Cambodia's AI Future: Building the Network Infrastructure to Power Innovation

OLARN ONGBORIRUKKUL

30 November 2025

Agenda

01	Cambodia's AI Readiness & Why Infrastructure Matters NOW
02	AI Data Center Architecture

Cambodia's AI Readiness

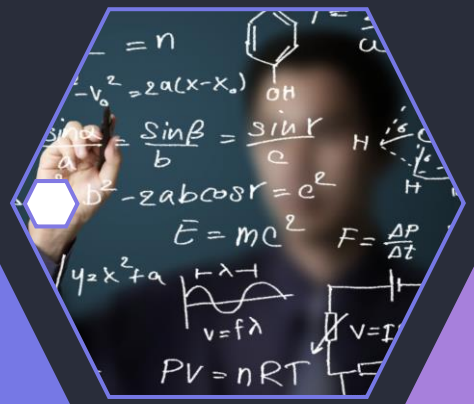
Why Infrastructure Matters NOW

National’s Milestone

In July 2025, Cambodia became the fourth country in South-East Asia to complete UNESCO's AI Readiness Assessment

Challenge

Scientific & Educational



Technical & Infrastructural

Critical Gab

Economic progress

Social & Cultural

Challenge

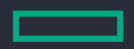
Legal & Regulatory progress

https://www.unesco.org/en/articles/cambodia-launches-ai-readiness-assessment-report-guide-ethical-and-inclusive-digital-transformation

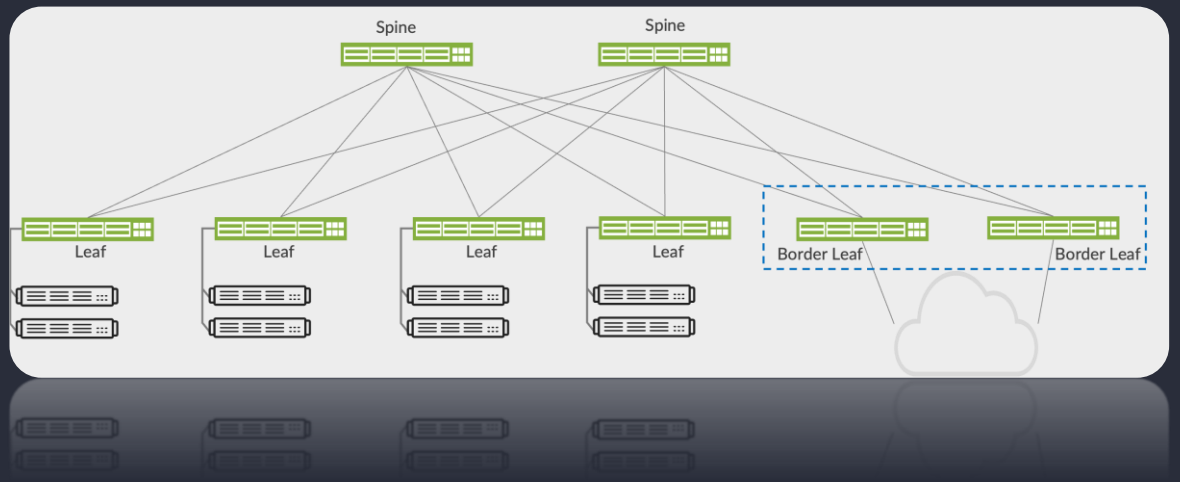
Challenge!

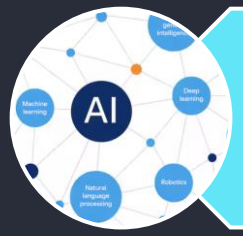
- High Cost cloud service and limitted scale
- Latency preventing realtime AI
- Power sufficiency
- Training workload locally

AI Data Center Architecture



Why AI Data Centers are Emerging





AI Workloads have changed traffic patterns and infrastructure requirements.

Traditional DCs: North-South, transaction workloads.

AI DCs: Massive, Sustained EAST-WEST GPU<>GPU flows that demand ultra-low latency, near-zero packet loss and predictable throughput

Compute Density (GPU Server) drive new network, power and cooling designs.

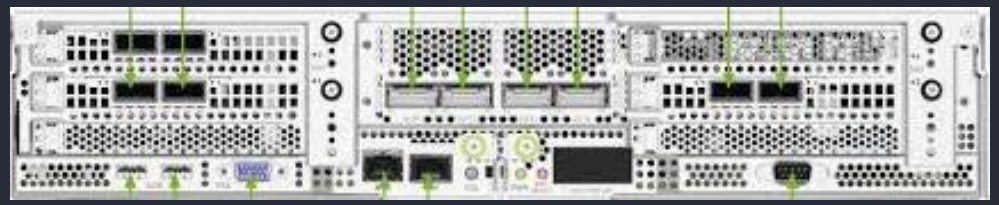
AI servers and training model requirements

