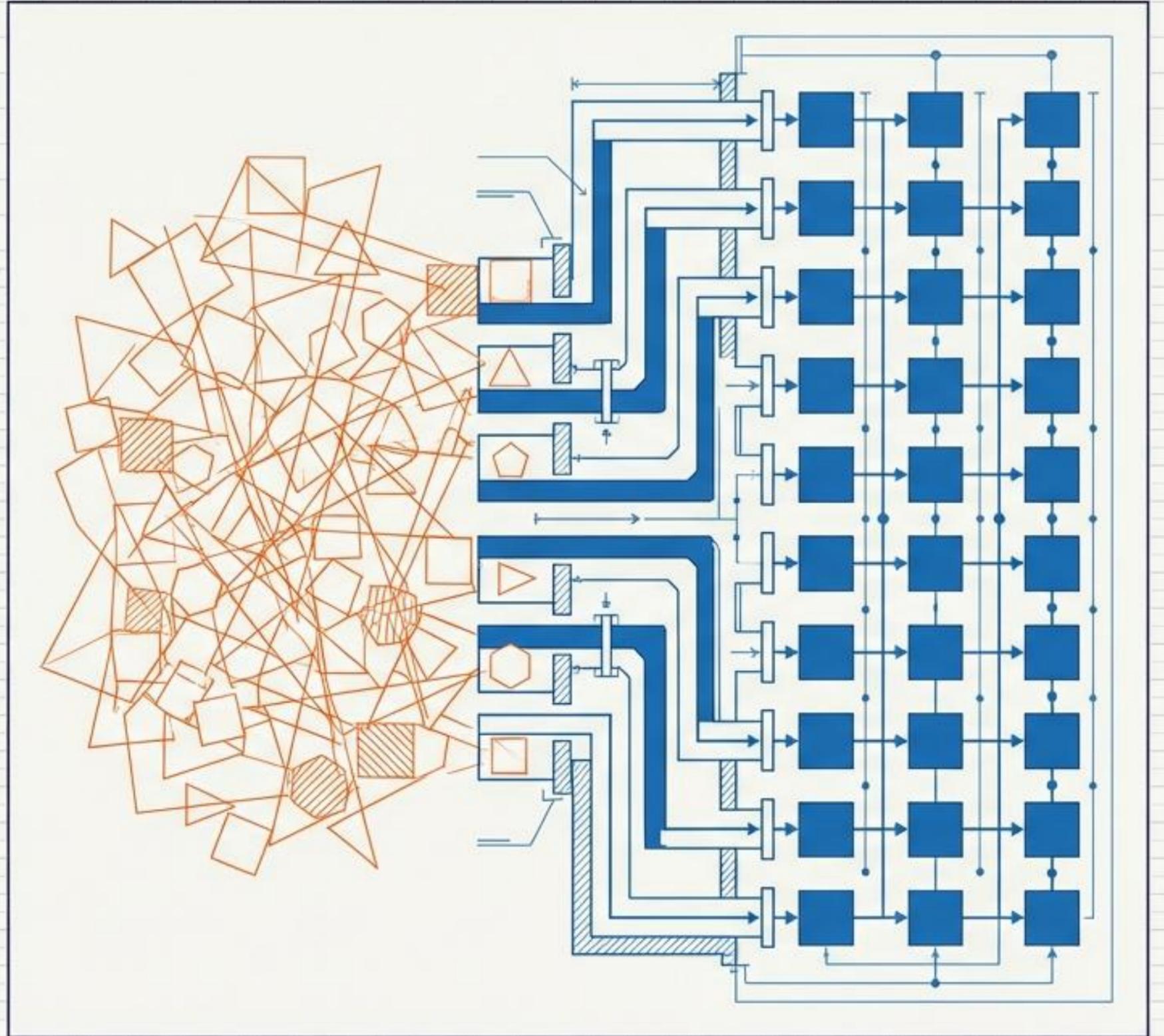
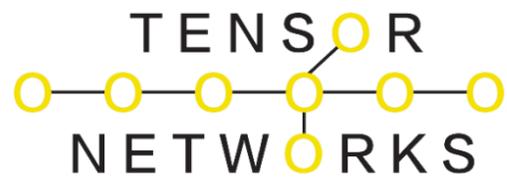


The Predictive Blueprint: Orchestrating the HPC Environment

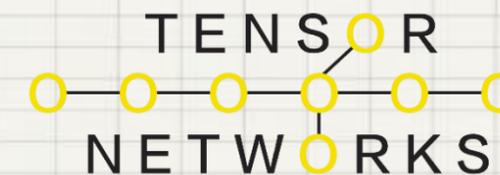
How the Predictive Tensor Control Plane (PTCP) transforms observability into bounded, pre-emptive orchestration for EDA, CFD, and FEA.



