Ultra Ethernet Consortium

IEEE-CNSV

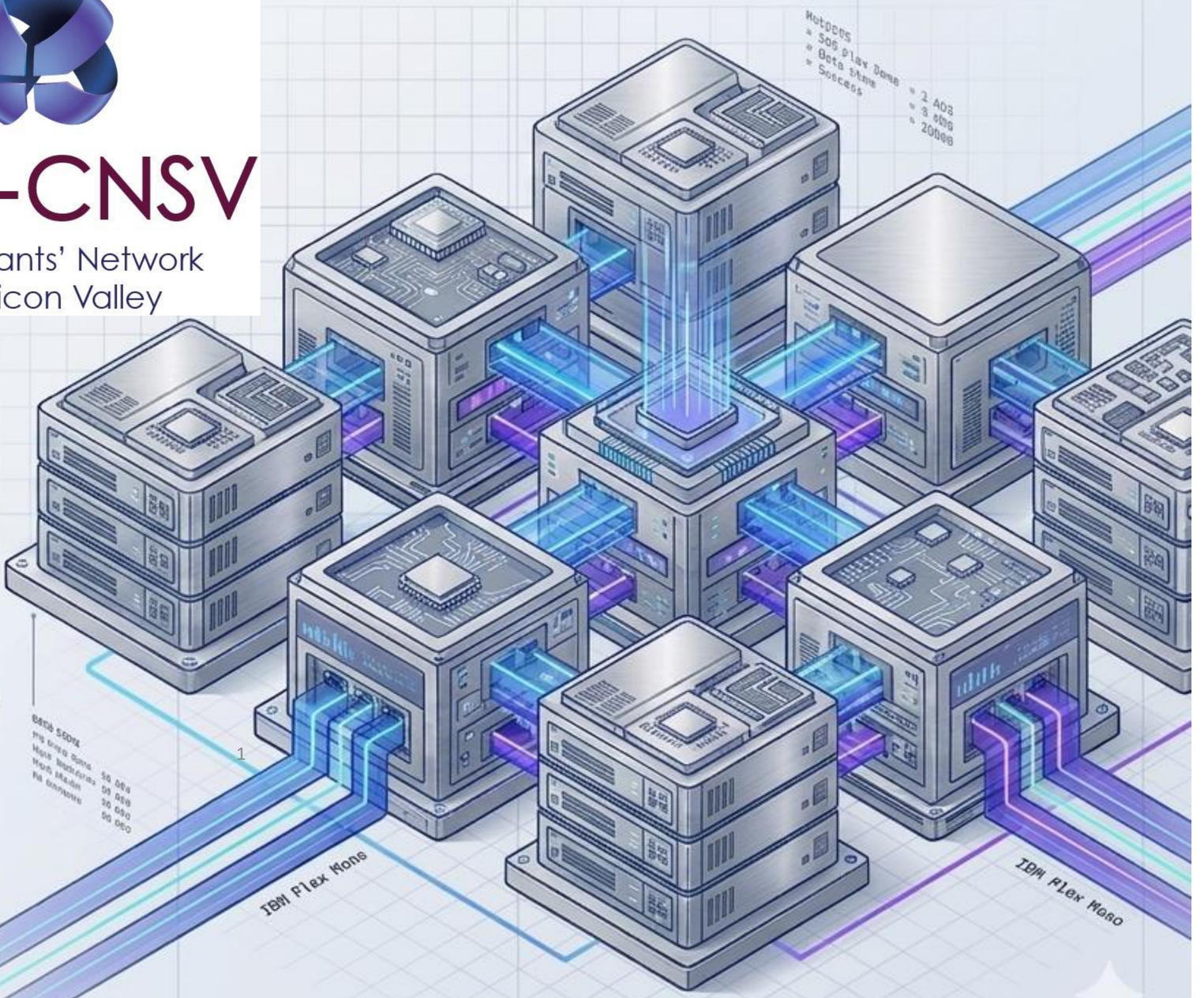
Consultants' Network of Silicon Valley

A **Primer** on **UltraEthernet**

Architecting the open standard for the future of AI and HPC networking.

Rip Sohan

Tuesday 8th September 2026



NotebookLM

The Efficiency Gap Today

Why Modern AI Outpaces the Network

The Scaling Disparity

GPU performance has scaled **1,000×**, while per-node network bandwidth has scaled only **10×** in the same period.

The Tail Latency Tax: slowest flow dominates

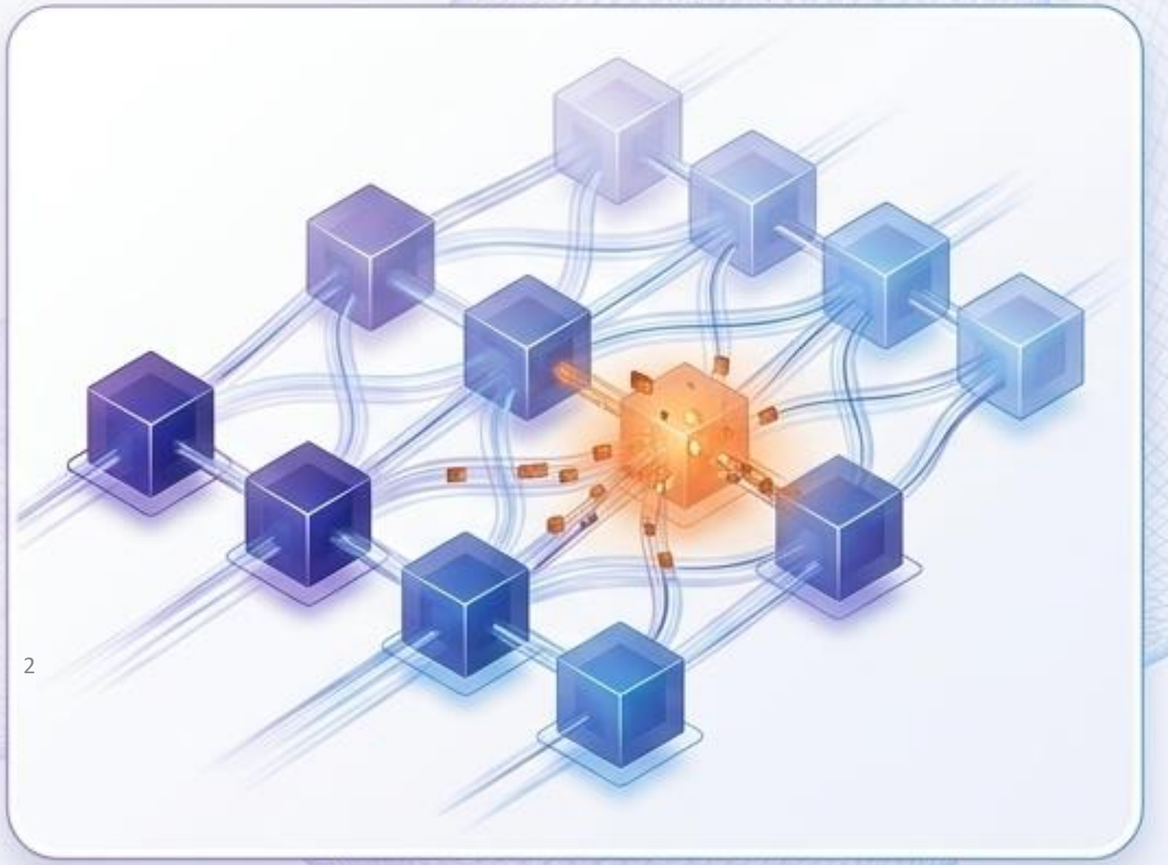
Synchronized operations force the entire cluster to idle while waiting for the **single slowest packet** in the transport "tail."

