CISCO Engage | Tech Day

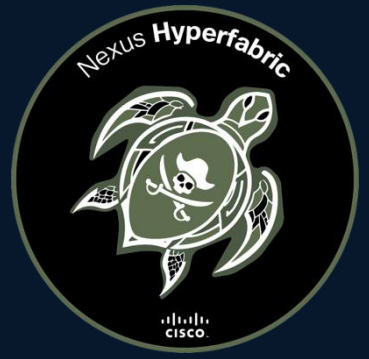
Nexus Hyperfabric

Dan Backman dbackman@cisco.com

Product Architect, Distinguished TME

2025-12-11

Agenda



1. Introducing Nexus Hyperfabric
2. Demo 1
3. The Hyperfabric Experience
4. Demo 2
5. Hardware Support
6. Multi-Site & Roadmap
7. Hyperfabric AI

cisco

Disclaimer

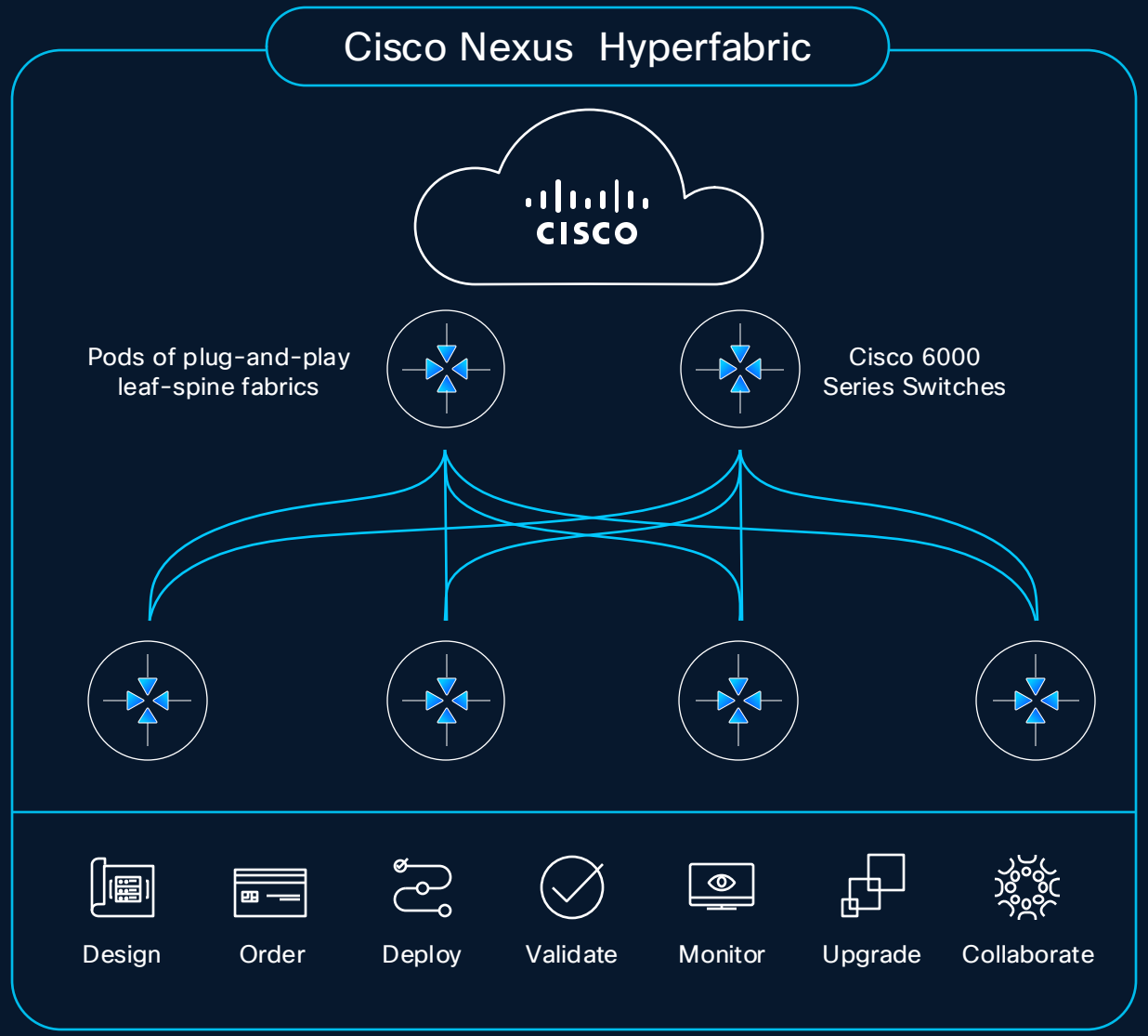
Some of the features described herein remain in varying stages of development and will be offered on a when-and-if-available basis.

