AI Networking

Empower AI Workload with Secure Cisco Datacenter Networking

Thomas Torlen

TSA Dataceter

hpalm@cisco.com

Deployment and scaling AI with AI PODs & more

AI POD

NVIDIA | AMD | NeMo | NIMs | Blueprints

Compute
Cisco UCS® w/ NVIDIA HGX or RTX PRO 6000 Blackwell (C885A M8 and C845A M8 respectively)
NVIDIA BlueField®-3 DPUs
Managed with Cisco Intersight®

Networking
Cisco Nexus 9300 w/ Silicon One®
Switches w/ NVIDIA Spectrum-X*
Managed w/ Cisco Nexus Dashboard

Storage
NetApp | Pure Storage | VAST Data

Atomic unit of Secure AI Factory

Scale the Secure AI Factory to meet your business needs

Cisco Secure AI Factory

NVIDIA | AMD

Compute

Storage

Networking
Cisco Nexus 9300 w/ Silicon One®
Switches w/ NVIDIA Spectrum-X*
Managed w/ Cisco Nexus Dashboard

AI POD | AI POD | AI POD

Security
Cisco AI Defense
Cisco Hybrid Mesh Firewall
Isovalent
Secure Firewall
Hypershield

Observability
IT Service Intelligence
App-Dynamics
O11y
Splunk Platform
Unified Data Platform

Deployment and scaling AI with AI PODs & more

AI POD

NVIDIA AMD NeMo NIMs Blueprints

Compute
Cisco UCS® w/ NVIDIA HGX or RTX PRO 6000 Blackwell (C885A M8 and C845A M8 respectively)
NVIDIA BlueField®-3 DPUs
Managed with Cisco Intersight®

Networking
Cisco Nexus 9300 w/ Silicon One® Switches w/ NVIDIA Spectrum-X*
Managed w/ Cisco Nexus Dashboard

Storage
NetApp | Pure Storage | VAST Data

Atomic unit of Secure AI Factory

Scale the Secure AI Factory to meet your business needs

Cisco Secure AI Factory

NVIDIA AMD

Compute

Storage

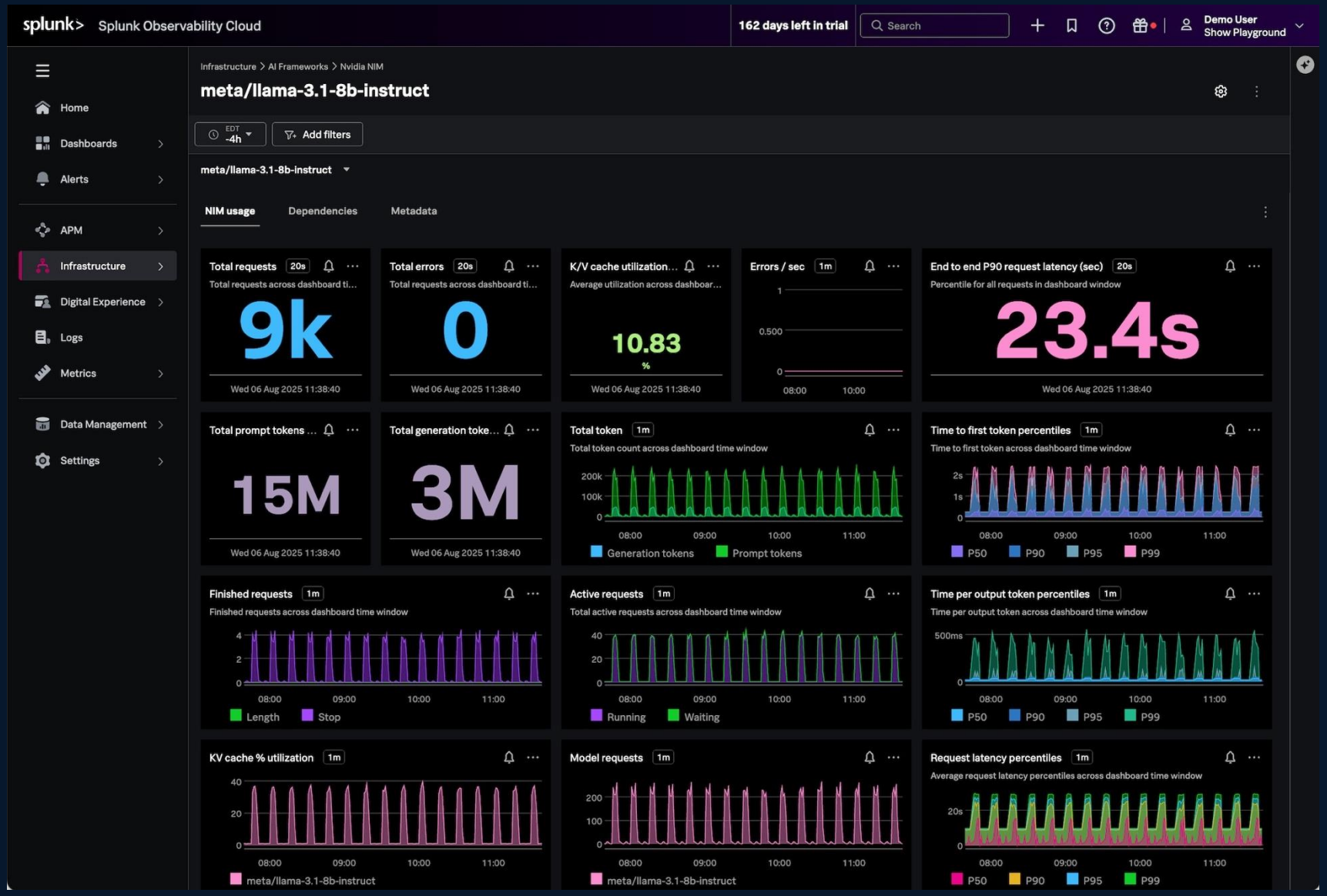
AI POD AI POD AI POD

Networking
Cisco Nexus 9300 w/ Silicon One® Switches w/ NVIDIA Spectrum-X*
Managed w/ Cisco Nexus Dashboard

Security
Cisco AI Defense
Cisco Hybrid Mesh Firewall
Isovalent
Secure Firewall
Hypershield

Observability
IT Service Intelligence
App-Dynamics
O11y
Splunk Platform
Unified Data Platform

Splunk AI Agent and AI Infrastructure Monitoring



Ethernet vs Infiniband

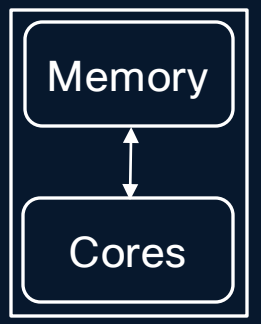
WWT testing made use of industry-standard MLCommons benchmarks, specifically the MLPerf Training and MLPerf Inference Data center problem sets. These enabled an apples-to-apples analysis of how network transport affects generative and inference AI performance.

BENCHMARK	MODEL	ETHERNET	INFINIBAND	ETH/IB RATIO
MLPerf Training	BERT-Large	10,886 s	10,951 s	**0.9977**
MLPerf Inference	LLAMA2-70B-99.9	52.362 s	52.003 s	**1.0166**

https://www.wwt.com/blog/the-battle-of-ai-networking-ethernet-vs-infiniband

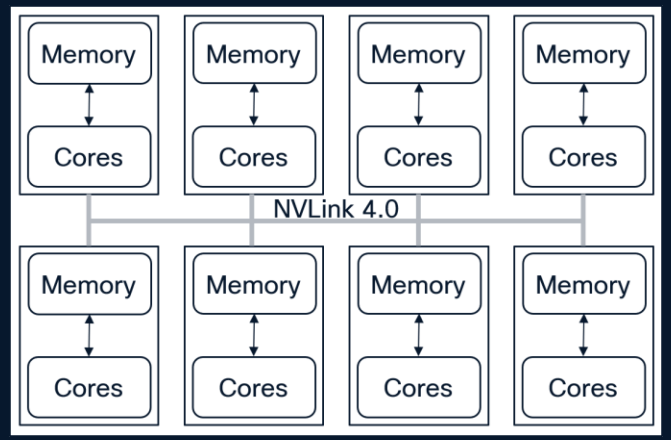
https://mlcommons.org/benchmarks/training/

AI Workloads Have Unique Network Requirements

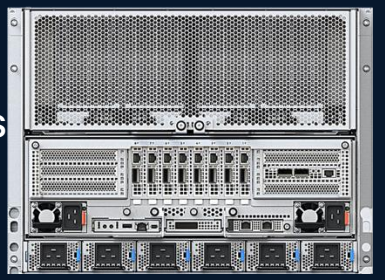


4.8TB/s

NVIDIA H200

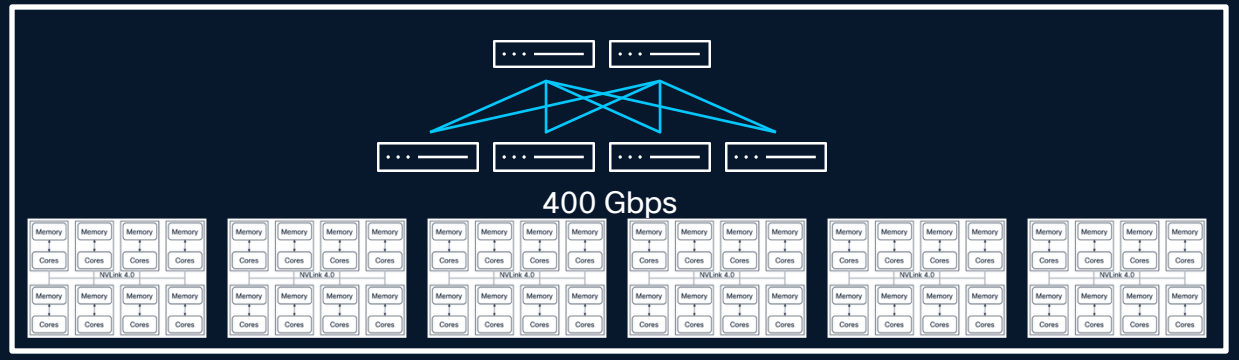


900GB/s



Cisco UCS C885A M8

GPU Fabric



Key network challenges and requirements

Training requirements

High bandwidth: Movement of massive datasets across the network demands high bandwidth networks

