CISCO Connect

Data Center Networking for the AI Era

Brenden Buresh- CCIE #2073, Distinguished Solutions Engineer

April 7, 2026

Cisco Powers How People and Technology Work Together Across the Physical and Digital Worlds

AI-Ready Data Centers

Future-Proofed Workplaces

Secure Global Connectivity

Digital Resilience

Accelerated by Cisco AI

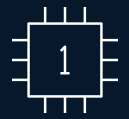
Cisco Data Center Strategic Priorities

Cisco AI Networking

Increase Performance

Simplify Operations

Security Fused into Every Layer



Silicon

Systems

Software

Optics

Operating Model

cisco

Cisco AI Networking Portfolio

Architectures Based on Business, Technical, and Operating Model Requirements

Nexus Dashboard

SONiC

On-Premises Managed Network
(Policy or Switch-Based)

Nexus Hyperfabric

SONiC

Cloud Managed Network
(Blueprint-Based)

Bring Your Own Management

IOS XR

OPEN Compute Project®

SONiC

Disaggregated Architectures

Cisco 8000 | 8500

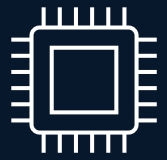
Common Cisco Nexus 9000 Systems

Built with Reference Architectures and Validated Designs

For Hyperscalers, Neoclouds, and Enterprises
Any Scale, Any GPU

cisco

Silicon and Systems Focus Areas



Silicon Diversity

- Cisco Cloud Scale and Silicon One ASICs powering Nexus
- Port speeds up to 400G / 800G
- SerDes speeds up to 25G / 50G / 100G

Smart Switches

- Built-in DPUs for services acceleration
- Multiple service functions at ASIC performance – security, NAT, telemetry
- Programmable pipelines for service firmware uprev

Ultra Low Latency

- Next-Gen ULL ASIC
- Silicon One based, 3.2T
- Enable 10G to 25G market transition

Cisco Silicon One: Breakthrough Innovation in Performance, Flexibility, and Efficiency

Get Performance at Scale

Protect your Investment

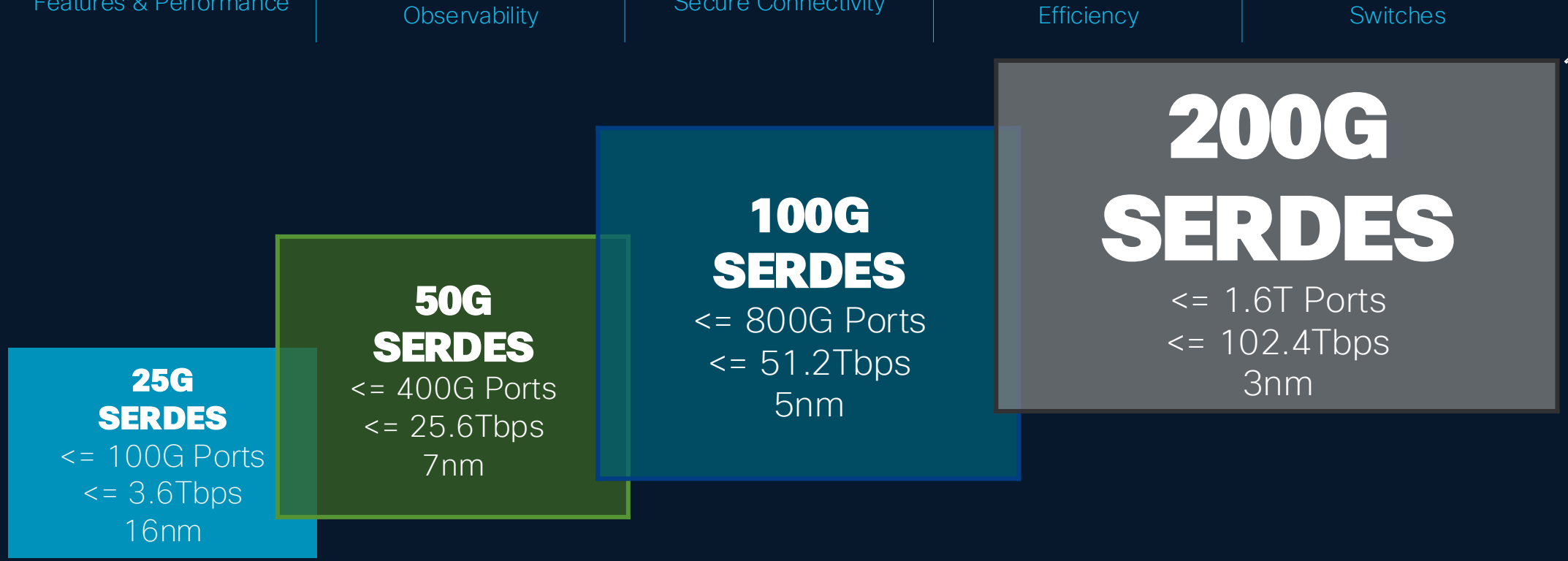
Boost Efficiency

Manage and Secure

Cisco ASIC Innovations – Nexus 9000 Series

Fully Programmable Multi-Slice Architecture

Features & Performance | Deep Visibility & Observability | Secure Connectivity | Low Latency & Power Efficiency | Modular & Fixed Switches



Cisco Nexus 9000 Series Switches Evolution

Shipping | *New*

Cisco N9300
48p 25G + 6p 100G
36p & 64p 100G

Cisco N9300
48p 50G + 4p 400G
32p & 64p 400G

Cisco N9300 Smart Switch
24p 100G &
48p 25G + 2p 100G + 6p 400G

Cisco N9300 SG2
64p 800G (QDD & OSFP)

Cisco N9100 Spectrum-4
64p 800G (QDD & OSFP)

Cisco N9300 SG2X
64p 800G (QDD & OSFP/QSFP)

Cisco N9300 SP2R
64p 800G (QDD & OSFP)

Nexus Fixed

Nexus Modular

Cisco N9500
Up to 576
100G ports

Cisco N9400
64p 400G (or)
128p 200G

Cisco N9800
Up to 288
400G ports

Cisco N9800 SP2R
36x800G OSFP LC

Port Speed: 100G | 400G | 800G

ASICs: <= 3.6T 25G Serdes | 4.8 – 25.6T 50G Serdes | >= 51.2T 100G Serdes

Cisco Silicon One G200

Uniquely Efficient and Capable / Optimized Enhanced Ethernet for AI/ML

One Architecture
Simpler and easier to maintain network

High Performance
2x higher performance than G100

Sustainability via Technology
2x more power efficient than G100

Ultra Low Latency
2x lower latency than G100

Optimal Network Design
512-wide radix enables
flatter, more efficient networks

Fully Shared Packet Buffer
Optimal fairness, burst performance, JCT

51.2 Tbps

Ultra Ethernet Consortium
READY

Advanced 112 Gbps SerDes
Cisco designed next-generation ADC SerDes
Support for Optics, 4-meter DAC, LDO and CPO

Intelligent Packet Flow
Non-correlated WECMP avoids hash polarization
Congestion-aware stateful load balancing
Congestion-aware packet spraying

Link Failure Avoidance
HW based traffic link failure redistribution optimizes real-world large-scale deployments

Programmable Processor
Deterministic ultra-low latency processor
with run to completion for ultimate flexibility

435B+ Lookups per Second
Enables advanced features like SRv6 uSID

Deep Visibility & Analytics
In-band telemetry including emerging protocols
Hardware analyzers enable post event debuggability

Cisco Nexus 64-Port 800G Portfolio

Built for Ethernet-Based AI Clusters that Scale

Silicon Diversity | NX-OS or SONiC | Common Operating Model

Shipping Today

Cisco Silicon One G200 (N9364E-SG2)

64 800G ports | 256MB packet buffer

Cisco Cloud Reference Architecture (CRA)

NVIDIA Spectrum-X Technology Enabled

AI Frontend and Backend

Orderable Now

Cisco NVIDIA Spectrum 4 (N9164E-NS4)

64 800G ports | 160MB packet buffer

NVIDIA Cloud Partner (NCP) Reference Architecture

NVIDIA Spectrum-X Technology Built-In

AI Backend

NEW

Introducing

Cisco Silicon One G300

102.4 Tbps
AI Network Switch

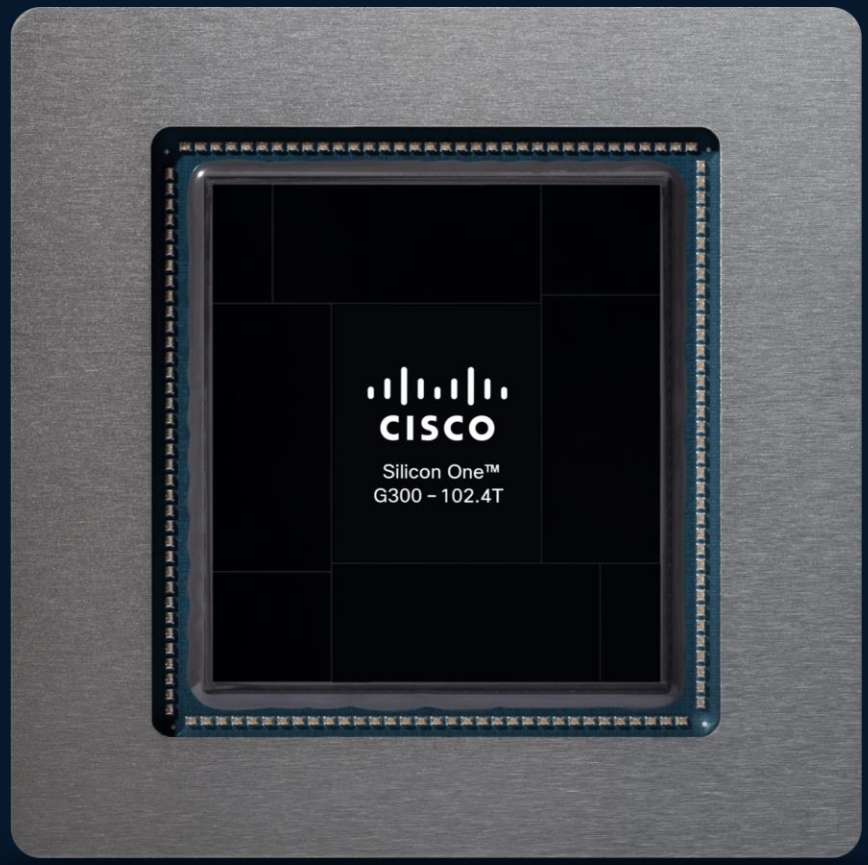
200 Gbps
Integrated Serdes

1.6T Ethernet
Port Switch

512 ports
High Radix Scaling

256MB
Fully-Shared Packet Memory

Optimizing TCO
for Scale AI Deployments



NEW

Introducing H2’26

Cisco Advanced 102.4T Systems Portfolio

Seamless Deployment Flexibility for Optimized AI Performance

Cisco N9300 Cisco 8133

Air Cooled

NX-OS

Features

3RU, 64p 1.6T OSFP

EIA 19"