The Incast Phenomenon

Synchronized many-to-one bursts overwhelm standard switch buffers, triggering **catastrophic packet drops**.

The Penalty of Loss

A mere **0.1% packet loss** results in a **50% drop** in effective throughput due to legacy retransmission timeouts.



2

ULTRA ETHERNET CONSORTIUM

The Efficiency Gap Tomorrow

Scaling Challenges for Next-Generation AI Factories

- Cluster Scale : Million-node AI Factories
- Collective Intensity : TP/CP parallelism bottlenecks
- Inference Shift : Sustained global tokens/second
- Resilience Churn : MTBF < training iterations

3

Vanilla networking assumptions no longer apply to AI and HPC

	Vanilla Data Center	AI & HPC Workloads
Traffic Profile	Many independent flows, asynchronous.	**Few massive flows, highly synchronized.**
Bandwidth Requirements	Low bandwidth per individual flow.	**Endpoints are incredibly fast; load is exceptionally high.**
Latency Profile	Variable Round-Trip Times (RTTs).	**Ultra-short RTTs are mandatory.**
Completion Metric	Independent flow completion.	**Strict tail-latency dependency (completion time is determined exclusively by the slowest flow).**

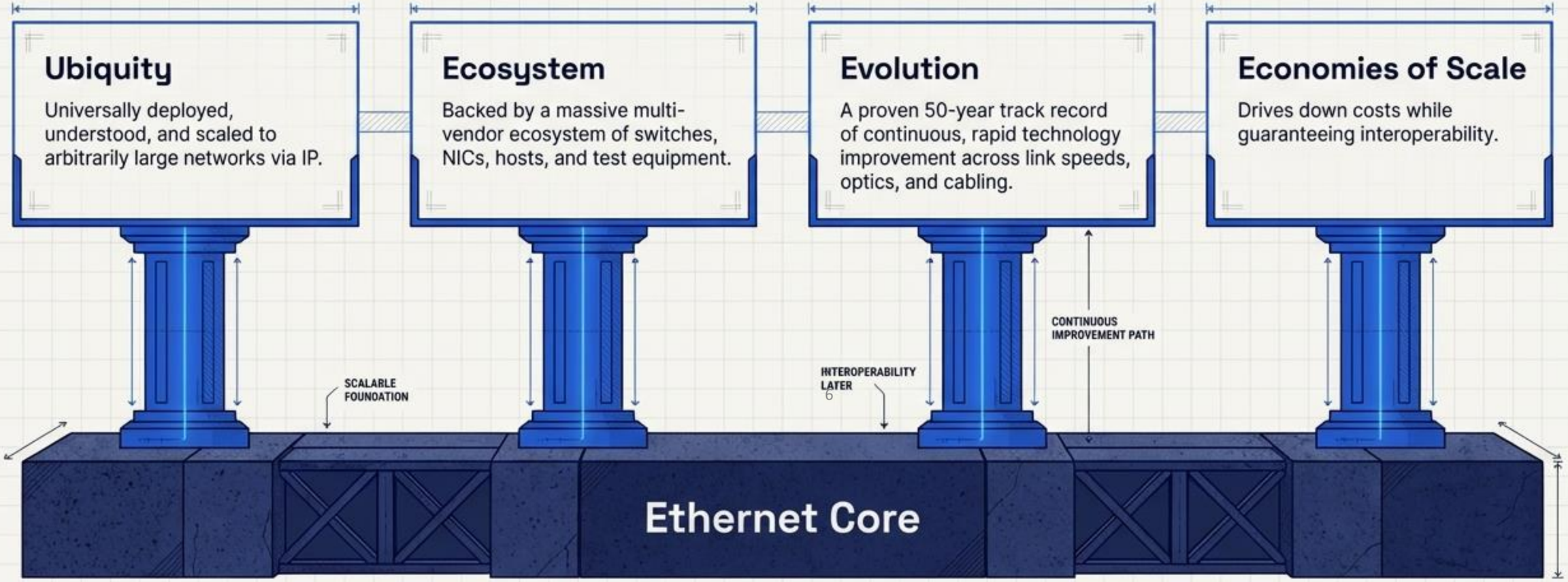
NotebookLM

Galvanizing the Industry

>100 Member Companies | >1,300 Active Participants

Why build on Ethernet?