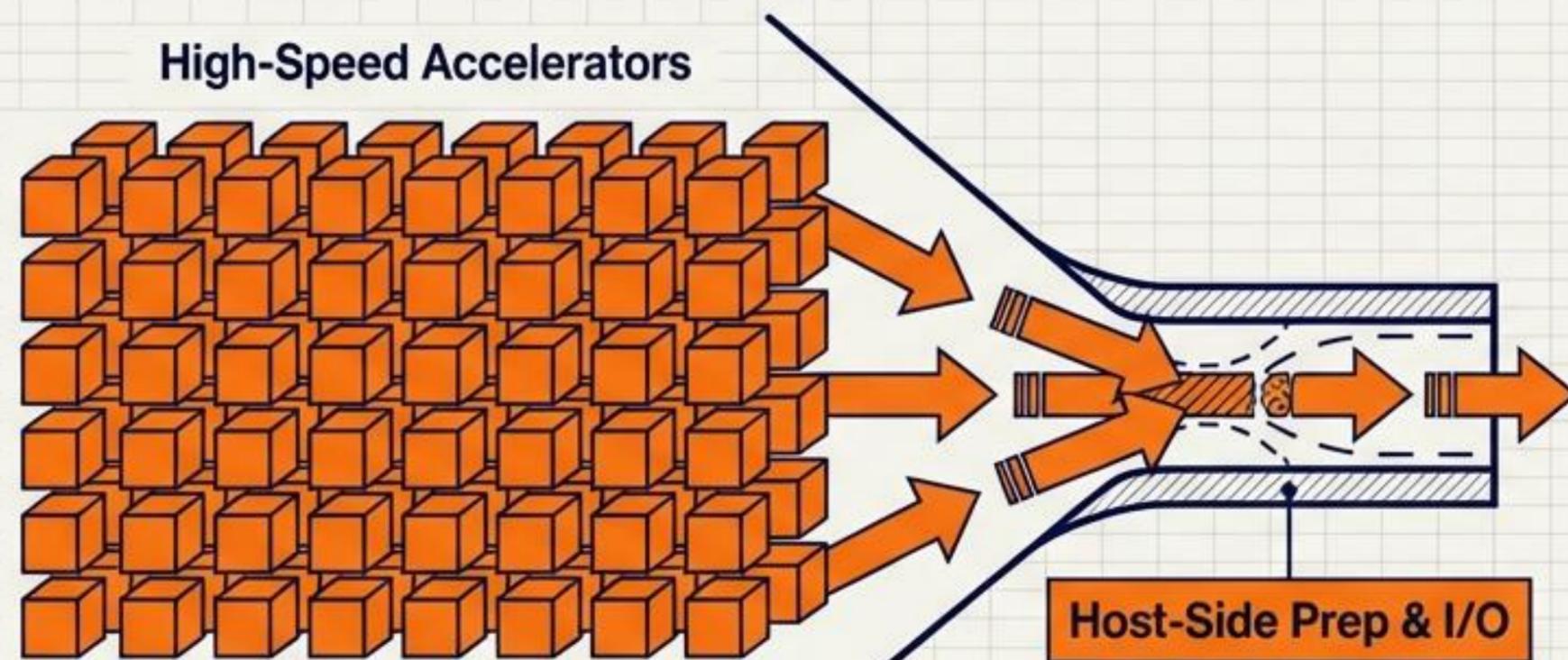


Disclaimer: Forward-Looking Statements This presentation may contain forward-looking statements regarding future product plans, expected features, or anticipated release dates. These statements outline our general product direction and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for our products remain at our sole discretion and are subject to change without notice. These statements do not represent a commitment, promise, or legal obligation to deliver any specific material, code, or functionality

Compute outpaces infrastructure: The host-side data bottleneck



The Core Dilemma



HPC workloads rely heavily on commercial off-the-shelf (COTS) hardware.

As compute accelerators and network fabrics scale, standard systems act reactively.