This roadmap is subject to change at the sole discretion of Cisco, and Cisco will have no liability for delay in the delivery or failure to deliver any of the products or features set forth in this presentation.

Available Now

Cisco Nexus Hyperfabric

- Design, deploy, and operate on-premises fabrics located anywhere
- Streamlined operations for IT generalists, application, and DevOps teams
- Outcome driven using purpose-built vertical stack

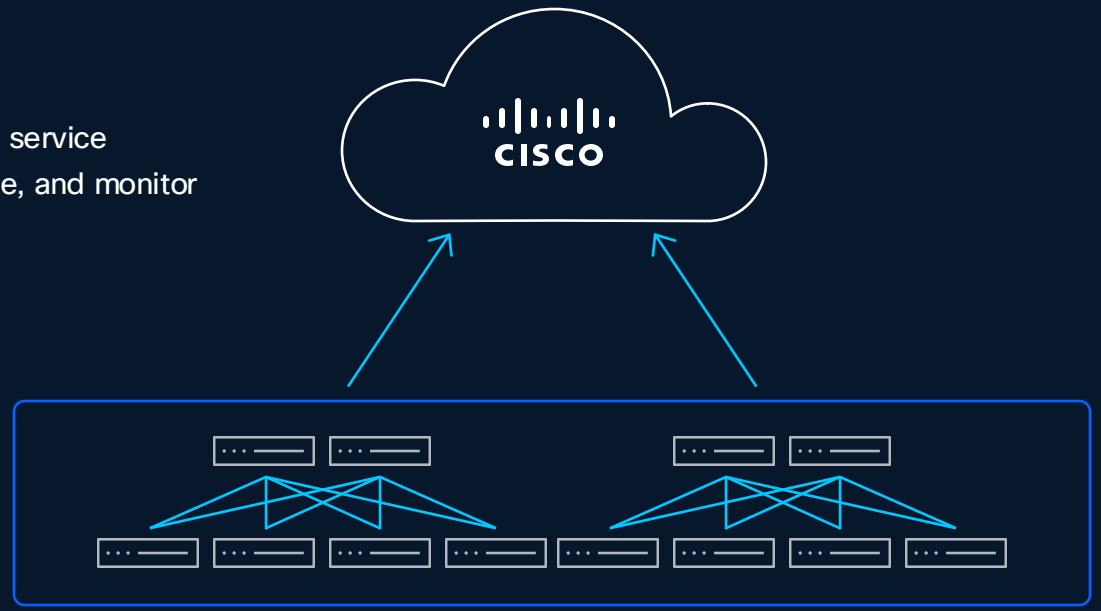


cisco

Nexus Hyperfabric components

Cloud controller

- Scalable, distributed multi-tenant cloud service
- Design, plan, configure, deploy, upgrade, and monitor
- Browser, API, and mobile access



Cloud-managed switches

- Boot-strapped from cloud
- Full visibility and control from the cloud
- Silicon One®-based performance and capabilities

High-performance fabrics

- I0/25/100/400/800 GbE connectivity
- Standards-based EVPN/VXLAN fabric with IPv4/IPv6 routing
- Mesh and/or spine-leaf fabric designs
- Horizontal scale

On-site web portal

- Step-by-step deployment tasks
- Registration and cabling
- Real-time validation

Use cases

Single global GUI/API endpoint for all owned fabrics

Data center modernization

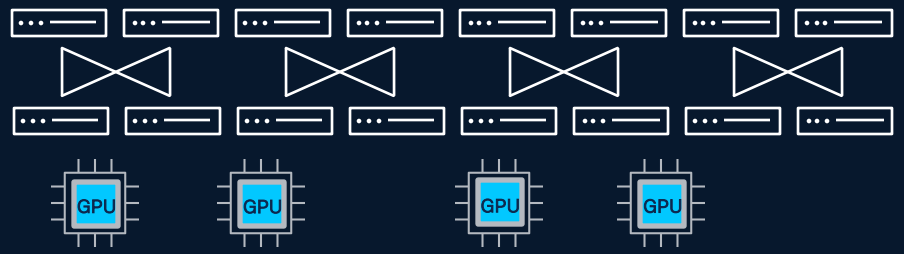
- Simplified operations, less complexity
- Scalability and flexibility
- Cost efficient, rapid time to value