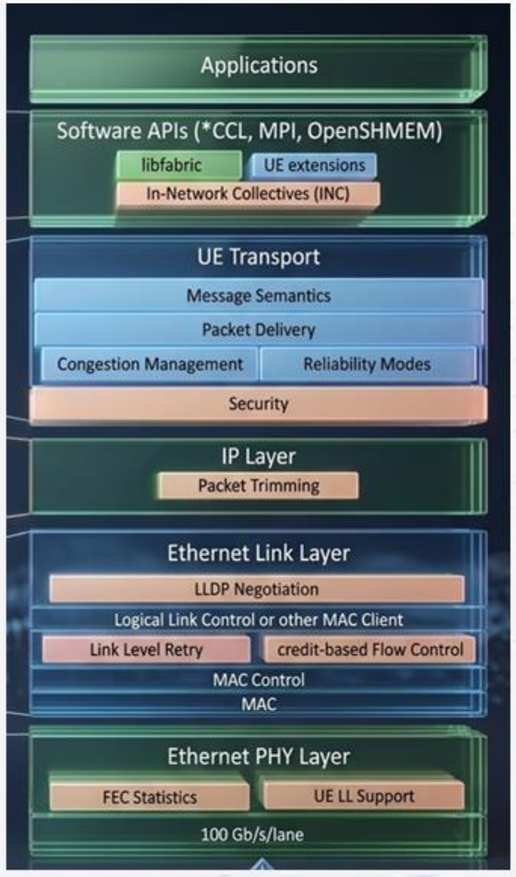
We are not changing the core; we are elevating it.

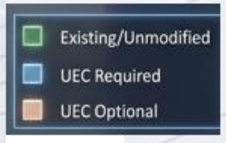


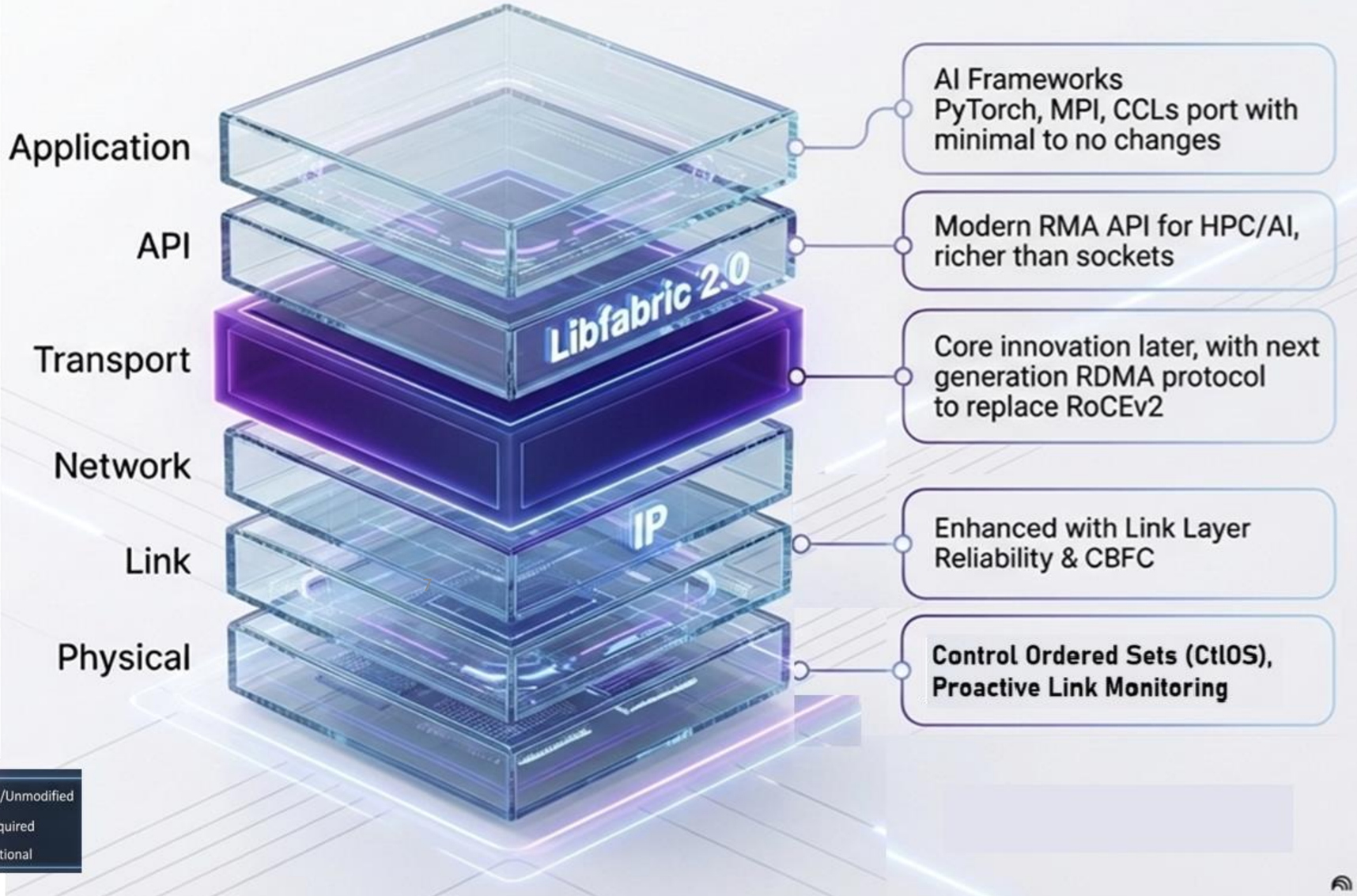
NotebookLM

Ultra Ethernet: Up and Down the Stack

Delivering data straight from the network into application memory.