Redundant Pluggable Power

102.4Tbps

Cisco N9300 Cisco 8132

Fully Liquid Cooled

NX-OS

Features

2OU, 64p 1.6T OSFP

ORv3N 21” blind mate

Redundant Integrated Power

Advanced Leak Detection

102.4Tbps

Hyperscale | Neocloud | Service Providers | Sovereign Cloud | Enterprise

Cisco Expands Silicon One P200 Across the Portfolio

With New Cisco N9000 and Expanded Cisco 8000 Systems

Cisco N9364-SP2R and Cisco 8223
51.2T Fixed Systems

51.2T, 3RU,
64p 800G Fixed
OSFP800 and
QSFP-DD800

cisco NX-OS

cisco ACI

cisco IOS-XR

SONiC

Cisco N9836-SP2R and Cisco 88-LC2
28.8T Modular Line Cards

28.8T, 36p 800G
OSFP800

Cisco 800G
ZR/ZR+ optics

1000km+ reach

Industry Leading Performance and Scalability

Smart Adaptive Processing

Quantum Resilient Line-Rate Encryption

Scale-Across | DCI | Universal Spine | Core & Peer Routing

cisco

Cisco and NVIDIA Partnership for AI Infrastructure

Quick Recap

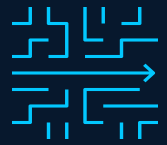
AI-Ready Data Center

1

Optimized and Dense GPU servers with NVIDIA (HGX/MGX/RTX) architectures on Cisco UCS

Cisco Networking solutions (On-prem or Cloud Managed)

Cisco Security and Observability



Cisco Networking with NVIDIA Spectrum-X

2

Cisco Silicon Switches with Spectrum-X Technology via software (NX-OS 10.6.1F onwards)

Cisco Switches with NVIDIA Spectrum ASIC (NX-OS and SONiC) (1HCY26)

The only NVIDIA partner providing Networking systems with Spectrum-X capabilities on native silicon

NVIDIA Reference Architecture Compliant

3

Cisco Secure AI Factory based on NVIDIA Enterprise Reference Architecture (ERA)

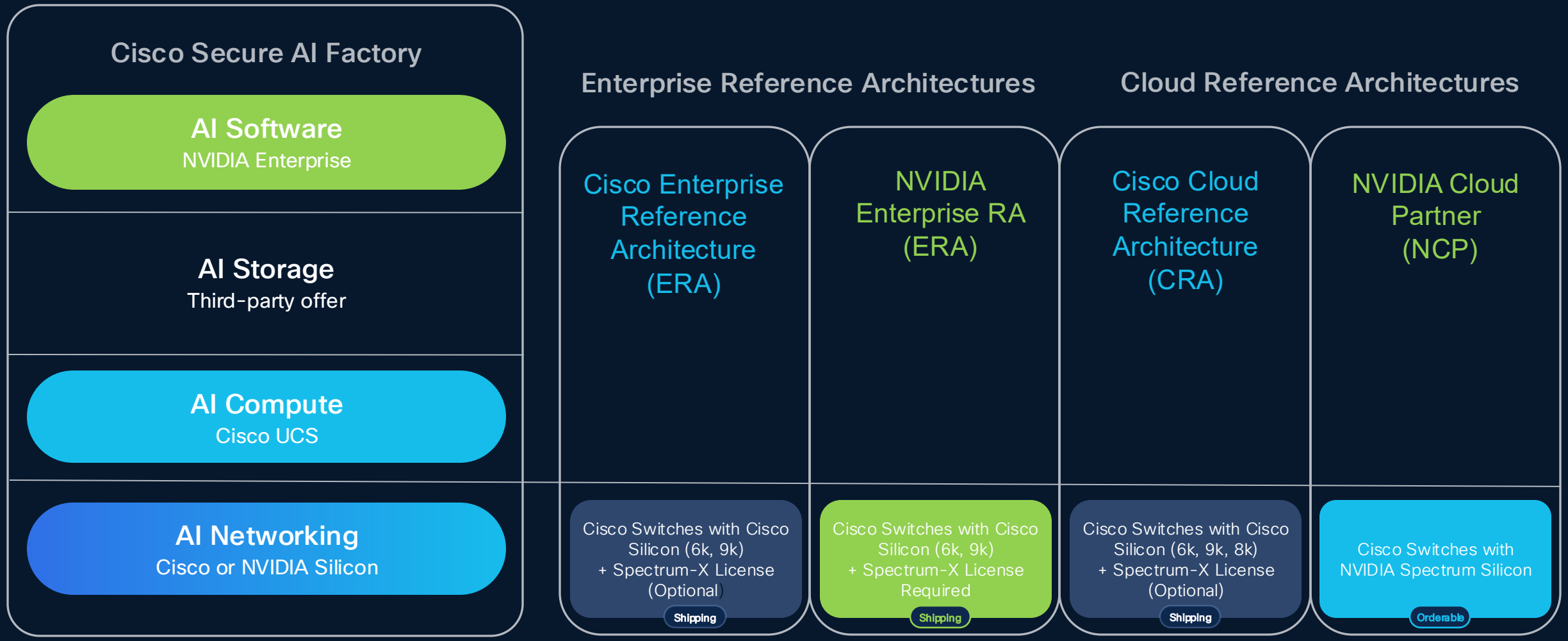
Cisco's Cloud Reference Architecture (CRA) based on design tenets of NVIDIA Cloud Partner (NCP) Reference Architecture and endorsed by NVIDIA using Cisco Silicon

Or NCP compliant with Cisco Switches based on NVIDIA Spectrum ASIC

cisco

Cisco Reference Architectures

Aligned with NVIDIA Reference Architectures (RA)

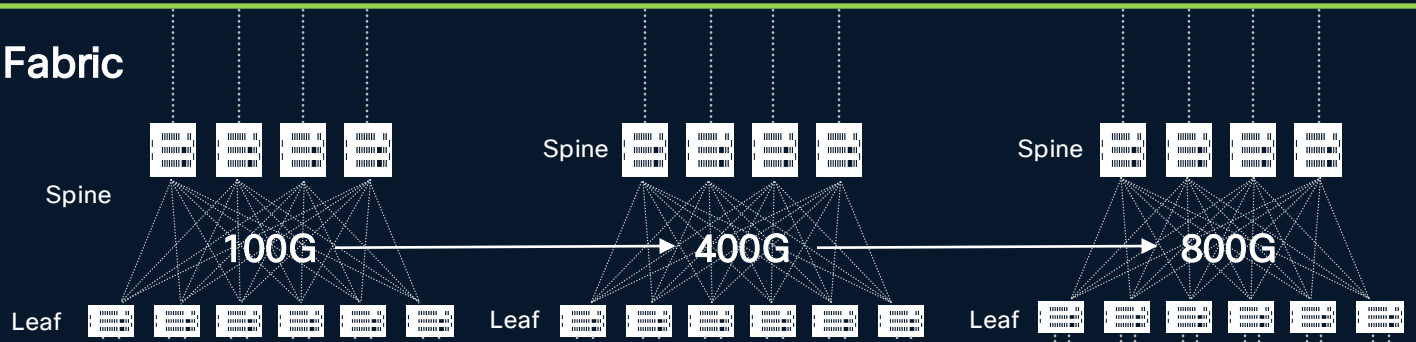


Speed Evolution in the Data Center

Inter-Data Center

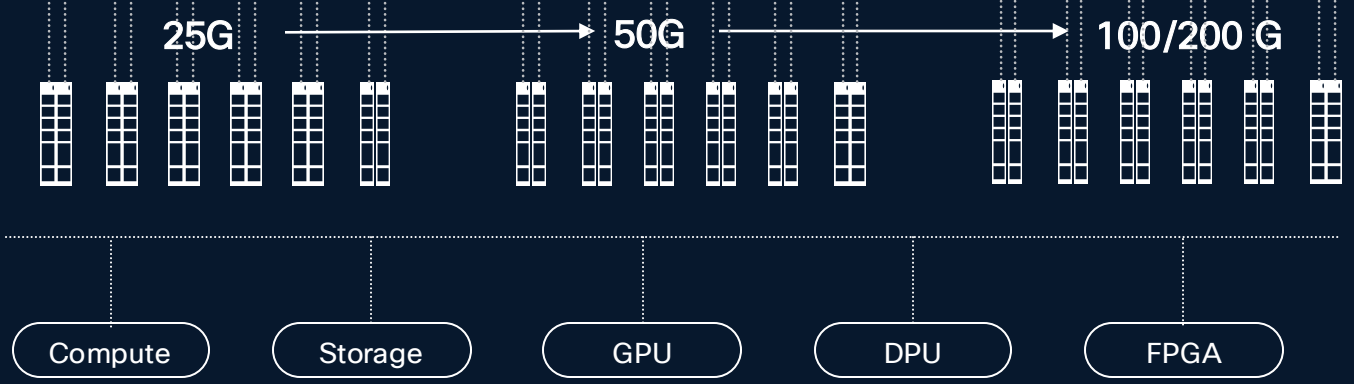
- Significant pluggable DCI (DWDM coherent)
- Open Line System

Switch Fabric



- Switch silicon bandwidth growing due to higher Radix and faster Serdes speeds
- Switch ASIC throughput growing: 6.4 Tbps to 12.8 Tbps to 25.6 Tbps to 51.2 to 102.4 Tbps
- Optics increasing from 100 Gbps to 400 Gbps to 800 Gbps

Access



- Server network connectivity evolves with server processor upgrade cycles as data center traffic grows
- Server port speed is transitioning from 25 Gbps to 200 Gbps
- Storage, GPU, DPU, FPGA driving connectivity bandwidth, PCIe speed increase

100% Pluggable Optics

Cisco Optics for AI Data Center

To meet the needs of AI Datacenters, Cisco leverages innovative Silicon Photonics, a Diverse supply chain, ensures quality with extensive Performance & Reliability Testing

Technology Innovation

Silicon Photonics Platform Innovation

Single Mode and Multimode Optics

Advanced Manufacturing Automation

Optics Design

High Quality Components

Performance Margins Beyond IEEE

Supply Chain Diversity

GF and TSMC SiPh Wafer Fab

Multiple DSP Sources

Multiple laser Sources

Mfg in Thailand and Malaysia

Robust Module Testing Protocols

Robust System Testing Protocols

Advanced Stress Testing

Testing at Scale

Unique to Cisco

Reliability

Extensive L1/L2 Qual

Unique to Cisco

Field Firmware Upgrade

Module FW can be Downloaded from software.cisco.com

NEW Q3'26

Reliable Connectivity for AI Data Centers

Cisco OSFP 1.6T Pluggable Optics

High-Performance and Reliability

Cisco Silicon Photonics Technology

Advanced DSP-Based Diagnostics and Monitoring

High Bandwidth AI Scale-Out

1.6T switch-to-NIC connectivity and high-density fabrics for massive AI cluster builds (2x800G, 4x400G, and 8x200G links)

Diversified Supply Chain

Seamless AI cluster scaling with diversified supply chain and interoperability across Cisco and 3rd party DSPs

Efficiency and Flexibility

Optics for power-efficiency and deployment flexibility in liquid cooled and air-cooled platforms