The Resulting Symptoms

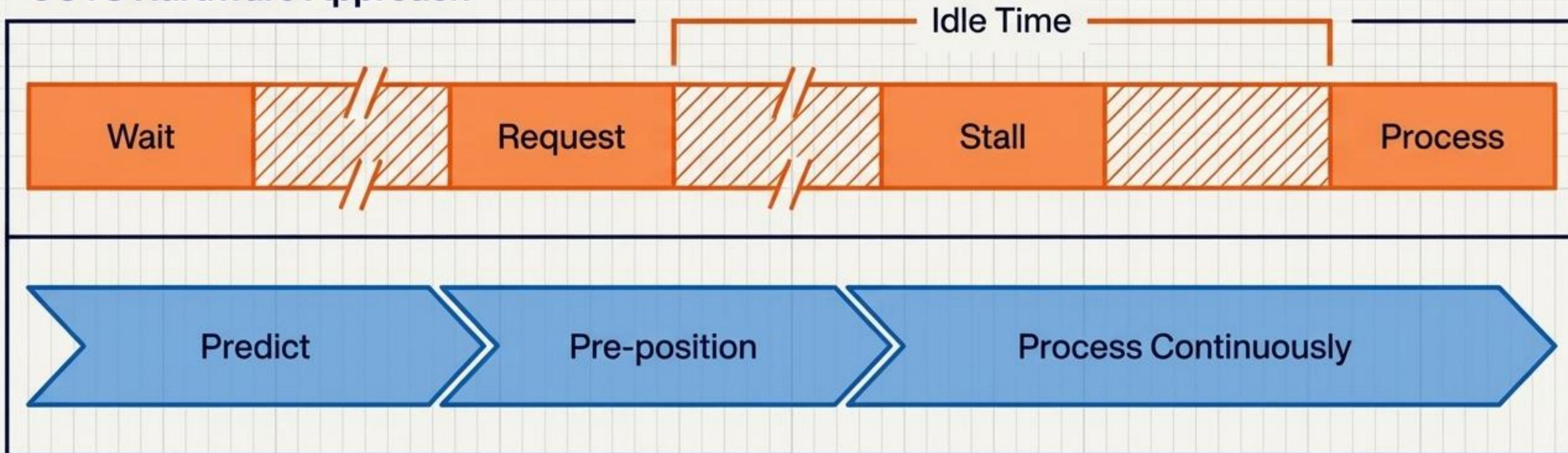
A blue and orange icon representing fragmented data blocks.	State Fragmentation: Data cannot be pre-positioned fast enough.
An icon showing a blue database cylinder, a clock, and arrows, representing heavy checkpointing.	Heavy Checkpointing: Multi-day runs overwhelm standard storage paths.
An icon of a network graph with a large orange arrow pointing through it, representing collective communication bottlenecks.	Collective Communication: All-reduce events create catastrophic network bottlenecks.

The system bottleneck has permanently shifted to host-side data preparation, CPU memory bandwidth, and I/O resources.

Anticipating data demand to eliminate compute stalls

Reactive data movement starves expensive compute cycles while waiting for network traversal or storage loads.