Edge data centers

- Zero-touch provisioning
- Scales from small mesh to large spine/leaf
- Built-in security and high availability
- Traditional services and BYO AI inference

Nexus Hyperfabric

New



Nexus Hyperfabric AI for AI/ML/HPC clusters

- Full stack: NVIDIA GPUs, Cisco UCS® and switches,
- optional VAST storage
- Aligned to NVIDIA Enterprise Reference Architecture
- High bandwidth, low latency
- Low operational overhead

Manage multiple customer data centers

- Rapid and predictable service delivery
- Flexibility to scale up or down easily
- Operational support system–friendly integration

cisco

Cisco Data Center Networking portfolio

	Nexus Hyperfabric	Nexus Dashboard
	 Hyperfabric	 ACI®  NX-OS
Operating model	Fabric-as-a-service Cisco® cloud-managed controller	Customer-managed On-premises controller
Flexibility and customization	Prescriptive ⟷	Customizable
IT staff network skill set	Generalist ⟷	Specialist

cisco

Demo

Hyperfabric Hardware

New!

Leaf: HF6100-60L4D

- 4x 100/400GbE QSFP56-DD (16x 100G breakout)
- 60x 10/25/50GbE SFP56

Spine/Leaf (400G): HF6100-32D

- 32x 100/400GbE QSFP56-DD
- 128x 100GbE via 400:4x100 breakout

AI Spine/Leaf (800G): HF6100-64ED

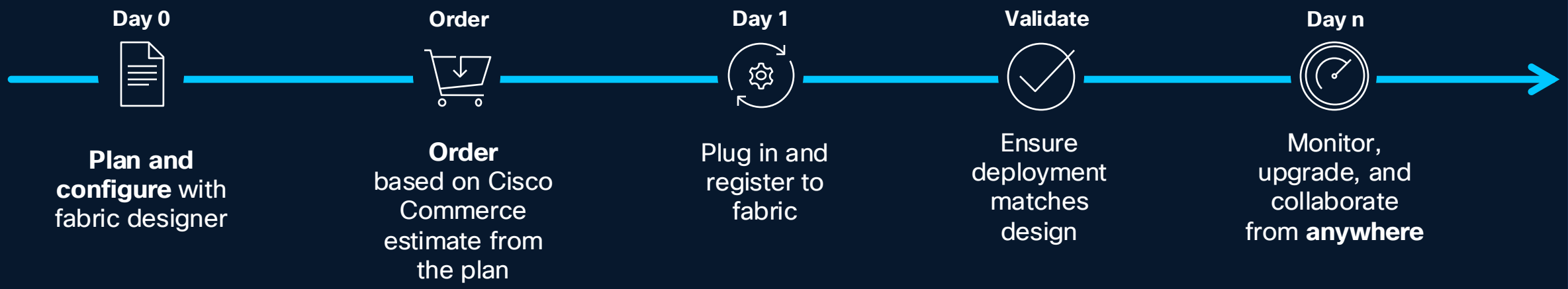
- 64x 800GbE OSFP800
- 128x 400GbE via 800:2x400 breakout

NEWS! Nexus 9300 HW in Hyperfabric

- Nexus C93108TC-FX3 (48x 1/10G Cu + 6x 100G QSFP+)
- Additional N9300 platforms to follow in CY2026+

cisco

How it works



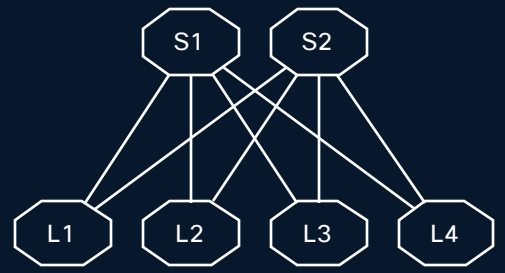
The Hyperfabric Lifecycle

- **Design:** Use the intuitive GUI to create validated fabric blueprints, generating BOMs and cabling plans
- **Deploy:** Switches auto-connect to the cloud controller for provisioning
- **Manage:** Manage fabrics through a single interface
- **Scale:** Manage updates, seamlessly expand or redesign fabrics

Fabric management

Internal network is opaque: no switch configs

- L2/L3 services described by service data model
- Cloud UI and API used to provision services
- Visibility and assurance are built-in