NEW Q3'26

Introducing AI Scale-Out Power Efficiency and Reliability for 800G Networks

Cisco 800G Linear Pluggable Optics (LPO) Solutions

CISCO L800G OSFP LPO

50% Less Power per Module

Powered by Cisco Silicon Photonics

Integrated Designs Reduce System Power Consumption 30%

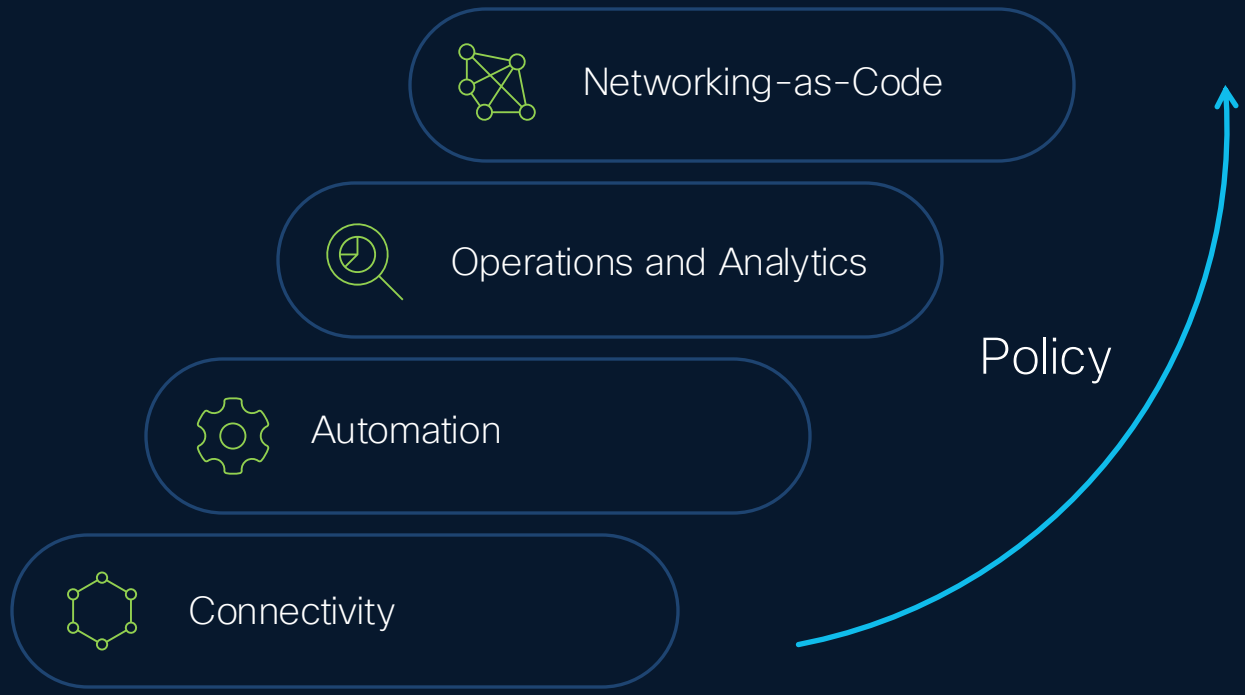
Shift power consuming optics DSP functionality to Cisco Silicon One platform

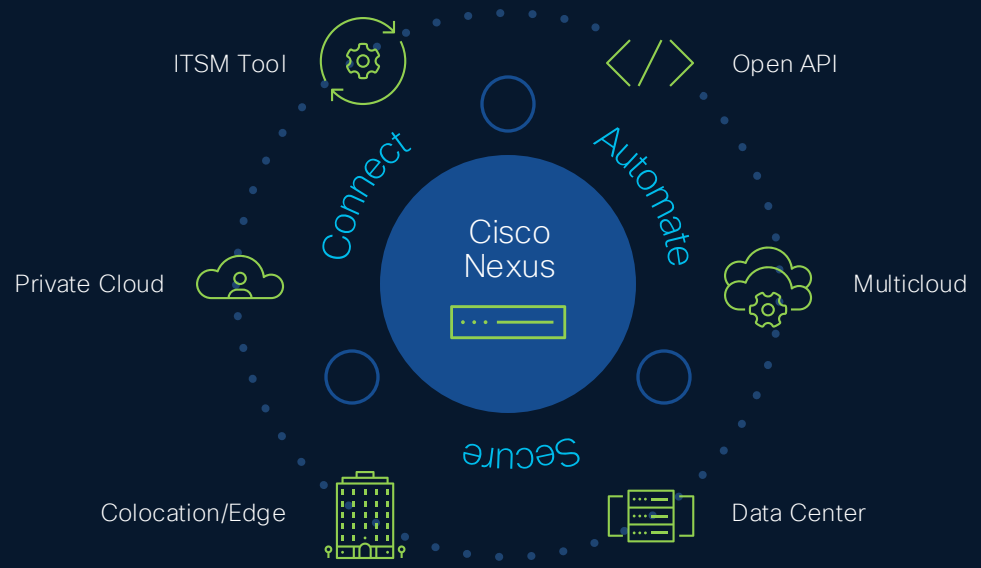
High signal integrity driven by Cisco Silicon Photonics and Cisco Silicon One

High reliability with lower thermal stress

Increase port density without PDU upgrades

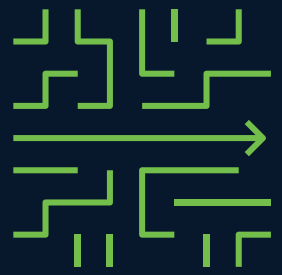
Cisco Data Center Networking with Nexus





Data Center | Colocation | SDWAN | Edge | Cloud

Cisco NX-OS Investment Areas

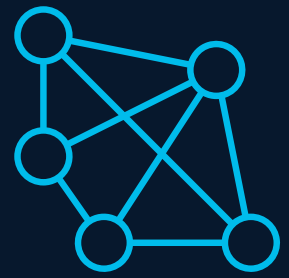


HW Enablement

Onboarding for diverse ASICs

Automation

Flexible Automation with Nexus Dashboard, open APIs, and DevOps



Fabric Innovations

Enhancements across various fabric architecture

AI/ML

AI/ML for networking blueprint

Security

Critical Security functionality and certifications support

Unified Fabric Experience with Nexus Dashboard

SOLUTION COMPONENTS

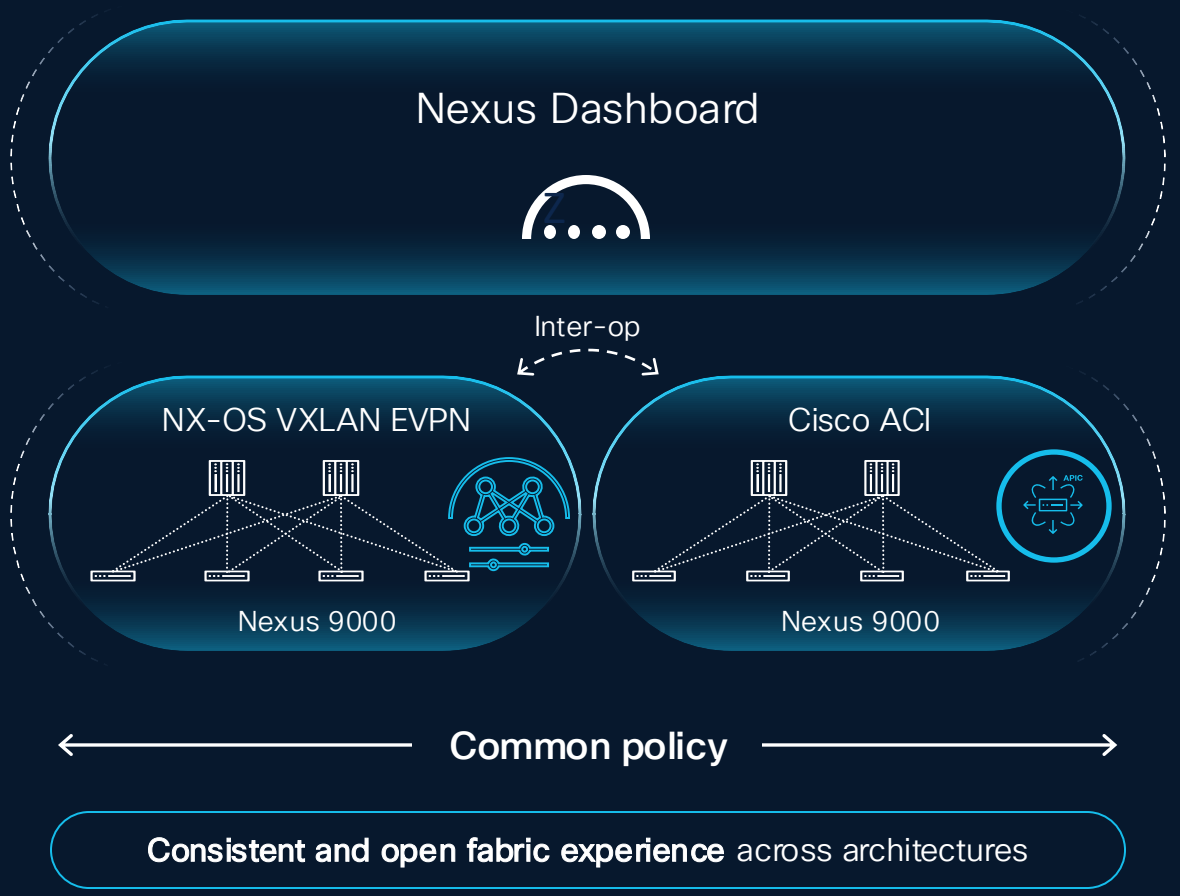
Zero-trust networking and micro segmentation based on endpoint / security groups

Advanced service chaining and service redirection

Standards based inter-operability with third-party networks using open networking protocols