1

Non-blocking lossless: Network inconsistencies can affect the accuracy and training time of AI models

2

Congestion management: Detect potential congestion and redistribute network traffic accordingly

3

Inference requirements

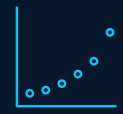
Low latency: Real-time AI applications require extremely low latency

4

General requirements

Visibility: Comprehensive visibility tools for real time monitoring, issue detection, and troubleshooting

5



Scalability: Dynamic nature of AI workloads mean that the network needs to be agile and scalable

6

cisco

Two NVIDIA Reference Architectures

Providing prescriptive designs for AI at scale.

ERA

Enterprise Reference Architecture

For Enterprise, Commercial, and Public Sector customers

Clusters of 32 to 1024 GPUs

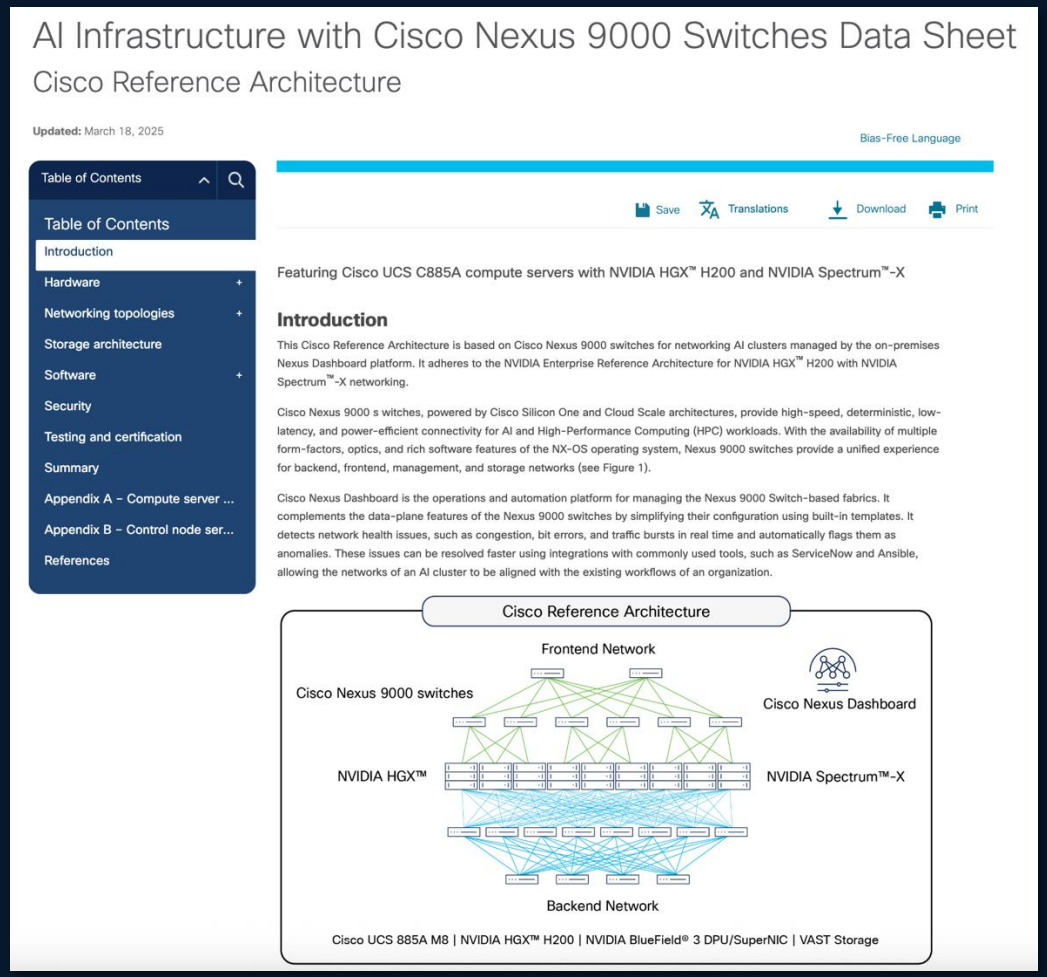
NCP

NVIDIA Cloud Partner

For cloud provider and web scale customers

Clusters of 1024 to 16,000+ GPUs

Published Reference Architectures



AI Infrastructure with Cisco Nexus 9000 Switches Data Sheet

Cisco Reference Architecture

Updated: March 18, 2025

Bias-Free Language

Table of Contents

Table of Contents

Introduction

Hardware

Networking topologies

Storage architecture

Software

Security

Testing and certification

Summary

Appendix A – Compute server ...

Appendix B – Control node ser...

References

Save Translations Download Print

Featuring Cisco UCS C885A compute servers with NVIDIA HGX™ H200 and NVIDIA Spectrum™-X

Introduction

This Cisco Reference Architecture is based on Cisco Nexus 9000 switches for networking AI clusters managed by the on-premises Nexus Dashboard platform. It adheres to the NVIDIA Enterprise Reference Architecture for NVIDIA HGX™ H200 with NVIDIA Spectrum™-X networking.

Cisco Nexus 9000 s witches, powered by Cisco Silicon One and Cloud Scale architectures, provide high-speed, deterministic, low-latency, and power-efficient connectivity for AI and High-Performance Computing (HPC) workloads. With the availability of multiple form-factors, optics, and rich software features of the NX-OS operating system, Nexus 9000 switches provide a unified experience for backend, frontend, management, and storage networks (see Figure 1).

Cisco Nexus Dashboard is the operations and automation platform for managing the Nexus 9000 Switch-based fabrics. It complements the data-plane features of the Nexus 9000 switches by simplifying their configuration using built-in templates. It detects network health issues, such as congestion, bit errors, and traffic bursts in real time and automatically flags them as anomalies. These issues can be resolved faster using integrations with commonly used tools, such as ServiceNow and Ansible, allowing the networks of an AI cluster to be aligned with the existing workflows of an organization.

https://www.cisco.com/c/en/us/products/collateral/data-center-networking/nexus-hyperfabric/nexus-9000-ai-era-ds.html

NVIDIA Certified Cisco Nexus Hyperfabric AI Enterprise Reference Architecture Data Sheet

Updated: May 14, 2025

Bias-Free Language

Table of Contents

Table of Contents

Introduction

Hardware

Networking topologies

Storage architecture

Software

Security

Testing and certification

Summary

Appendix A – Compute server ...

Appendix B – Control node ser...

Save Translations Download Print

Cisco Nexus Hyperfabric AI Enterprise Reference Architecture certified by NVIDIA, featuring Cisco® cloud-managed AI/ML networking of Cisco UCS® C885A M8 Racker Servers with NVIDIA HGX™ H200 and NVIDIA Spectrum™-X.

Introduction

Cisco Nexus® Hyperfabric AI is an on-premises AI cluster that is managed by a cloud-hosted controller. It empowers and simplifies your AI initiatives and accelerates AI deployments with a comprehensive, integrated, cloud-managed solution. Cisco Nexus Hyperfabric AI Reference Architecture is based on Cisco Silicon One® switches and adheres to the NVIDIA Enterprise Reference Architecture (Enterprise RA) for NVIDIA HGX H200 and Spectrum-X.

Figure 1 shows the key components of the solution. The key hardware components used in the cluster are described in the next section.

https://www.cisco.com/c/en/us/products/collateral/data-center-networking/nexus-hyperfabric/hyperfabric-ai-era-ds.html

