 2x potential

HPC *CPU Scale-out*

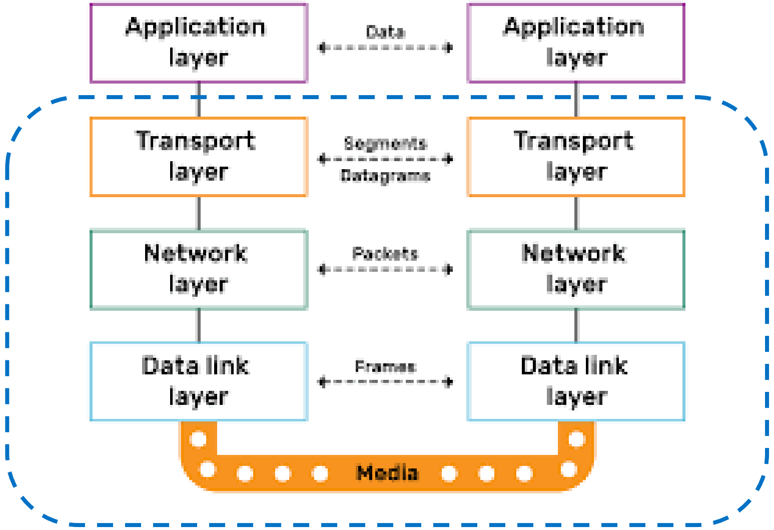


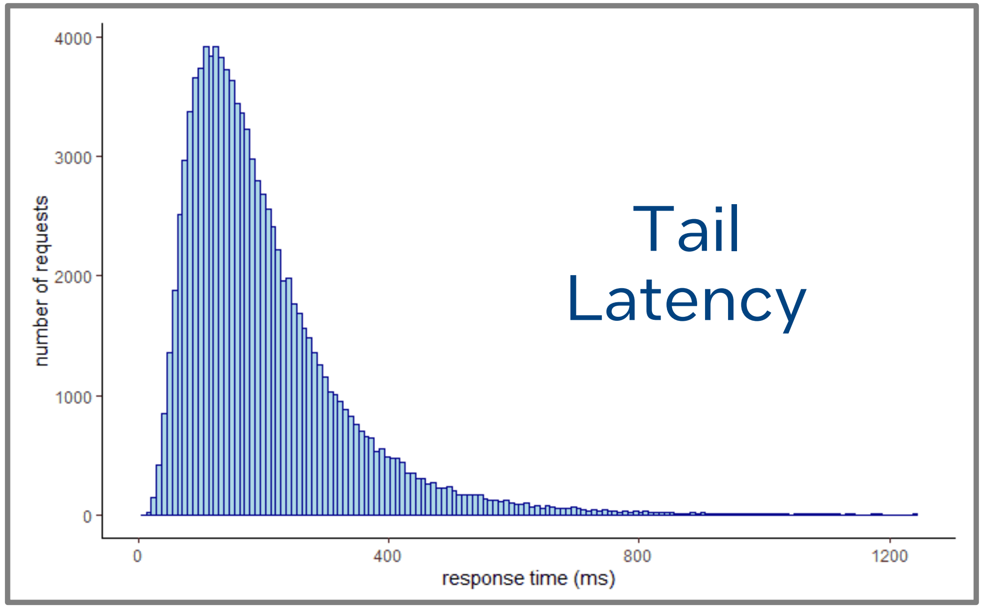
- MPI
- Small messages – Latency sensitive
- High packet rate
- Existing application support - required

https://youtu.be/miv5PExXTmc?t=782

https://mvapich.cse.ohio-state.edu/static/media/talks/slide/kawthar-slingshot-osu-booth-sc22_2.pdf

AI/HPC Common Requirements





- Transport primitives for
  - Large Scale
  - Multi pathing
  - Relaxed ordering
  - Modernized Congestion Control
  - Optimized RDMA
  - Performance – bandwidth, latency, tail latency, Packets/S
  - High network utilization
  - Stability and Reliability

Key goals: high utilization & low tail latency!