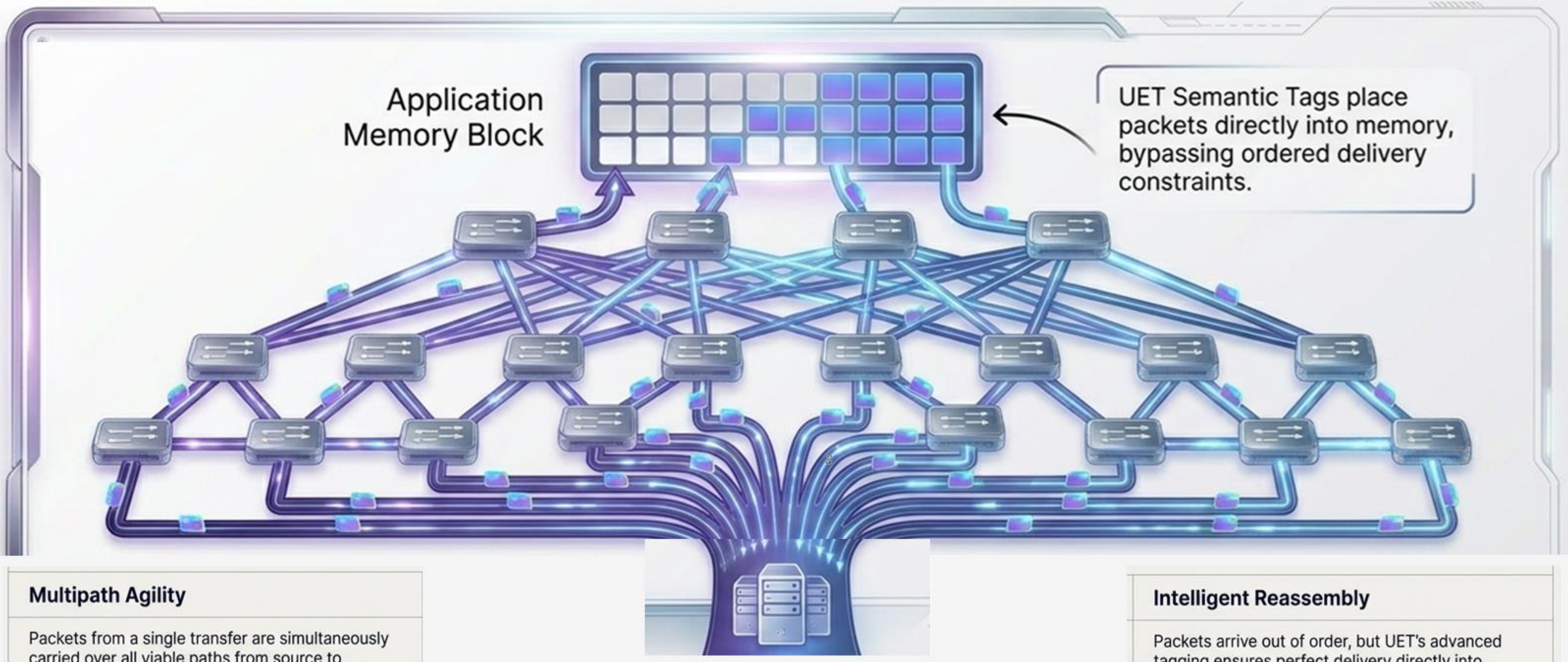






Packet Spray As a First Class Primitive

Eradicating hotspots via multi-path, out-of-order delivery.



Multipath Agility

Packets from a single transfer are simultaneously carried over all viable paths from source to destination.

Intelligent Reassembly

Packets arrive out of order, but UET's advanced tagging ensures perfect delivery directly into application memory, regardless of arrival sequence.

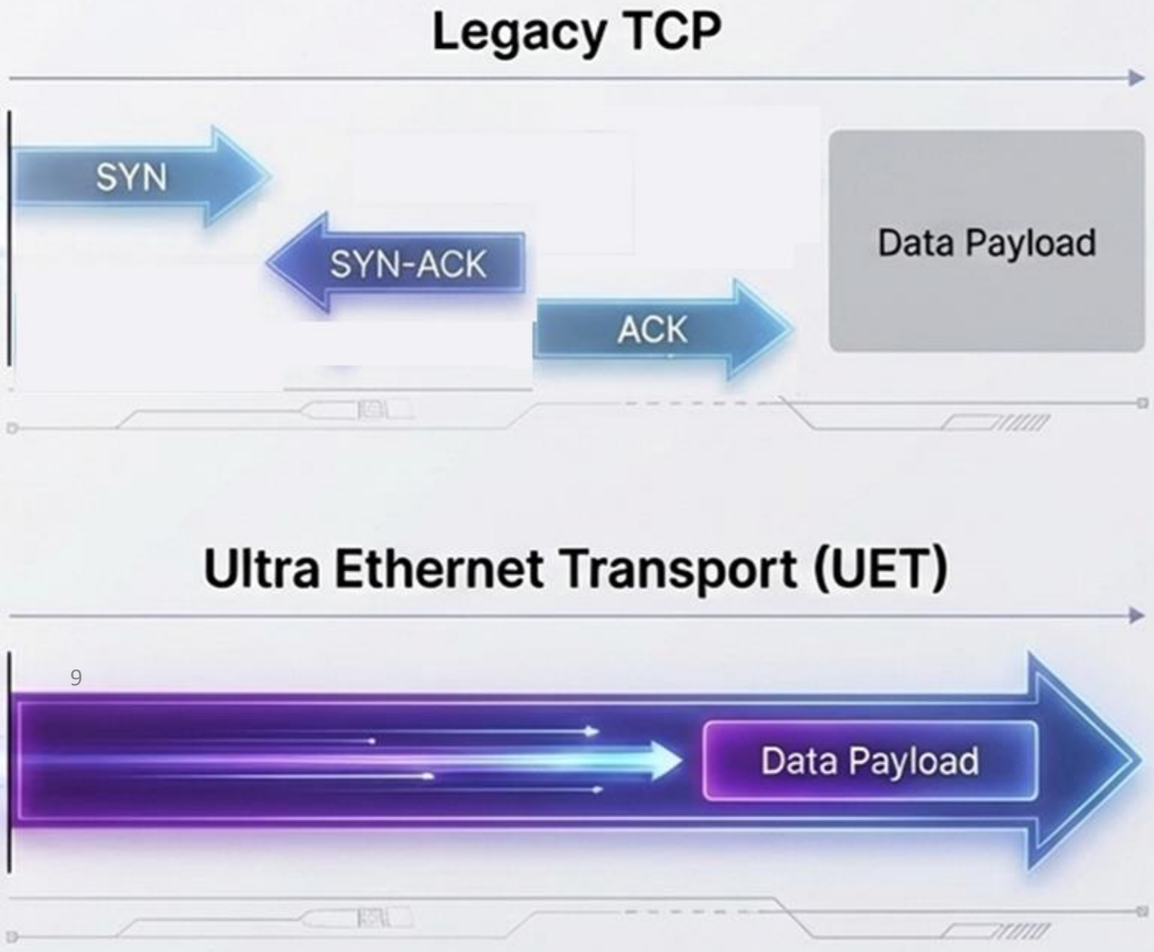
Dynamic Connection Support

Optimizing for extreme scalability and short-lived, high-velocity transfers.

Zero-Restart Penalty:
Rapid connection setup allows data transmission before the handshake even completes.

