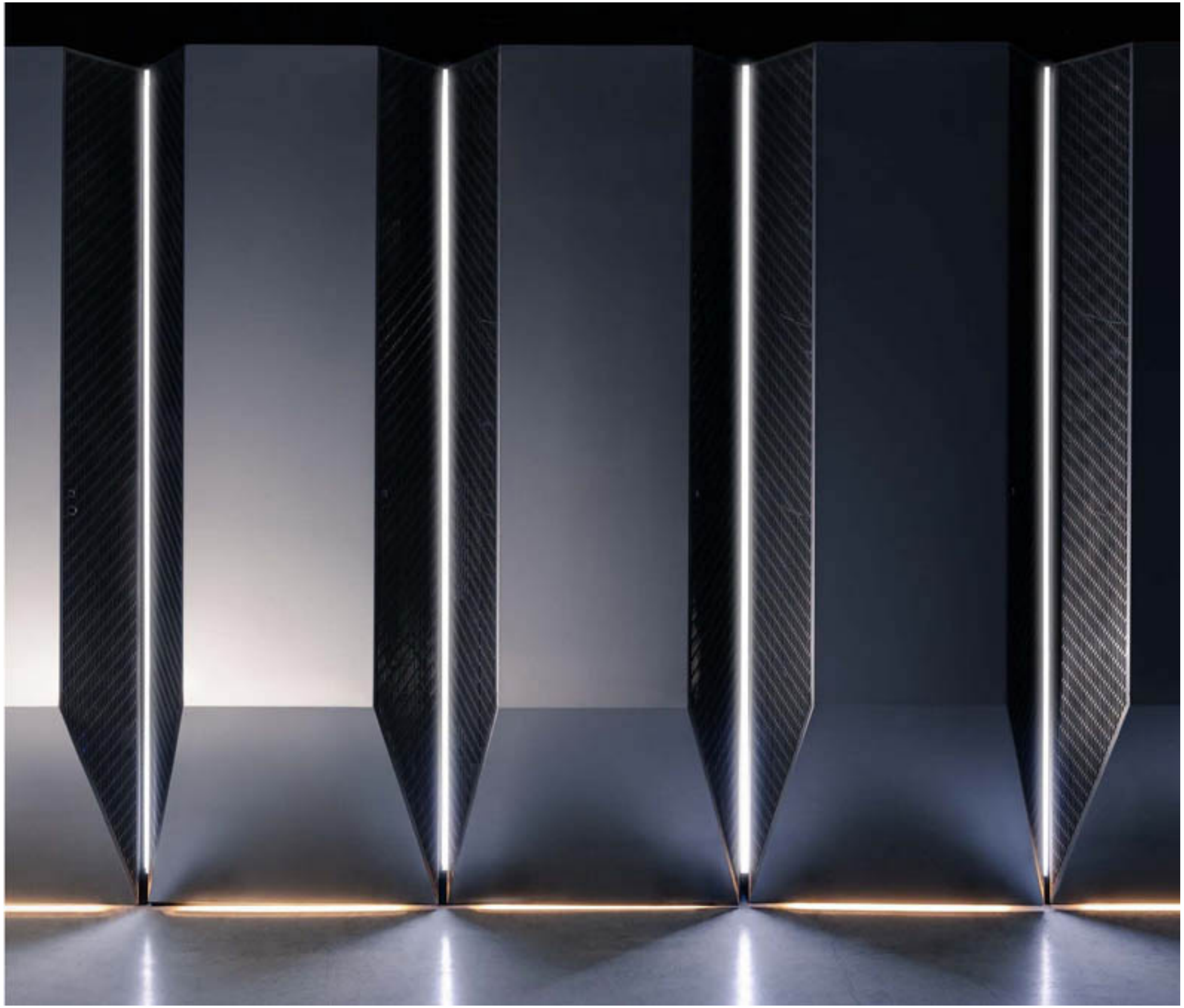
TESLA

Tesla Transport Protocol over Ethernet (TTPoE)

A new lossy, Exa-Scale fabric for the Dojo AI Supercomputer

Eric Quinnell, Ph.D.
Dojo Fabric Lead



Problem Statement

TCP/IP is too slow for scaled AI interconnect

- Bound by CPU SW kernel

Lossless fabrics are complex and brittle

- Priority Flow Control (PFC) affects the global network

Datacenter Ethernet and RDMA: Issues at Hyperscale

Torsten Hoefler
ETH Zürich and Microsoft

Duncan Roweth, Keith Underwood, Bob Alverson
Hewlett Packard Enterprise

Mark Griswold, Vahid Tabatabaee, Mohan Kalkunte, Surendra Anubolu
Broadcom

Siyan Shen
ETH Zürich

Abdul Kabbani, Moray McLaren, Steve Scott
Microsoft

Ideal Fabric:

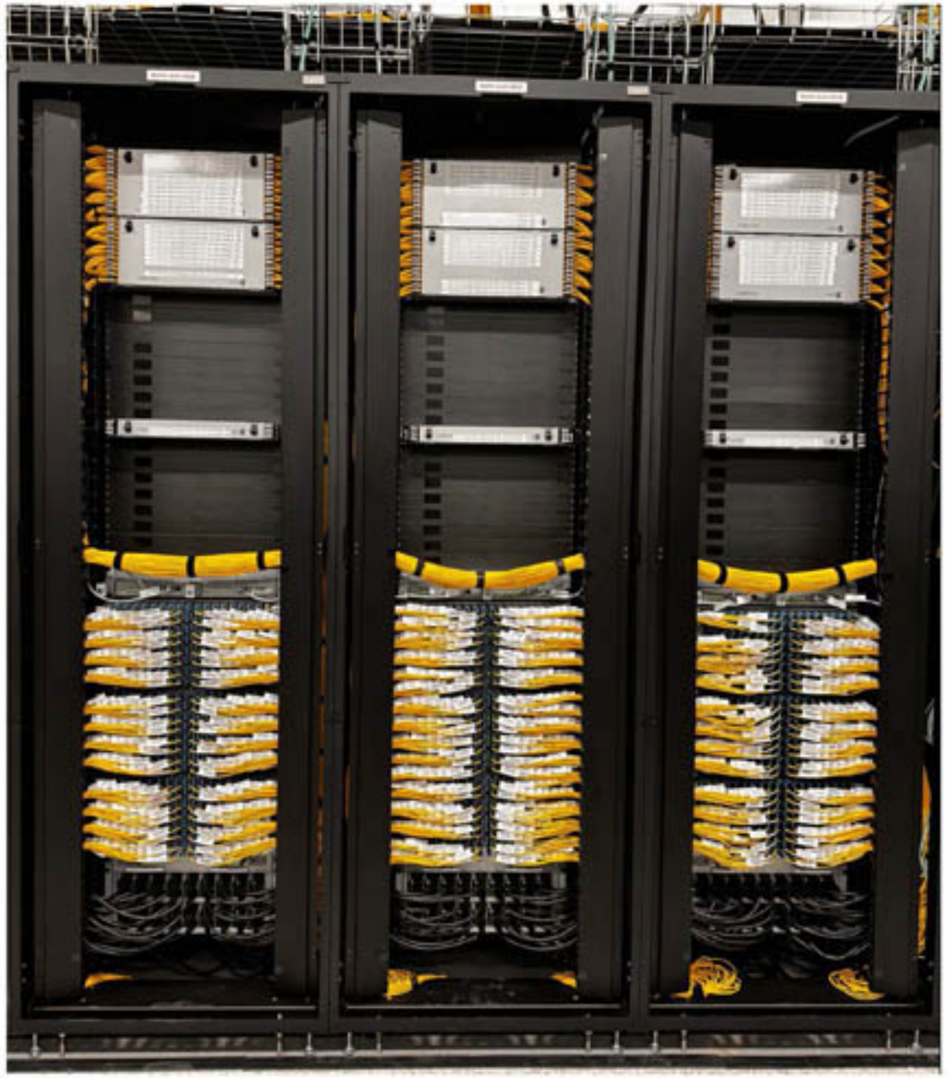
- Lowest latency
- Highest bandwidth
- Simple Software

For Tesla AI:

- Layer 2 only
- Collective communications and ingest
- Low congestion, single application

TTPoE

Tesla Transport Protocol over Ethernet (TTPoE)
is a peer-to-peer ethernet Transport Layer Protocol executed **entirely in hardware**.



Why a custom transport protocol?

1. **Vertical Integration –** *extend Dojo RDMA onto optical fabric*
2. **"Lossy" ethernet network –** *ease of scaling, cost, congestion mgmt.*
3. **Use 3rd party hardware –** *Ethernet II frames "Just Work"*

TCP got it right – just do it in hardware

Dojo OSI Layers

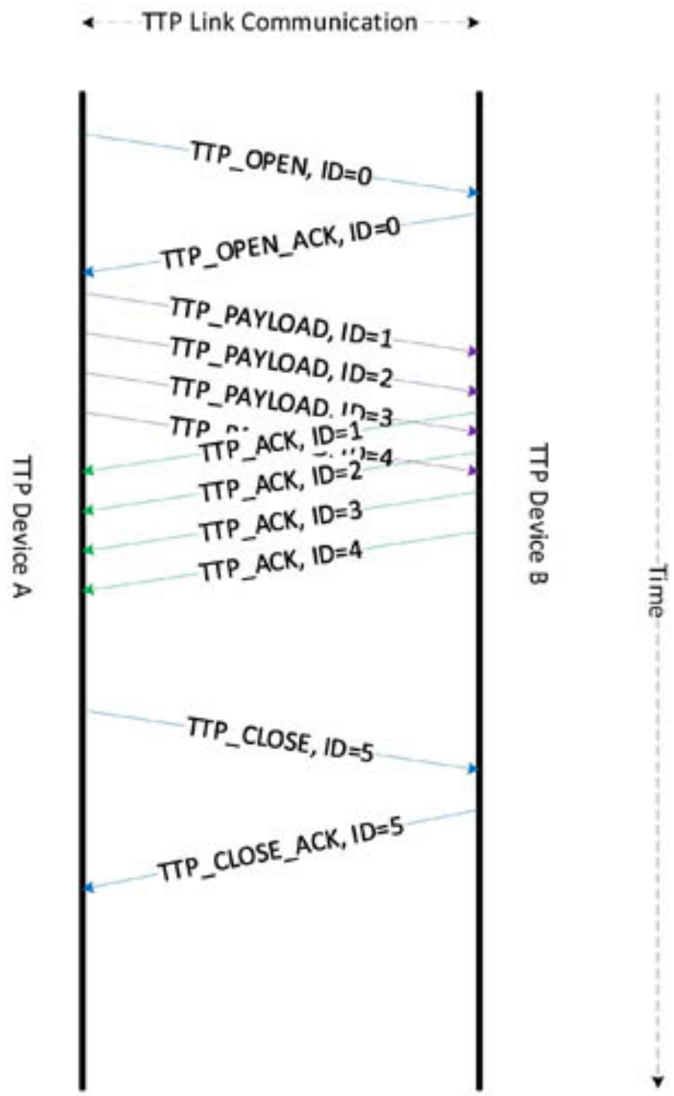
Standard Stack

OSI Layer	Example Protocols (TCP/IP)	TCP/IP Implementation
Layer 7 Application	HTTP, Telnet, FTP	Software
Layer 6 Presentation	JPEG, PNG, MPEG	
Layer 5 Session	NFS, SQL	
Layer 4 Transport	TCP, UDP	
Layer 3 Network	IPv4/IPv6	
Layer 2 Data Link	Ethernet Frames, MAC addresses, VLAN	Hardware
Layer 1 Physical	Data Encoding, Physical Specs	