General Purpose Server

2x25G still the most common configuration we sell

50G per server

NVIDIA DGX H100 Server



8x400G Backend GPU Network

2x100G Frontend Network

4x200G Storage Network

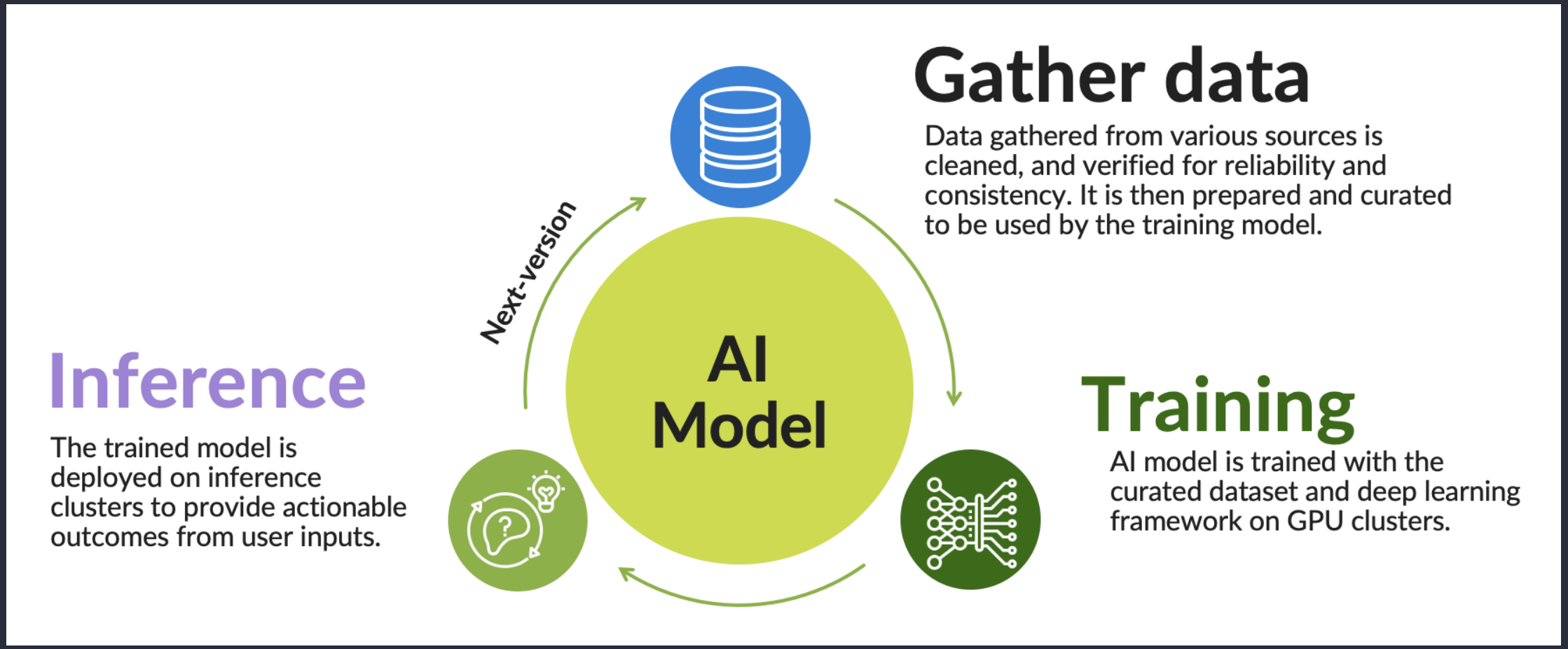
~4Tbps per server

➔ explosive growth in networking requirements

AI Data Center Architecture

Lifecycle of an AI Model

The AI/ML app lifecycle can be a continuous, iterative process of designing and developing models, training and validating them with curated data, and deploying them into production while monitoring their performance for constant refinement and improvement.



Anatomy of an AI/ML Cluster & Network

Cluster Components

Training Cluster

Inference Cluster

Shared Storage Pools

Dedicated Cluster Storage

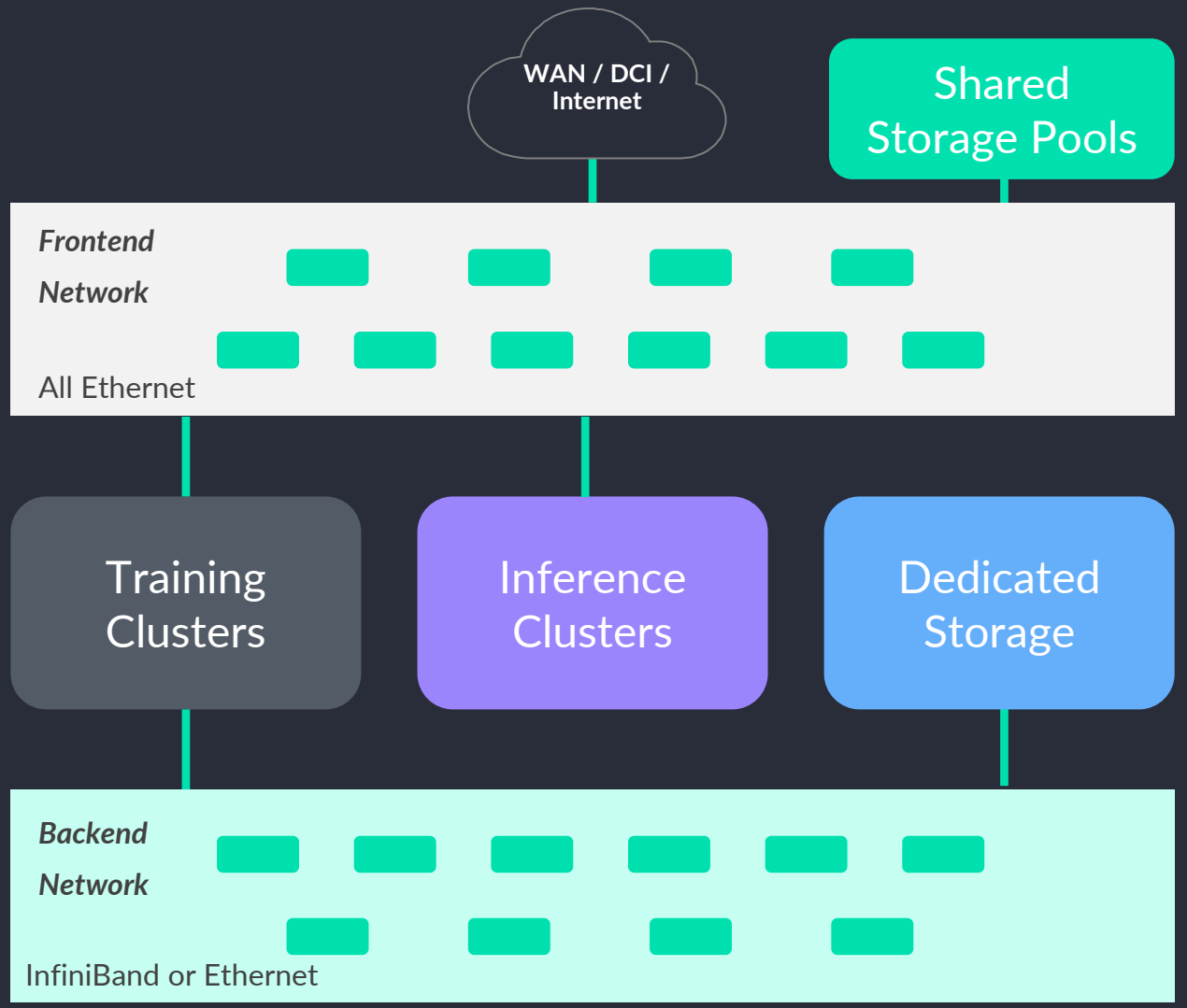
Cluster Networks

“Frontend”

- Inference clusters use this network
- Shared storage pools
- Management network for training