Ephemeral State:
Idle connections are aggressively torn down to minimize memory cost.

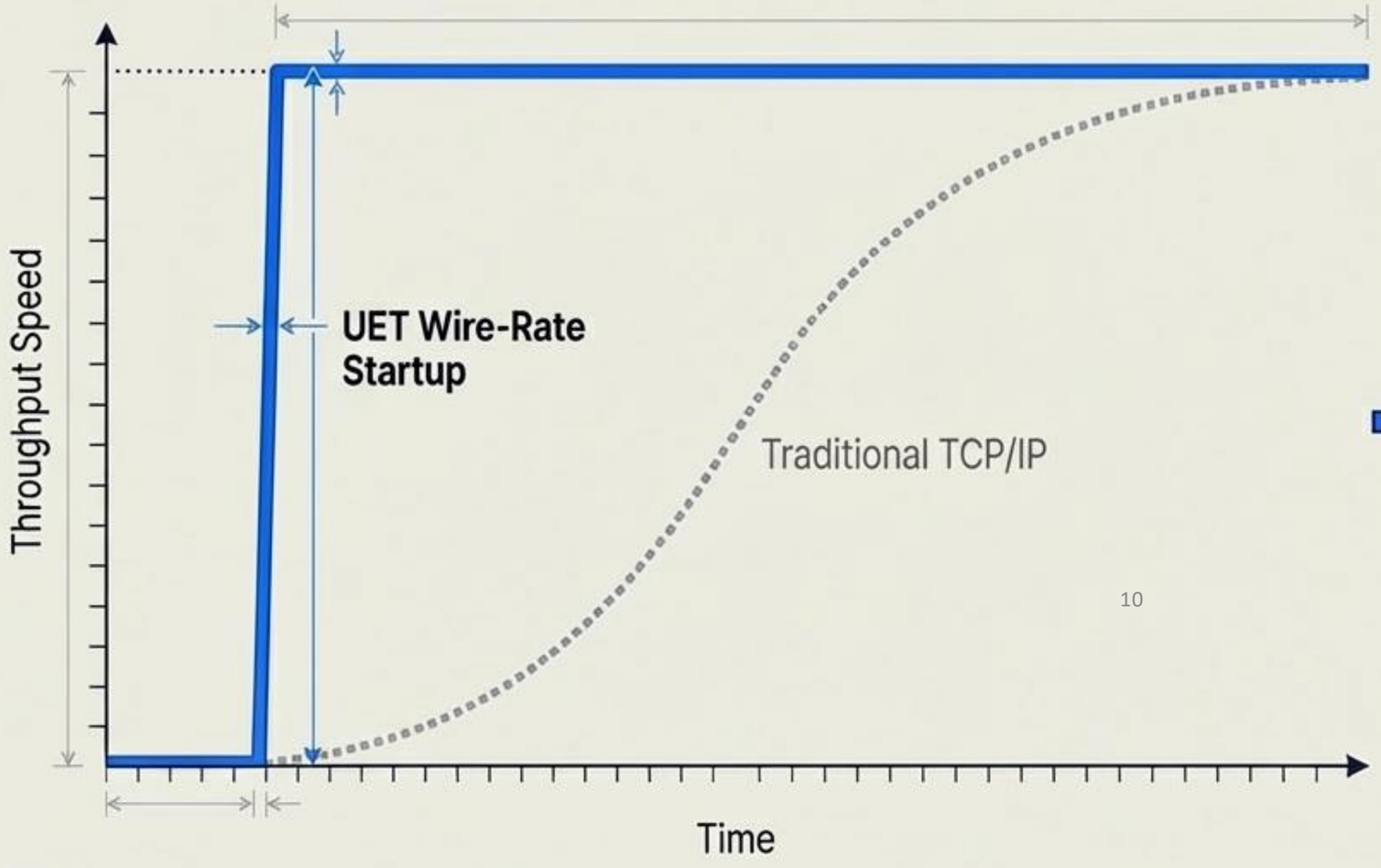
Massive Scale:
Avoids the fatal overhead of traditional TCP-like handshakes at AI cluster scale.



NotebookLM

Wire-rate startup: Eliminating the handshake penalty

For AI workloads, an entire massive data transfer might only last a few round-trip times. We cannot wait to ramp up.



Instant Throughput: UET allows data to be transmitted before a traditional handshake completes, starting off at full wire-rate.

Optimized for Bursts: Dramatically improves performance for the short, highly synchronized bursts typical of GPU collectives.

Zero Restart Penalty: Minimizes state cost by allowing idle connections to be torn down and instantly spun back up without latency penalties.

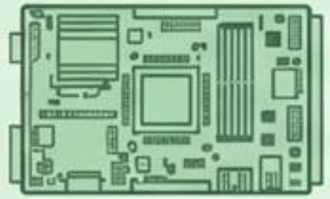
NotebookLM

Sender and Receiver Congestion Control

A flexible, multi-path, wire-rate architecture purpose-built to eliminate tail latency and incast in AI/HPC workloads.

Zone 1: Initiator (Host A)

SENDER-BASED CC

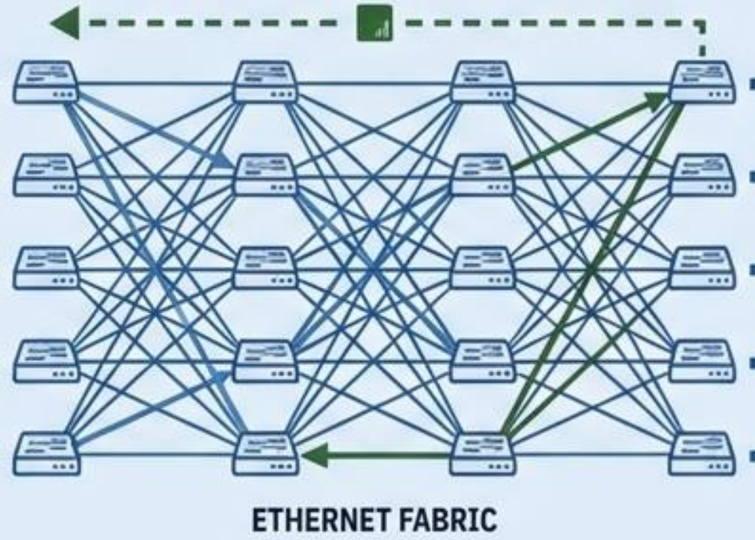


HOST A

- Fast ramp-up, wire-rate start.
- Telemetry-driven (ECN, delay).
- Optimal for high-bandwidth.