TODAY's RDMA TRANSPORT ARE DEFICIENT

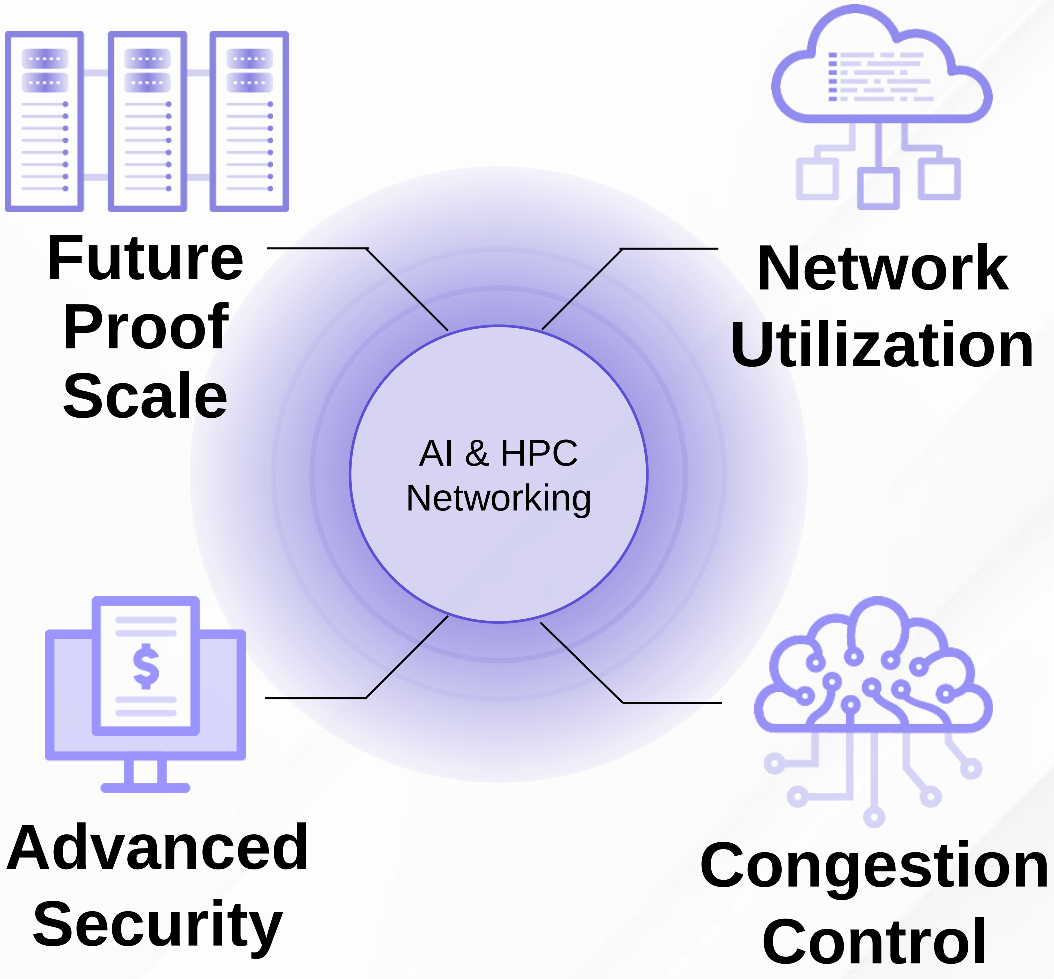
- **Lack of multipathing** makes load balancing difficult and solutions brittle
- Requires **in-order packet delivery**
- **Go-back-n:** massively inefficient for dropped packets
  necessitates a "lossless" network
- DCQCN congestion control is **brittle and hard to tune**
  - Specific to workload and network details

➔ **Challenges with scale**

It's time to modernize RDMA

UEC TRANSPORT ADDRESSES GRAND CHALLENGES

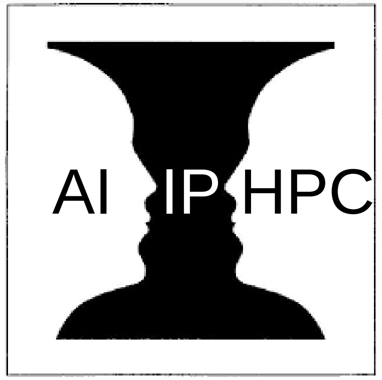
- Future proof system scale with up to 1M endpoints
- Improved network utilization with multi-path routing
- Lower tail latency with flexible packet ordering
- Faster congestion control response times
- Modernized & optimized RDMA operations and APIs
- Security built-in from the beginning
- End-To-End telemetry provides improved network visibility