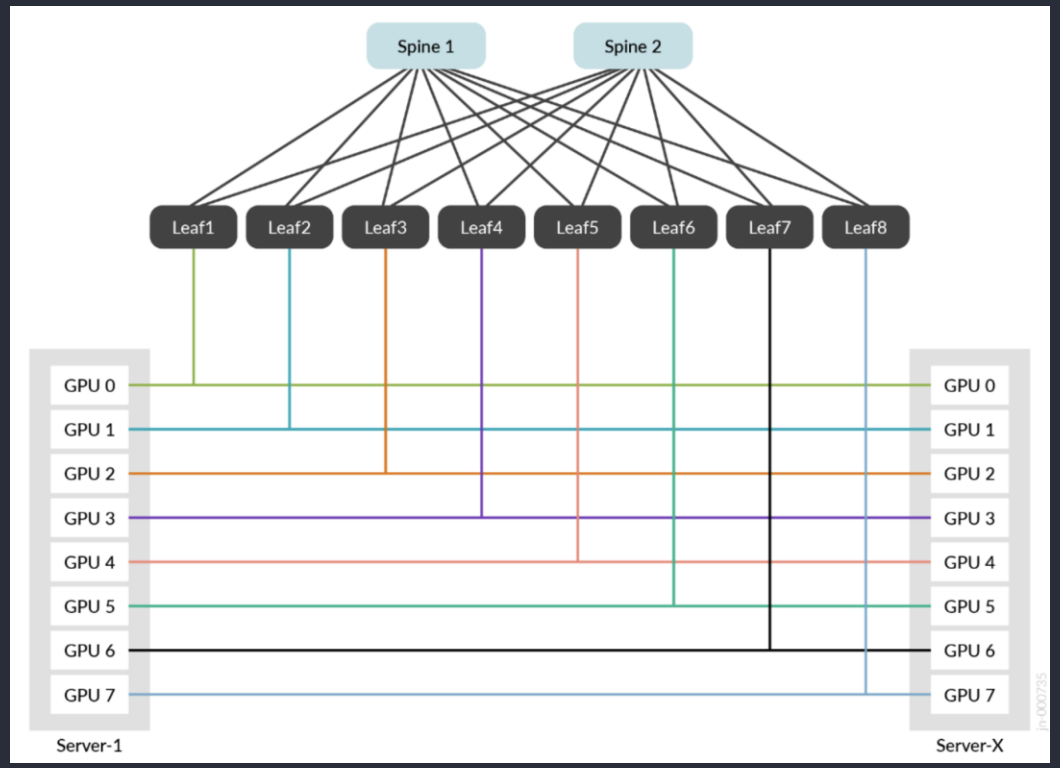
“Backend”

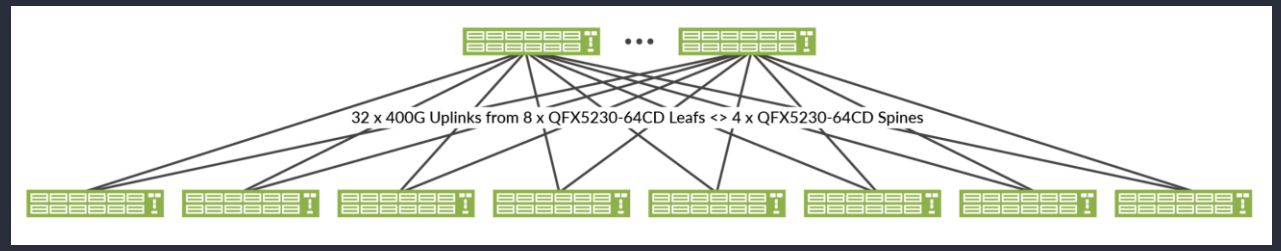
- GPU Compute Fabric
- Dedicated Storage Fabric (may be converged with compute)

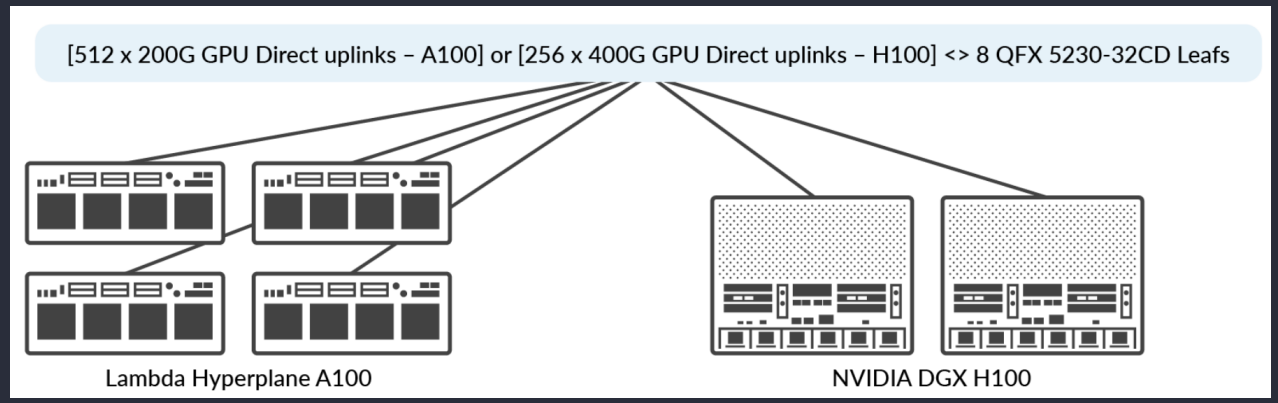


AI Data Center Architecture

AI DC Architectures







Rail Optimized Stripe

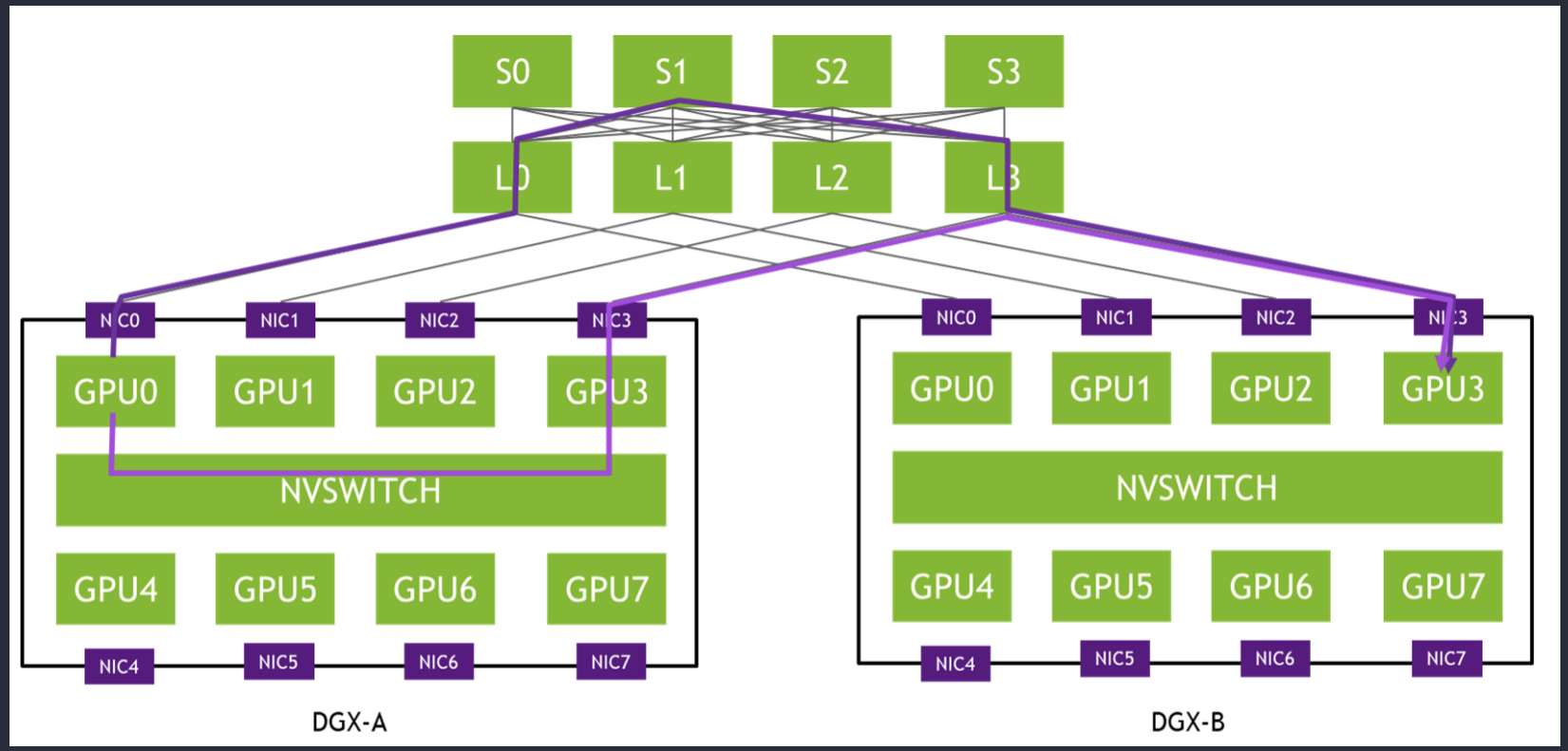
To provide maximum throughput,
Minimal latency
Minimal network interference for AI traffic flows