Cisco N9100 Series

Powered by NVIDIA Spectrum-X Ethernet Silicon

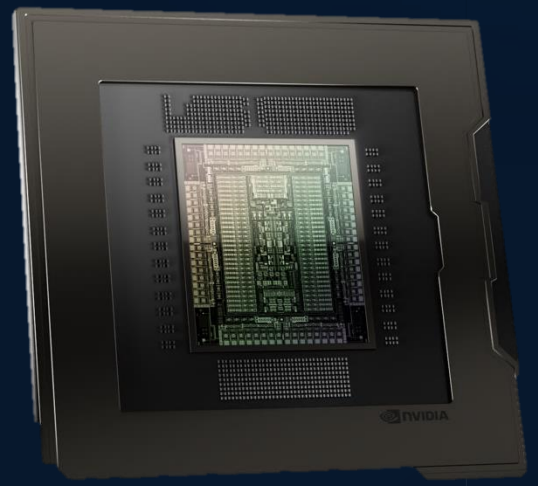
Introducing

Cisco N9100 Series Switches

AI Scale

NCP RA Compliant

Simplified Operations



64 x 800G OSFP Ports

NVIDIA

Powered by Spectrum-X Ethernet silicon

cisco

Operating system of choice NX-OS or SONiC

cisco

Automated and Operated by Nexus Dashboard

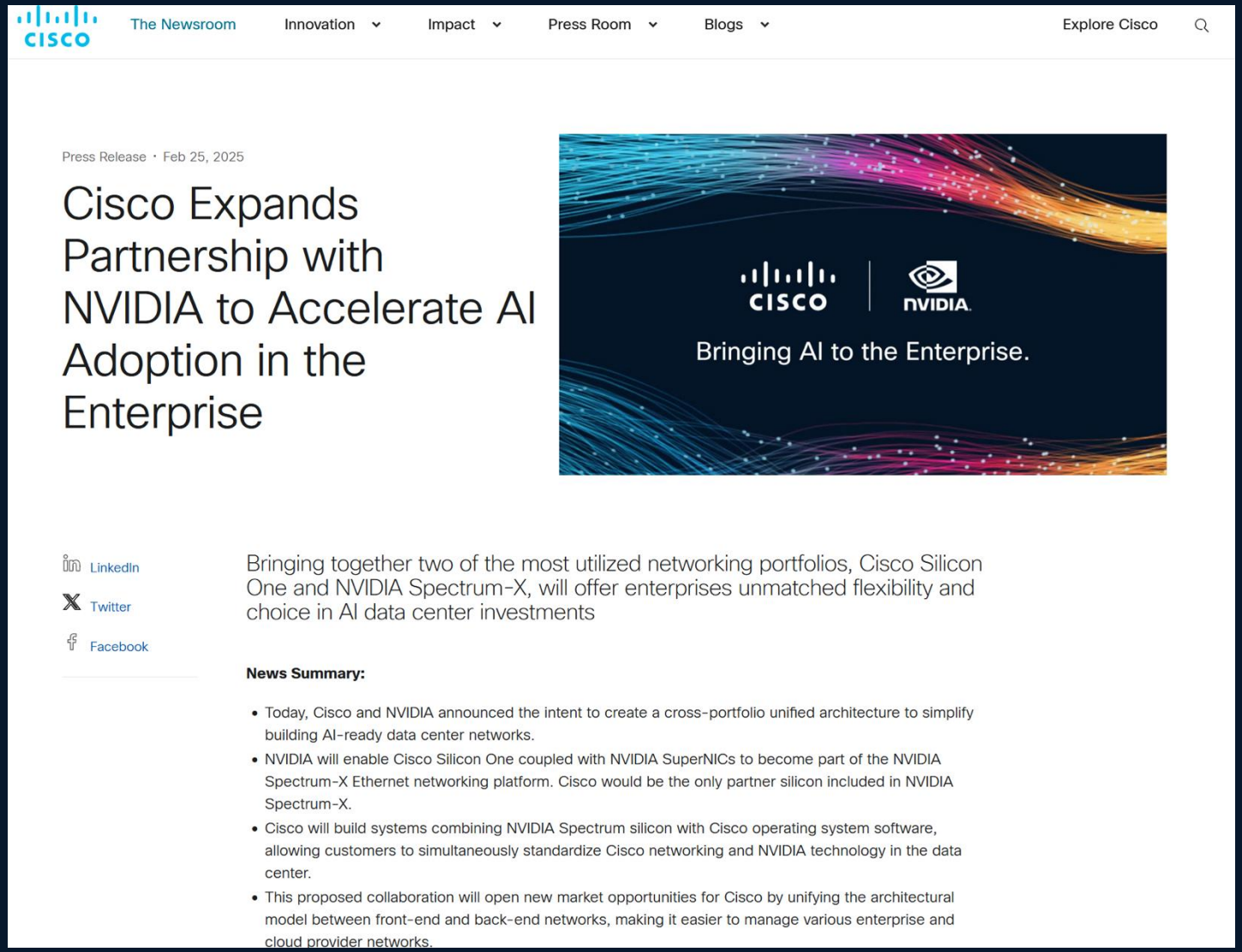
cisco

Cisco N9164E-NS4-O Switch

- Compact 2RU 51.2T Switch with OSFP ports
  64x 800GE | 128x 400GE
- 51.2T NVIDIA Spectrum-X Ethernet ASIC
  100G SerDes
- Ethernet scale-out for AI clusters
  NCP RA Compliant
- Supports Cisco NX-OS and Cisco SONiC
- Provisioning and management by Nexus Dashboard