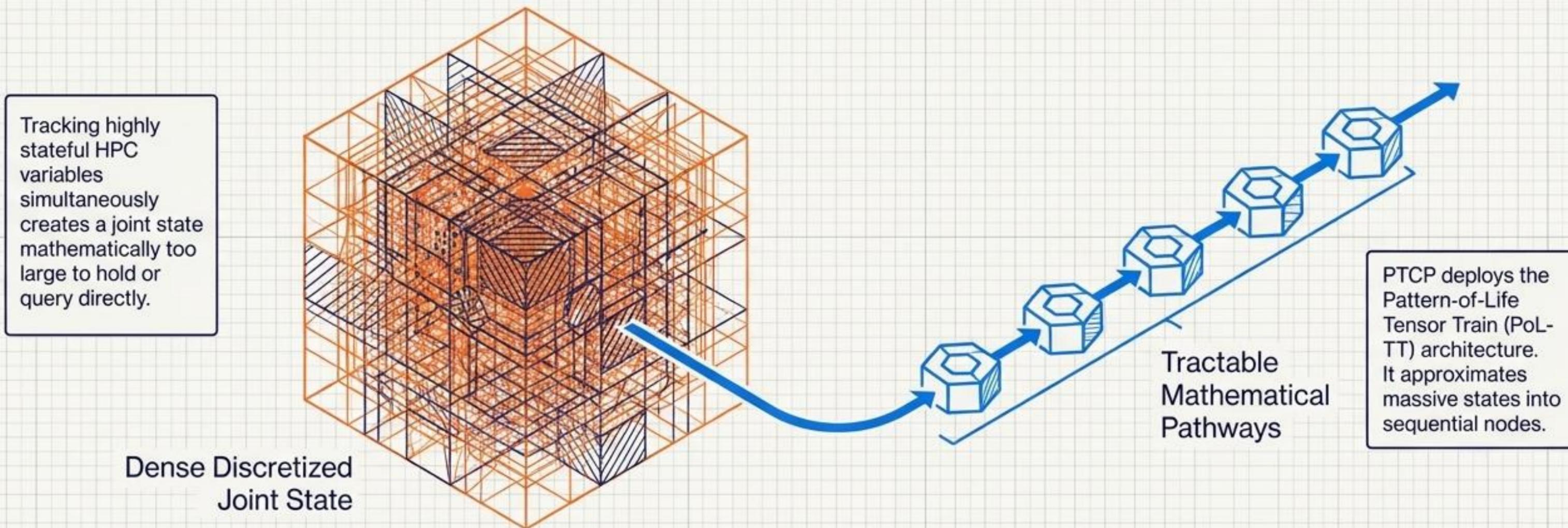
COTS Hardware Approach



PTCP Architecture

The Predictive Tensor Control Plane (PTCP) ensures continuous execution by anticipating demand before requests arrive.

Turning observability into bounded matrix operations



Dense Discretized Joint State

Tractable Mathematical Pathways

$$P[i_1, \dots, i_d] \approx \sum G^{(1)}[1, 1, a_1] G^{(2)}[a_1, i_2, a_2] \dots G^{(d)}[a_{(d-1)}, i_d, 1]$$

The HPC Workload Diagnostic Matrix

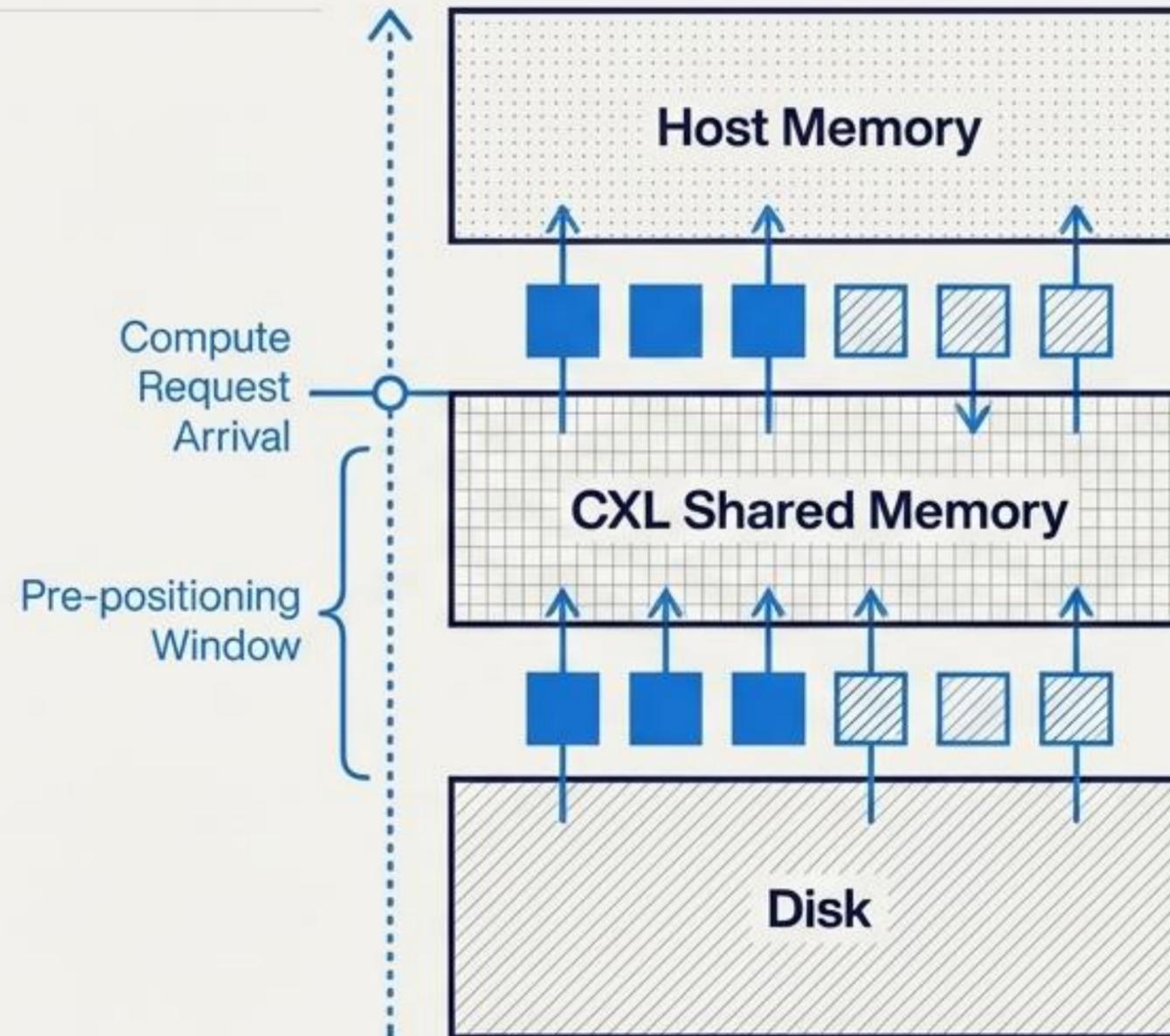
	EDA (Electronic Design Automation)	CFD (Computational Fluid Dynamics)	FEA (Finite Element Analysis)
Core Bottleneck	GT America Standard Multi-TB graph databases exceeding host memory; cache spills forcing heavy I/O.	GT America Standard Massive all-reduce communication events; frequent transient state checkpointing.	GT America Standard Sparse linear equation processing; multi-day non-linear runs bottlenecking PCIe.
PTCP Intervention Mechanism	Predictive Memory Tiering & Bounded Cache Retention.	GT America Standard All-Reduce Storm Mitigation (Rack Agents) & Cross-Layer Checkpoint Coordination.	GT America Standard Joint State Vector Querying & Compute-Near-Data Execution.
System Impact	GT America Standard Pre-positions hot data into CXL shared memory.	GT America Standard Pre-emptively paces network traffic; buffers checkpoints during storms.	GT America Standard Eliminates compute stalls and avoids dropped packets entirely.