1 Stipe ->

- 8 Leafs (64 x 400G) + 4 Spines (64 x 400G)
- 512 A100 GPU (64 Servers) or 256 H100 GPU (32 Servers)
- Ratio 1:1

NCCL

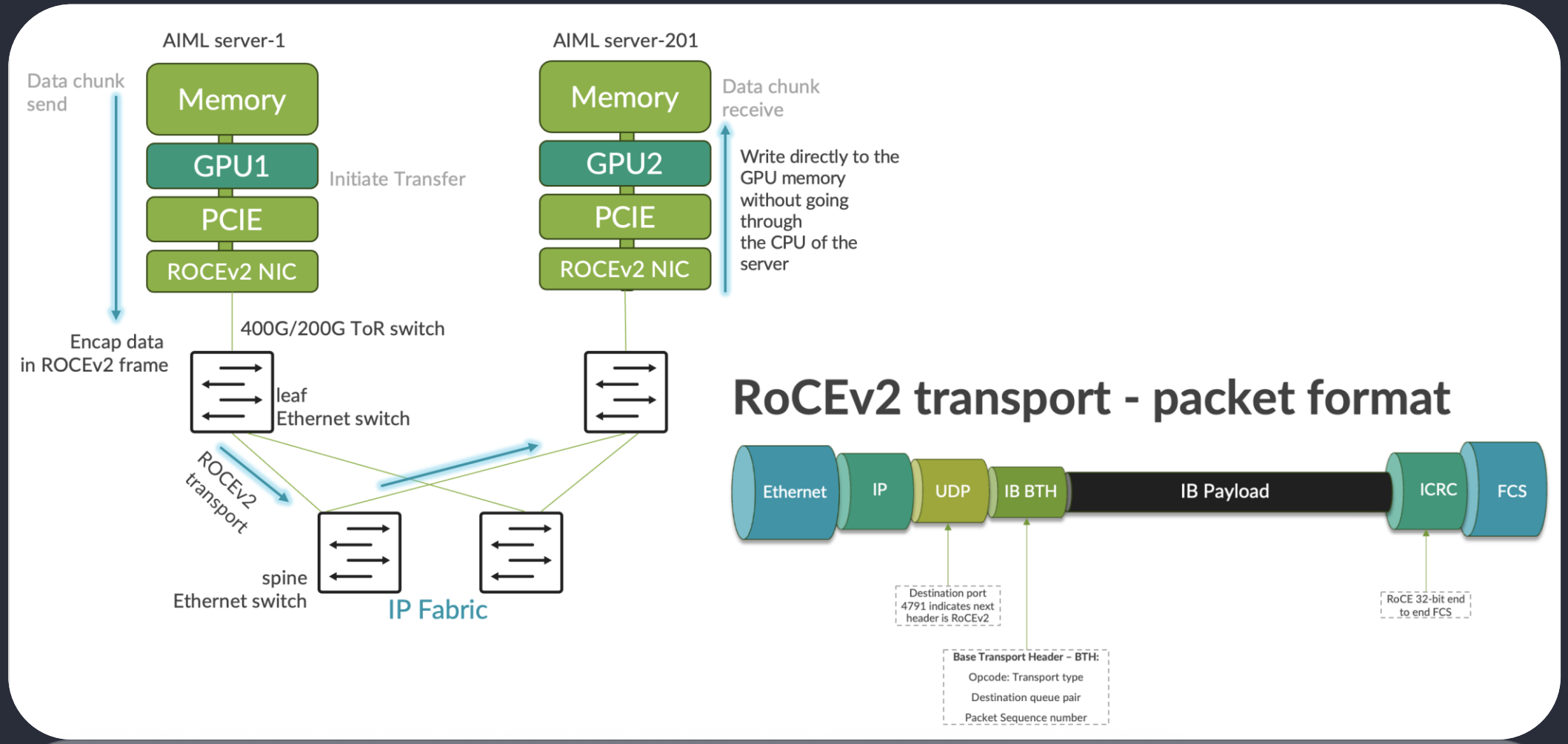
- NCCL (NVIDIA Collective Communications Library) acts as communication orchestrator for multi-GPU multi-node environments
- Topology-aware and is optimized to achieve high bandwidth and low latency.
- The new feature introduced in NCCL 2.12 is called PXN, as PCI × NVLink, as it enables a GPU to communicate with an NIC on the node through NVLink and then PCI. over PCIe, NVLink,



Source: https://developer.nvidia.com/blog/doubling-all2all-performance-with-nvidia-collective-communication-library-2-12/#:%7E:text=Rail%2Doptimized%20network%20topology%20helps,having%20lighter%20connections%20between%20rails