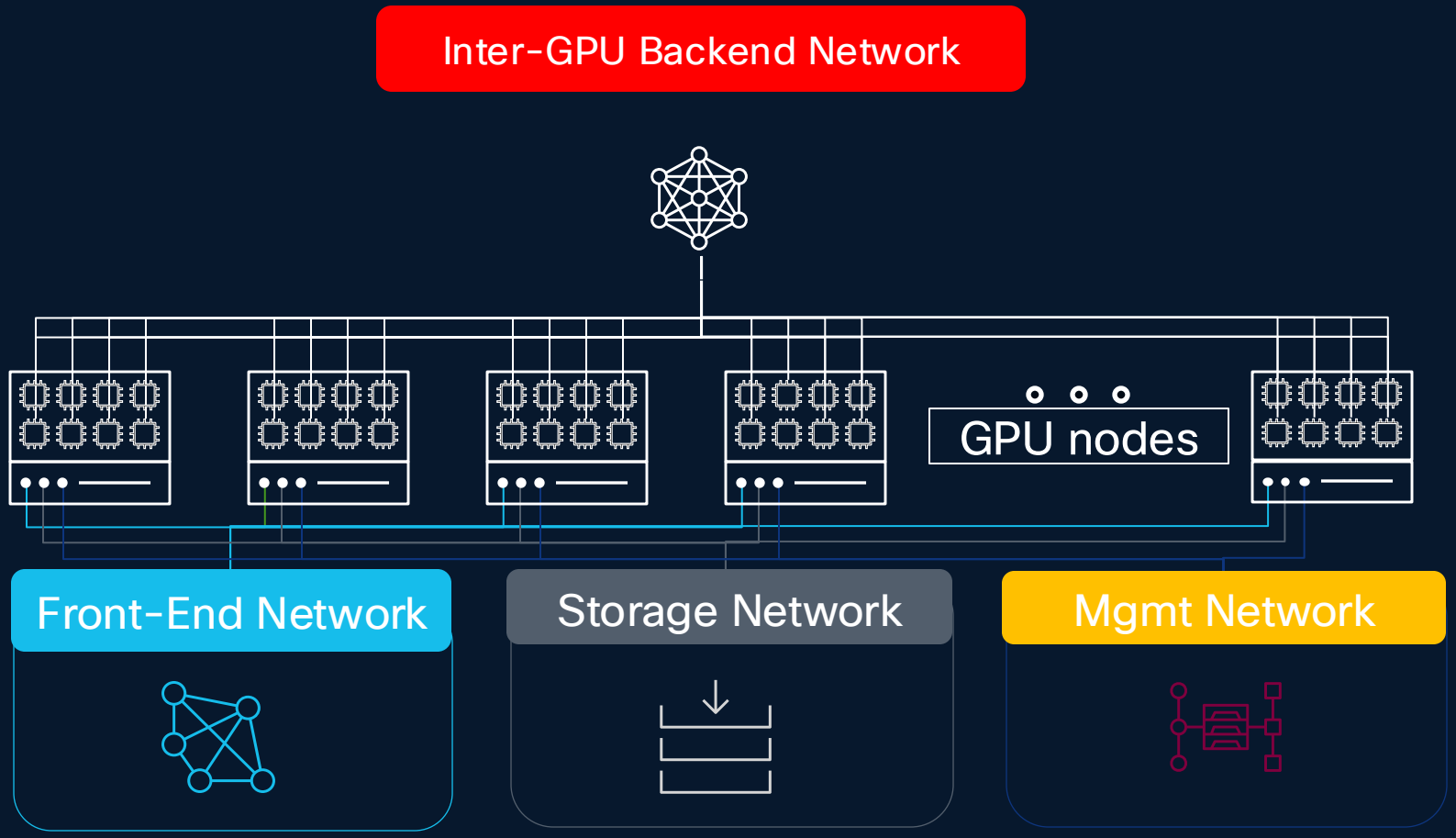
N9164E-NS4-O

Intelligent Packet Flow

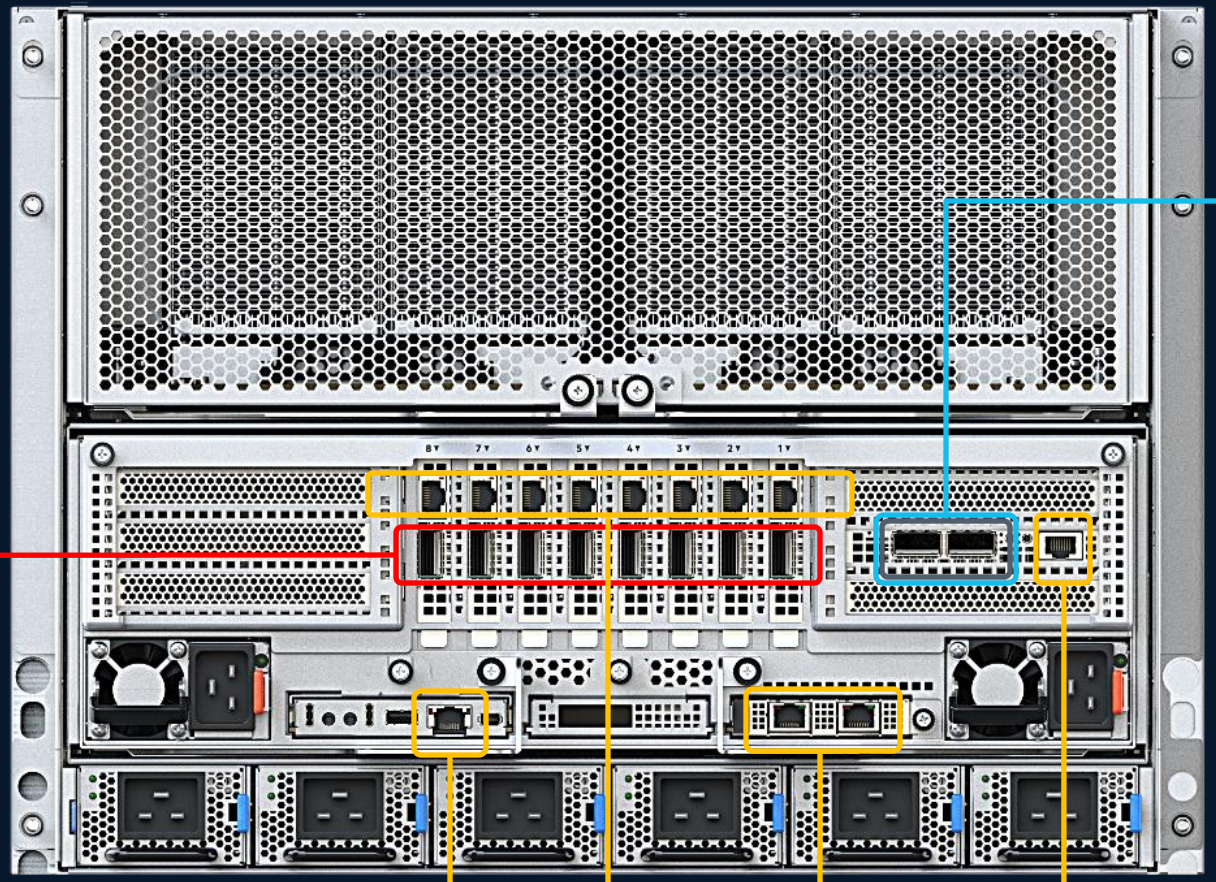


https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2025/m02/cisco-expands-partnership-with-nvidia-to-accelerate-ai-adoption-in-the-enterprise.html

Multiple networks for AI infrastructure



UCS C885A M8 External Storage Connectivity



Front-End Network

Storage Network

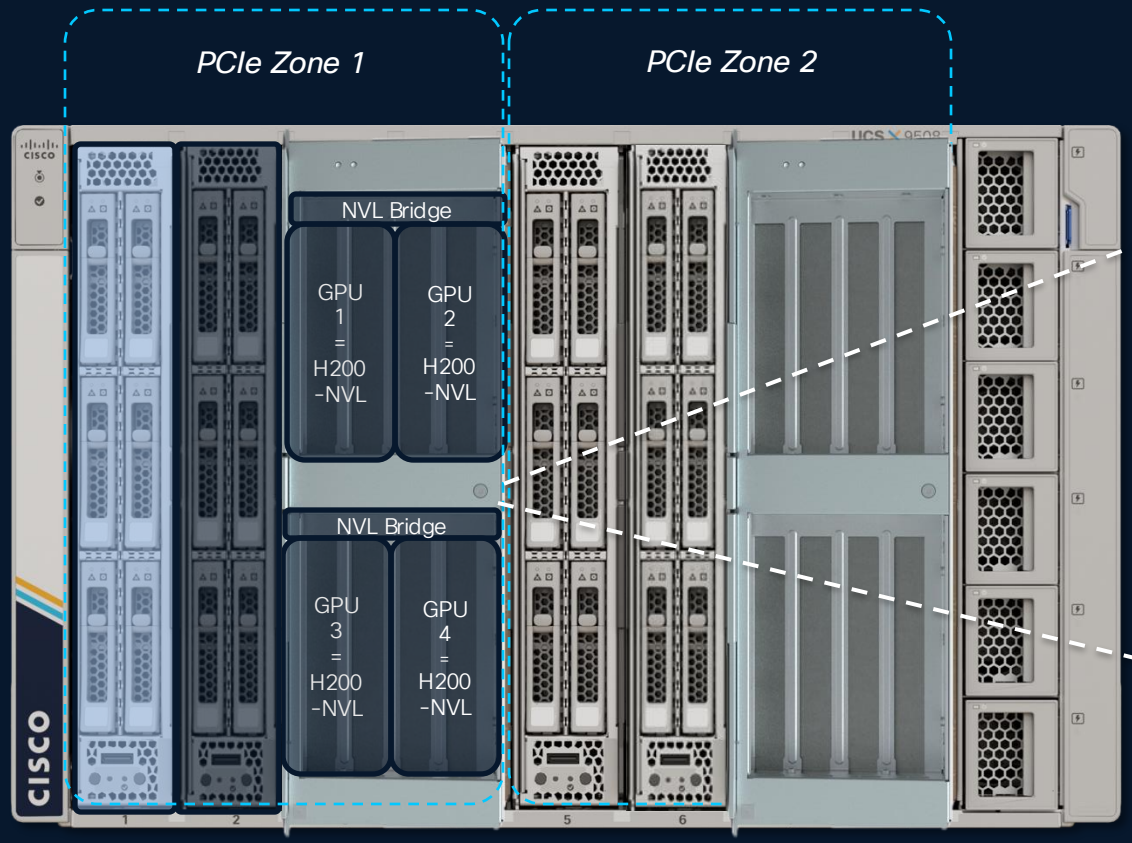
May be converged to a shared physical network

Inter-GPU Backend Network

Management Network

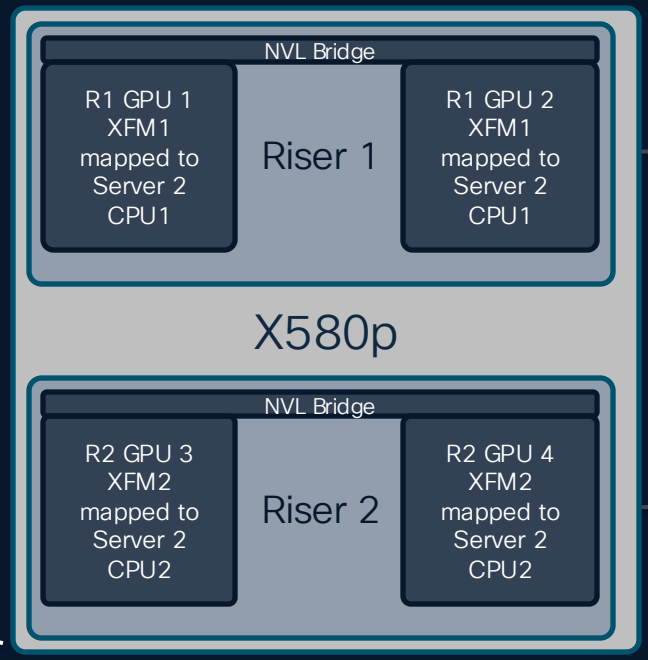
X580p – All GPUs Allocated to Server 2 w/NVL Bridge

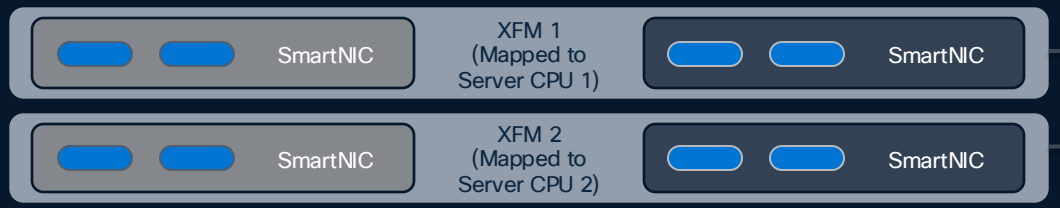
(2x NICs mapped to one server)