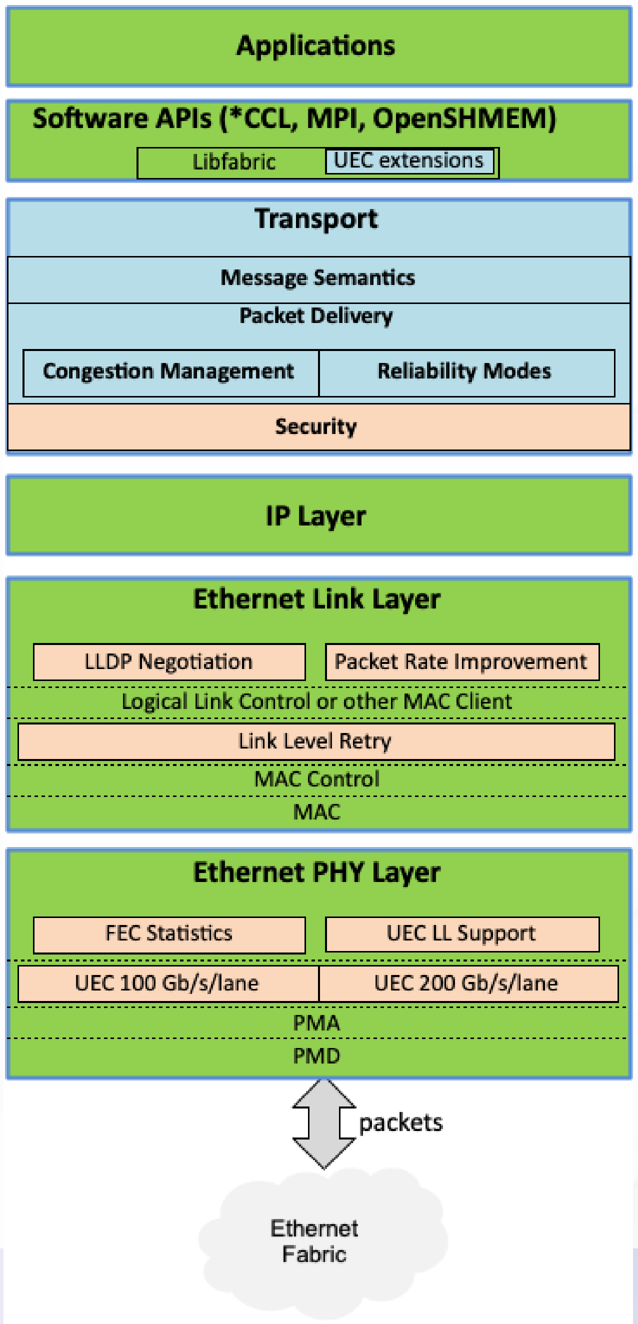
Highest infrastructure utilization
*at **ultra-high scale***

UEC Stack Overview
(partial feature list)

- Software API
  - Libfabrics 2.0 with extensions
- New Transport Layer
  - Multi-pathing
  - Packet spraying
  - Ordered (ROD) and un-ordered (RUD)
  - Lossy (no PFC) or Lossless
  - Congestion Control: Enhanced Tx and new Rx
  - Trimming
  - In Network Collective
- Network Layer
  - IP v4/v6
  - ECN
- Data Link Layer
  - Negotiation – LLDP
  - Link Level Retry - LLR
  - Header Efficiency Improvements
- Physical Layer
  - IEEE Compliant 100G Signaling
- AI and HPC Profiles