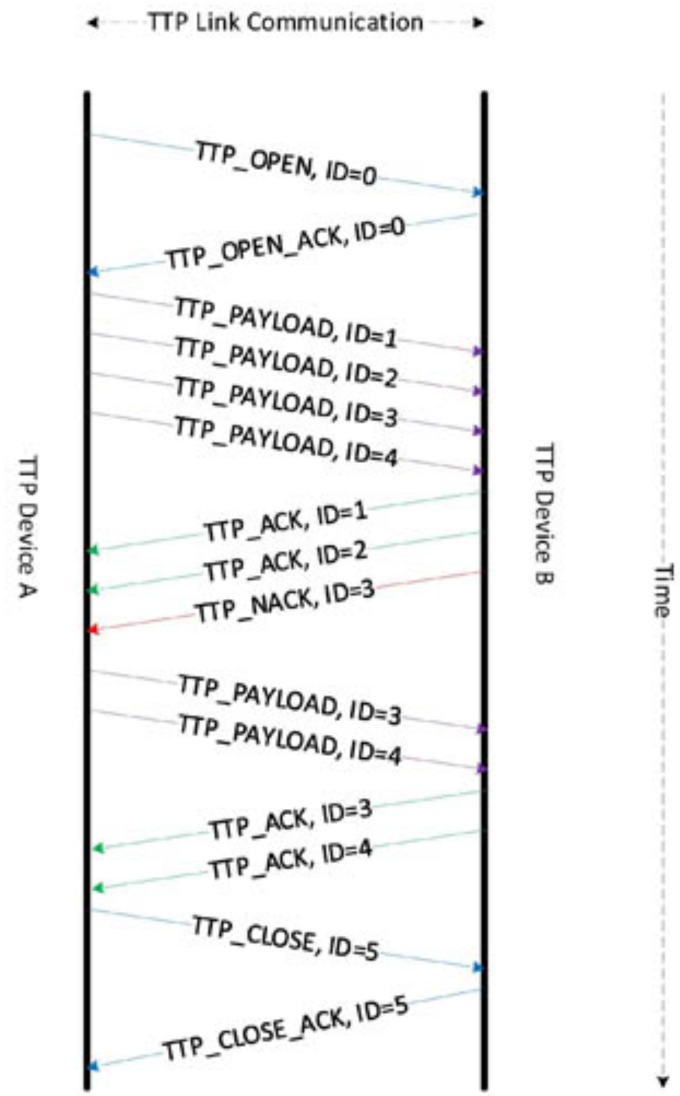
Dojo Stack

OSI Layer	Example Protocols	Dojo Implementation
Layer 7 Application	Pytorch, Dojotorch	Software
Layer 6 Presentation	FFMPEG, HEVC, YUV	
Layer 5 Session	Dojo RDMA Descriptors	
Layer 4 Transport	**TTP**	Hardware
Layer 3 (Optional) Network	IPv4/IPv6 (Optional)	
Layer 2 Data Link	Ethernet Frames, MAC addresses, VLAN	
Layer 1 Physical	Data Encoding, Physical Specs	

TTP transaction examples



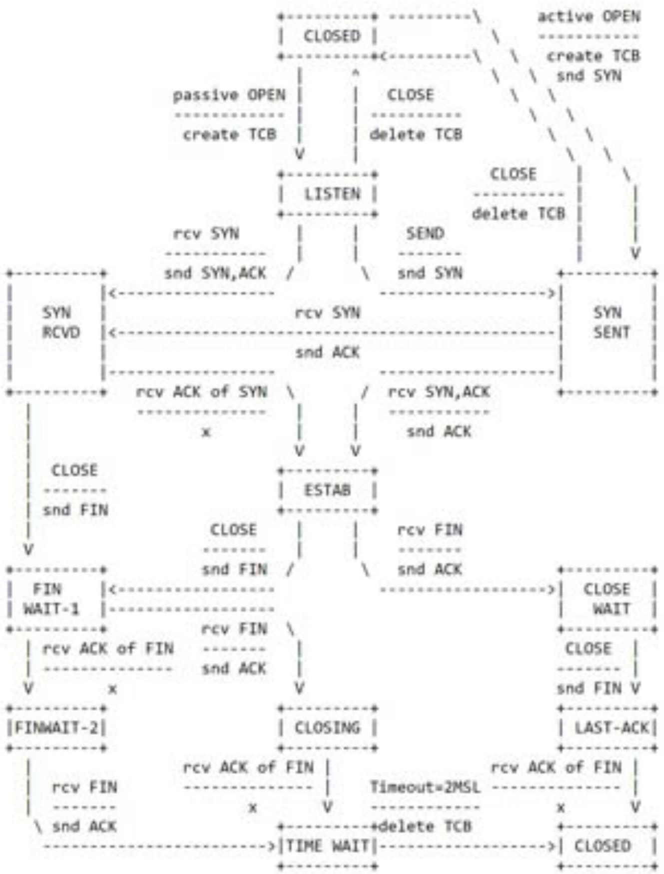
Clean TTP transfer Example



NACK TTP transfer Example.
TTP_PAYLOAD, ID=3 is either lost or out of order

Transport Layer State Machines

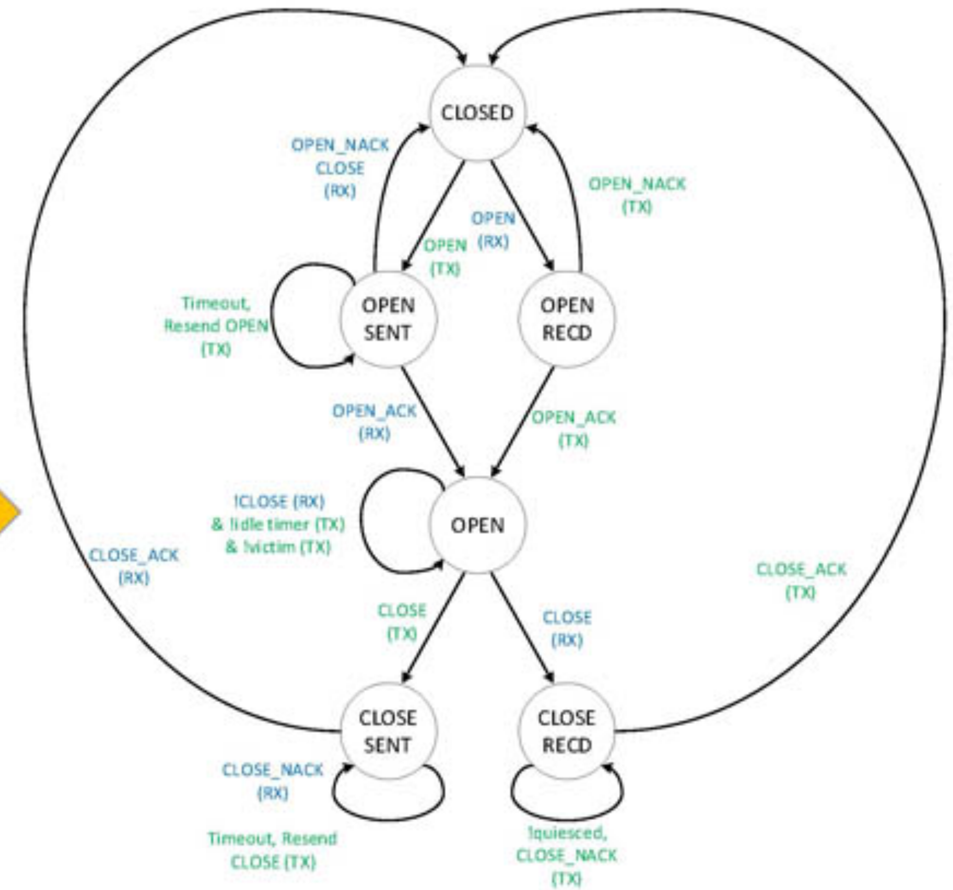
TCP STATE MACHINE



TCP Connection State Diagram
Figure 6.