Server 1: No GPU

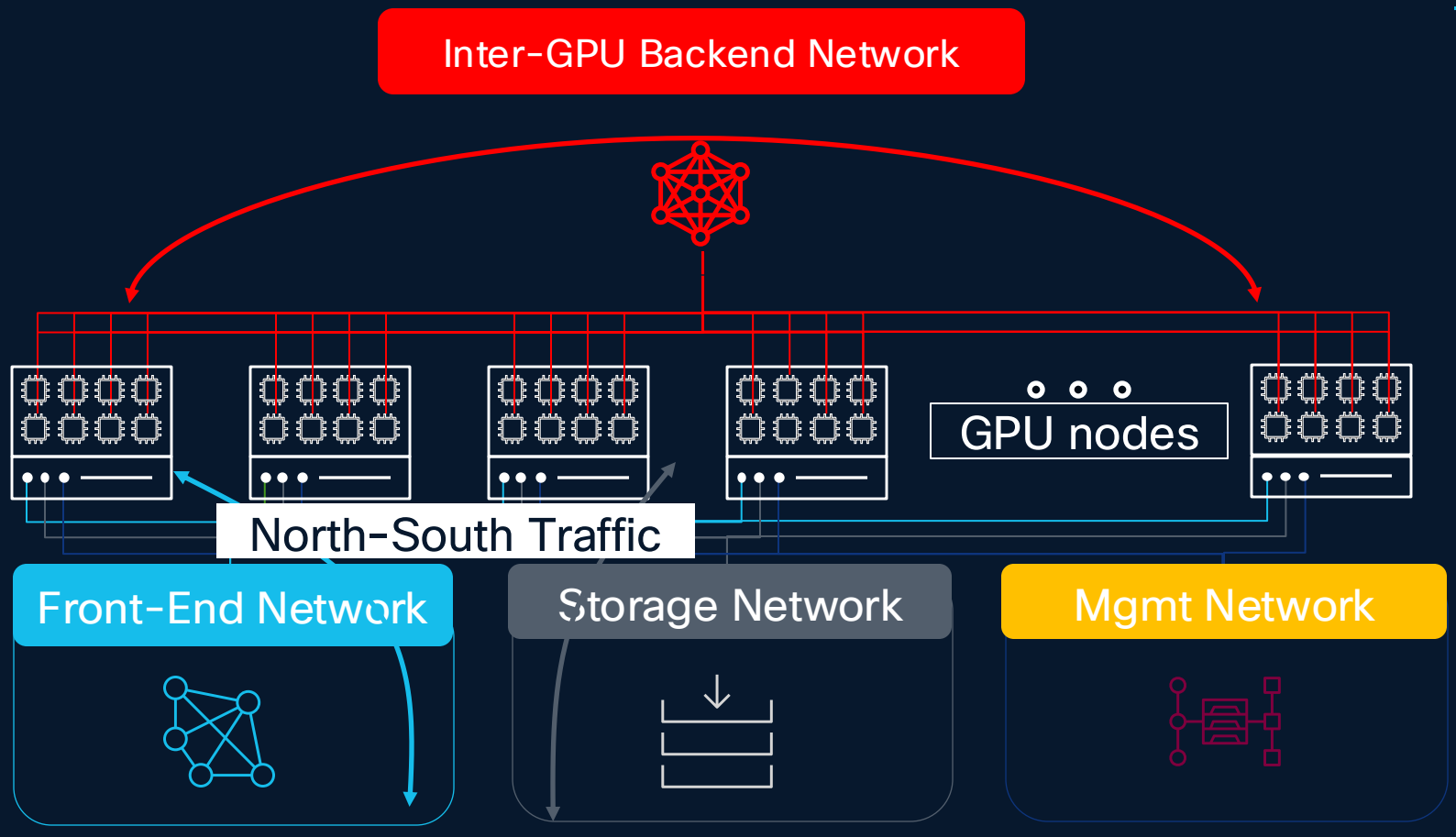
Server 2: 4x GPUs 2x NICs





Note: X580p default in 3/4 or 7/8 and is not supported in slots 2/3 and 6/7. It can be installed in slots 1/2 or 5/6 as well.

Multiple networks for AI infrastructure



Nexus Dashboard & NX-OS

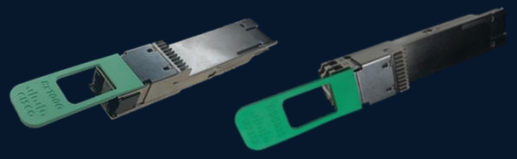
- Single management platform
- Single monitoring and visibility platform
- Consistent APIs – Reuse existing automation frameworks
- Reuse expertise of your teams

Cisco AI/ML approach

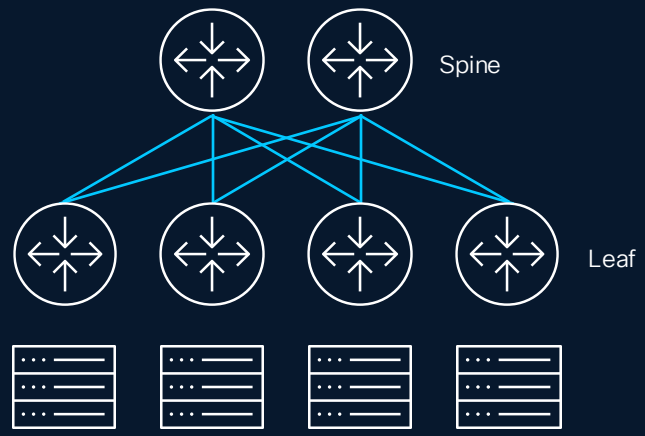
Building blocks

Nexus 9300 800/400G Switches

400G DR4/FR4 or 100G DR Optics



Non-blocking spine-leaf fabric



AI/ML for networking blueprint

RDMA over Ethernet (RoCEv2)

Lossless network (PFC + ECN)

Proven full network stack

- Fabric automation and visibility made simple
- Extensive interop testing with major NIC vendors
- Widely deployed OS with rich feature set
- 25.6T and 51.2T Silicon supporting 50G/100G SerDes
- Recent deployments with financial institutions, research schools, government agencies

Transform infrastructure for AI: CVD blueprints: Network + Compute + Storage

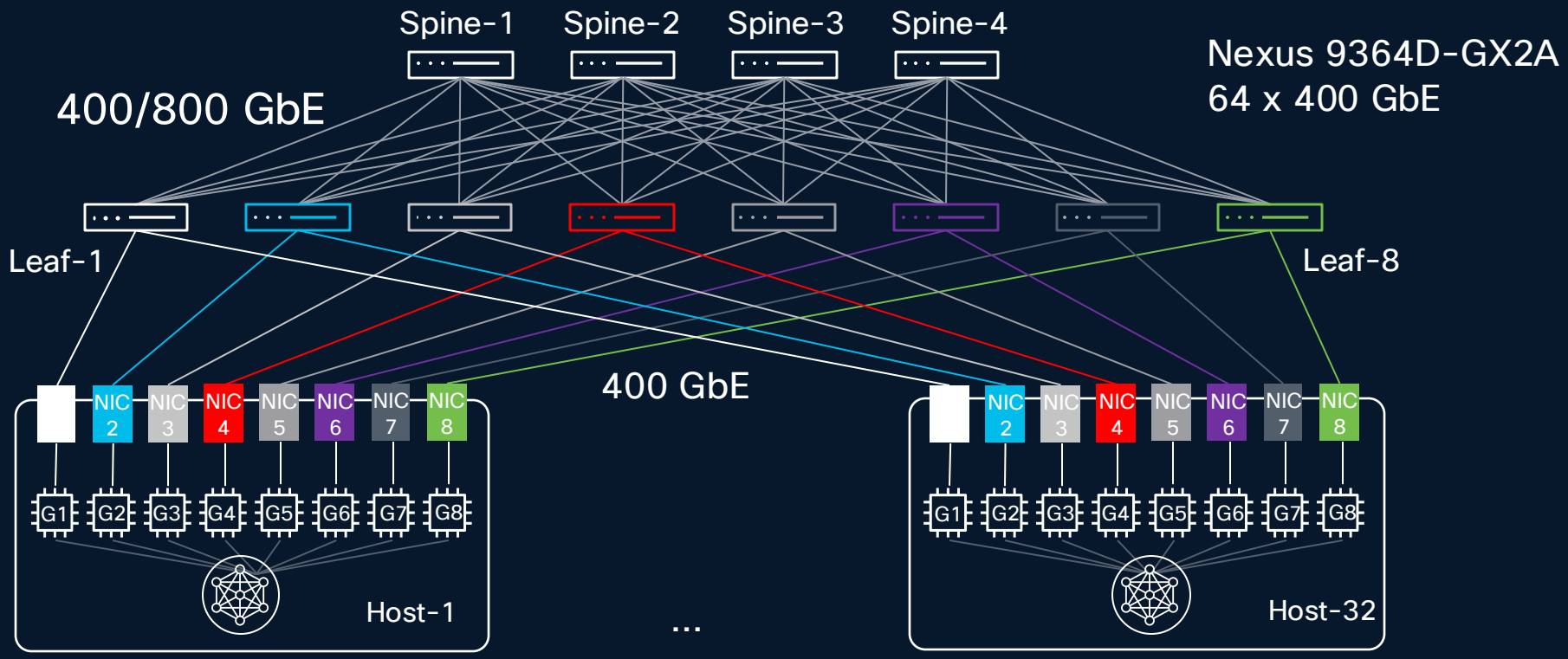
Inter-GPU back-end network design

Non-blocking and rails-optimized using Cisco Nexus 9000

Non-Blocking

32/64 ports connected to spine switches

32/64 ports connected to GPU hosts



Port 1 on all hosts connects to Leaf-1
Port 2 on all hosts connects to Leaf-2, and so on