Zone 2: Ethernet Fabric (The Network)

IN-NETWORK TELEMETRY & SPRAYING

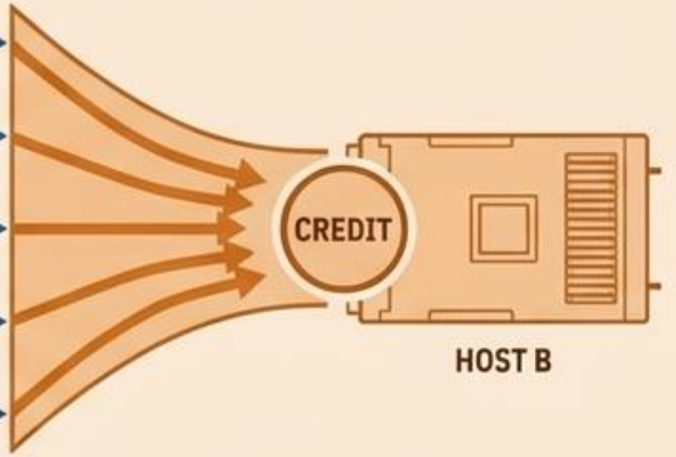


ETHERNET FABRIC

- Packet spraying across paths.
- Eliminates link polarization.
- Optional packet trimming.

Zone 3: Target (Host B)

RECEIVER-CREDIT BASED



HOST B

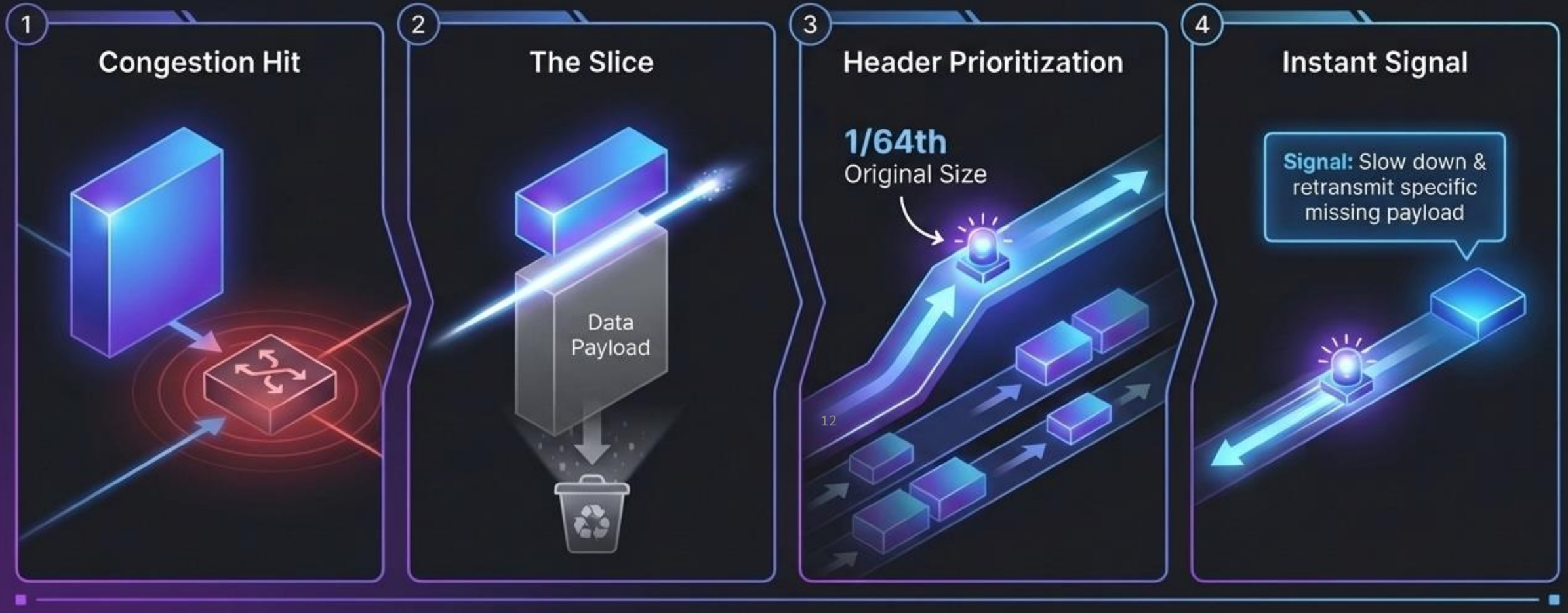
- Receiver manages capacity with credits. Mitigates incast.
- Ideal for oversubscribed networks.

The UEC Advantage: Replaces proprietary algorithms with an open standard. Thrives on best-effort Ethernet, no PFC required.