IETF RFC-793

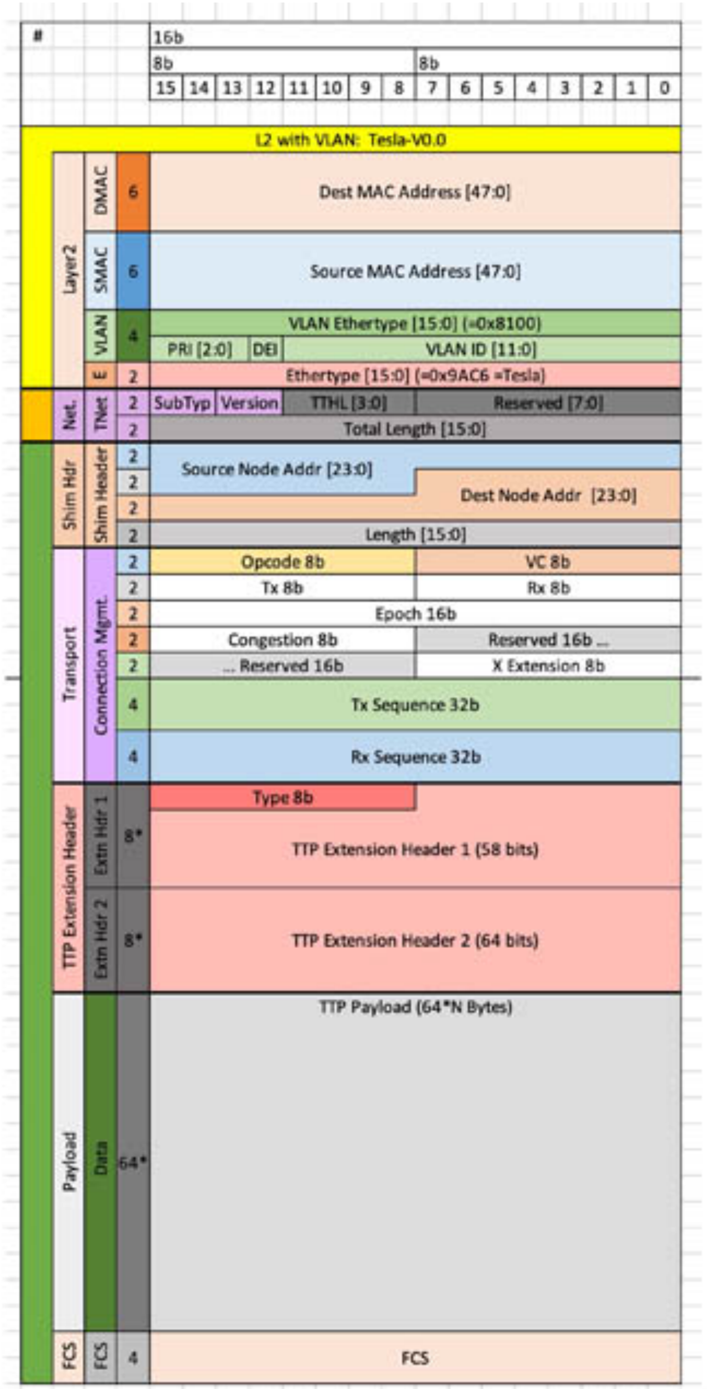
TTP STATE MACHINE



Modifications made for hardware-only execution

- 2 millisecond quiesce in a microsecond protocol is too long
- No reliance on virtual memory – physical memory only
- Automatic OPEN/CLOSE with no SW involvement

TTP Header Frame



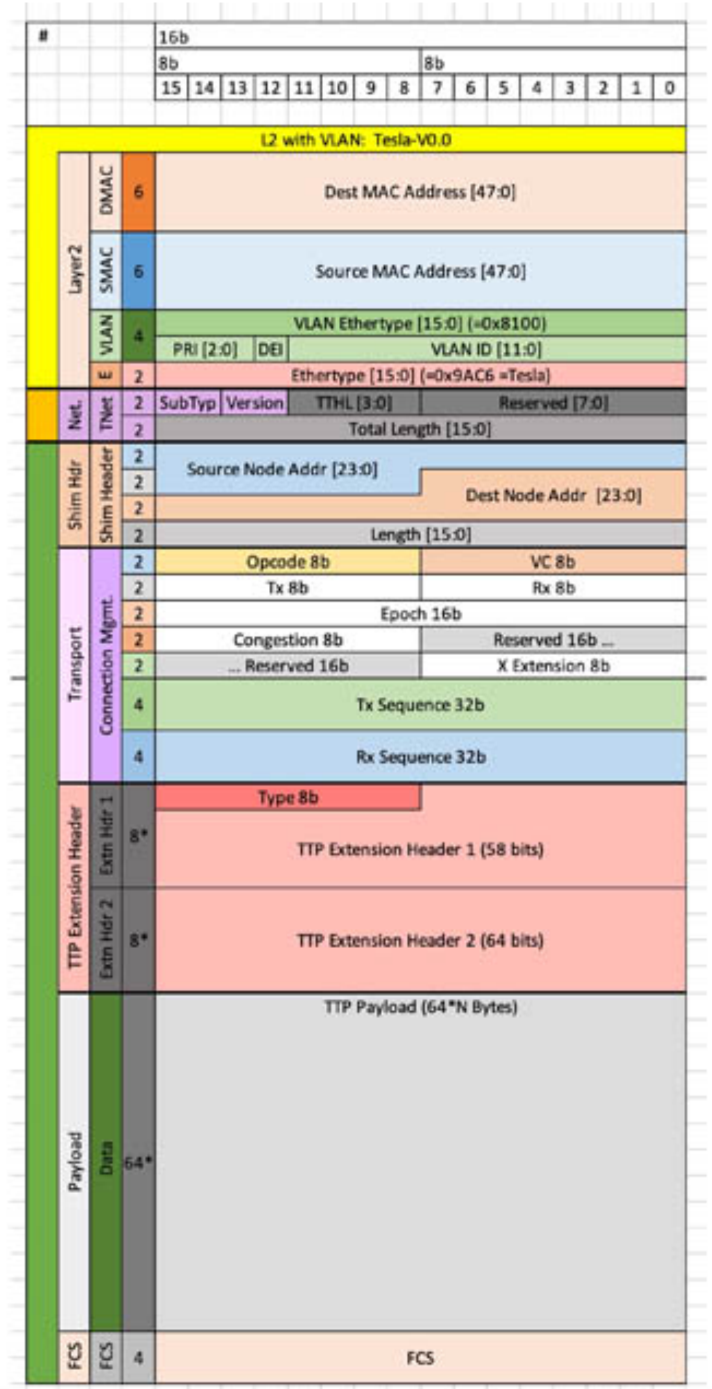
TTP uses Ethernet-II simple formats with optional standard Layers

- Dojo at scale uses only Layer 2, currently not using Layer 3
- MAC addresses are a hardware hash of the SOW Physical Address (PA)
- A TTP endpoint can concurrently handle 512 unique links, dynamically replaced via victimization and LRU
- Virtual channels (VCs) allow for non-blocking control, semaphore, completion, and data movement

Lossy Protocol

<u>TTPoE is a "lossy" transport protocol</u>

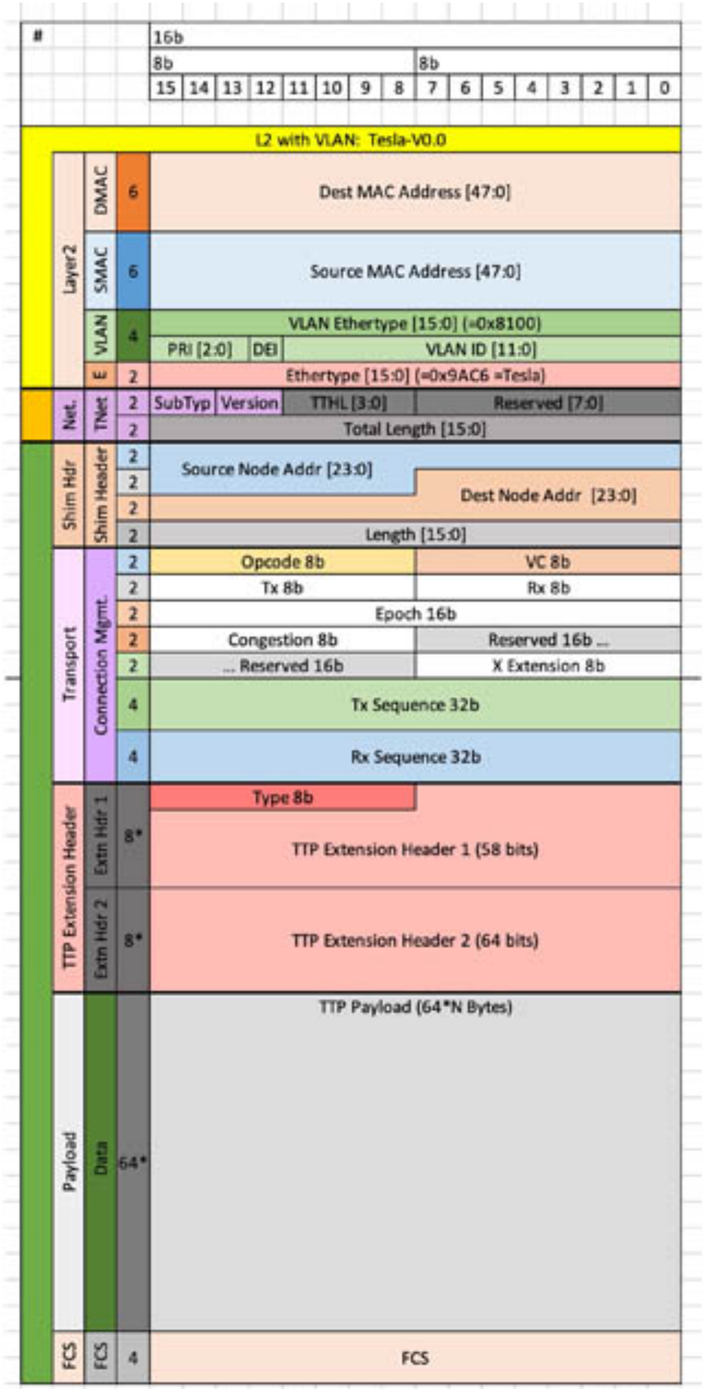
- "Lossy" transport meaning the underlying medium expects to lose packets and retry – full packet transmission is still guaranteed.
  - Similar to TCP and unlike UDP.
- TTP will default to packet drops and replays in corner cases of congestion, backpressure, or errors
- Speculative transmission is limited by SRAM size before a RTT ACK. This, in effect, forces a "TTP window size" beyond which bandwidth is lost
- Local SRAM lines are not retired/deallocated until the ACK comes back, allowing HW to replay the line.
- Replay amounts are also limited by SRAM, constraining the scale of replay storms