AI Data Center Architecture

ROCEv2 Transport in AI DC



AI Data Center Architecture

AI DC Features

Proactive

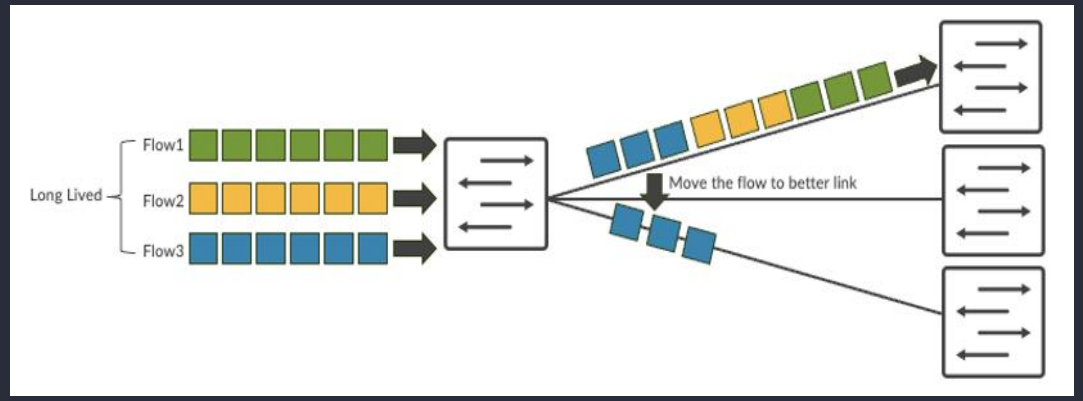
Efficient Load Balance

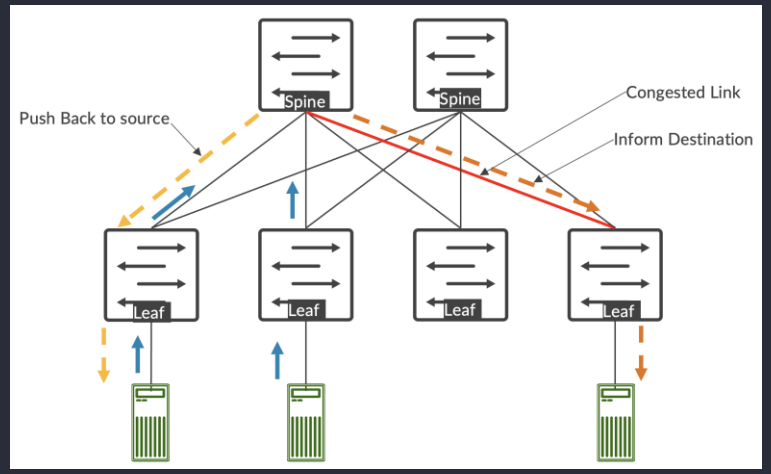
- DLB (Dynamic Load Balance)
- GLB (Global Load Balance)

Reactive

Congestion Management

- DCQCN (PFC + ECN)

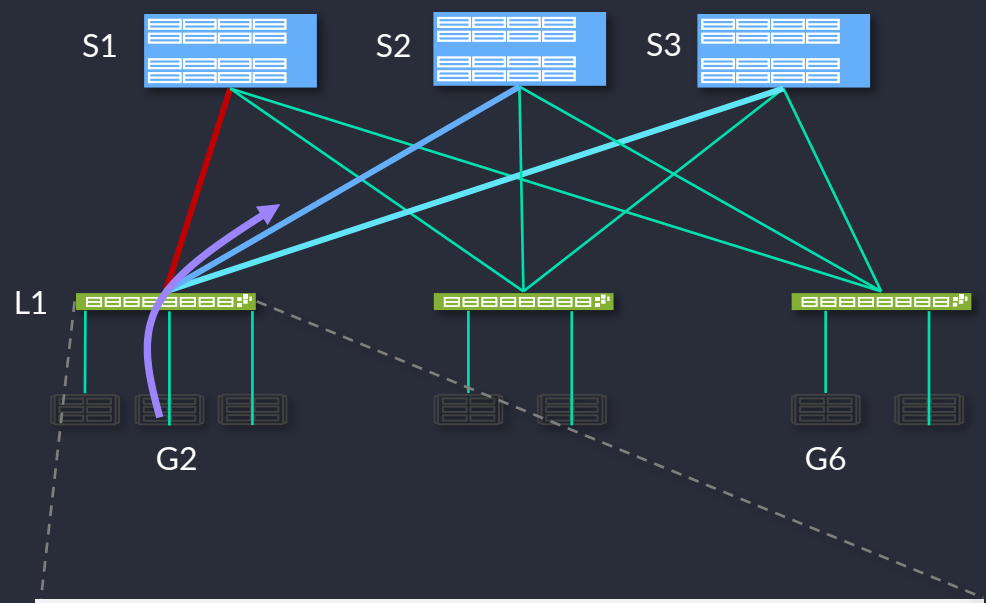




Dynamic Load Balancing (DLB)

- DLB monitors link quality on local ECMP paths based on real-time link utilization and queue depths of each link.
- DLB factors local link quality and forwards on the least loaded path.
- Minimizes flow collisions and the consequent packet drop.

Dynamic Load Balancing (Link quality based)



Forwarding a new flow on L1 (G2 -> G6)

Path to	Via (ECMP)	Local link quality (DLB)*	Path selection
G6	S1	Low	☐
	S2	High	☑
	S3	Medium	☐

* Colors & High, med, and low are used for simplification. Actual link quality is a numerical metric.

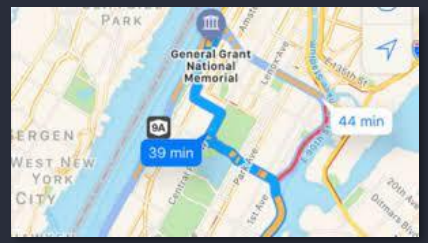
Global Load Balancing (GLB)

- Spines monitor local link quality to all leaves (DLB) and advertise these to all upstream leaves
- Leaves use received link quality from spines along with local link quality (DLB) for load balancing
- Upstream switches avoid downstream congested paths and select a better end-to-end path
- Like Google Maps traffic condition-based path selection

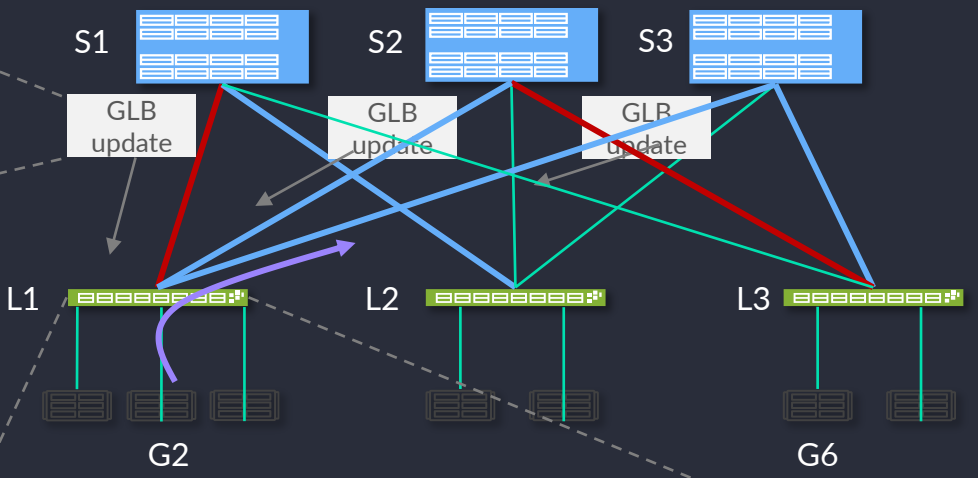
GLB from S1	
To L2	High
To L3	High

GLB from S2	
To L2	High
To L3	Low

GLB from S3	
To L2	High
To L3	High



Like google maps traffic



Forwarding a new flow on L1 (G2 -> G6)