NotebookLM

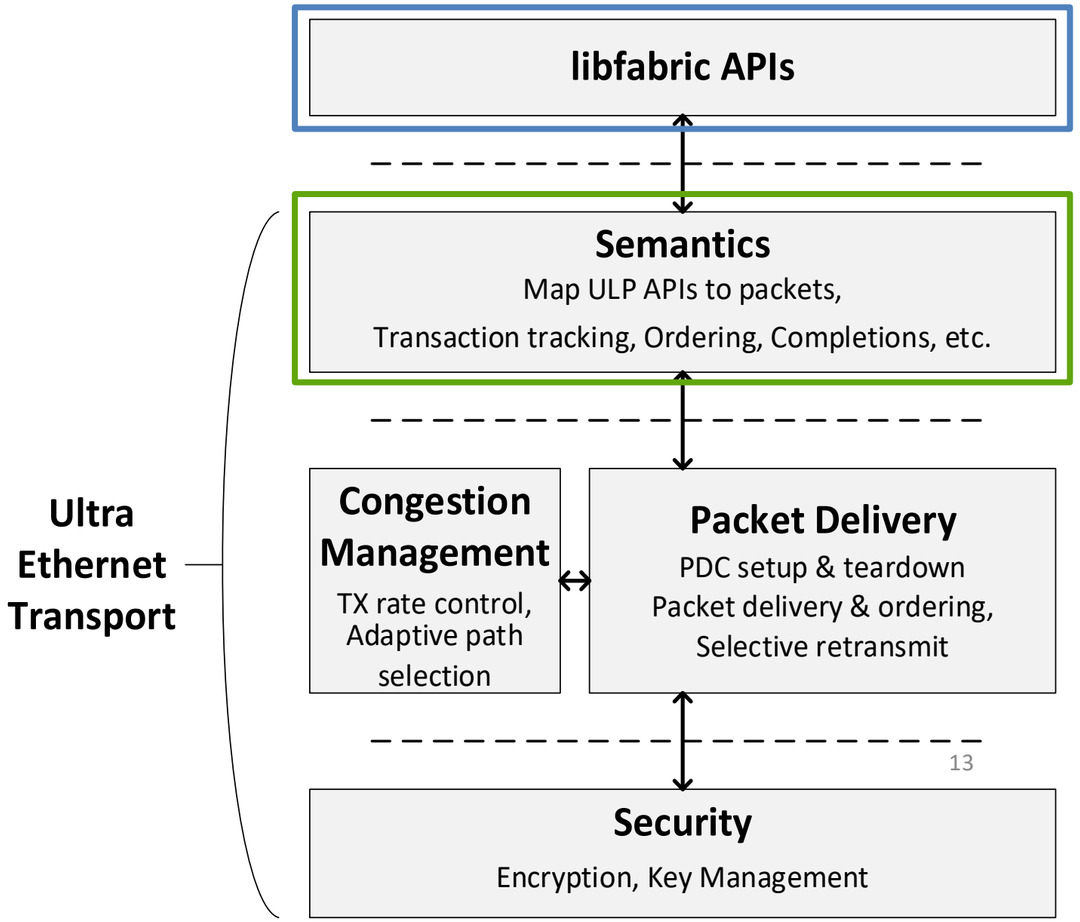
Packet Trimming

Fast, precise congestion signaling without black holes



NotebookLM

SEMANTIC LAYER: EXACT MATCH TAGS, DEFERRED SEND



- RDMA services: Send, Write, Read, Atomics
  - Focus on MPI and *CCL
- Scalable addressing
- Optimized extensions:
  - Deferred Send improves RNR
  - Exact match tags for HW offload of ordering between endpoints using shared receive queues

High scale, high utilization, low tail latency

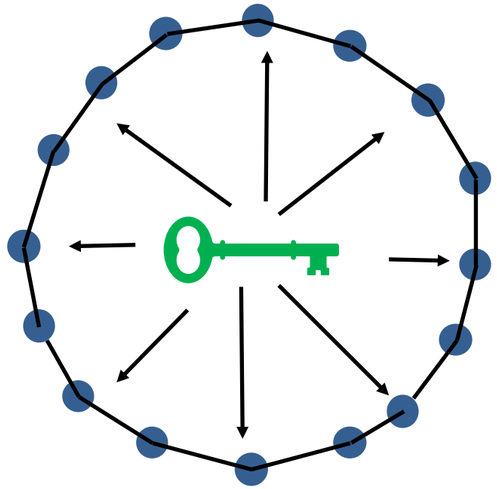
UET Security

Integrated Security in Ultra Ethernet Transport

Builds on core principles from IPSec and PSP

- AES-GCM, KDFs, IVs, Key Rotation, Anti-Replay
- Designed for high scale and group and client-server communication
- Includes a model for host-level security and authorization

- **Security for group applications: Security Domains**
  - Group Keying
  - Jobs exist in Security Domains



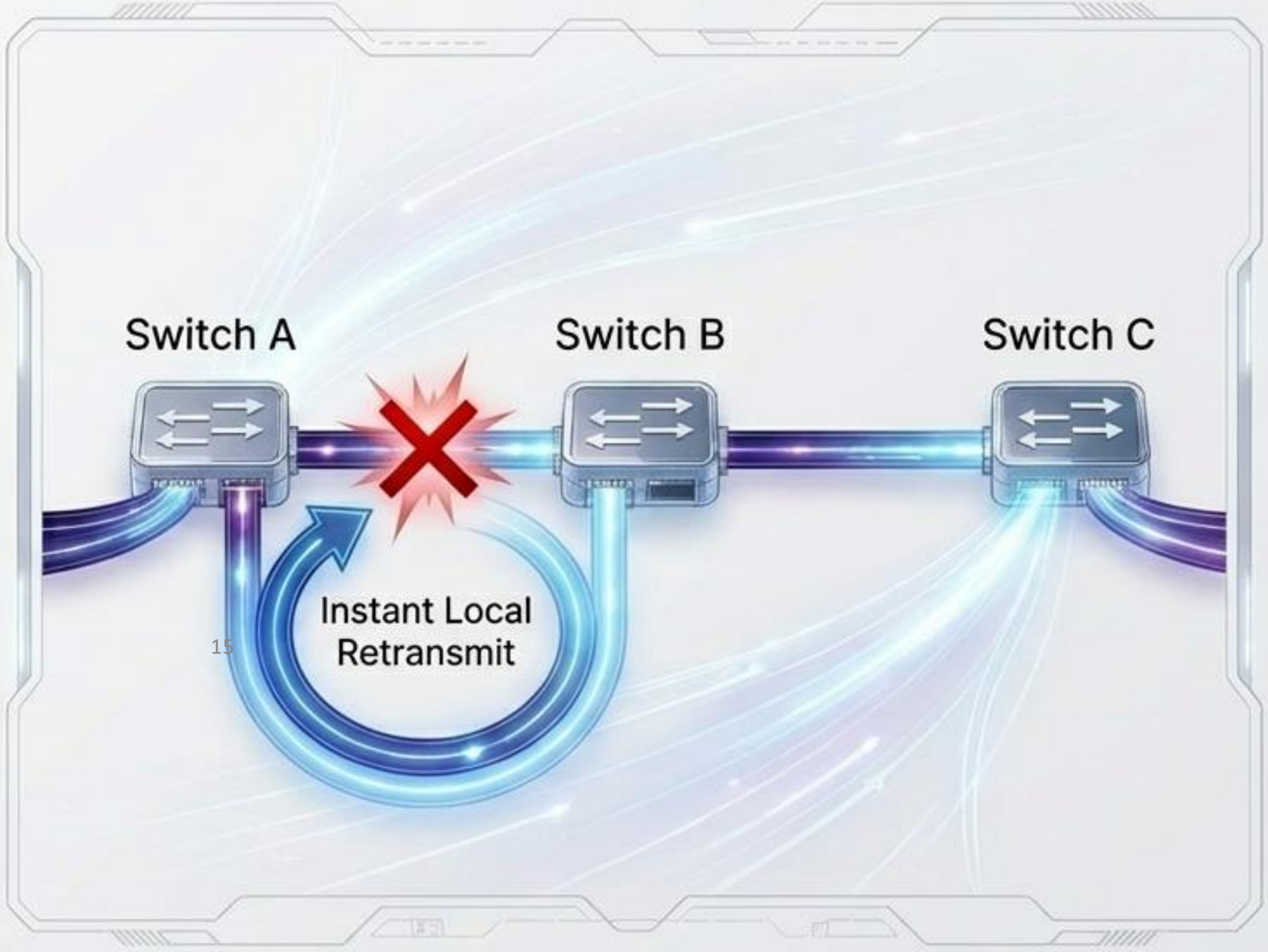
- **Members trust others in the same group**