Integration with CI/CD pipelines with DevOps-ready APIs

Administrative multi-tenancy

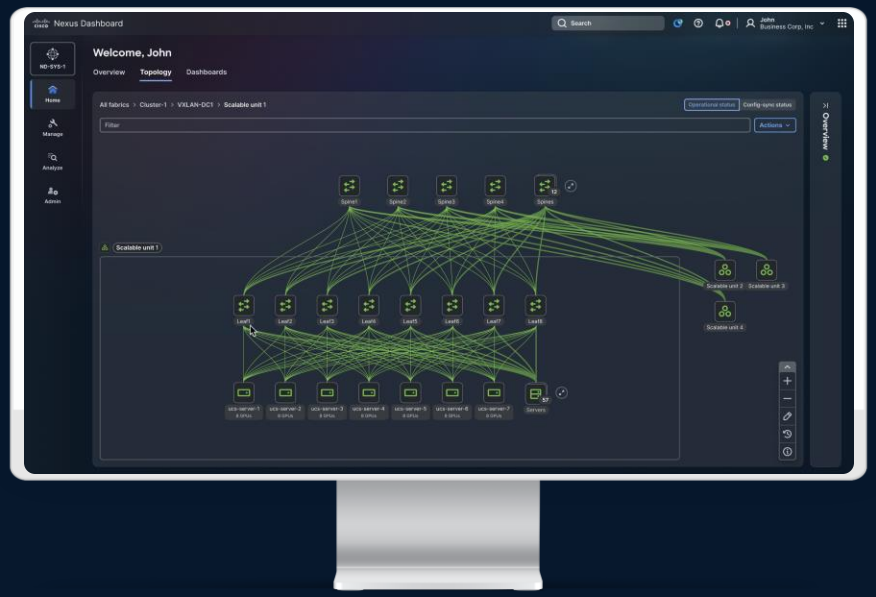


Choice of Operational Model

Purpose Built and Prescriptive AI/ML Fabric Builder

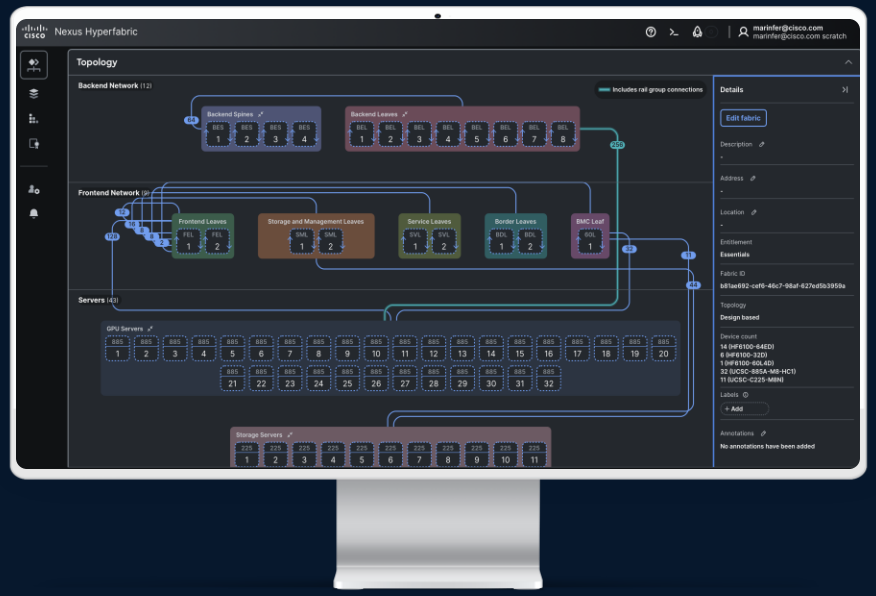
Nexus Dashboard

On-premises managed network



Nexus Hyperfabric

Cloud managed network



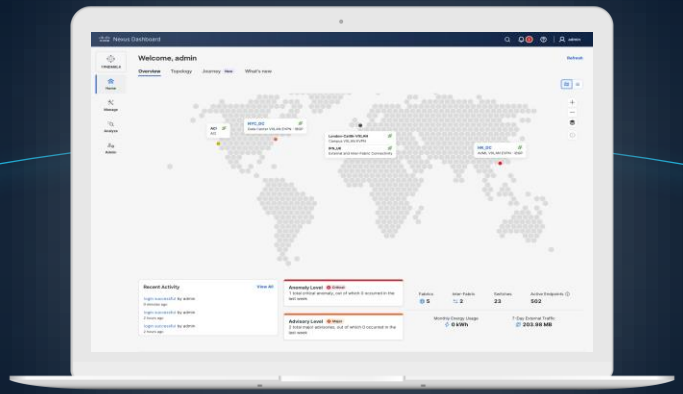
Fully automated fabric configurations with embedded best-practices for lossless networks

Topology-aware visualization and proactive troubleshooting

SLURM integration for job monitoring, correlation of job health with GPU metrics, NIC metrics, and network anomalies.

Cisco Nexus Dashboard

Simple, Secure and Sustainable



All Managed by Cisco Nexus Dashboard

Highly Secure Encryption
With end-to-end Segmentation

Flexible Fabric Deployments
Cisco ACI, NX-OS, SAN & Data Broker

Unparalleled Network Visibility
Within Nexus Dashboard

Enabling AI/ML workloads
With reduction in power utilization

Simple

Secure

Sustainable

Cisco Nexus Dashboard

Provisioning

Classic
2- or 3-tier NX-OS architectures

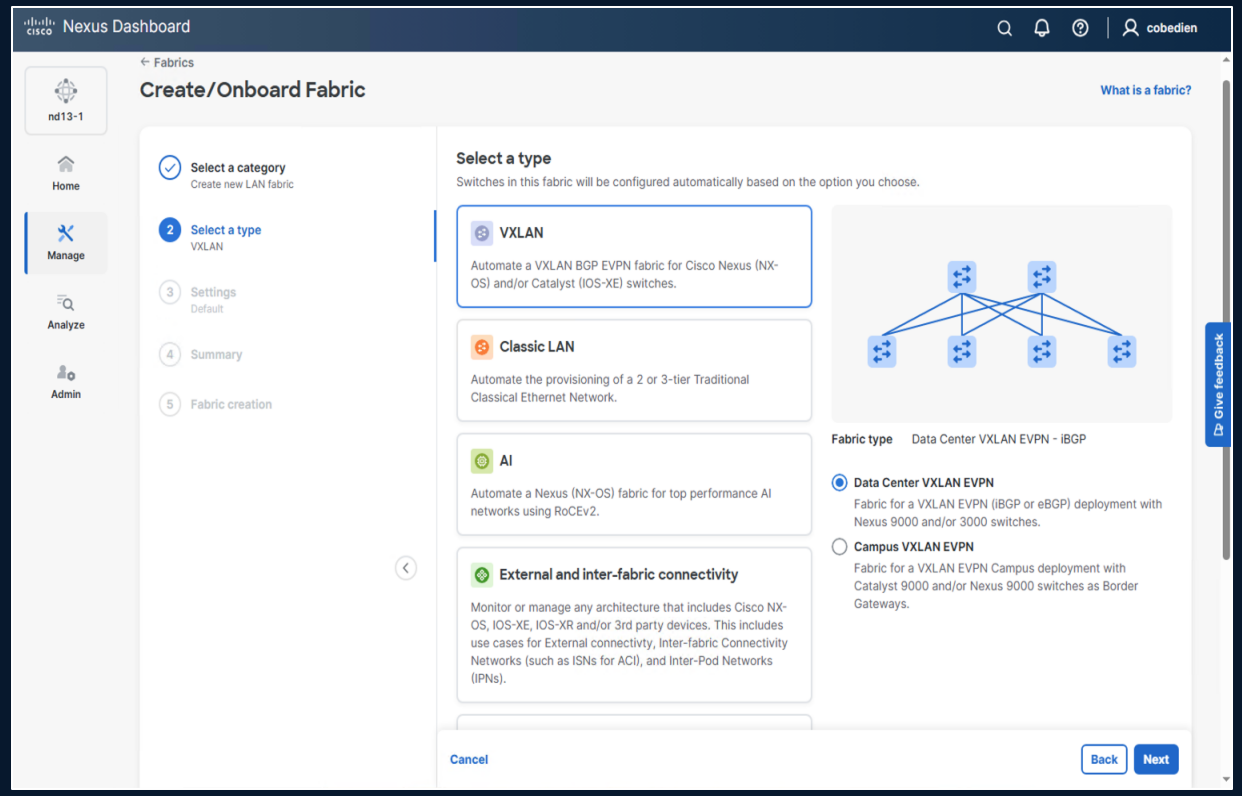
VXLAN
BGP EVPN fabrics for Nexus & Catalyst (IOS XE)

AI/ML
Fine-tuned network provisioning to transport AI/ML apps with the best performance

External and Inter-Fabric Networks
NX-OS, IOS XE, XR, third-party networks mainly used to interconnect ACI, NX-OS, and Campus

IP Fabric for Media
Fine-tuned network provisioning and monitoring for broadcasting and media

Routed
BGP-based CLOS NX-OS fabrics



Provision all your fabrics from a single place

Cisco Nexus Dashboard

Managing

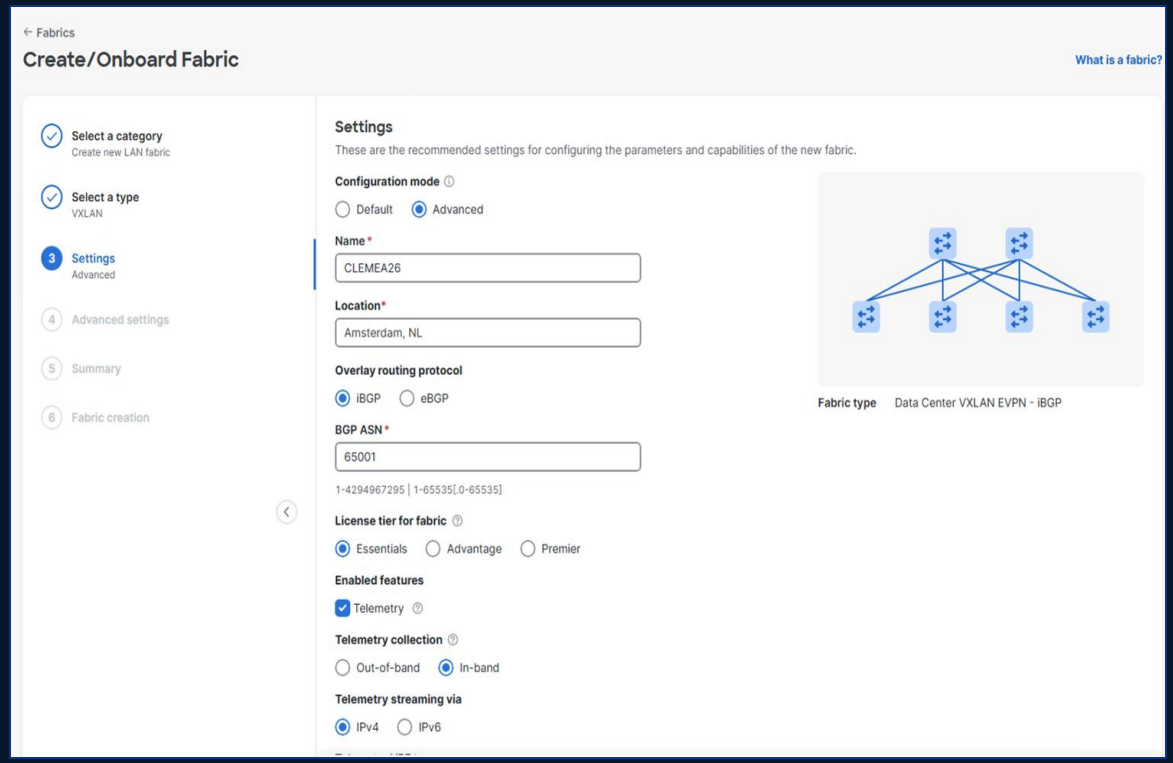
Configure once, deploy anywhere: Define your configuration or network policy and select where you want to apply it, accelerating the time to provision connectivity

Excellent User Experience: Set up connectivity to routers, switches, and other devices (incl. VRF-lite, MPLS) in a few clicks

Guided Workflows: A step-by-step guide detailing how to build the connection, what is required, and where to implement

Traffic steering and redirection: Send traffic to firewalls, load balancers and other L4-L7 services or service chains by using Policy-Based Redirection (PBR)

Synchronize changes: After you perform a change or if the configuration running locally on a switch or fabric is not what you expect, you can re-sync and deploy to maintain consistency