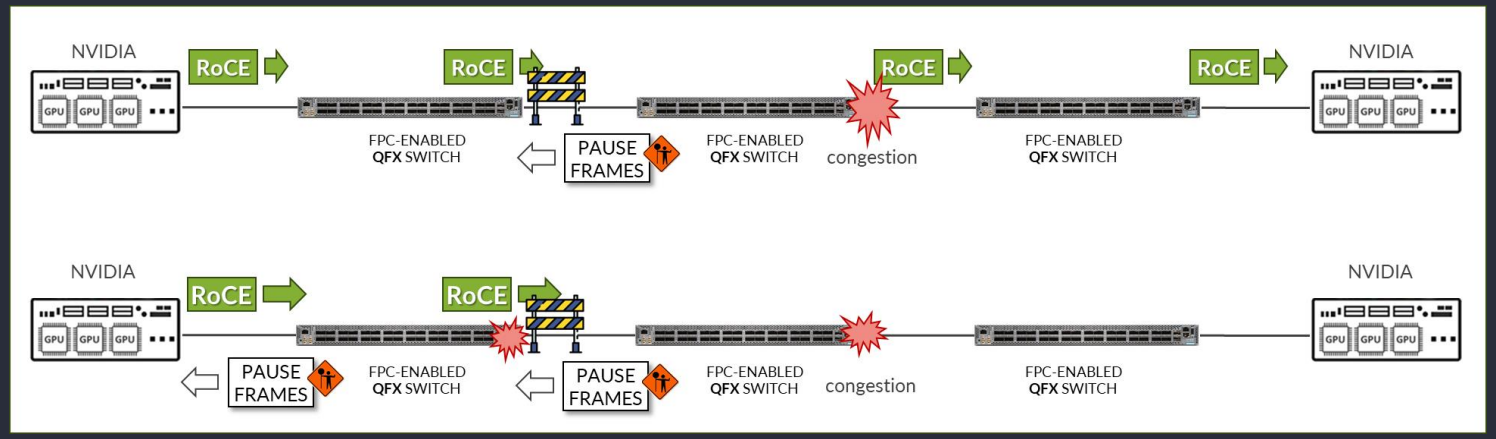
Path to	Via (ECMP)	Local LQ (DLB)*	Global LQ (GLB)*	Path selection
G6	S1	Low	High	☐
	S2	High	Low	☐
	S3	High	High	☑

* Colors & high, med, and low are used for simplification. Actual link quality is a numerical metric.

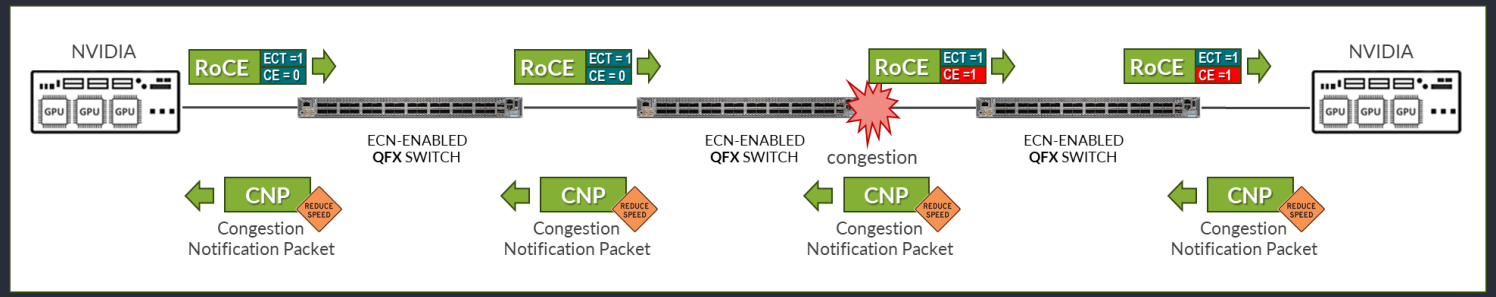
Congestion Control – lossless fabrics

(DCQCN : PFC + ECN) in the IP Fabrics

PRIORITY-BASED FLOW CONTROL (PFC)



EXPLICIT CONGESTION NOTIFICATION (ECN)



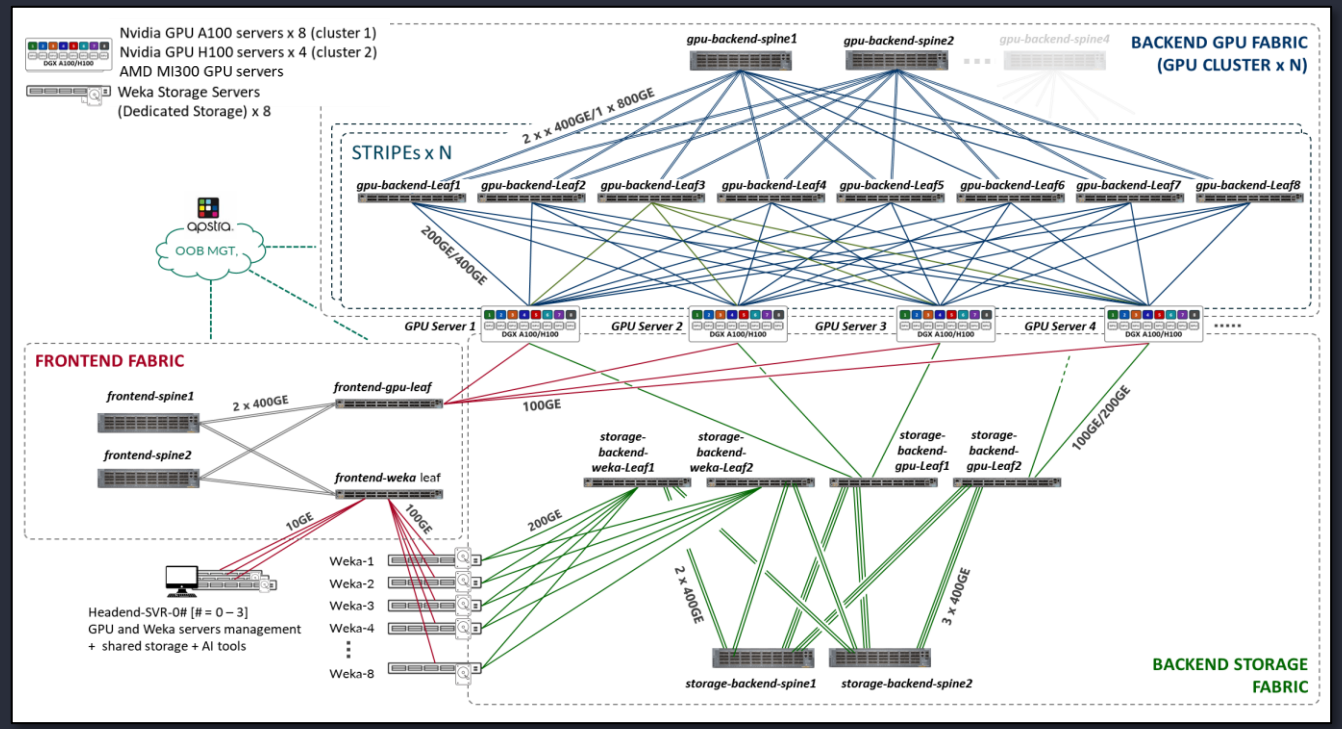
Congestion Control aims to deliver **lossless Ethernet**.