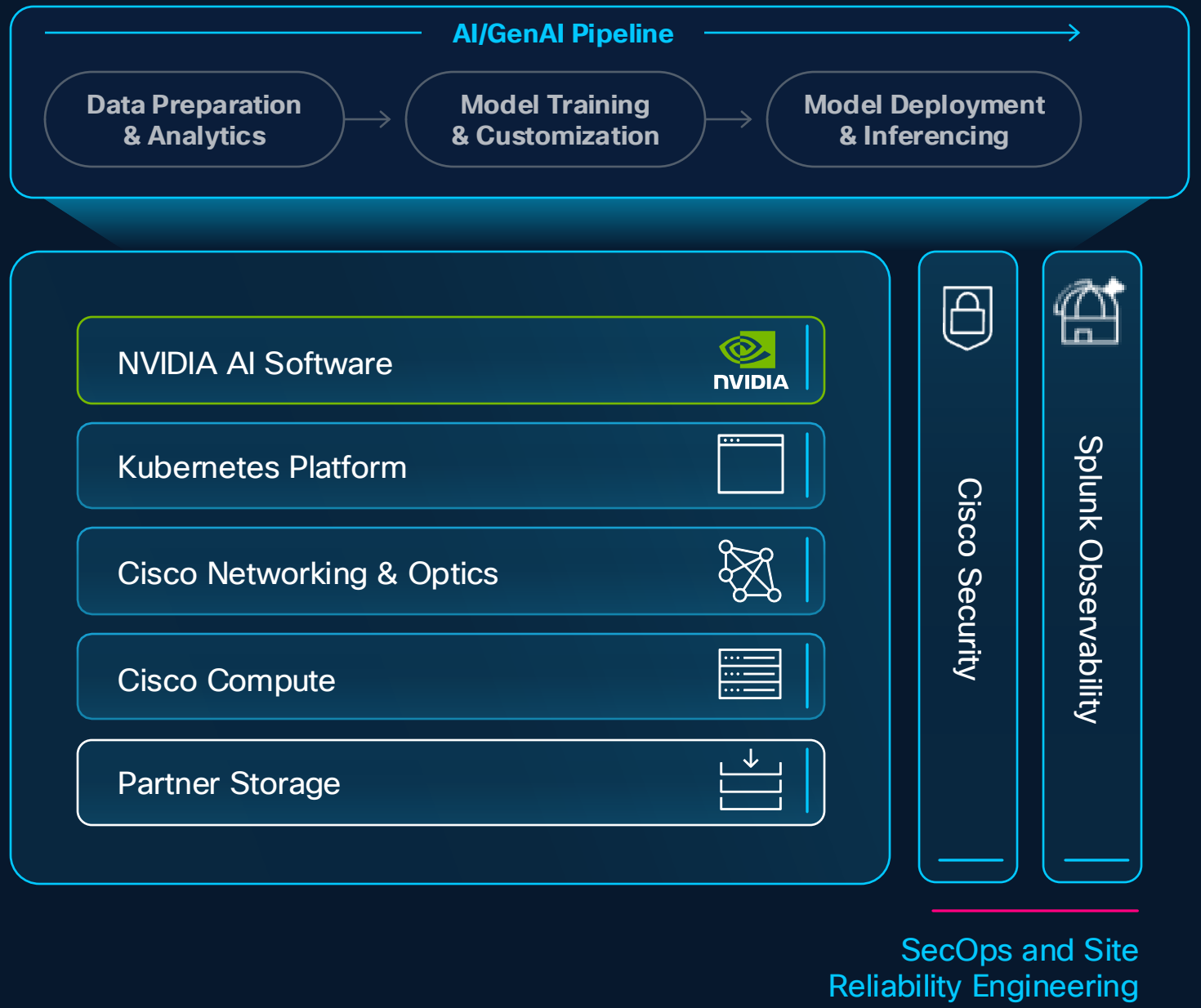
Cisco Secure AI Factory with NVIDIA

Delivering Trusted AI Outcomes

AI Practitioners

A reference design with validated architectures to accelerate AI adoption for enterprises with integrated AI infrastructure and software solutions