Congestion Management

<u>Congestion management is distributed</u>

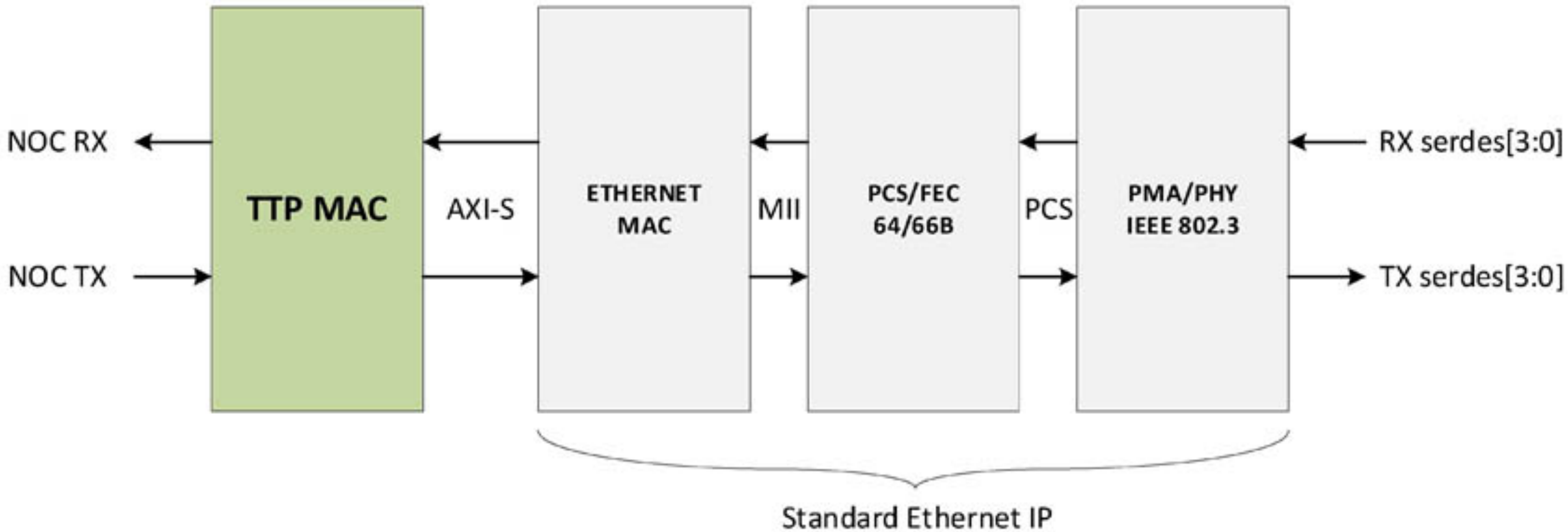
- Exponential backoff, rate control, and algorithms are handled by local link TX channels, not by central network or switch.
- Fault Tolerant flow “flushes” the TTP network and removes a bad link before continuing training
- No PFC, no Nagel Algorithm, no QoS, no tokens, no lossless artifacts



TTP MAC IP

The Transport Layer hardware is an IP block between a NOC and an Ethernet standard MAC

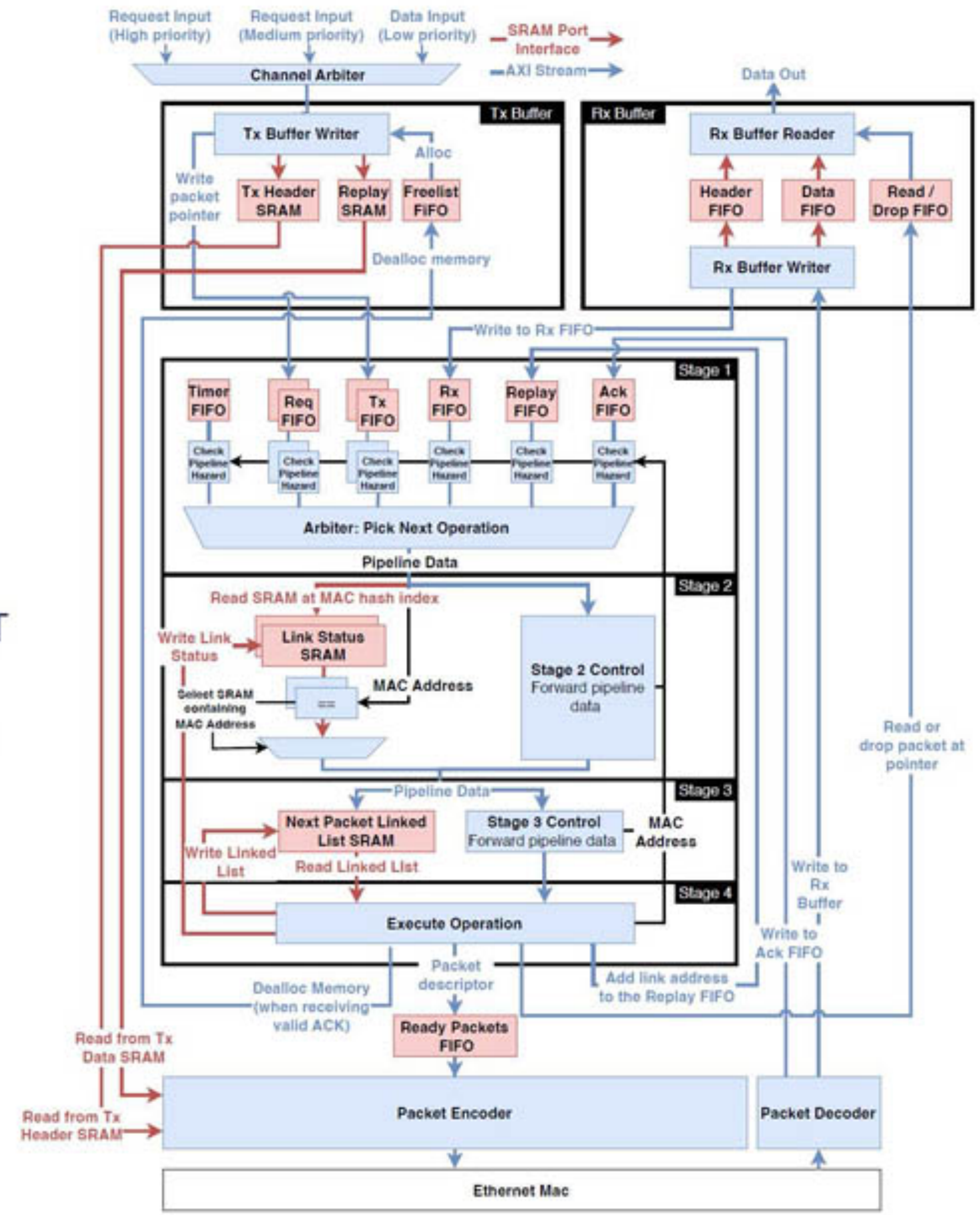
- Translates and coalesces 64B/cycle NOC packets into up to 1kB TTP Ethernet packets
- Speaks AXI-S or SOP/EOP formats
- Optionally activates standard MAC features – pause packets, counters, stats, LLDP
- IP block instantiated in FPGA and Silicon implementations



TTP MAC Micro-Architecture

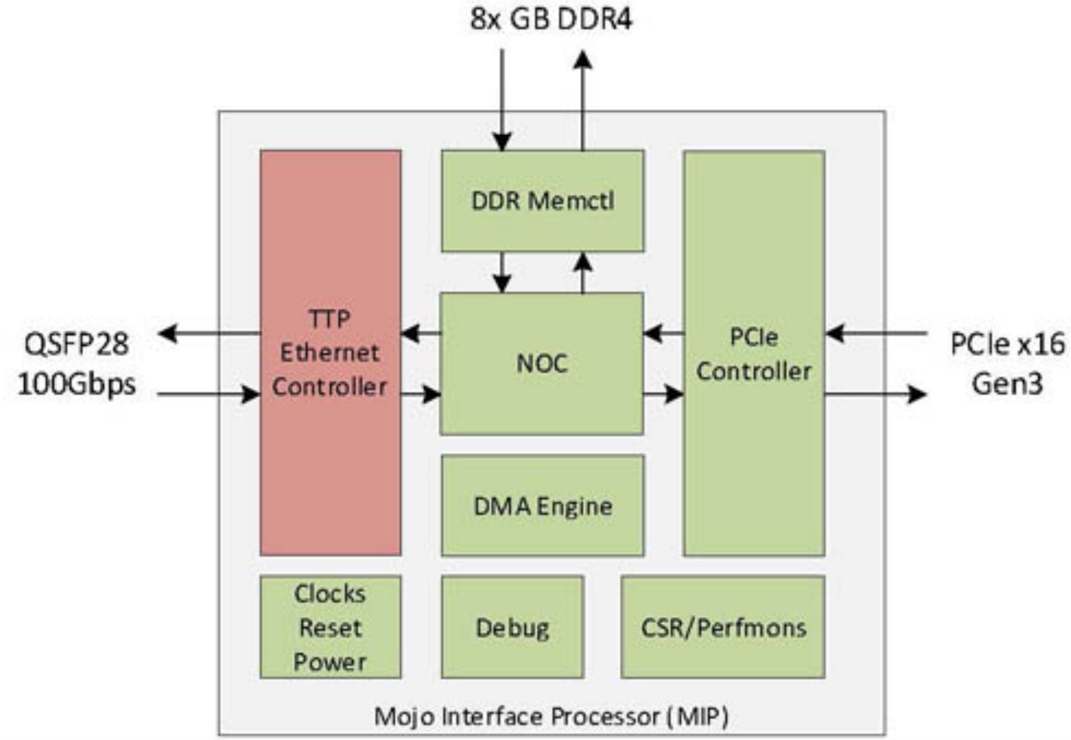
TTP's Micro-Architecture uses techniques from SMP Caches, Snoop Filters, CPUs