Cisco Nexus Dashboard

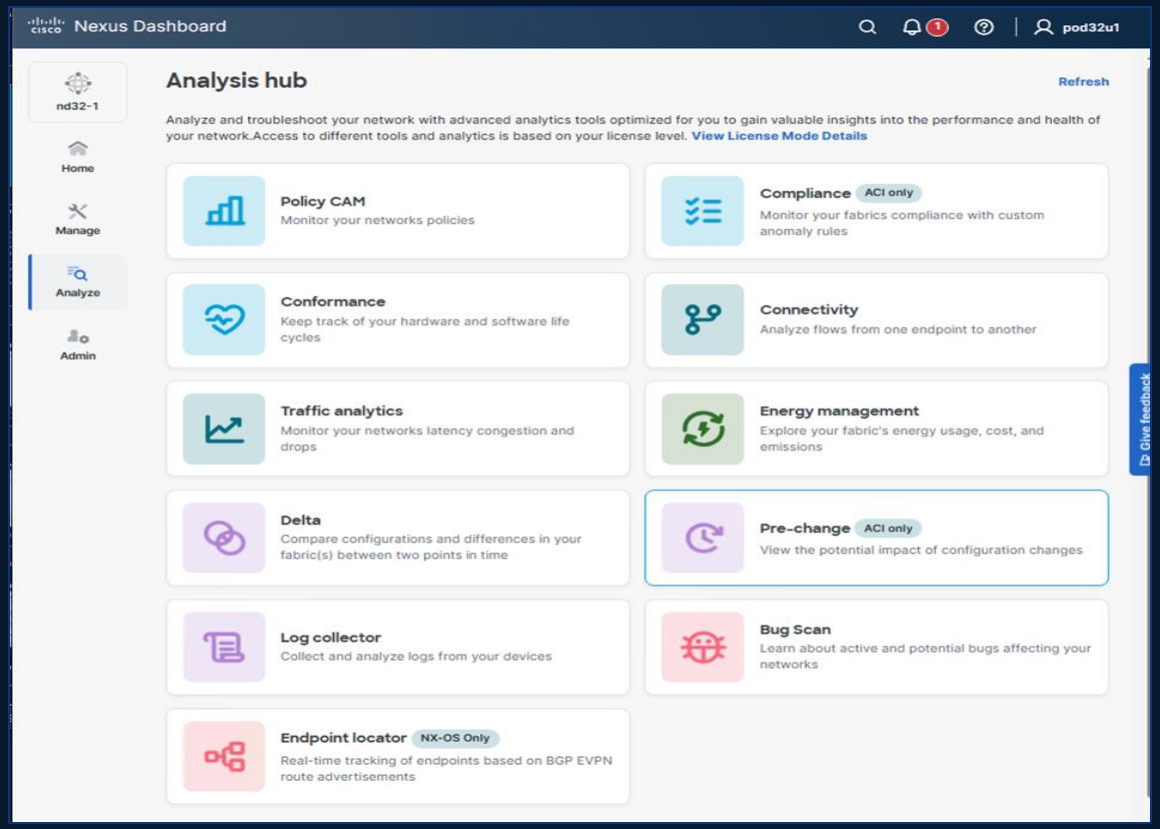
Analyzing

Updated topology : View fabrics, switches, interfaces, and endpoints with their corresponding anomaly scores

Granular visibility for every connection: From overall fabric score to category, service, and connection, Traffic Analytics can monitor individual client-to-service sessions and allows you to “tap-in” by capturing flow records on demand

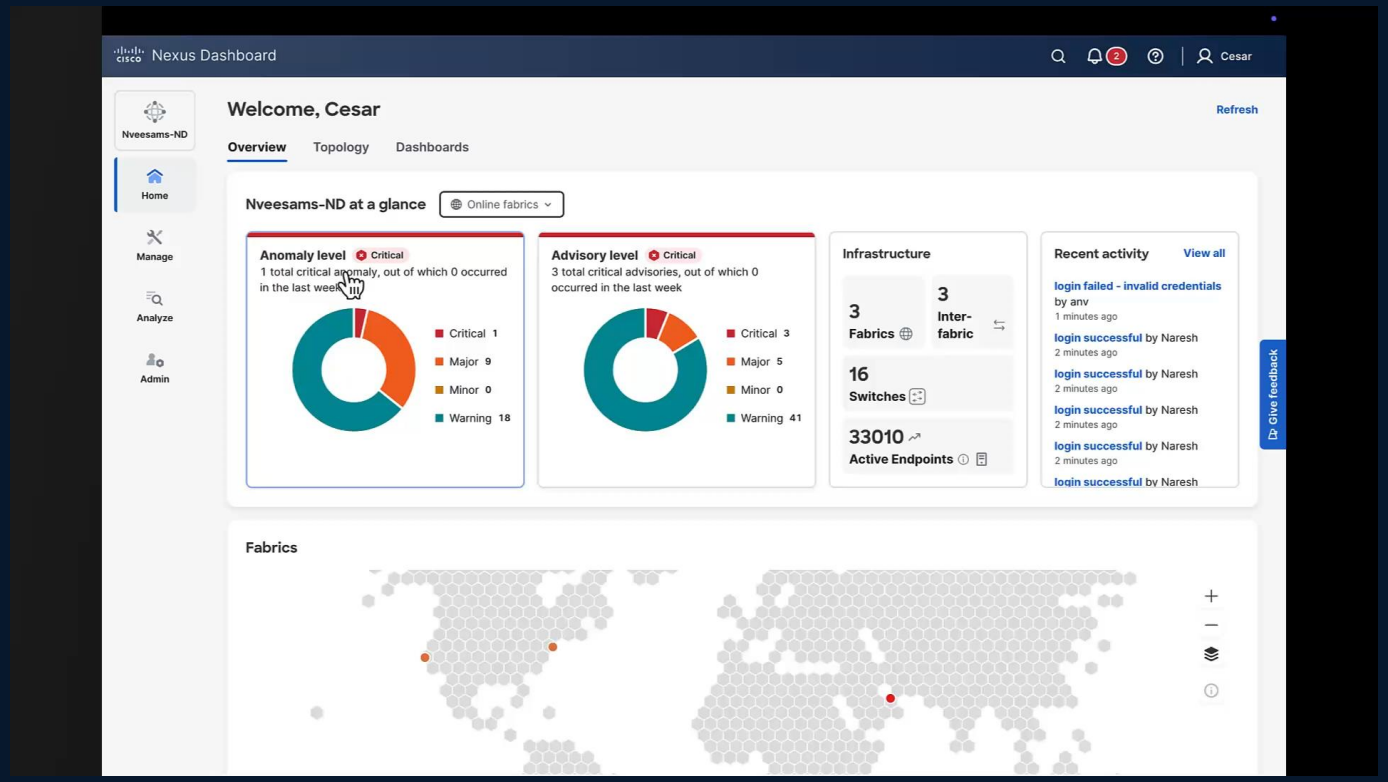
Host and flow path visualization: View discovered hosts, end-to-end flows, and multicast NAT information to aggregate multicast flows per sender/receiver

Conformance: Keeps track and automatically checks for Hardware and Software support and Verified Scalability.



Cisco Nexus Dashboard

Traffic Analytics



Monitor Networks Latency, Congestion and Drops

Cisco Nexus Dashboard

Optimize Experiences with End-to-End Visibility

splunk> a CISCO company

Observability & Security

For networking, IT , security and engineering teams

Cross domain full stack observability with Application, Networking, Security and AI

Airgap / Cloud native including application dependency mapping, IT service intelligence and Enterprise security

Single DCN app for Splunk Enterprise and Splunk Cloud for ND,ACI and N9K

Nexus Dashboard with Splunk Embedded

ACCEDIAN

Assurance

For network and IT teams

Cloud-native AI service assurance for service provider/telco using nexus Dashboard and Provider connectivity assurance

Network performance, enriched connectivity analysis, DC to external services

Accedian SFP recognition, Automation of agent installation using Nexus Dashboard

ThousandEyes

Assurance

For network and IT teams

CloudAI driven closed loop assurance for Airgap, cloud, and hybrid Enterprise networks

DC configuration correlation, active synthetics, proactive recommendations for Inter DC, Intra DC, and Hybrid insights

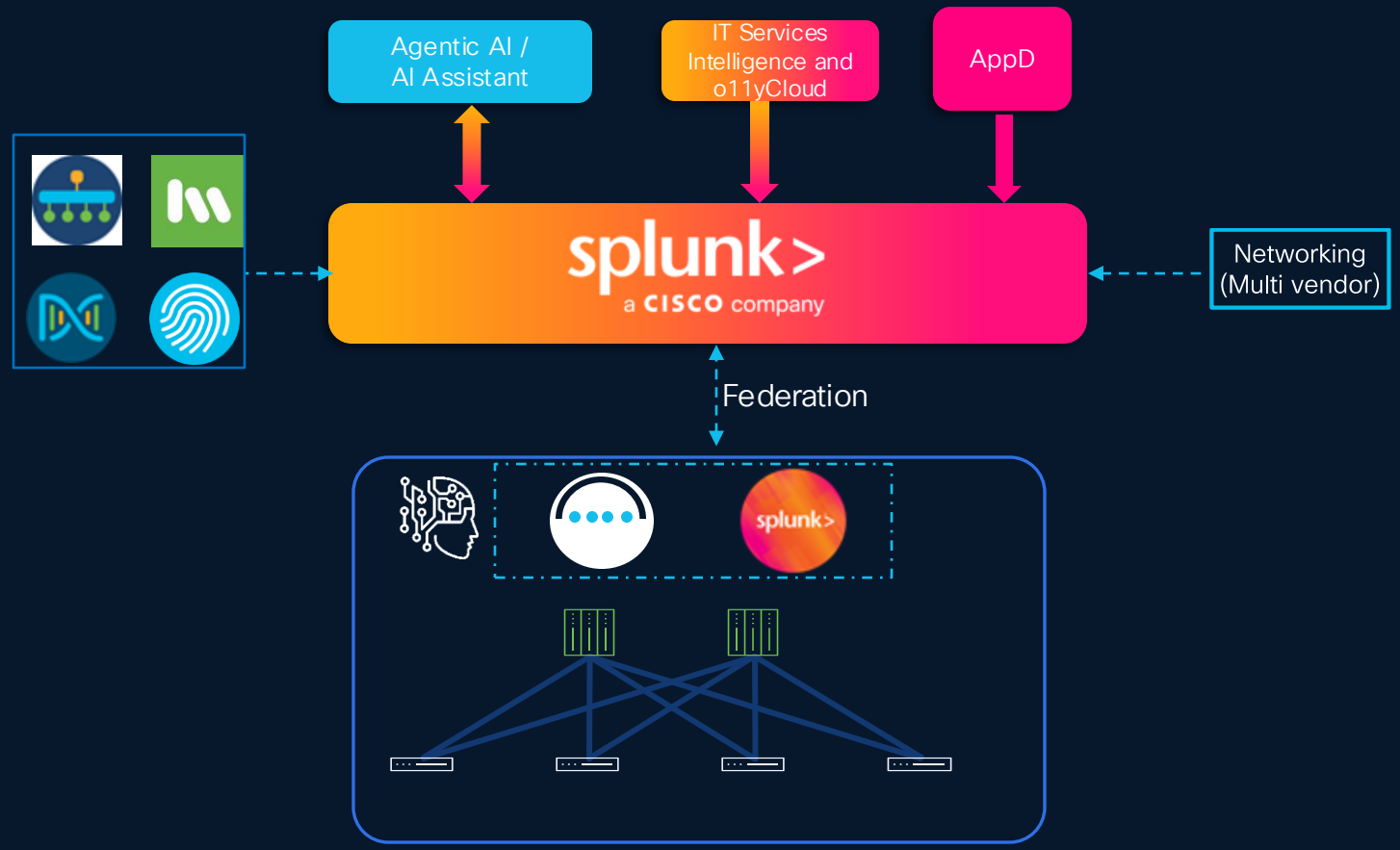
Automated ThousandEyes agent installation, Internet insights, application performance

AI driven assurance and observability: reactive to proactive

cisco

Committed

Embedded Splunk with Nexus Dashboard



Unified Operations Management: Configuration and Analytics across on prem, cloud and hybrid deployments

Unified analytics platform : Realtime custom dashboarding , reporting and alerts and custom action. Federation with Splunk platform

AI Observability: On Prem AI predictive analytics for data center networking data including AI fabrics

Audit and Compliance : Incremental storage of data for audit, compliance and historical data