DC QUANTIZED CONGESTION NOTIFICATION **(DCQCN)**

Juniper AI DC

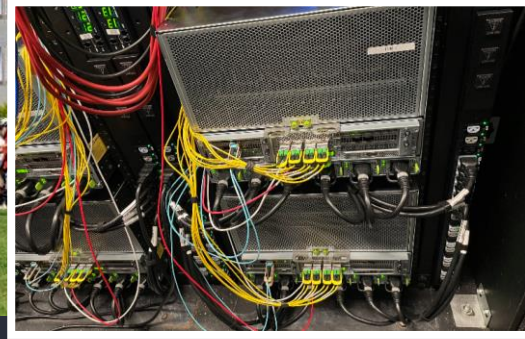
Innovation Lab

AI JVD



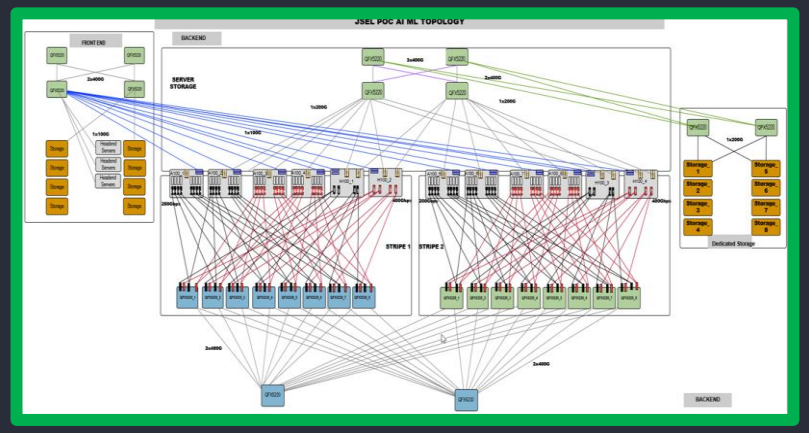
- Comprehensive, end-to-end template for deploying AI networking solutions.
- Rigorously pre-tested, validated.
- Reduce complexity and risk.
- Tee shirt sized for small, medium, and large AI DCs.

Juniper AI Lab



AI-optimized Ethernet: Put to test

We achieve comparable results with InfiniBand in MLPerf benchmarks



- *Ran Bert-Large, DLRM, Llama2, and GPT-3 training models (MLCommons)*
- *Able to run customer models*

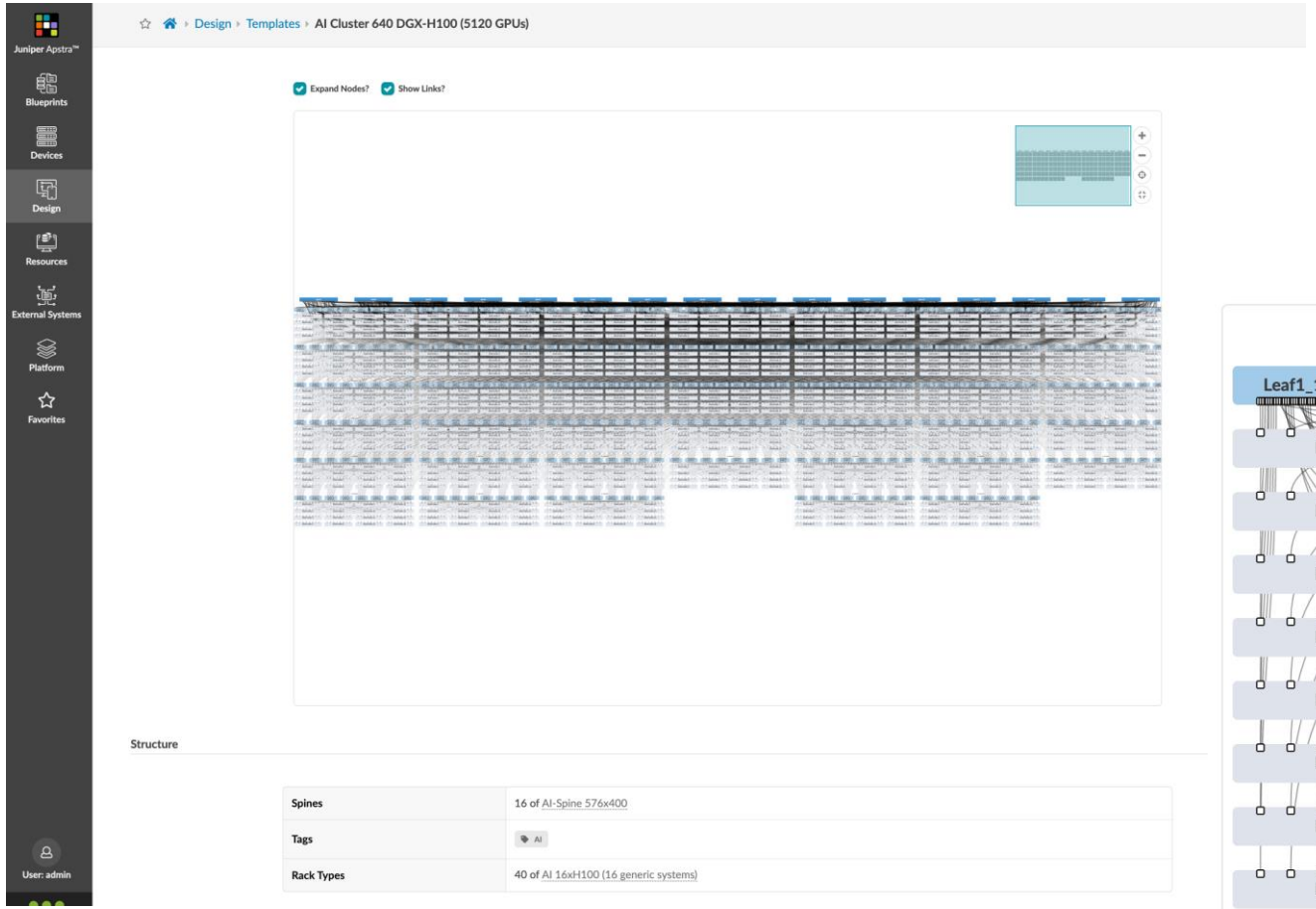
	Juniper Ethernet Training Time	InfiniBand Training Time
BERT-Large (Training)	~2.6 min (64 A100 GPU)	2.5-3.3 min (64 A100 GPU)