EDA: Eliminating I/O burdens for massively stateful workloads

Context: Physical design, parasitic extraction, and timing analysis.

Predictive Memory Tiering

- PTCP executes a probabilistic tiering policy via its TT-compressed behavioral model.
- Forecasts data reuse and pre-positions hot segments into CXL shared memory, preventing CPU/GPU starvation from disk reads.



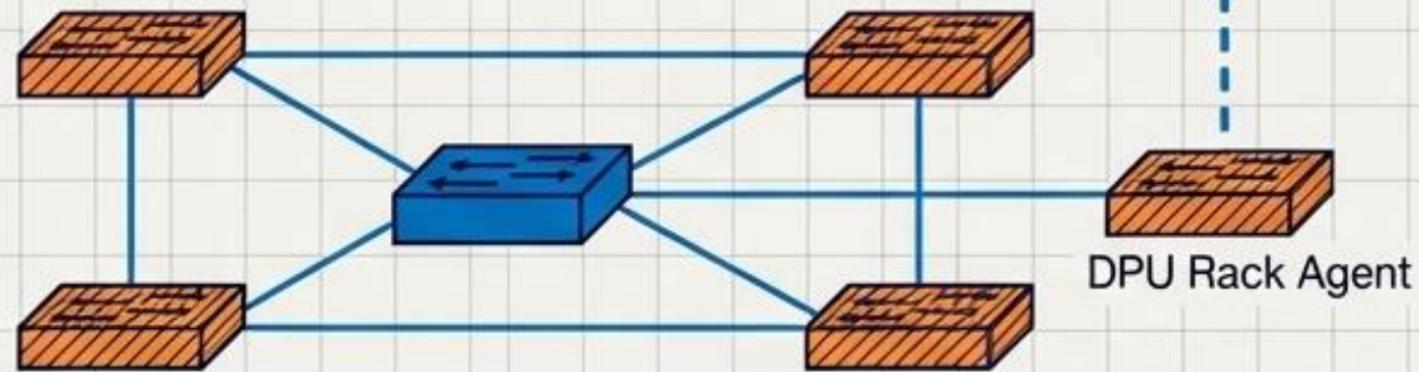
Bounded Cache Retention

- Manages bursty state access via strict cache envelopes.
- Manages bursty state access via strict cache retention envelopes.
- Utilizes precise prefetch caps to guarantee mispredictions never overwrite currently active data.

CFD: Mitigating all-reduce storms and orchestrating checkpoints

Context: Navier-Stokes equations and transient aerodynamics analysis.