Nexus Dashboard + Splunk = Better Together

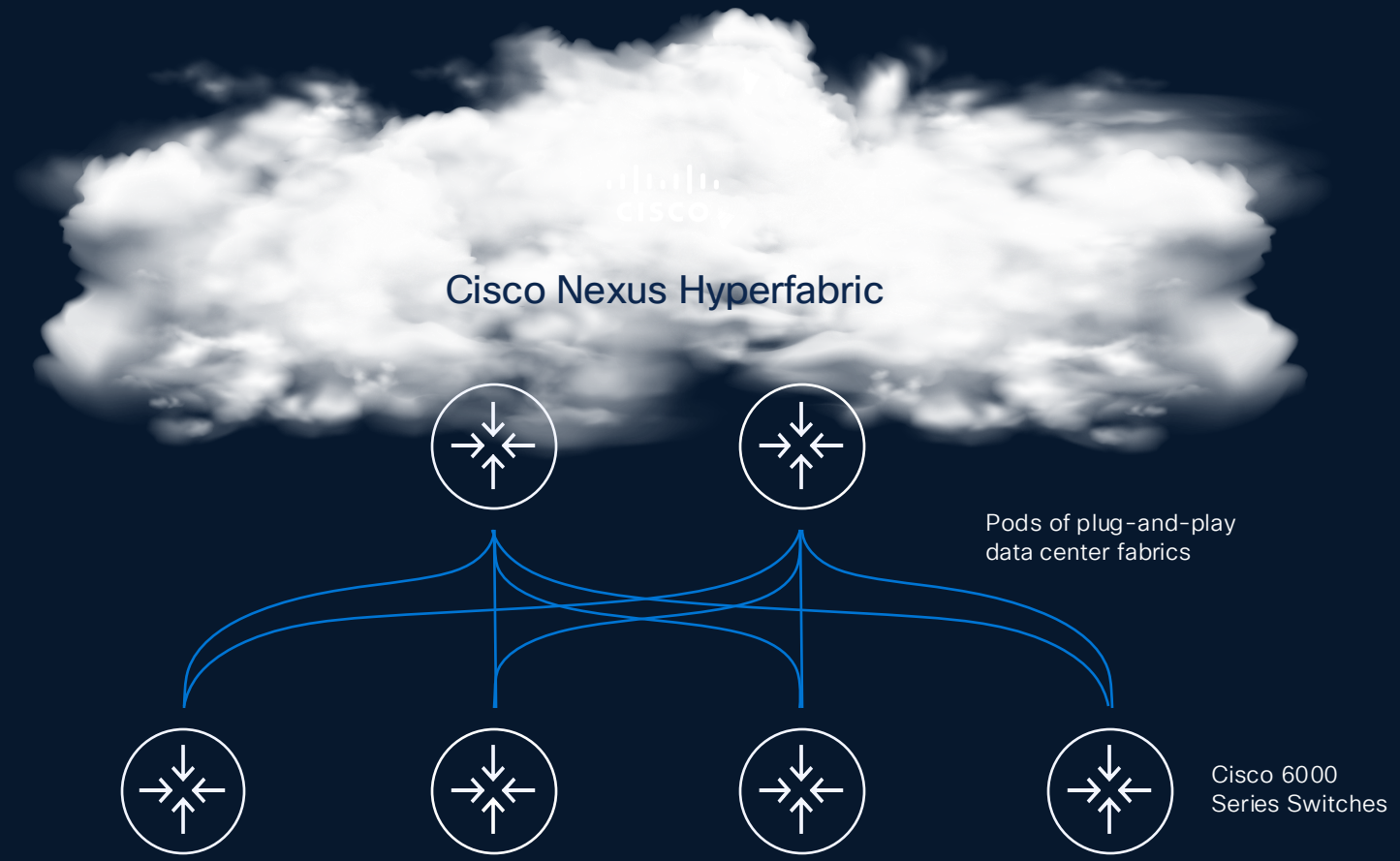
Cisco Nexus Hyperfabric

Data Center Network Fabric-as-a-Service

Design, deploy and operate on-premises fabrics located anywhere

Easy enough for IT generalists, application and DevOps teams

Outcome driven by a purpose-built vertical stack



Design | Order | Deploy | Validate | Monitor | Upgrade | Collaborate

Network for AI

High-Performance AI fabric

UEC Ready, high density 400/800G fabric

Intelligent Packet Flow

AI-aware traffic management suite

Architected for Sustainability

Responsible deployments

Liquid cooled devices

A Holistic Approach to AI Networking

AI for the Network

AI Powered Network Operations

Cisco AI Assistant in Nexus Dashboard

Cross-Domain Insights

Unified insights from network to application through platform integrations

AI Job Observability

Real-time visibility and monitoring for proactive troubleshooting

Front-end network | Back-end network | Storage network

Cisco Brings Open, Flexible Infrastructure to the Secure AI-Ready Data Center

Cisco AI Networking

AI Models

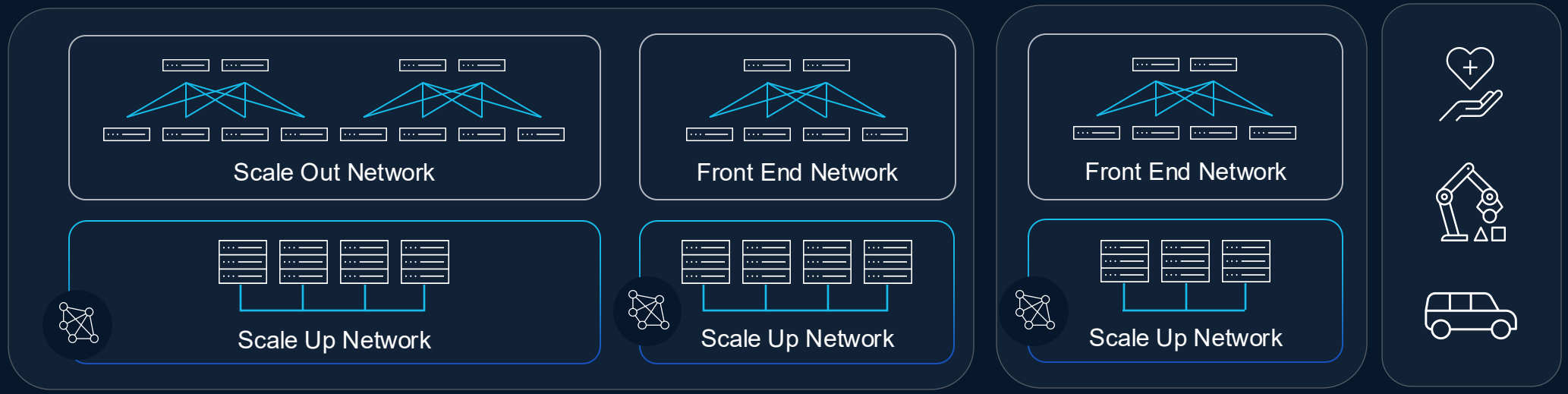
Foundational Models, Small Models, Agentic Applications, Intelligent Edge

AI Connectivity & Control

Agentic Networks, AI Gateways, MCP Proxies, AI Security

Scale Across

AI Network Infrastructure



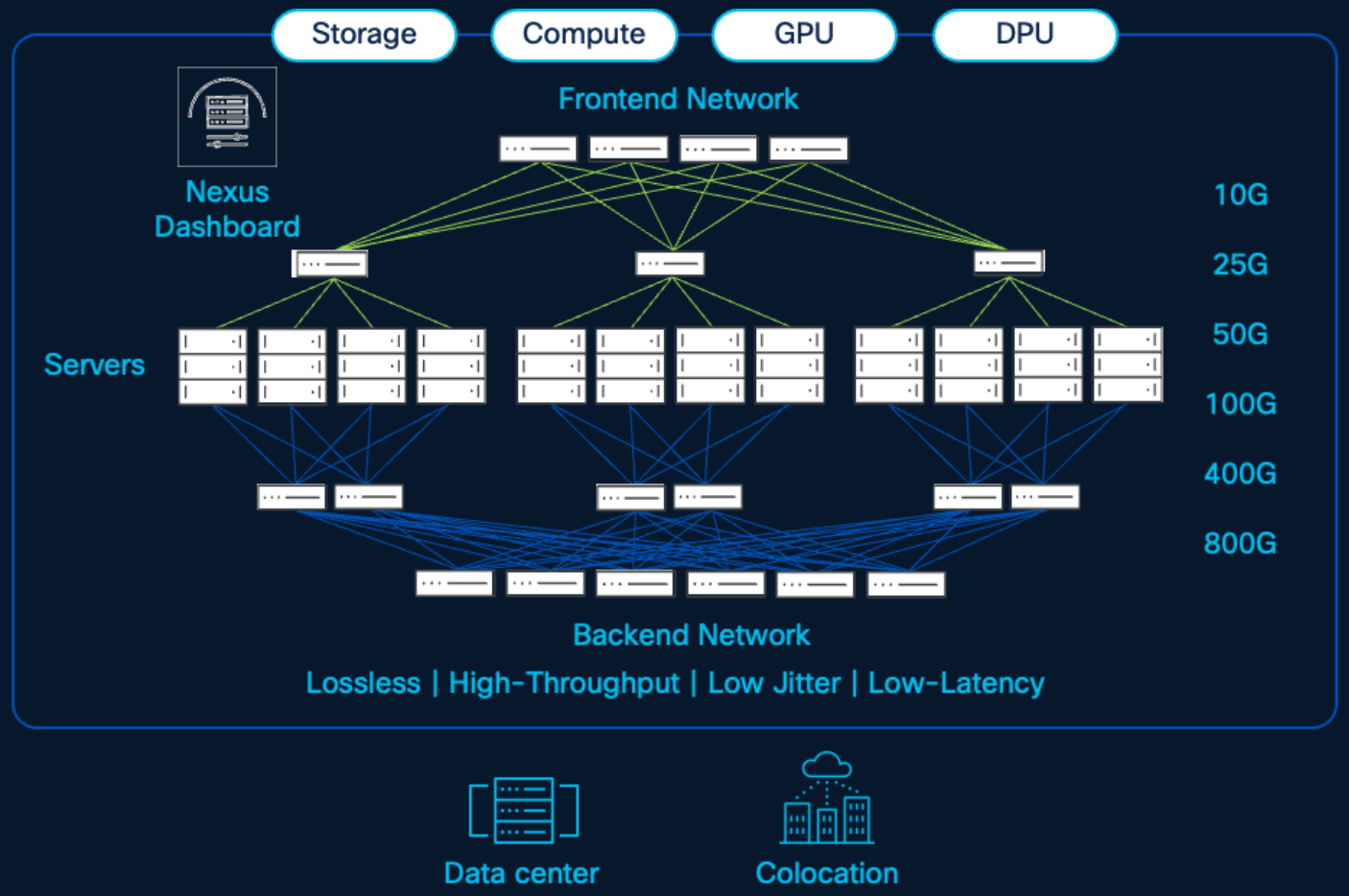
Cloud DC

Edge DC

IOT Edge

Cisco's AI/ML Approach with Nexus

Silicon, Systems, Software, Operations



400/800G Ethernet Transition(25.6T & 51.2T switches)

High-bandwidth fabrics with reduced footprint and energy savings

---

RDMA over Ethernet (RoCEv2)
Non-Blocking Lossless network (PFC + ECN)

Powered By Intelligent Packet Flow
Advanced congestion aware Load Balancing, hardware accelerated telemetry & fault-aware recovery