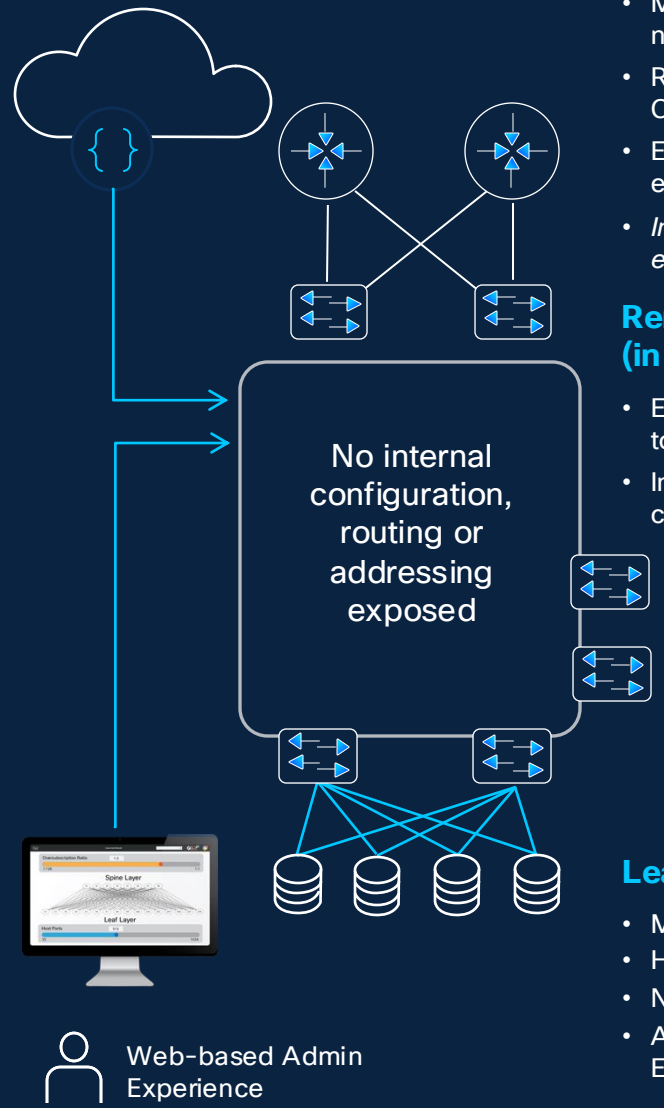


- BGP/EVPN signaling
- Interface addressing
- Internal VTEP reachability

Cloud SaaS controller manages all internal pod configurations

(internal routing and interface addressing)

Cloud-based provisioning/ telemetryAPI



Gateway peering config

- Multi-VRF peering and dot1q to upstream network
- Routing and peering configuration through Cloud UI
- External routing/peering configuration exposed
- *Internal routing/addressing not exposed to external network*