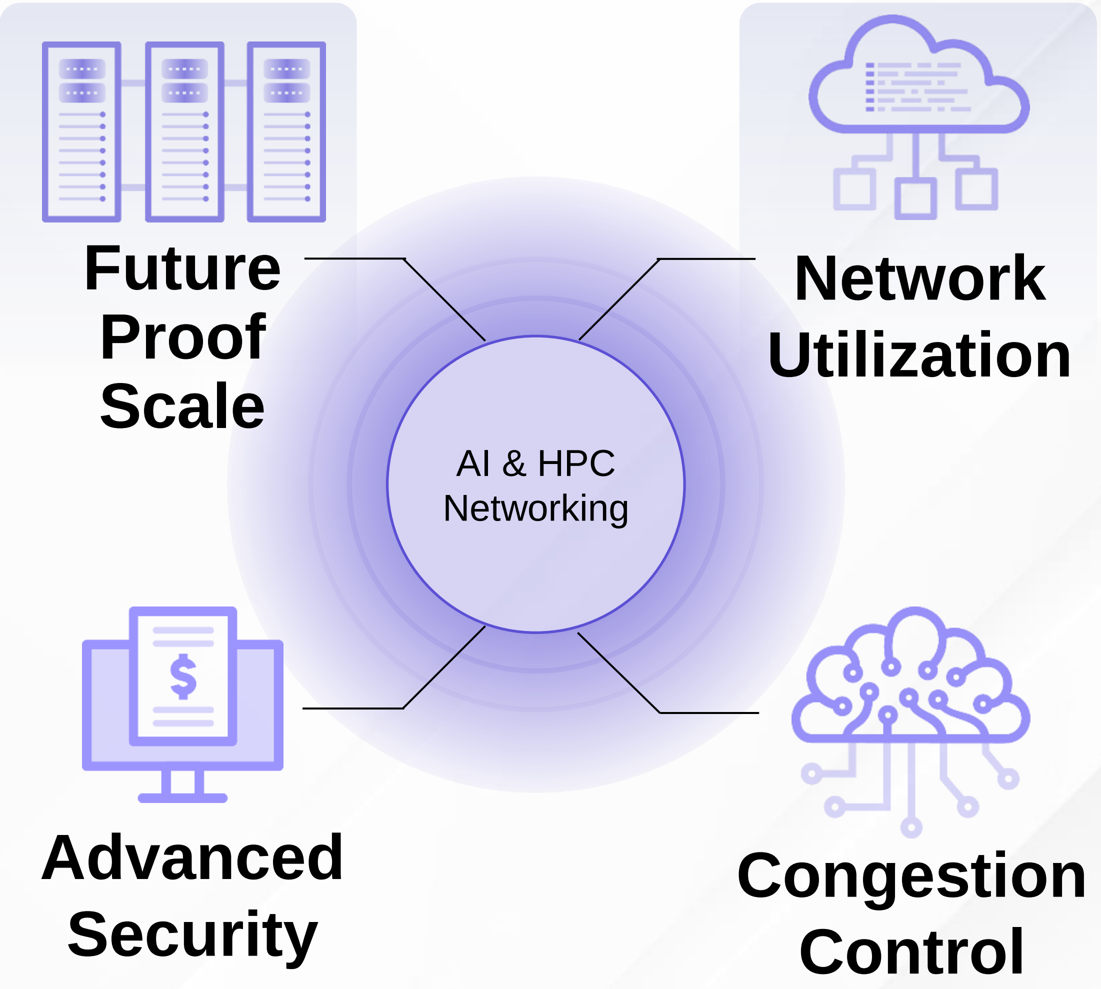
UEC Stack



UltraEthernet

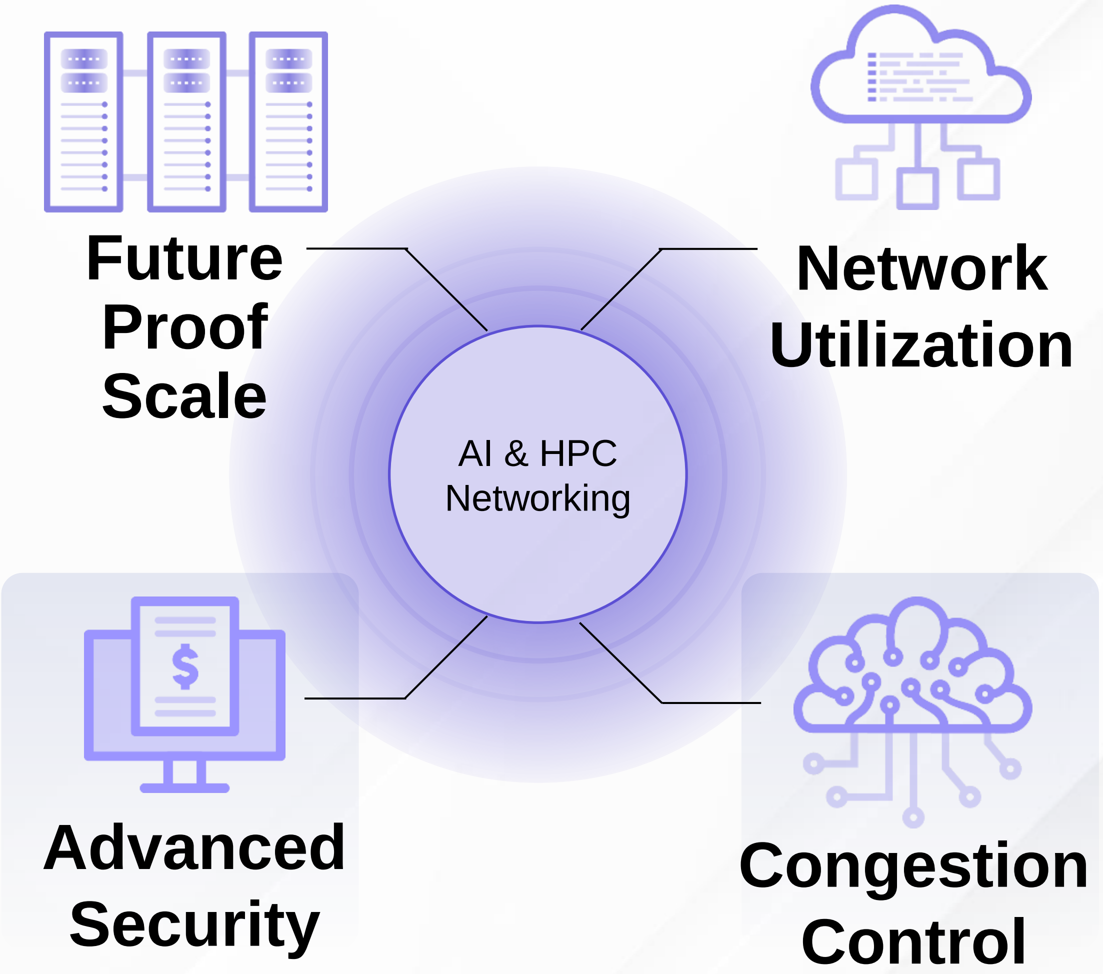
UEC ADDRESSES SYSTEM SCALE & NETWORK UTILIZATION

- Determinism and predictability become more difficult as systems grow
  - SCALE: Network Stability, Fairness, re-convergence times, deadlock avoidance are part of the design
- “Packet spraying” - every “flow” may simultaneously use all paths to the destination, vs. flow using a single path
  - Network Util: Achieves more balanced use of the entire network
- From Rigid to Flexible Ordering – suites AI workloads
  - API: Supports modernized RDMA operations and APIs and relaxing packet ordering, per workload guidance!
  - TR: Flexible ordering enables packet-spraying in bandwidth-intensive large collective operations; without reorder buffers
  - TR: Rigid packet and message ordering uses "go-back-n" for loss recovery, but restricts network utilization and increases tail latencies
  - HW: Minimize state and complexities of Initiator and target
  - JCT: Critical to curtail tail latencies