- 4-stage Read-Modify-Write (RMW) Pipeline
- TX Buffer size determines maximum outstanding packets before stall/backpressure
  - ACK packets "retire" a packet from the common buffer
  - 1MB TX Buffer allows for ~80 microseconds latency tolerance RTT
- Virtual Channels to prioritize and avoid livelock/deadlock
- Multi-channel "coherent" arbitration to update link and use the TX Physical Channel
- DMA descriptors issue to TTP MAC
  - Can be PUSH for implicit pass-thru local-to-remote
  - Can be explicit HBM2HBM fabric memcpy



"Mojo" 100Gbps Dumb-NIC

Feature	Spec
Ethernet Speed	100Gbps QSFP
PCI-e	Gen3 x16
Memory	8GB DDR4
Power	<20W max
Reliability	5-year tested
DMA engine	Dojo DMA
CPU+OS	None
Active Links	512 unique, 2-way, LRU



First integration box - D1 Die

TSMC 7nm, 645mm^2

Physically and logically arranged as a 2D array

- 354 DOJO processing nodes on die

Extremely modular design

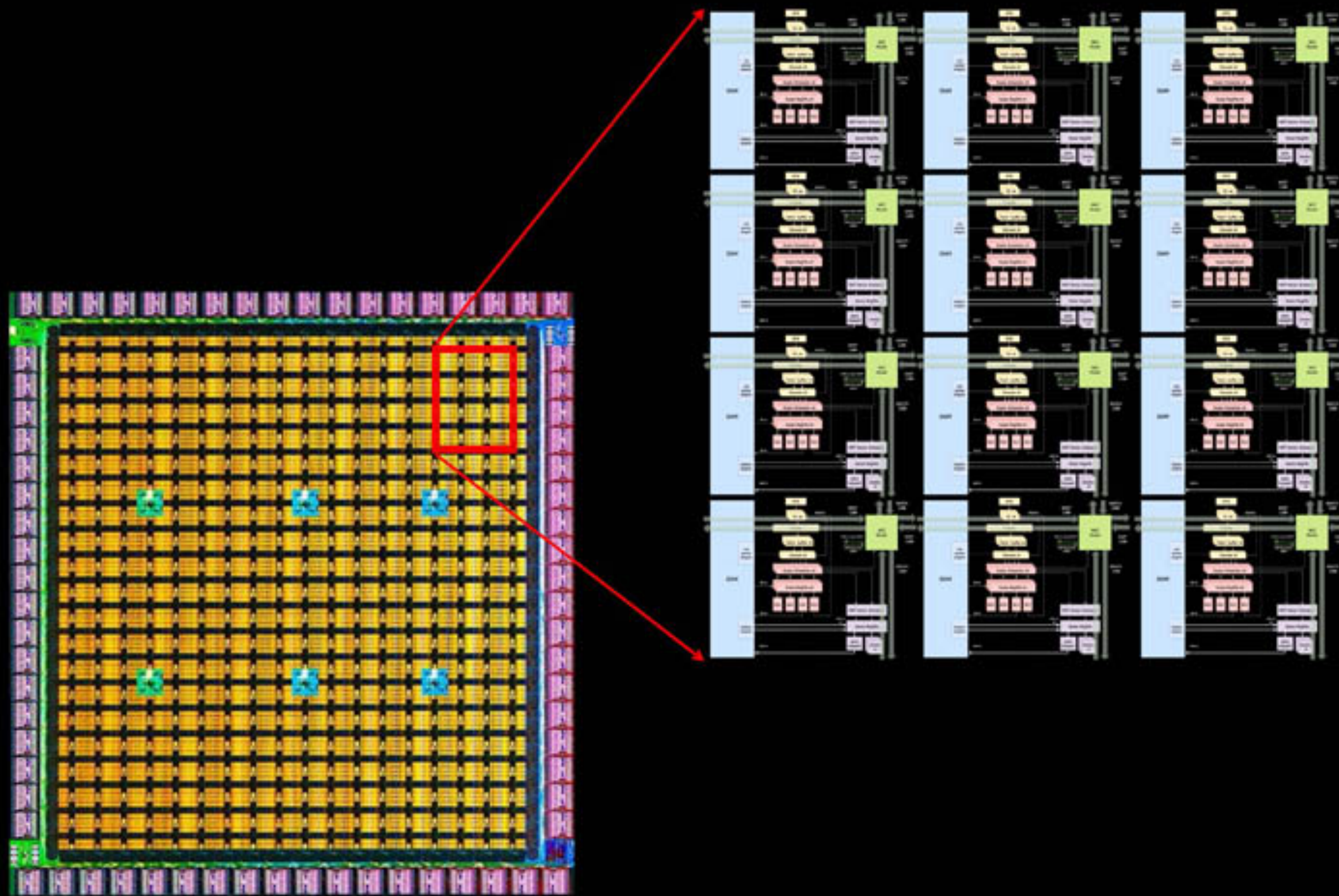
362 TFlops BF16/CFP8, 22 TFlops FP32 @2GHz

440 MB SRAM

Custom low power serdes channels on all edges

- 576 bidirectional channels
- 2 TB bandwidth on each edge

Seamless connection to neighboring dies



Second integration box – Dojo Training Tile

5x5 array of known good D1 chips
- 4.5TB/s off-tile bandwidth per edge
  - Half of in-tile bandwidth

Fully integrated module
- Electrical + thermal + mechanical
- 15kW of power delivery

Custom power delivery
- Horizontal data communication plane
- Vertical power delivery and cooling
- 15kW per module

Custom high-density connectors
- Seamless connection to neighboring training tiles

V1 Dojo Interface Processor

32GB High-Bandwidth Memory

- 800 GB/s Total Memory Bandwidth

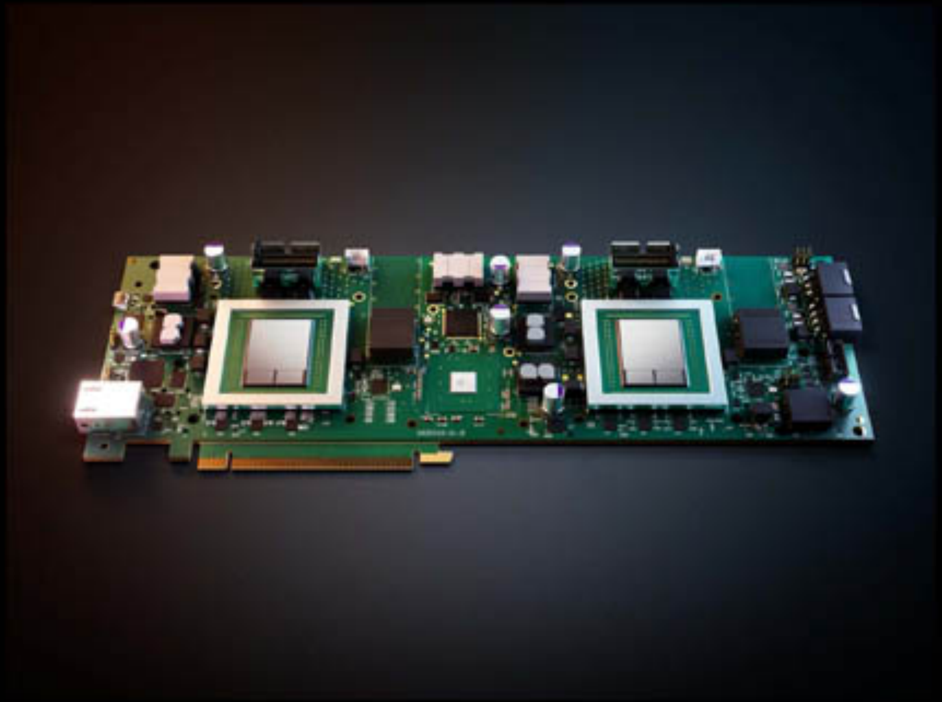
900 GB/s TTP Interface

- Tesla Transport Protocol (TTP) - Full custom protocol
- Provides full DRAM bandwidth to Training Tile

50 GB/s TTP over Ethernet (TTPoE)

- Enables extending communication over standard Ethernet
- Native hardware support