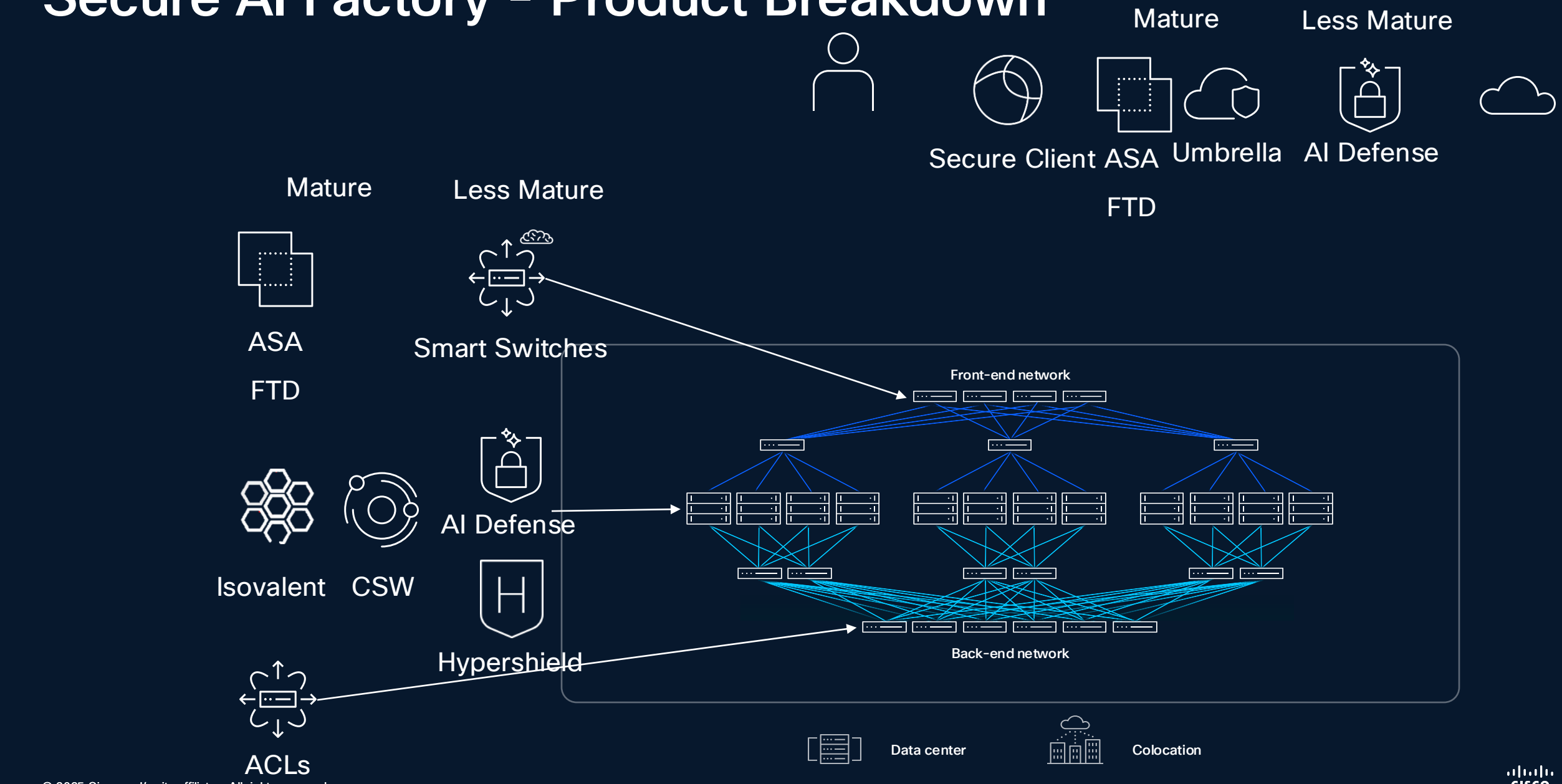
IT Infrastructure & Operations



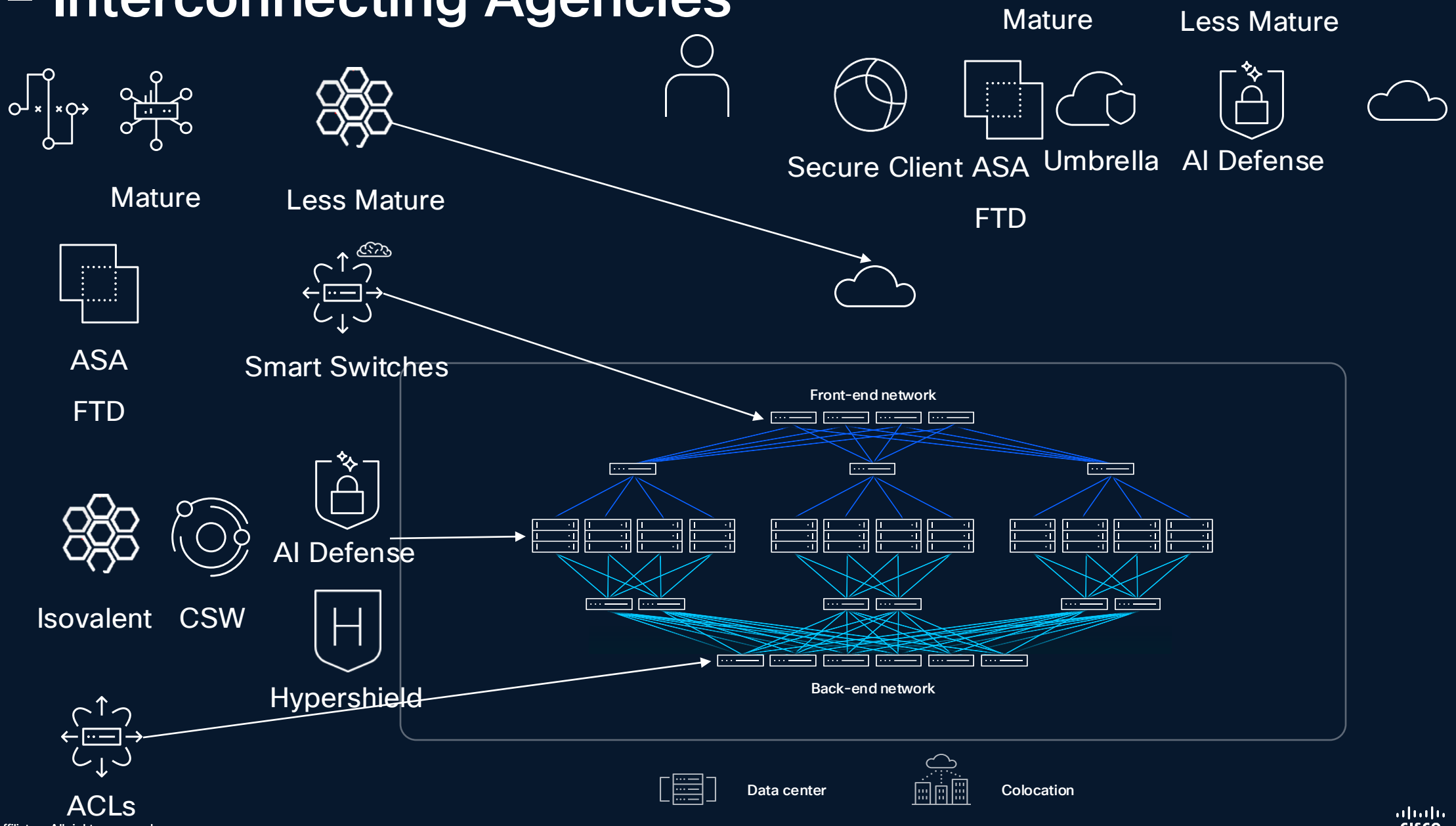
SecOps and Site Reliability Engineering

cisco

Secure AI Factory - Product Breakdown



WAN – Interconnecting Agencies



Cisco Nexus 9300 Series – 400G Fixed Switches

N9K-C9364D-GX2A

25.6T ASIC

2 RU 64 400G ports

MACsec on 16 ports

120MB on-die packet buffer

N9K-C9348D-GX2A

25.6T ASIC

2 RU 48 400G ports

MACsec on 48 ports

120MB on-die packet buffer

N9K-C9332D-H2R
N9K-C9332D-GX2B

12.8T ASIC

1 RU 32 400G ports

MACsec on 32 ports (H2R)
MACsec on 8 ports (GX2B)

80MB on-die packet buffer + 8GB HBM (H2R)
120MB on-die packet buffer (GX2B)

Cisco Nexus 9364E-SE2 - 64p 800G Fixed Switch

Compact 2RU 51.2T Switch

G200 ASIC (5nm) | 112G SerDes | 256MB packet buffer

64 800G ports | Up to 128 line-rate 400G ports (2x400G breakout)

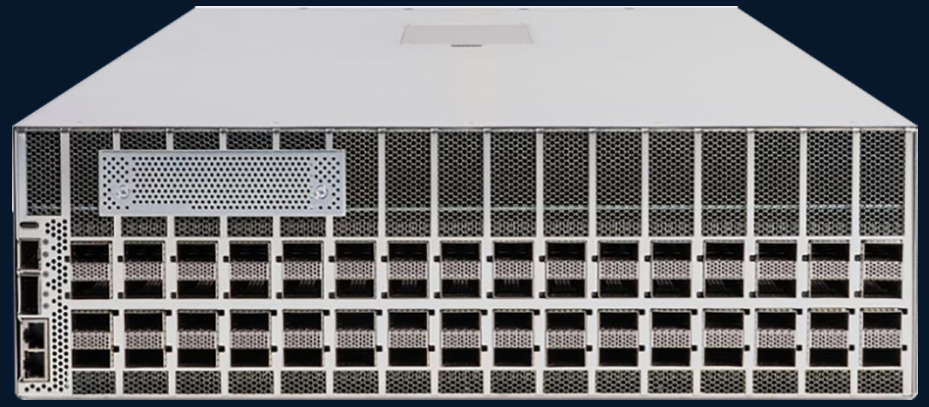
Choice of QSFP-DD800 or OSFP ports

QSFP-DD800 ports are backward compatible with QSFP-DD, QSFP28, QSFP+

Quad Core x86 CPU | 32GB RAM | 128GB SSD

Cisco NXOS spine and AI/ML spine/leaf capable

Cisco 8223-64E-M/8223-64E-MO



Compact 3RU 51.2T Router

P200 ASIC (5nm) | 112G SerDes | 144MB packet buffer | 16 GB HBM

64 800G ports | Up to 128 line-rate 400G ports (2x400G breakout)

Choice of QSFP-DD800 or OSFP ports

QSFP-DD800 ports are backward compatible with QSFP-DD, QSFP28, QSFP+

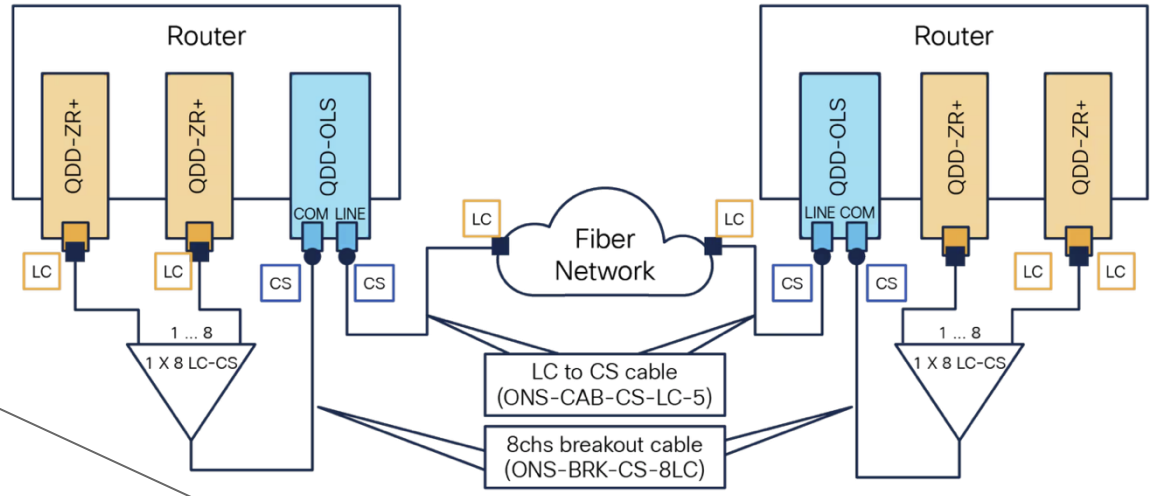
Octa Core x86 CPU | 64GB RAM | 128GB SSD

SONiC and IOS-XR (Roadmap)