Networking Topology



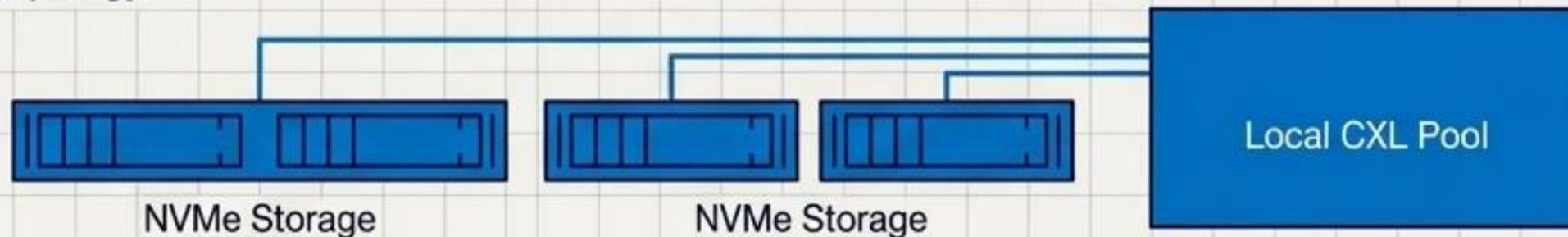
Fabric Manager

Congested

Cross-Layer Coordination

All-Reduce Storm Mitigation: DPU Rack Agents evaluate the PoL model to forecast critical-path collective communication bottlenecks. Emits bounded hints for pre-emptive pacing, replacing solely reactive explicit congestion notification (ECN).

Storage Topology

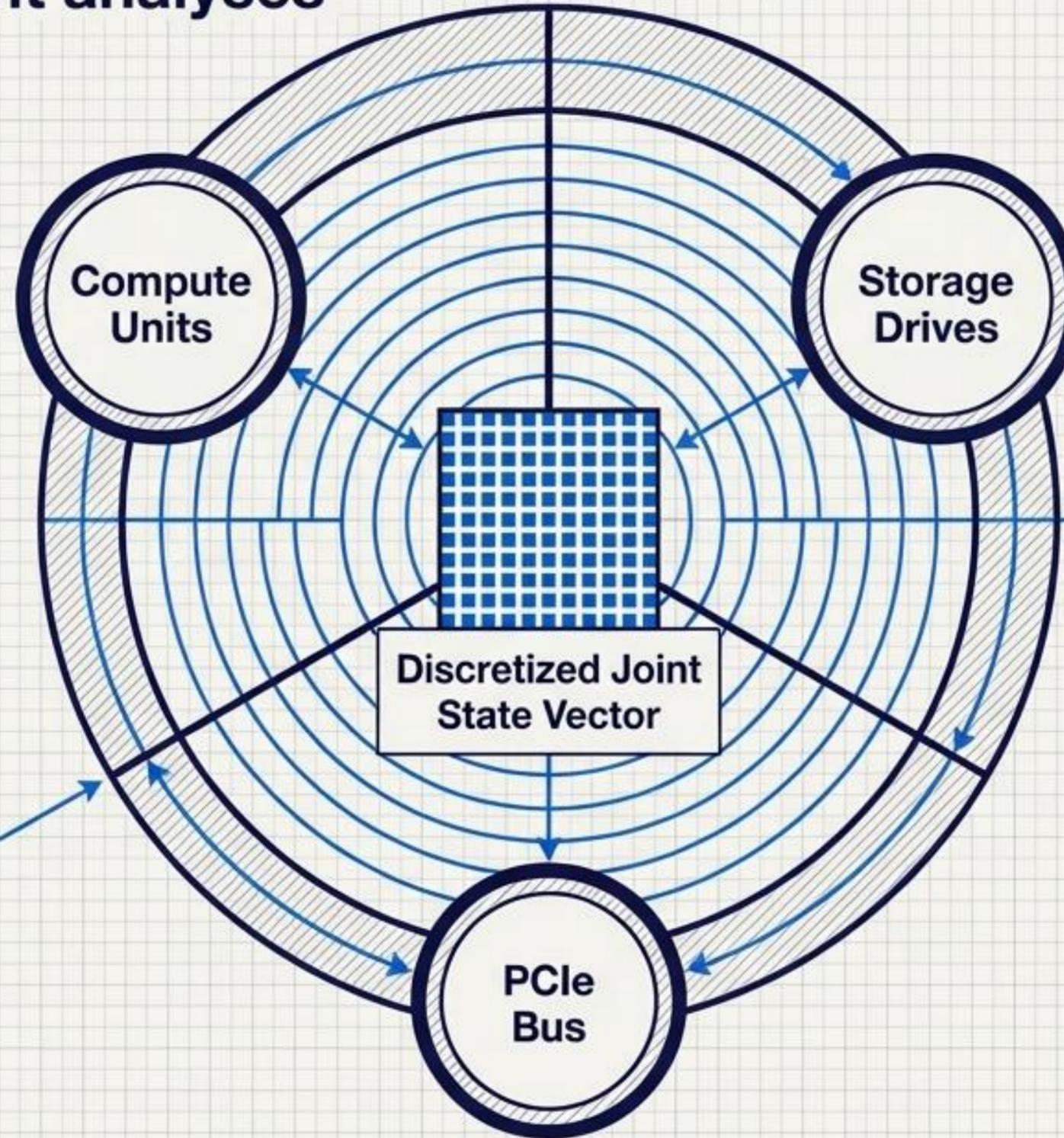


Buffered/
Managed

Cross-Layer Checkpoint Coordination: Storage and network views converge. If the network is congested, the local NVMe agent queries the model and temporarily buffers simulation checkpoints in CXL/DRAM until the collective communication window closes.