---

AI fabric templates, AI analytics, telemetry, congestion scores

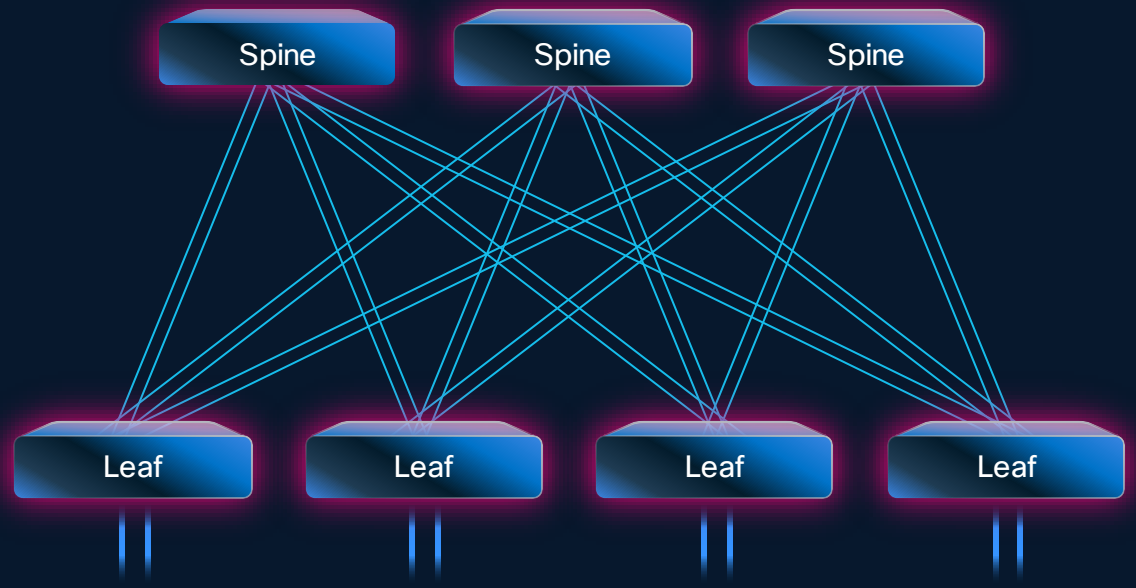
---

Validated designs for networks and ecosystem partners
AI/ML Blueprint

Powering AI Fabrics with Intelligent Packet Flow

Unified Architecture

Ultra Ethernet READY



Cisco Intelligent Packet Flow

- Advanced Load-Balancing
- Hardware-Accelerated Telemetry
- Fault-Aware Recovery

cisco

Compute AI Portfolio

Address AI Workloads with Visibility, Consistency, and Control

Validated solutions for AI with compute, network, storage, and software

Build the Model
Training

Optimize the Model
Fine-Tuning and RAG

Use the Model
Inferencing

RTX PRO SERVER
Supporting RTX PRO 6000 Blackwell Server Edition GPUs

Cisco UCS®
GPU-dense servers
PCIe and NVLink Servers

Cisco UCS blade (with GPU extensions) and
rack servers

Enterprise AI edge

Dense Compute for Demanding AI

Full-Stack AI with Compute and Networking

cisco

Cisco AI PODs

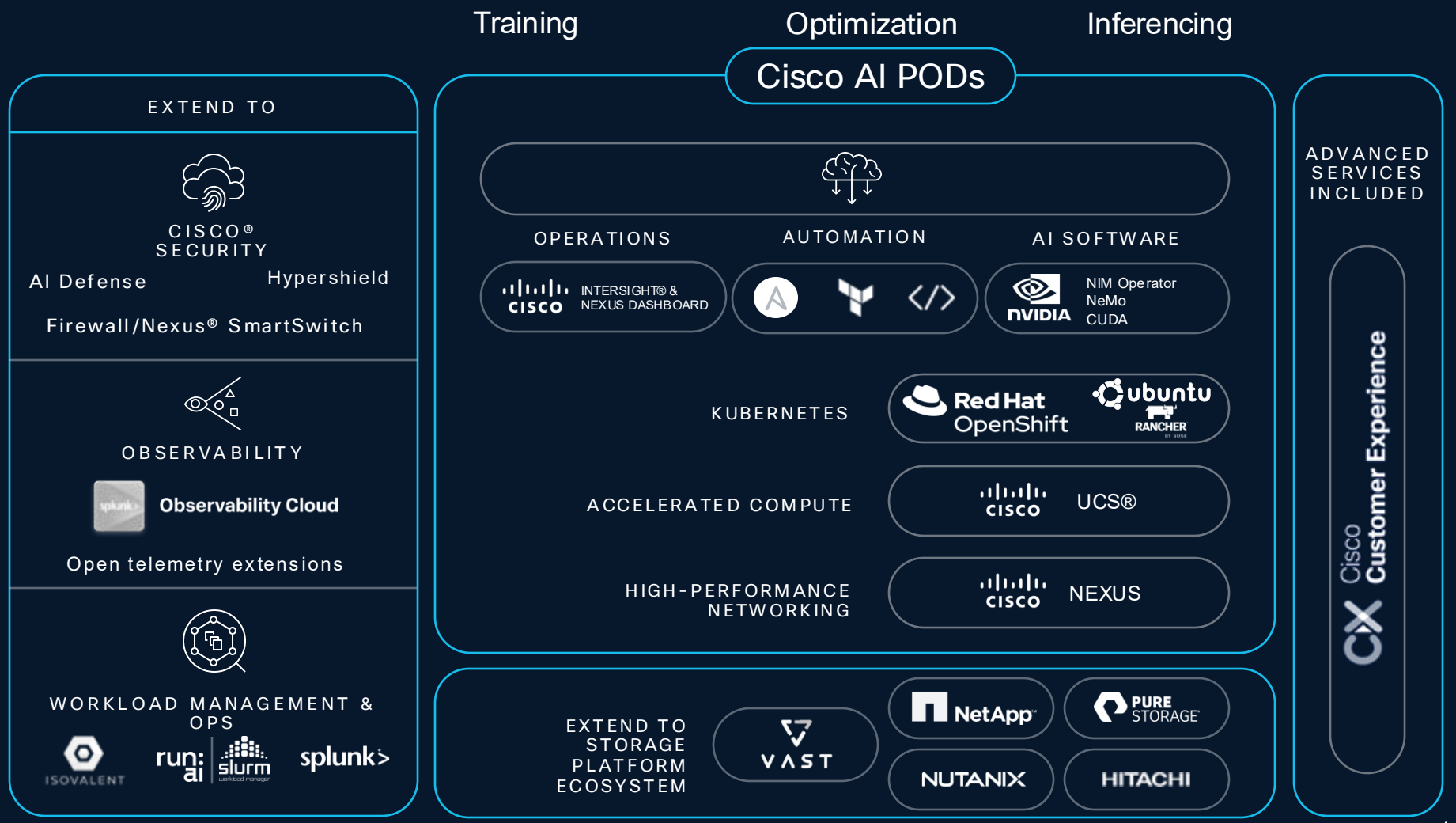
A Scalable Architecture, Built to Support Any AI Workload Simply & Efficiently

Deploy AI with confidence
Cisco CVD, NVIDIA ERA

Fully supported stack including Cisco and 3rd party components
Cisco CX Success Track

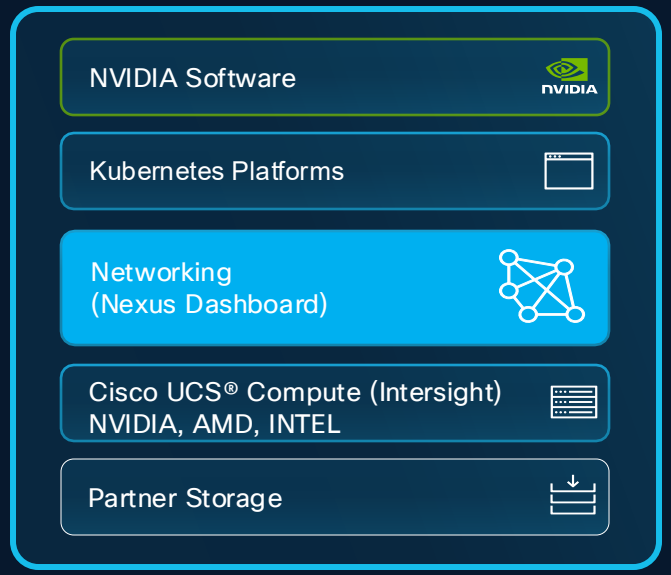
Orderable, use case driven AI-ready infrastructure stacks
Inferencing. Optimization. Training.

Incremental, atomic-level -or- fabric-based cluster scale



Cisco AI PODs: Flexible Operating Models

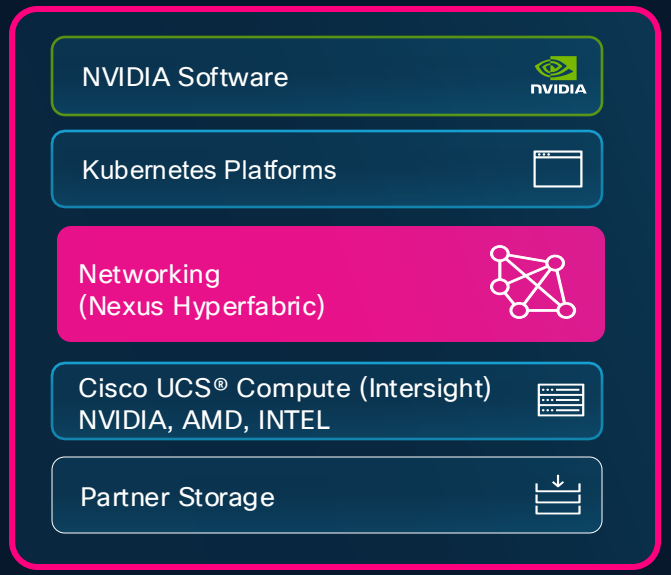
AI POD w/ On-prem management



Modular, pre-validated infrastructure:

- Full stack, buy & deploy
- Nexus Dashboard: On-prem networking management

AI POD w/ Cloud management

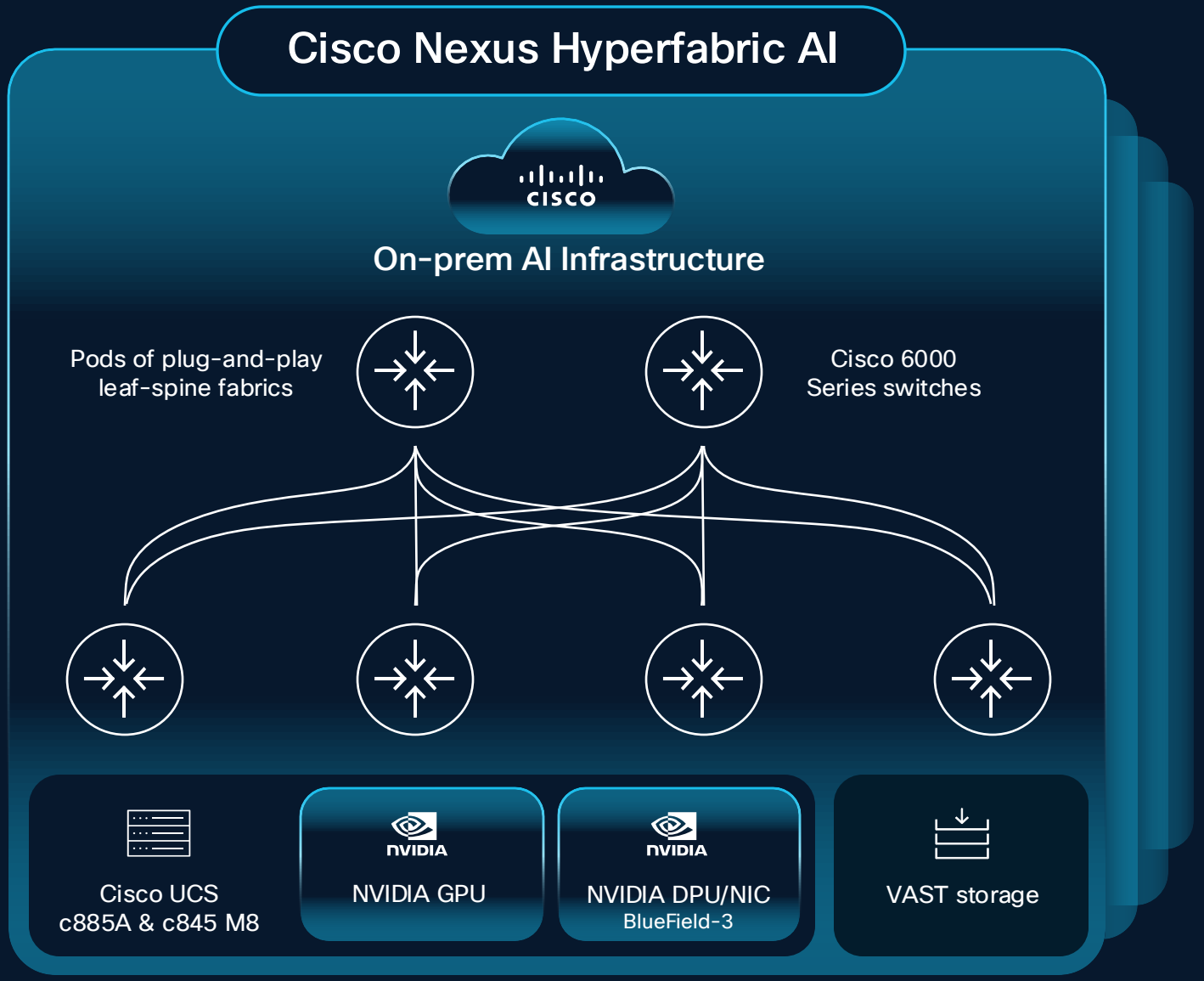


Turnkey infrastructure:

- Full stack, buy & deploy
- Nexus Hyperfabric: Cloud-managed Networking
- Nexus Hyperfabric AI: Cloud-managed physical infrastructure

Cisco Nexus Hyperfabric AI

- High-performance Ethernet
- Cloud-managed operations
- Unified stack, including NVAIE
- AI-native operational model
- Democratize AI infrastructure
- Visibility into full-stack AI



Cisco AI PODs

Included in Cisco Secure AI Factory with NVIDIA

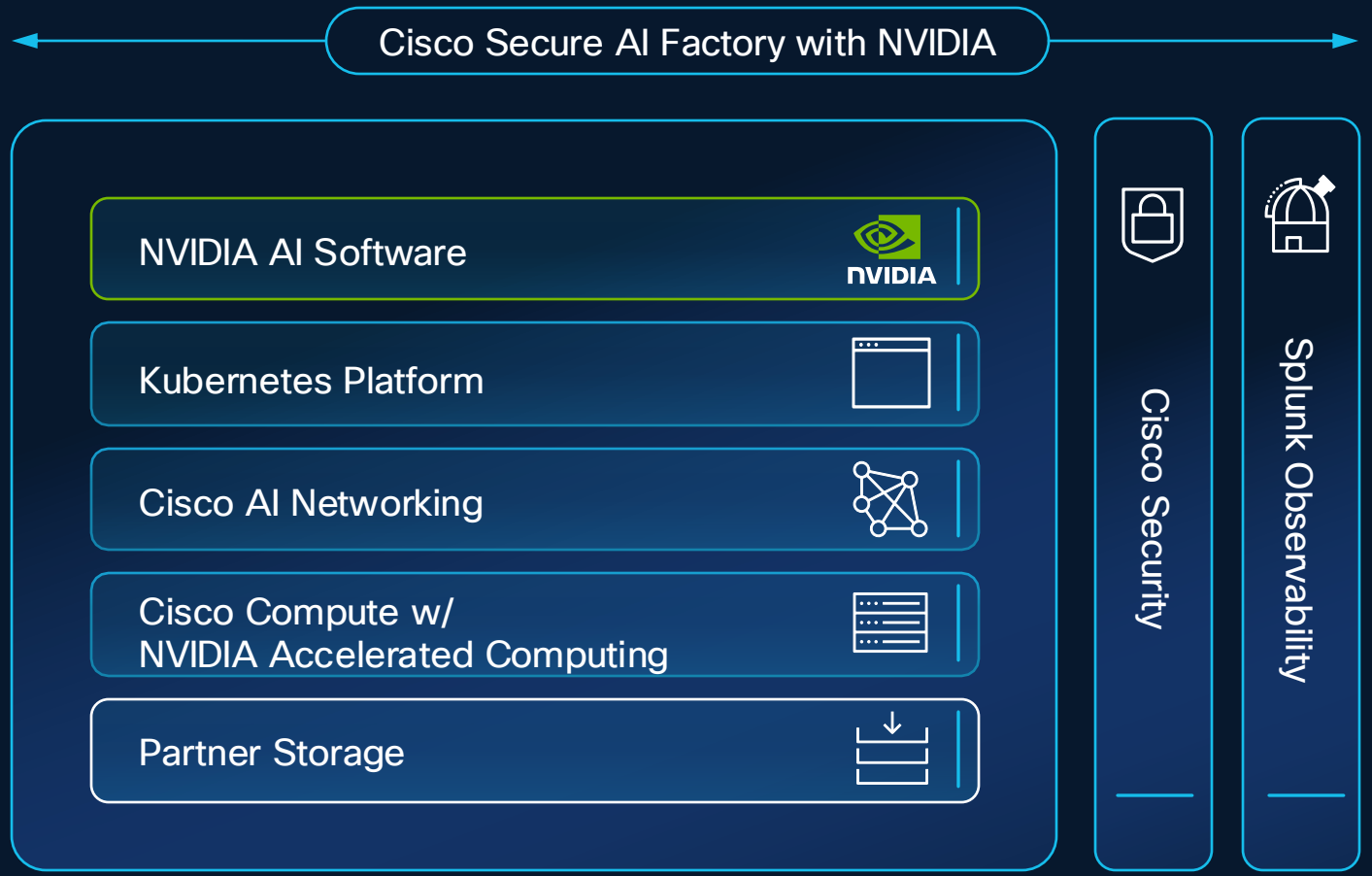
Why Cisco AI PODs?

- Security-first architecture enables safe enterprise AI
- Unmatched performance AI infrastructure enables efficient model training, customization, and inferencing
- Pre-validated AI infrastructure stack for simplified deployment drastically reduces set-up time

Cisco Secure AI Factory with NVIDIA

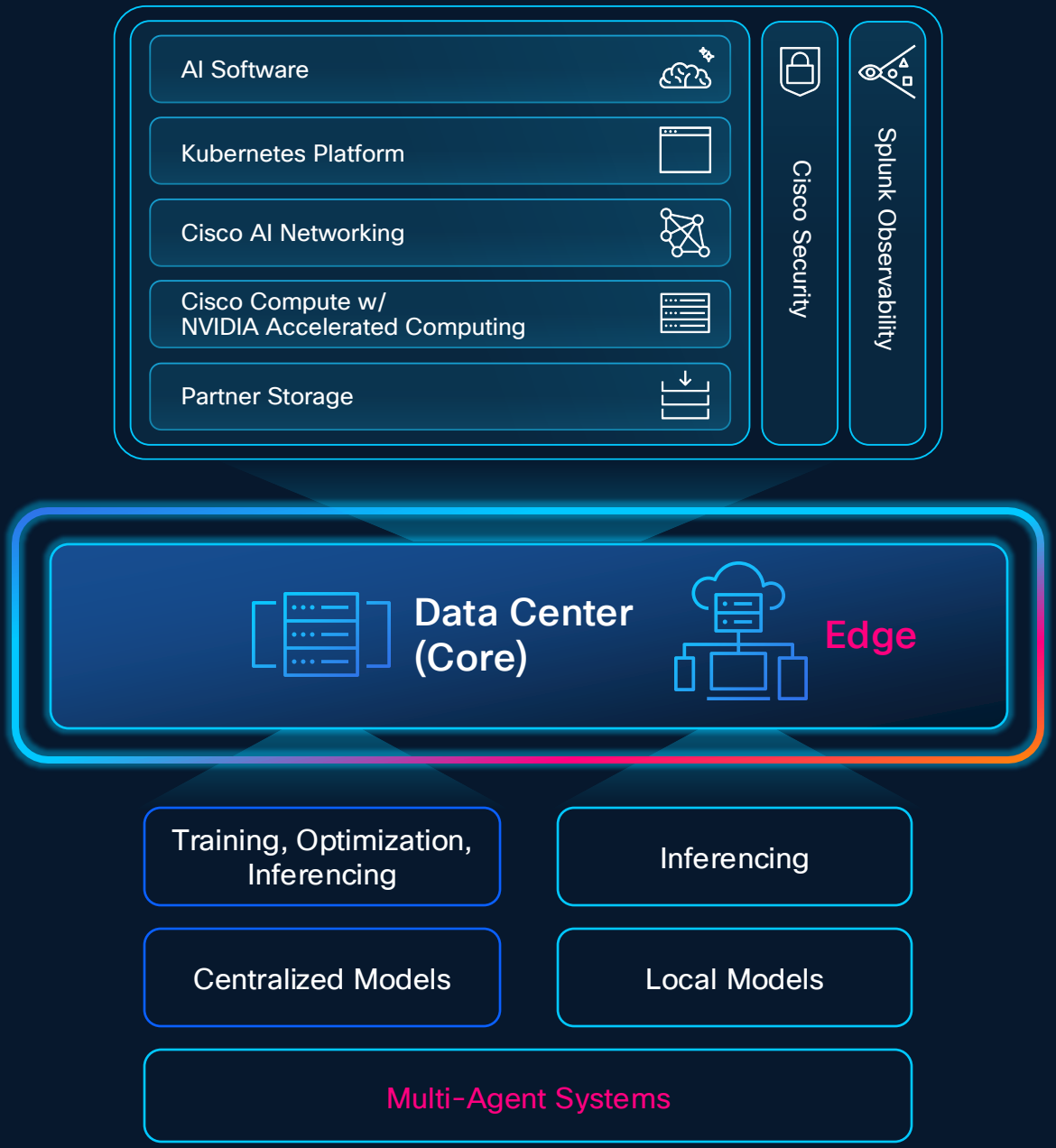
Delivering Trusted AI Outcomes

A modular reference design that combines high-performance infrastructure with full-stack security and observability



Evolution of Cisco Secure AI Factory with NVIDIA

Extend Secure AI Factory from Core to the Edge



Cisco Differentiation

The Security

Security-first architecture enables safe enterprise AI

The Network

High-performance integrated AI networking enables efficient model training and inferencing

The Assurance

Pre-validated AI infrastructure stack with flexible deployment options improves data scientists and developer productivity

Thank you


CISCO