Technology choices for GPU backend fabric

	InfiniBand	AI Optimized Ethernet
Vendor Lock-In	Yes	No
Easy to Operate	No	Yes
Cost	Higher	Lower
Performance	Yes	Yes

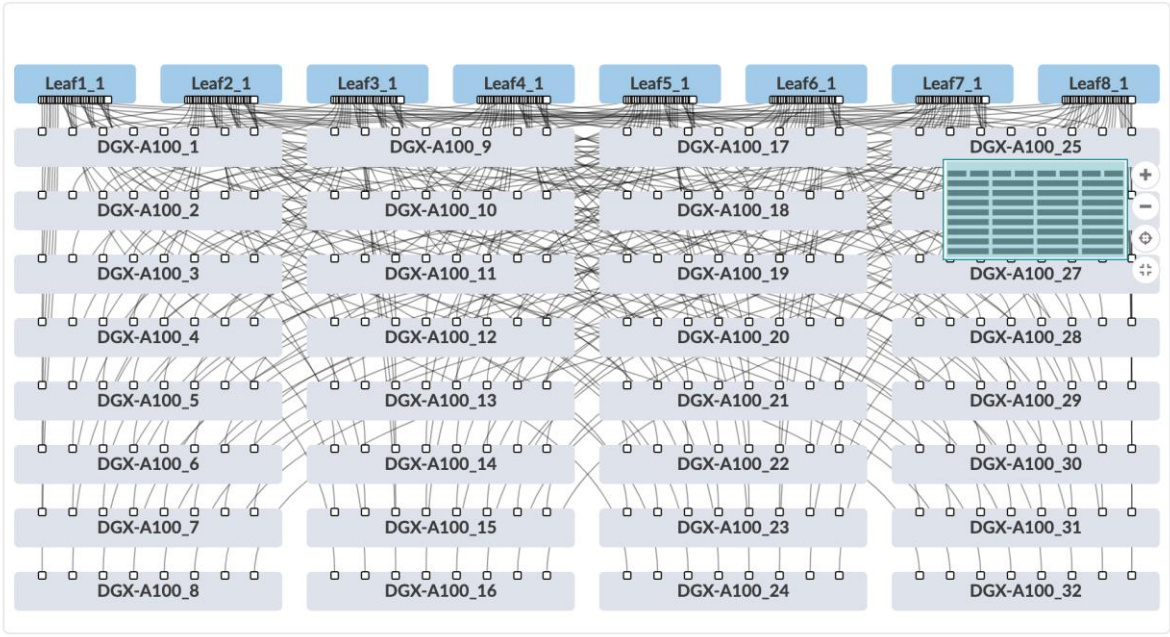
Juniper Apstra™

Design and Deploy AI Clusters with Apstra



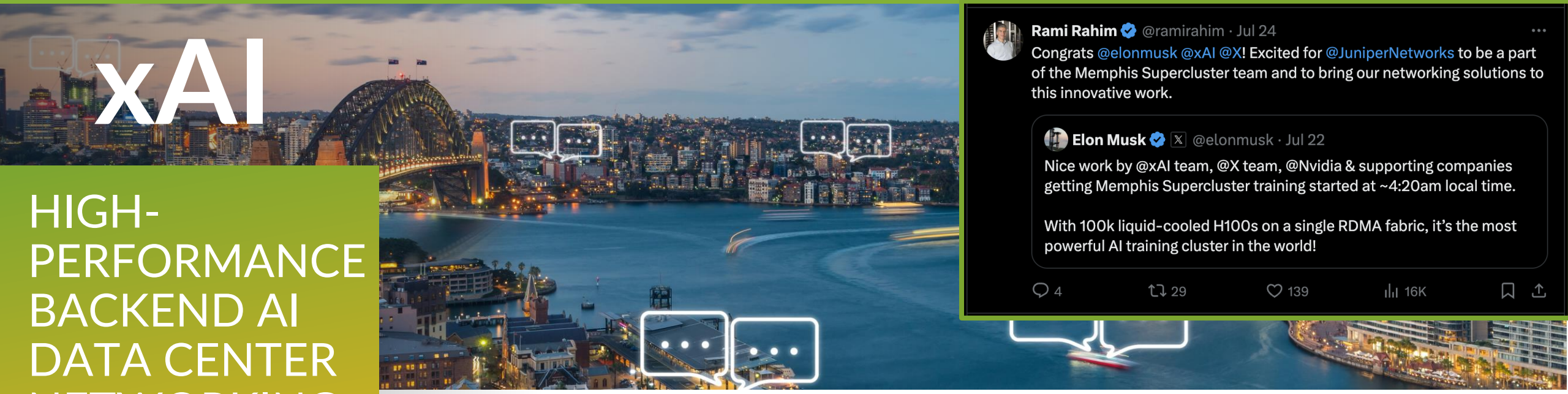
Example of 5120 GPU cluster

- Rail-optimized design accommodated with custom Rack of 8-way leafs and 8-interface servers
- Juniper guided design with Terraform automation of all fabrics and necessary configs



Example of rail-optimized GPU group or stripe

xAI



HIGH-PERFORMANCE BACKEND AI DATA CENTER NETWORKING

THE CHALLENGE

- Accommodate large-scale and efficient GPU cluster and server connectivity.
- AI training is compute-intensive , with traffic flows that break traditional data center networking
- Efficient use of GPU cycles. (Job completion time)

THE SOLUTION

- Juniper QFX5240 for 800G leaf spine connectivity
- Enabling the same form factor for both Front and Back-end AI clusters
- Efficient Load balancing and congestion control for RDMA traffic

WHY JUNIPER?

- **Junos feature richness** for congestion management with ROCEv2
- **AI Lab for POC** for various load balancing scenarios
- **Lightening responsiveness to needs**
- **Better supply chain** for switch and optics with Juniper

JUNIPER NETWORKS

Summary and Key Takeaways

1. AI has a very clear future
2. Network Infrastructure is the Critical part
3. AI DC Architecture with DC Switch port folio and features rich

1. Rich DC portfolio:
   - 400G/800G/1.6T Ethernet/IP
   - Apstra Fabric Manager
2. Advanced Software Features (Load Balancing, DCQCN, EVPN, ...)
3. JVD – AI DC & ENT DC

Simplified & validated designs

+

Scale & Performance

+

Automated DC operations

Thank You