ADVANCED SECURITY, CONGESTION CONTROL & TELEMETRY

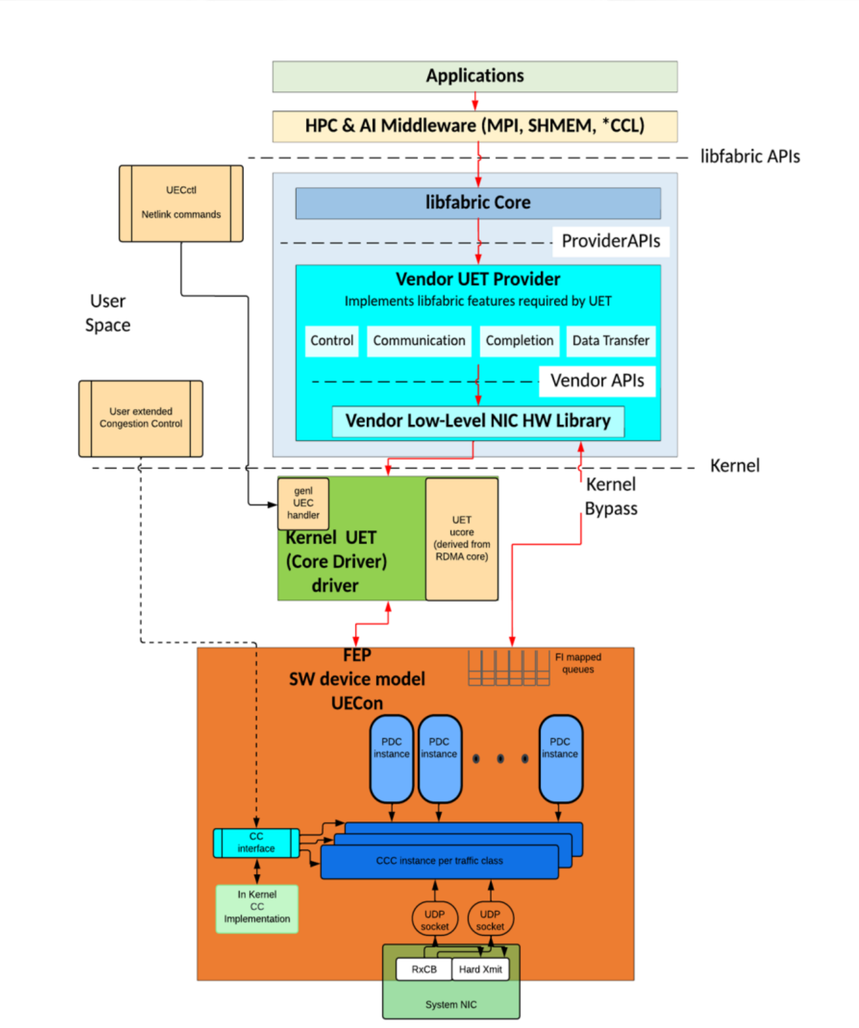
- Congestion
  - Support packet spraying
  - Address incast (e.g., as a result of All-to-All)
  - Optimized response time while maintaining high utilization
  - 1st RTT
- In Network Collective
  - Leverage switch offload to gain higher effective bandwidth, lower latency
  - Applicable to some highly common collective e.g. AllReduce
- Advanced Security
  - Encryption support that doesn't balloon the session state in hosts and network interfaces
  - Similar conditions in AI and HPC



Modern Transport and RDMA Services Needs for AI and HPC

Requirement	UEC Transport	Legacy RDMA	UEC Advantage
Multi-Pathing	Packet spraying	Flow-level multi-pathing	Higher network utilization
Flexible Ordering	Out-of-order packet delivery with in-order message delivery	N/A	Matches application requirements, lower tail latency
AI and HPC Congestion Control	Workload-optimized, configuration free, lower latency, programmable	DCQCN: configuration required, brittle, signaling requires additional round trip	Incast reduction, faster response, future-proofing
In Network Collective	Built-In	NONE	Faster Collective operation, lower latency
Simplified RDMA	Streamlined API, native workload interaction, minimal endpoint state	Based on IBTA Verbs	App-level performance, lower cost implementation
Security	Scalable, 1st class citizen	Not addressed, external to spec	High scale, modern security
Large Scale with Stability and Reliability	Targeting 1M endpoints	Typically, a few thousand simultaneous end points	Current and future-proof scale

A Strawman for Kernel Support Model



Basic interface similar to IB/RDMA devices

- netlink for device initialization/creation
- OOB communication for connection establishment
- Kernel bypass using direct mapped HW queues
  - State in a Packet Delivery Context
  - Congestion rate managed by a Congestion Control Context
- Core libraries similar to RDMA core
  - possibly derived/refactored
- SW device model along with initial specification

UEC TARGET TIMELINES

July 2023

Public announcement

Q4 2023

New members start joining

Q4 2024

V1.0 Specification published

2025 ⇨

First products available in the market

Linux

It's Time to Join! Time to Interact !

intel®