FEA: Sustaining bandwidth for massive massive finite element analyses

Context: Sparse linear equations for crash simulations and structural integrity.



Compute-Near-Data Execution

Massive FEA output data triggers intelligent storage agents.

Executes local preprocessing and compression before writing to flash media.

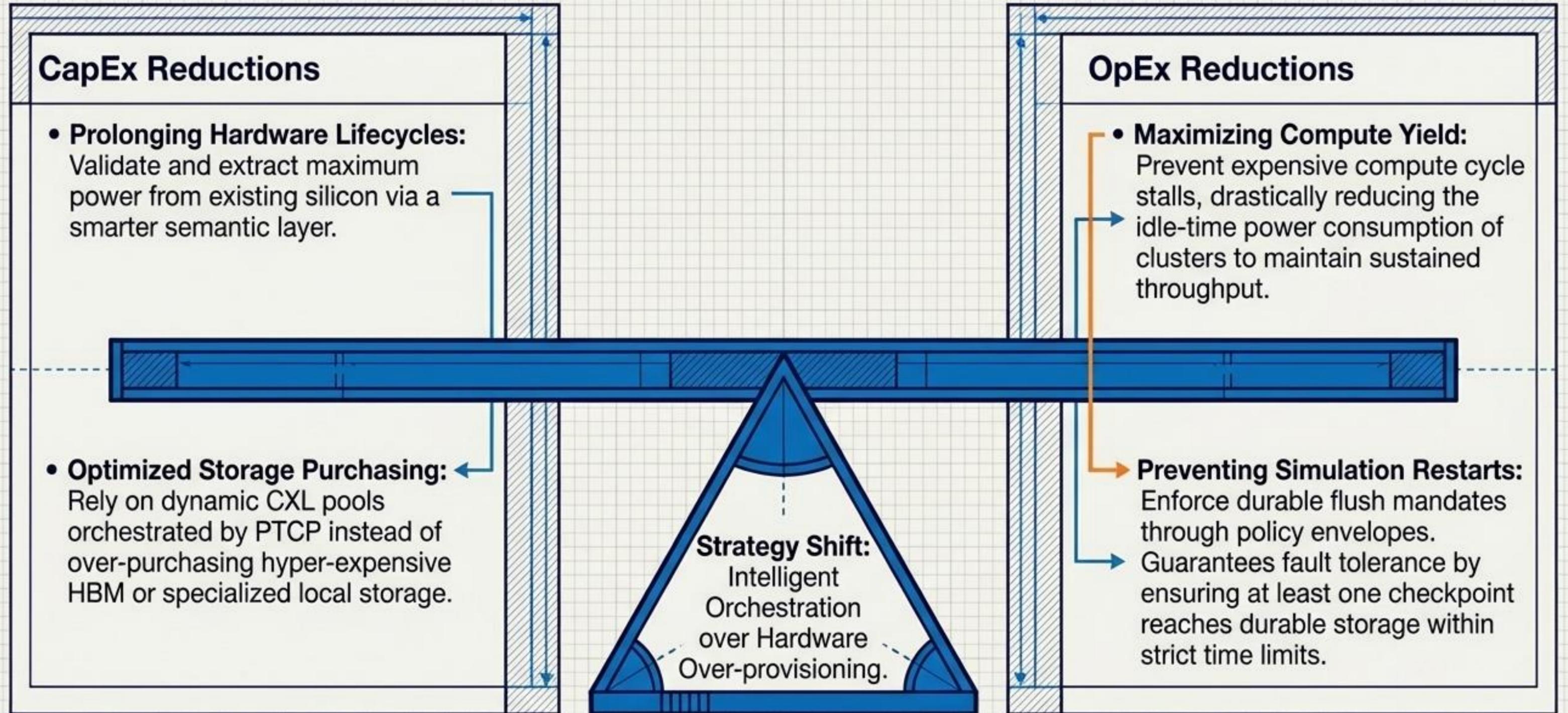
Eliminating PCIe Stalls

By querying the exact same state vector, components operate in lockstep. Network avoids dropping packets.