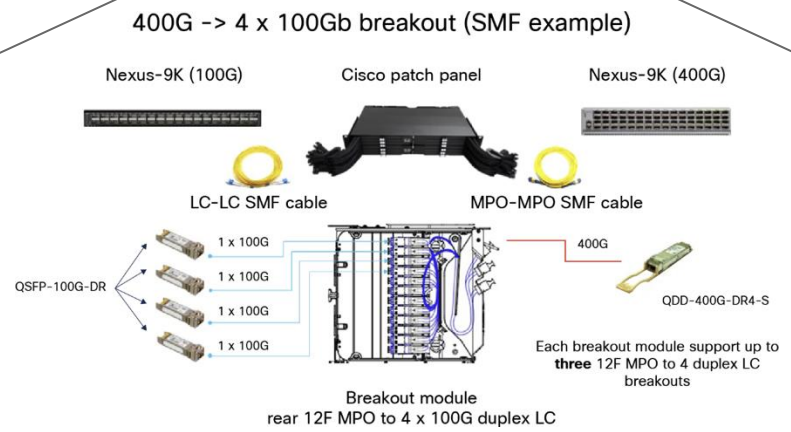
Transceivers

DC Transceivers

DCI Transceivers



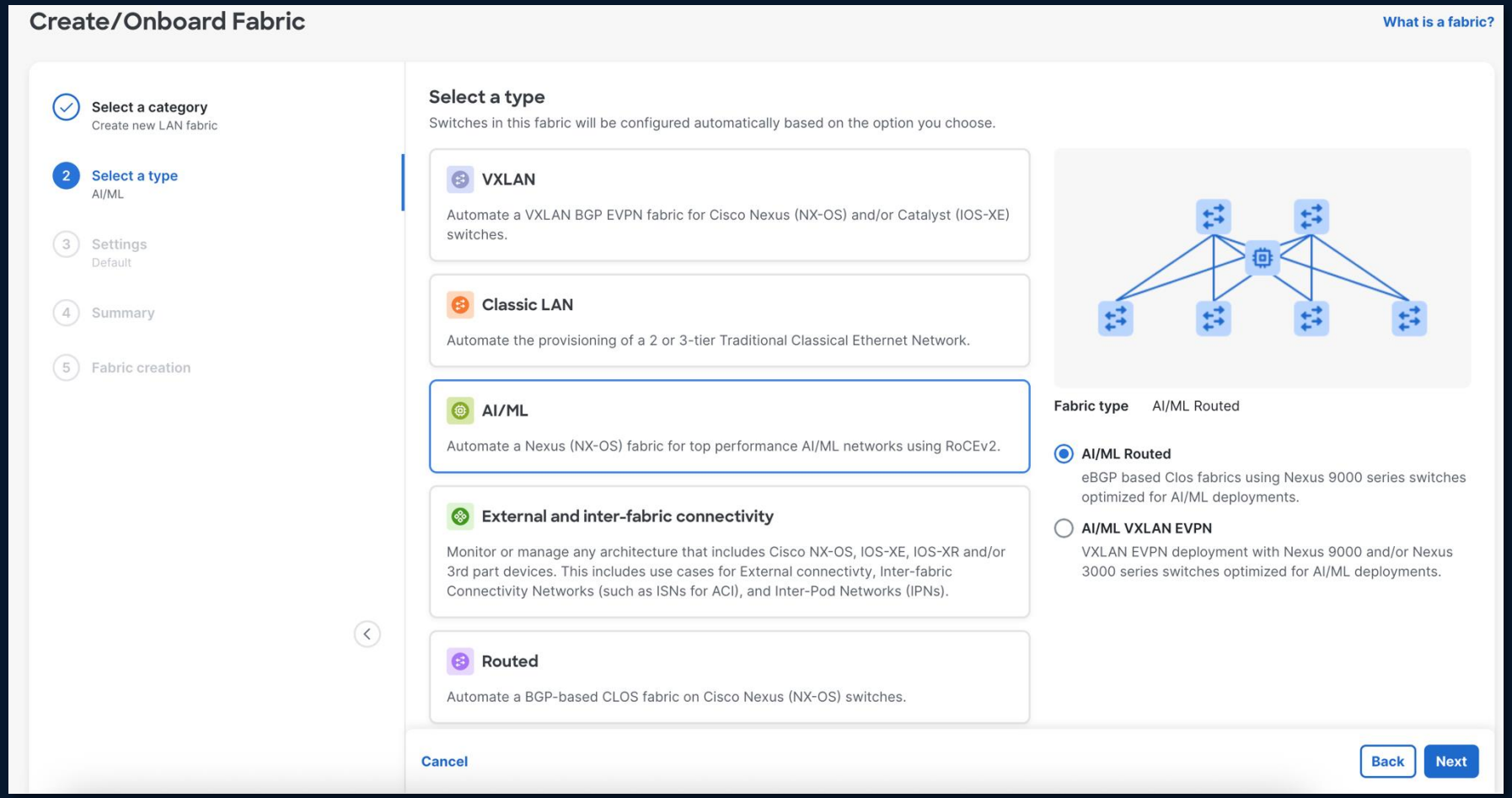
Breakout Panels



Automation and Operational Simplicity

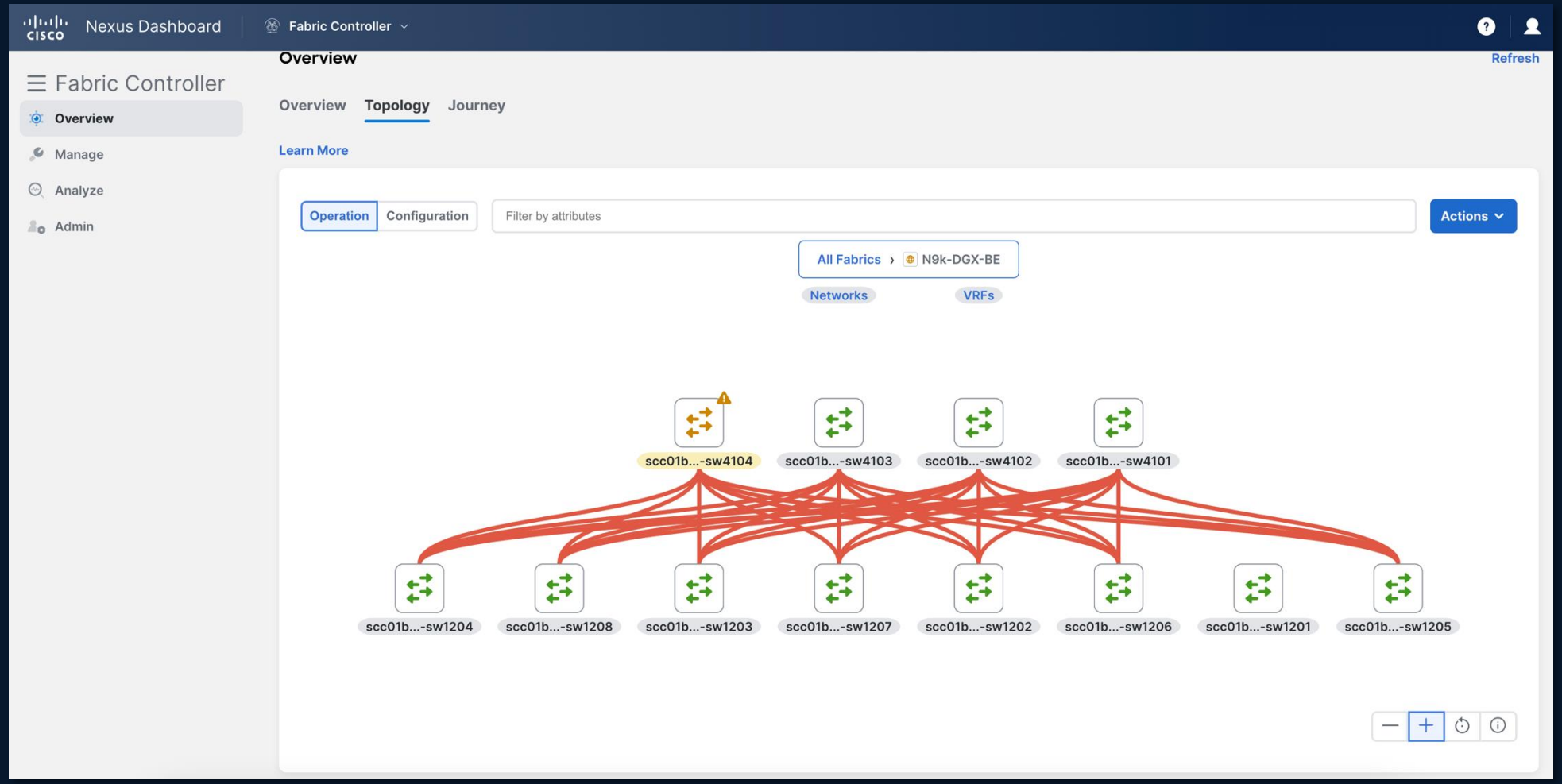
Nexus Dashboard for AI/ML Networks

Prescriptive AI/ML Template



Nexus Dashboard for AI/ML Networks

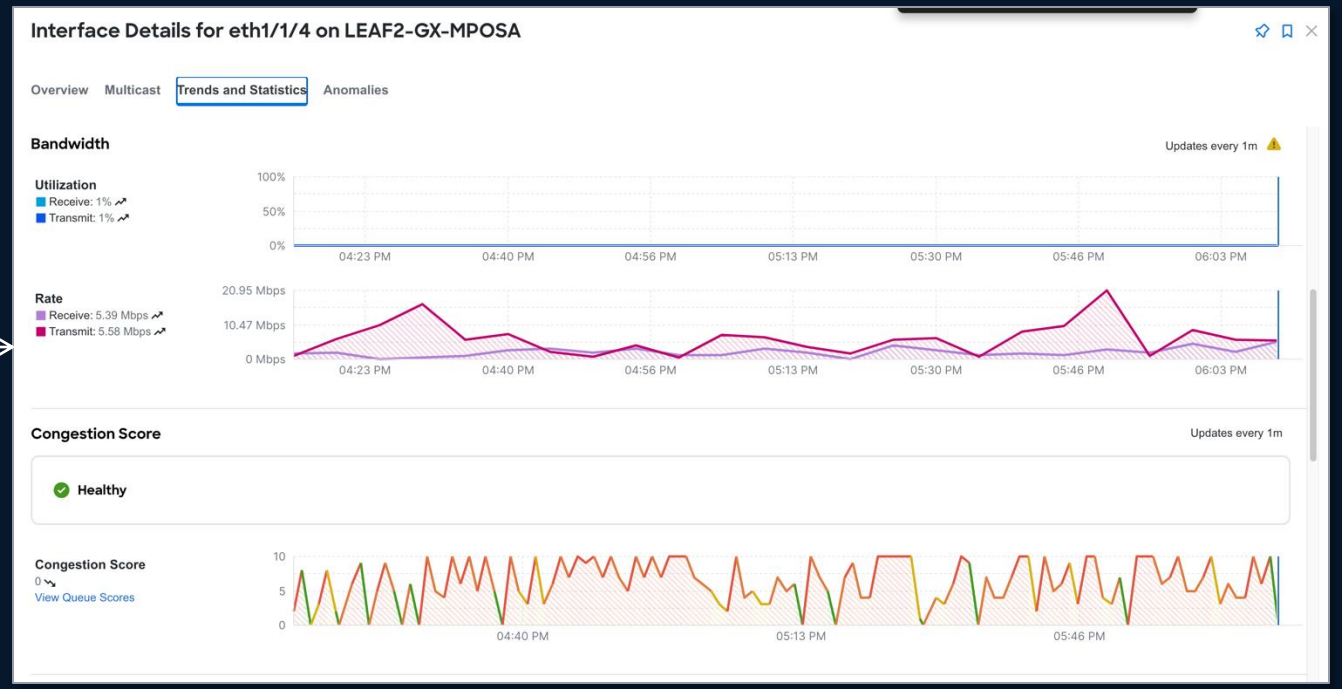
Rails-Optimized, 8-Leaf and 4-Spine Network Design Connecting 256 GPUs



Nexus Dashboard: Analytics

Simplifying network operations

- AI network visibility
- UX/UI dashboard
- Visibility – lossless Ethernet
- Monitoring (ECN, PFC)
- Congestion score
- Application to network performance correlation
- Telemetry and NetOps



With the granular visibility provided by Cisco Nexus Dashboard Insights the network administrator can observe drops

Tune thresholds until congestion hot spots clear and packet drops stop in normal traffic conditions

This is the first and most important step to ensure that the AI/ML network will cope with regular traffic congestion occurrences effectively


CISCO