14

Link Layer Enhancements: LLR

Highly efficient, hop-by-hop reliability.

- Standardized optional extension for the Link Layer.
- Localizes error recovery to the specific physical link where the drop occurred.
- Radically reduces the burden on end-to-end transport recovery, protecting tail latency.

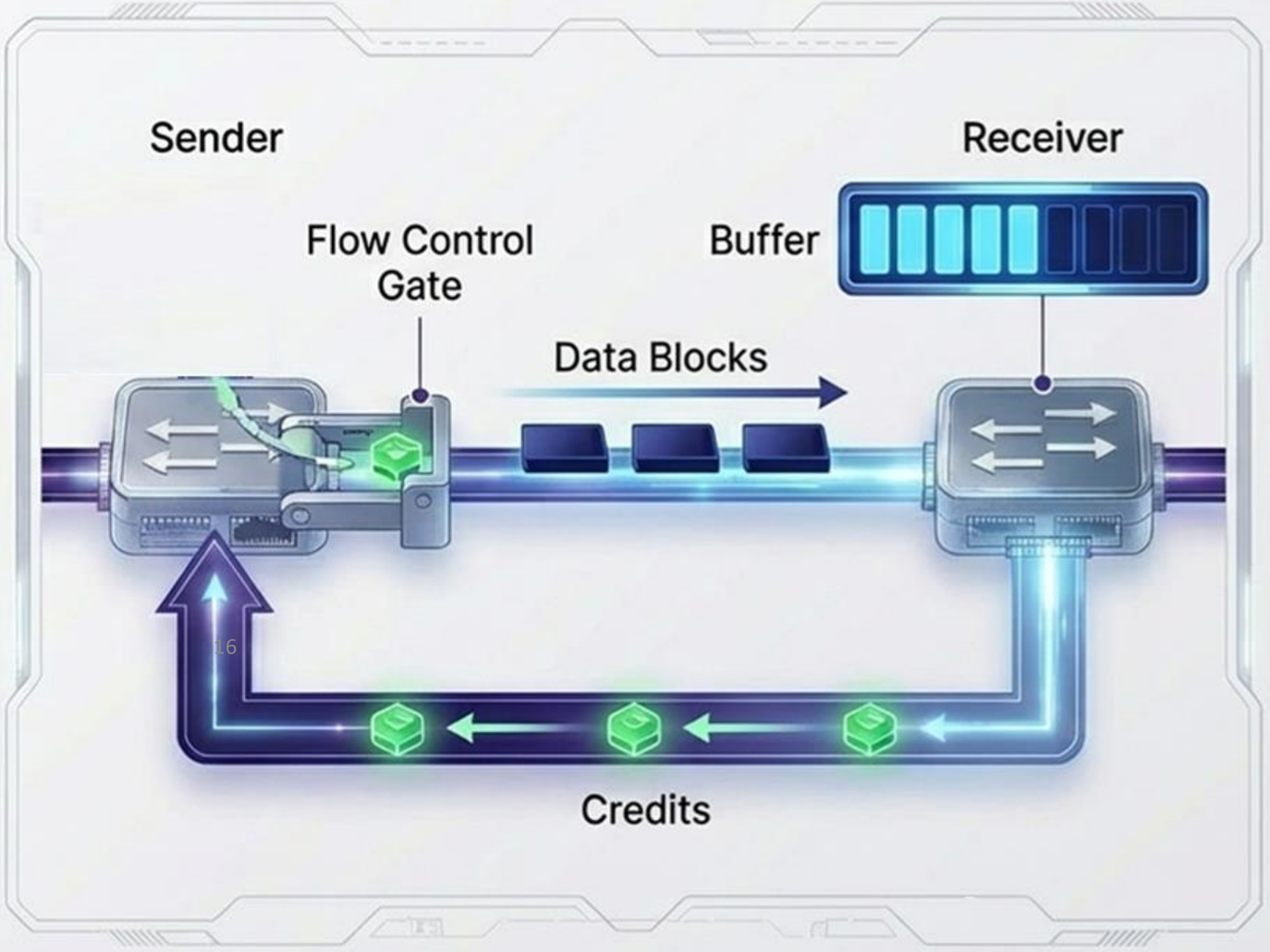


NotebookLM

Link Layer Enhancements: CBFC

Deterministic buffer management at the switch level.

- Standardized optional extension for Credit Based Flow Control.
- Prevents buffer overruns before they occur.
- Ensures lossless characteristics across the entire switch fabric.



NotebookLM

UPCOMING IN ULTRAETHERNET 1.1

Estimated release date: September 2026

1. Support for scale-up (memory semantic) networks
2. Unified Forwarding Header
3. Idle Reduction
4. Scale-out In-network Compute (INC):
   a) *Reduce, All-Reduce, Broadcast collectives*
   b) *Min, max, sum, sum-scale, logical and bitwise operations (or/and)*
5. Programmable Congestion Management
6. Congestion Signaling
   a) *Maximum Delay*
   b) *Min Available Bandwidth*
7. Verbs mapping

Help us define the future of networking

The fabric is open. The standard is ready.

Ultra Ethernet Consortium

Download the 1.0 Specification

Join the Working Groups

Build the Future with UEC

NotebookLM