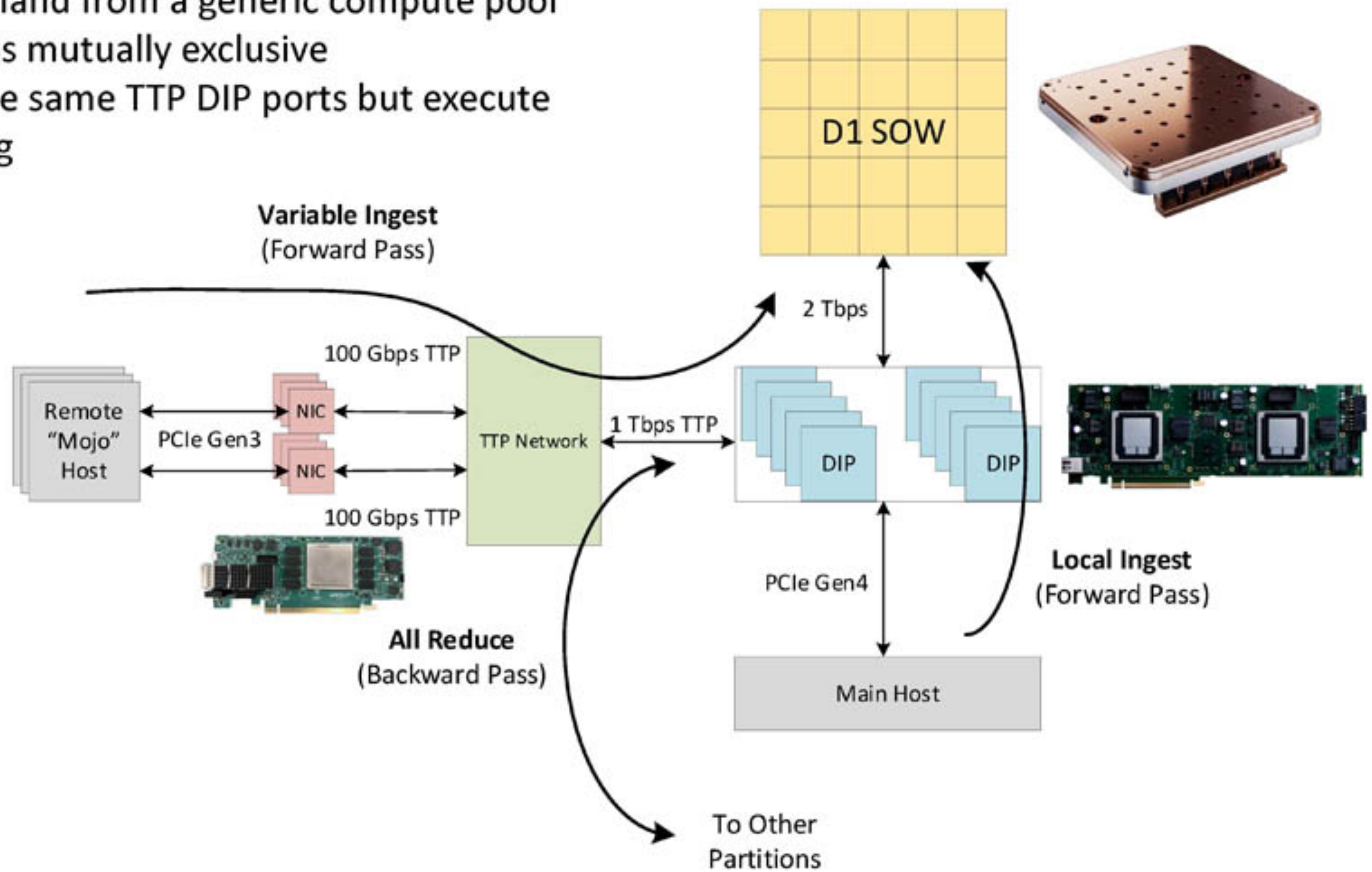
32 GB/s Gen4 PCIe Interface



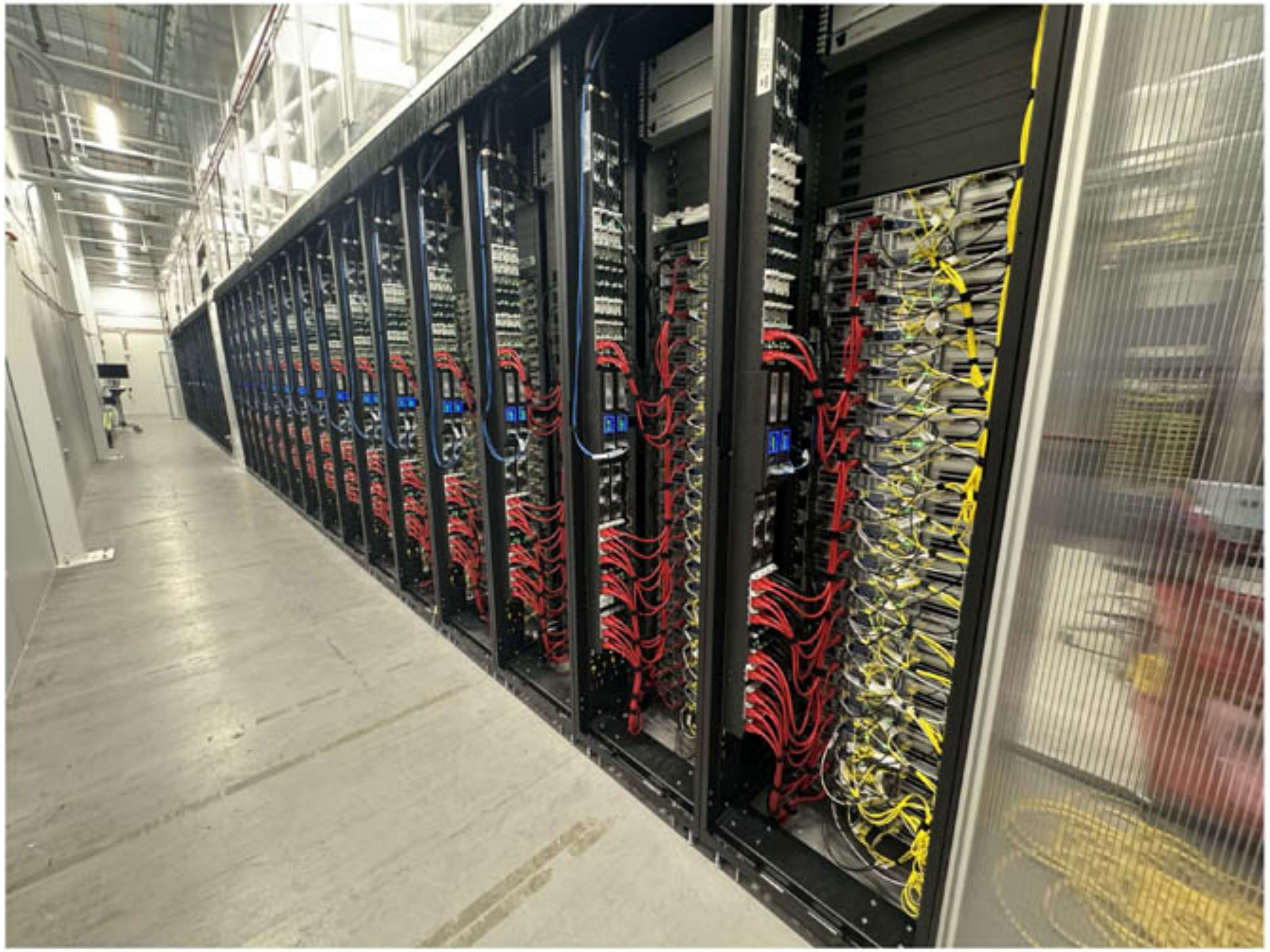
“Mojo” Hosts – Variable Ingest via TTP Network

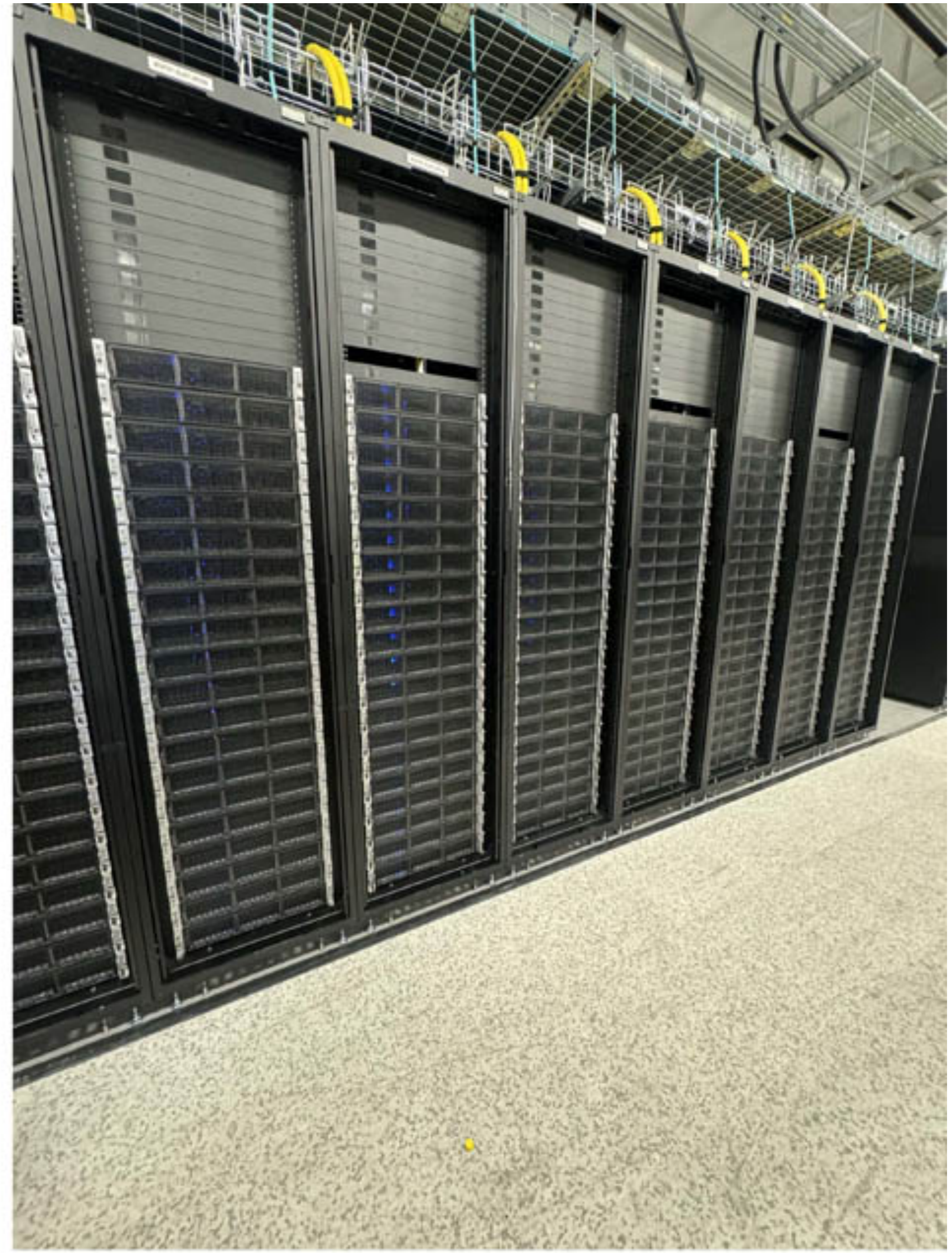
Vision networks can be heavily ingest limited

- Vision-based tensors and training clips in GBs
- “Mojo” Hosts are scheduled on demand from a generic compute pool
- Forward/Backward pass TTP traffic is mutually exclusive
  - i.e. ingest and all-reduce share the same TTP DIP ports but execute during different phases of training