Compute units avoid I/O wait times.

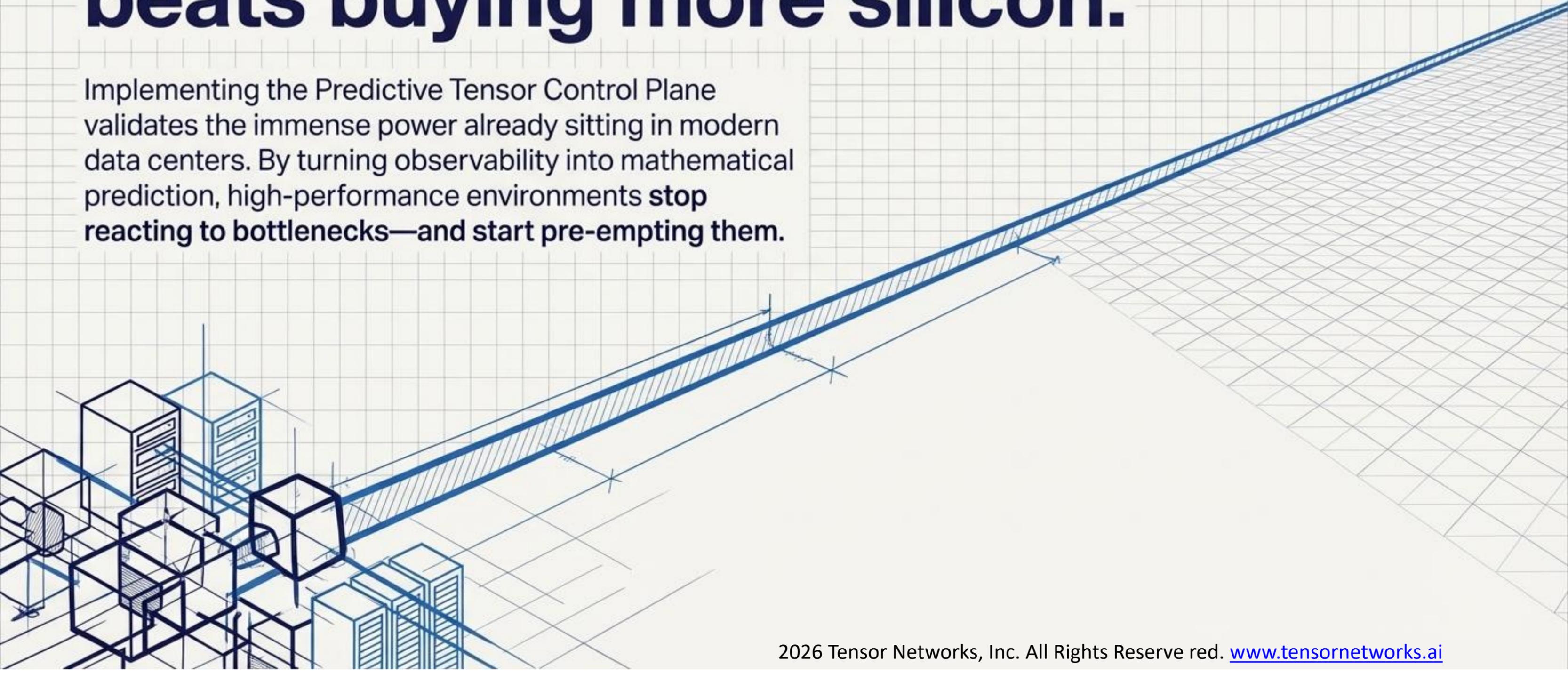
Storage drives bypass PCIe bus bottlenecks entirely.

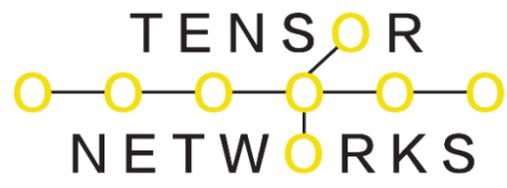
The Economic Balance Sheet: Transforming the HPC financial profile



Smarter orchestration beats buying more silicon.

Implementing the Predictive Tensor Control Plane validates the immense power already sitting in modern data centers. By turning observability into mathematical prediction, high-performance environments **stop reacting to bottlenecks—and start pre-empting them.**





Thank You for Your Partnership We look forward to building the future of networking together. Please don't hesitate to reach out with any questions.

Tensor Networks, Inc. Contact Us:
information@tensornetworks.net