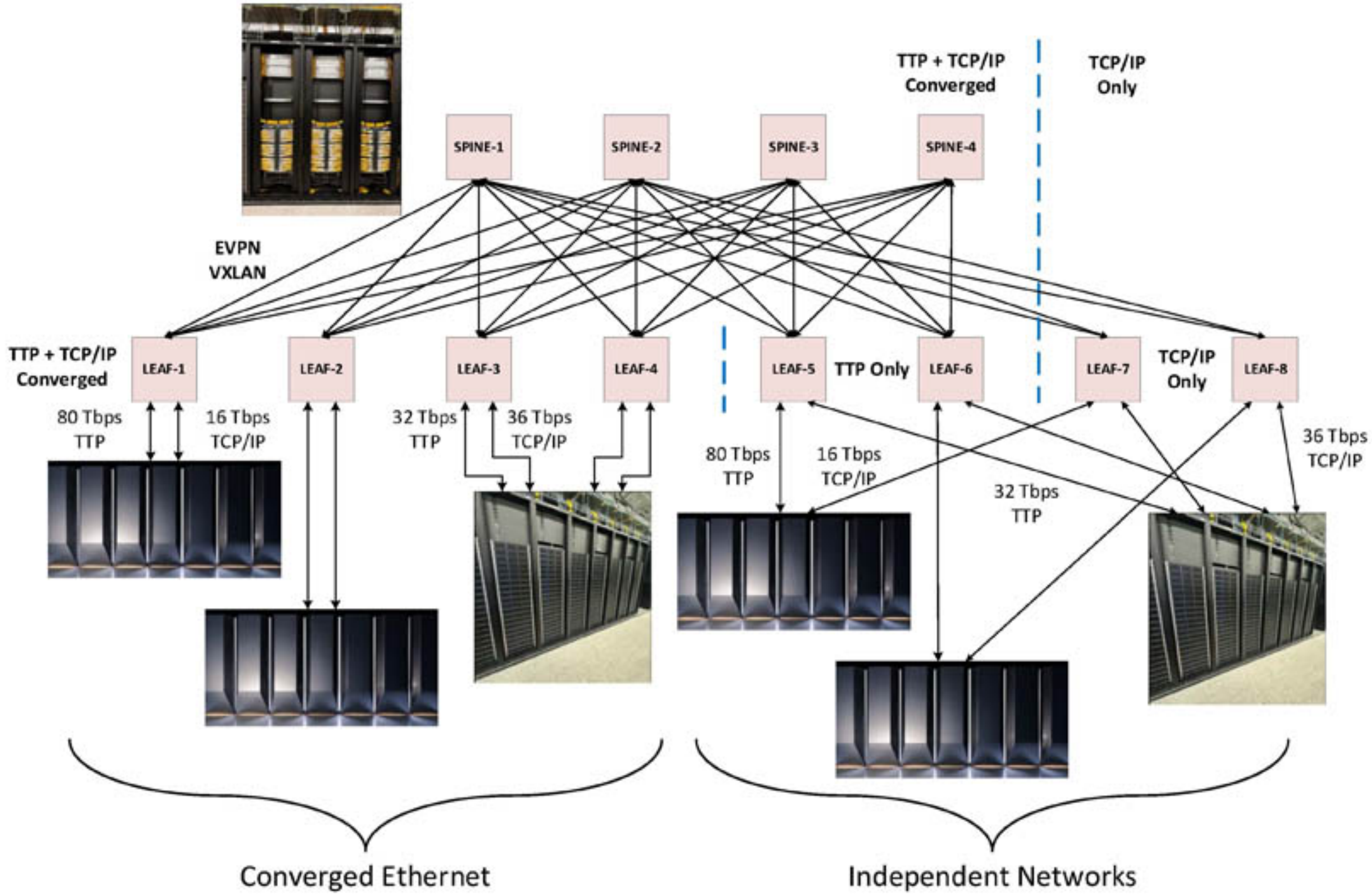
MDCH – Mojo Dojo Compute Hall





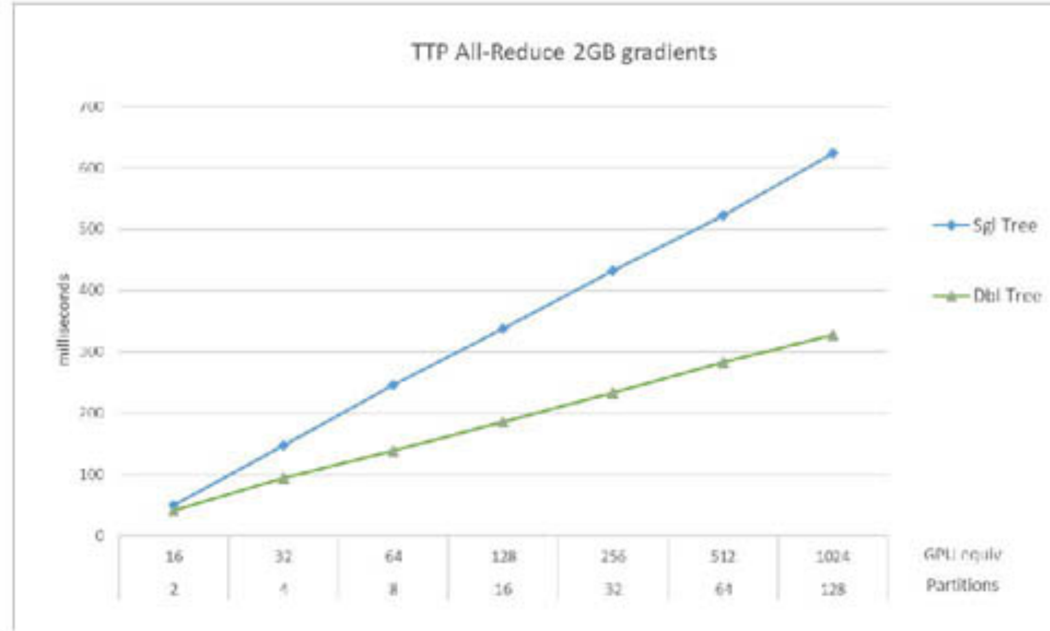
Dojo Engineering System

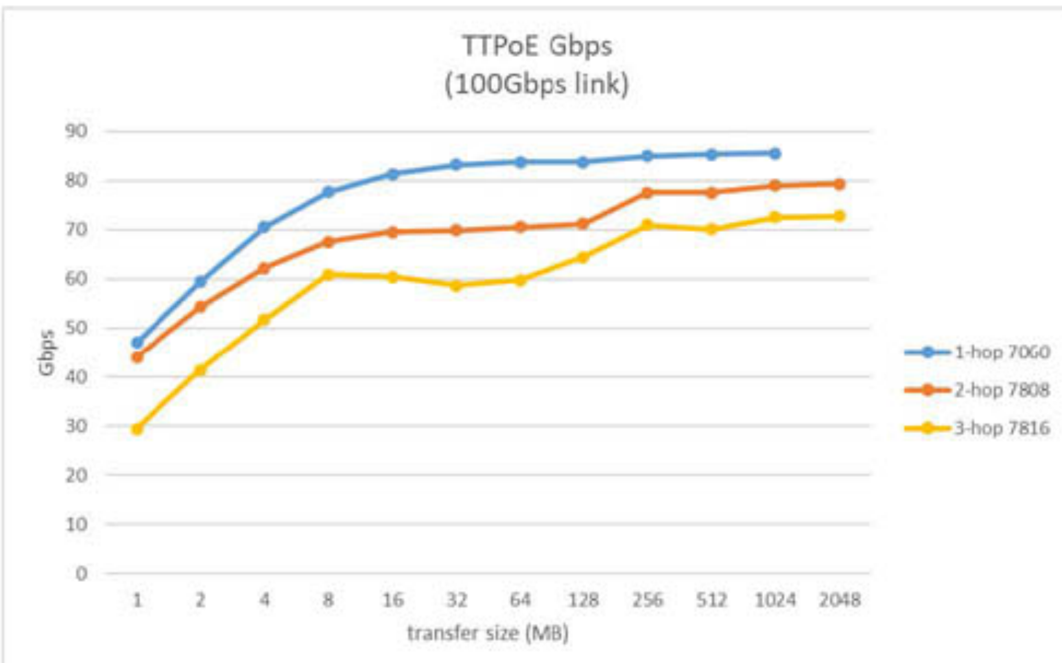
- 4xExaFLOP BF16/FP16 Cluster
- 40 PB Local Storage
- 40,960 Main Host Cores
- 61,440 Mojo Host Cores
- 320 Tbps TTP All-Reduce I/O (endpoint)
- 128 Tbps TTP Ingest I/O (endpoint)
- 208 Tbps TCP/IP (endpoint)
- Converged and non-Converged network experiments

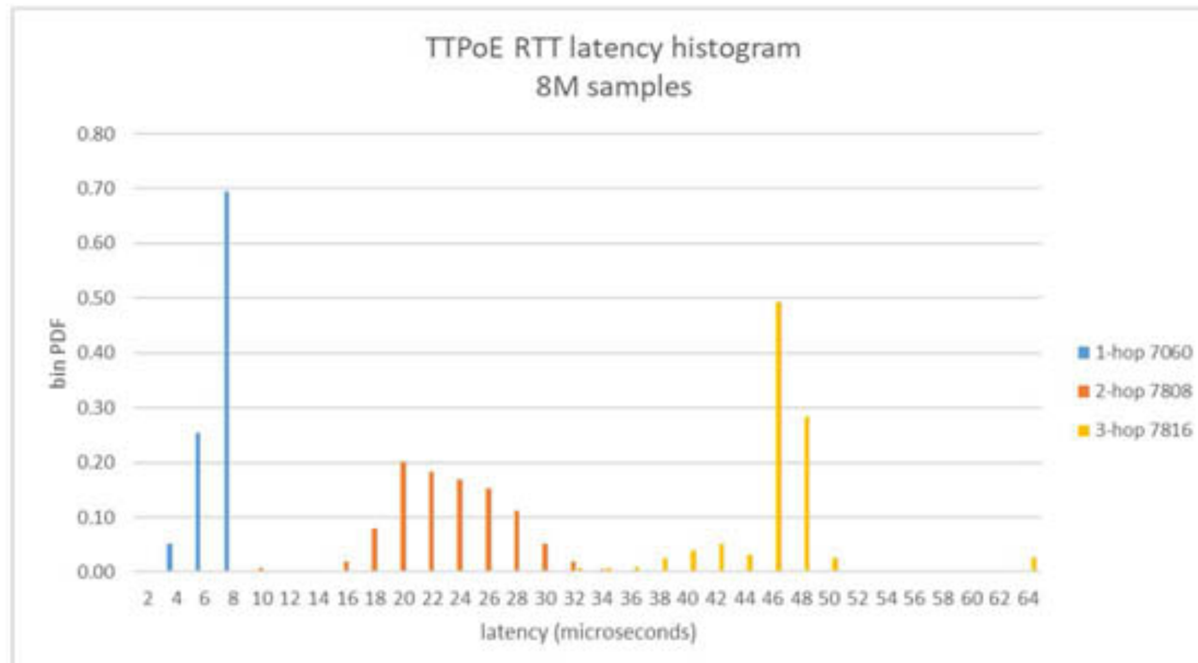


Results

- Measured on Arista 7060, 7808, and 7816 switches
- RTT latency is random sampling of in-flight packets + ACK return
- Gbps is wall time real-data movement
- All-reduce measure is network only, non-pipelined
  - SOW has all-reduce not shown (pre-network)
- All-reduce throughput is determined by the slowest node in system





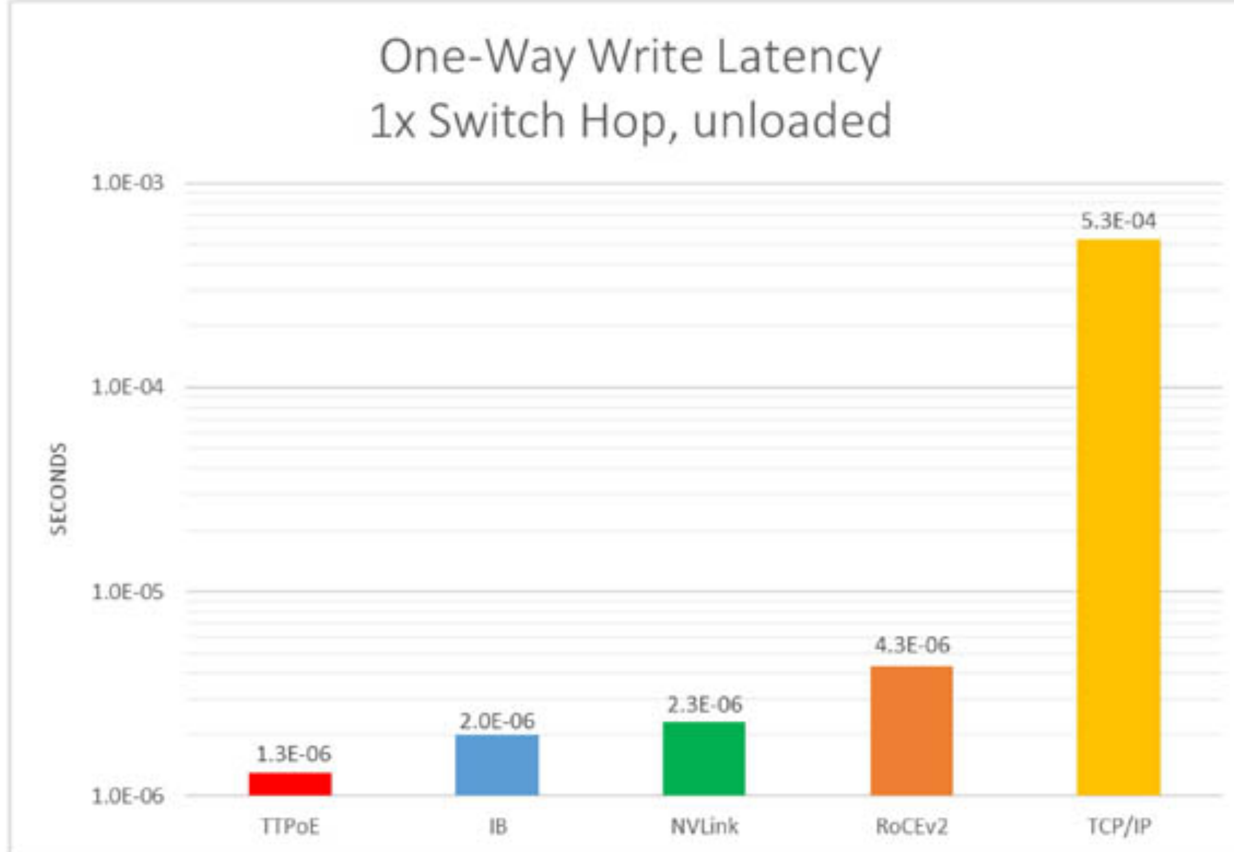


Backup – Latencies

Intended de-emphasis on synthetic latency measurements

Differences of greater consequence:

- lossy vs lossless
- centralized vs distributed congestion
- proprietary vs open source
- sustained bandwidths at scale



TTPoE, TCP/IP – Spectrum3 SN4700
IB – Spectrum 9700 IB
Nvlink – DGX-H100 NvSwitch level1 (internal)
RoCEv2 – 7812 R3

Inconsistent methodology and hardware, not at scale

TTPoE in Ultra Ethernet Consortium (UEC)

Ultra Ethernet Consortium

https://ultraethernet.org/

Steering Members

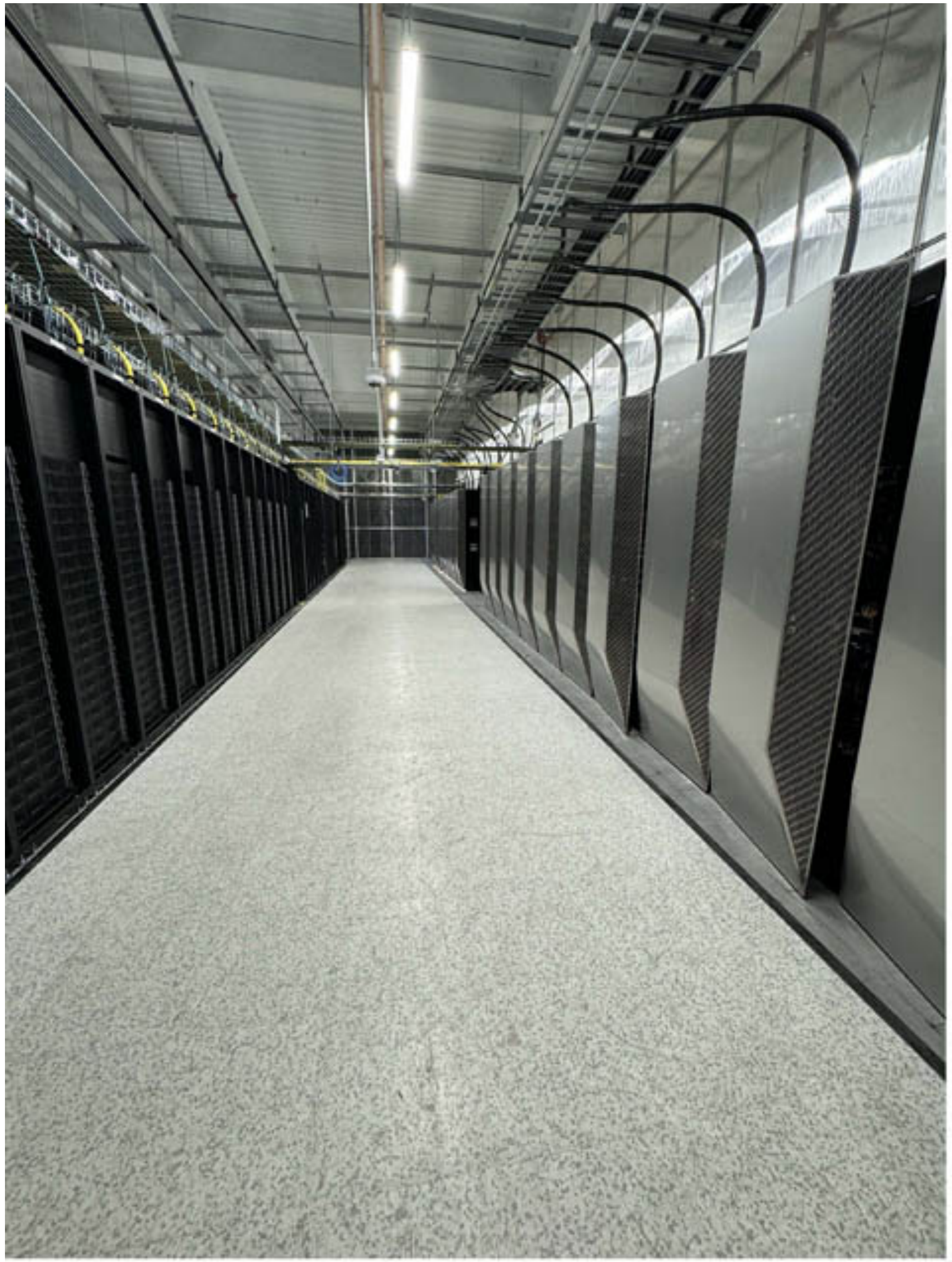
AMD, ARISTA, BROADCOM, CISCO, EVIDEN an atos business, Hewlett Packard Enterprise, intel, Meta, Microsoft, ORACLE

While large lossless RoCE networks can and have been successfully deployed, they require careful tuning, operation, and monitoring to perform well without triggering these effects. This level of investment and expertise is not available to all network operators and leads to a high TCO. A transport protocol that does not depend on a lossless fabric is needed.

https://ultraethernet.org/wp-content/uploads/sites/20/2023/10/23.07.12-UEC-1.0-Overview-FINAL-WITH-LOGO.pdf

Tesla has achieved Exa-scale with a lossy fabric, executing real training runs deployed in FSD

Tesla is joining the UEC and offering the TTPoE protocol publicly



Team Acknowledgements

Prototyping is Easy. Scaling is Hard

Thanks to the

TTPoE Original Inventors, Network Deployment Team, Silicon Design Team, System and Infrastructure Team, SW and Drivers Team, Linux Patch Team, SDN Team, DevOps Team, QA Team, DC Tech Team, Supply Team, and all TTP/Mojo Interns

Tesla Transport Protocol over Ethernet (TTPoE)