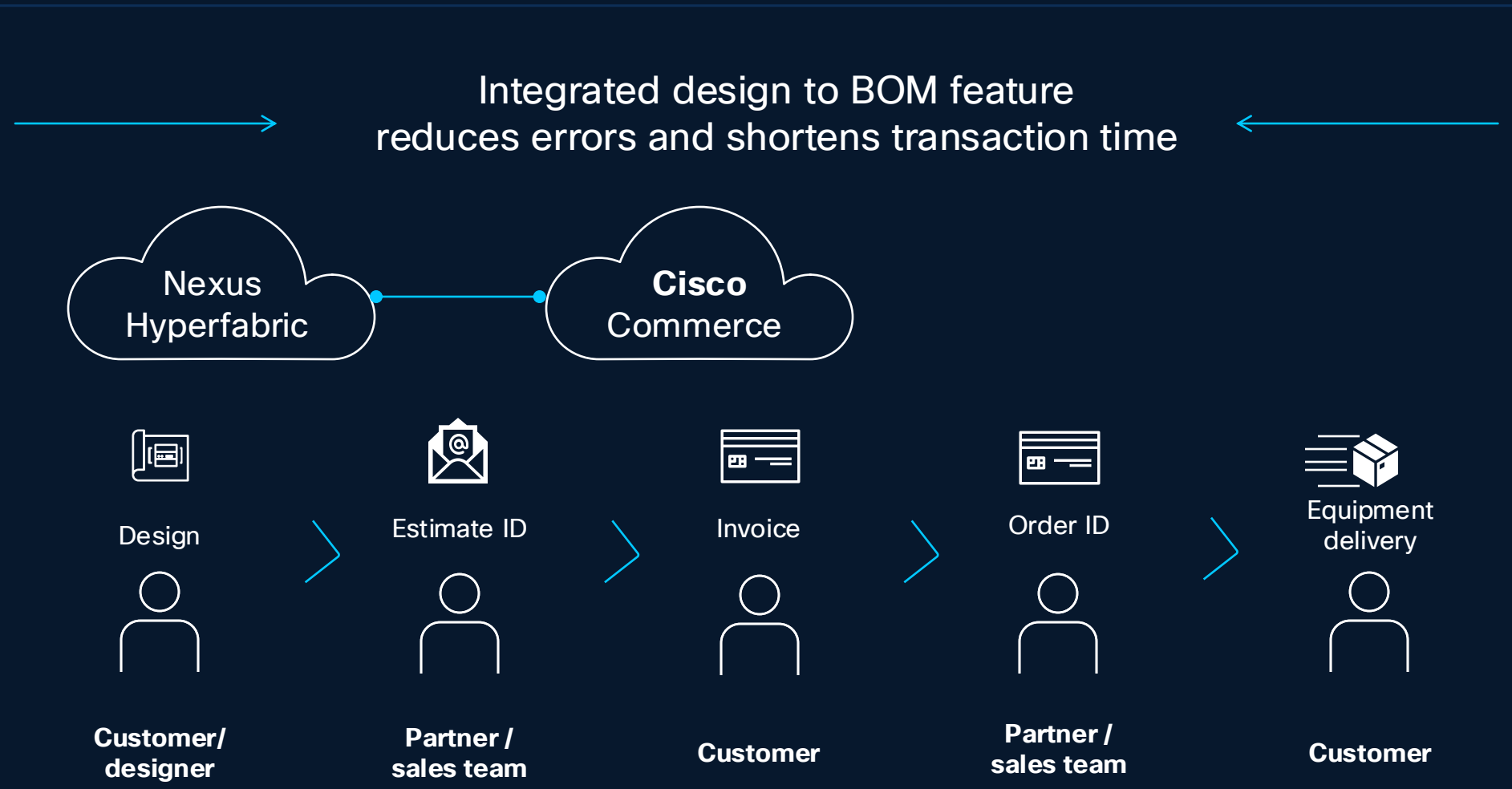
Remote gateway / pod connect (in development)

- Encrypted tunnel connectivity to remote Nexus® Hyperfabric™ pods
- Internal signaling managed by cloud SaaS controller

Leaf: host/edge port config

- MLAG and host peering with dot1q VLAN
- Host connectivity services data model
- Network-based service data model
- Automated provisioning though API: Ex: Ansible/Terraform providers

Design to BOM

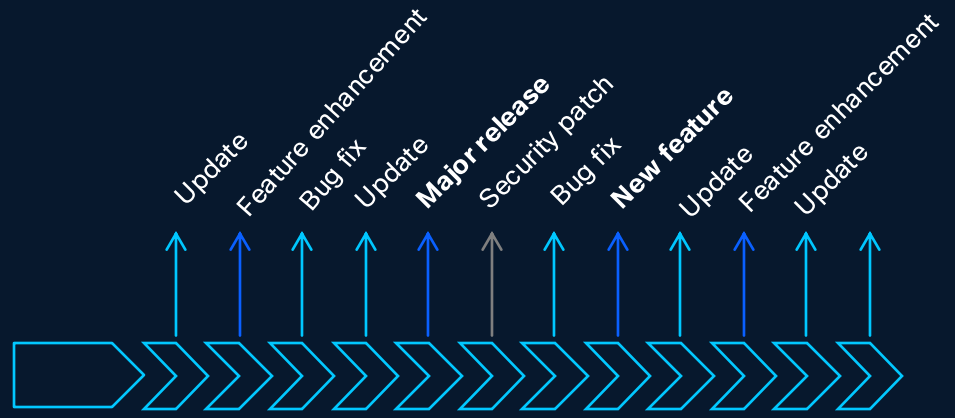


Software lifecycle management

Cloud SaaS controller:

Continuous delivery model: always up-to-date

- Continuous delivery of new features and software updates to the production cloud service
- No user testing or software maintenance required



On-premises switch software

Cloud-delivered software upgrades: user-driven update schedule

- Schedule switch software updates
- Software rollback support
- Intelligent sequencing of fabric upgrades

Flexible architectures

Deploy any fabric, anywhere

Mesh / spine-less fabrics

A Fabric of One”
Lab and API

2-switch fabric
120 host ports

4-switch fabric
240 host ports

2 spine, 2 leaf

2- or 4-way spine, 2-32 leaf
Nearly 2000 host ports

Leaf-spine DC fabrics

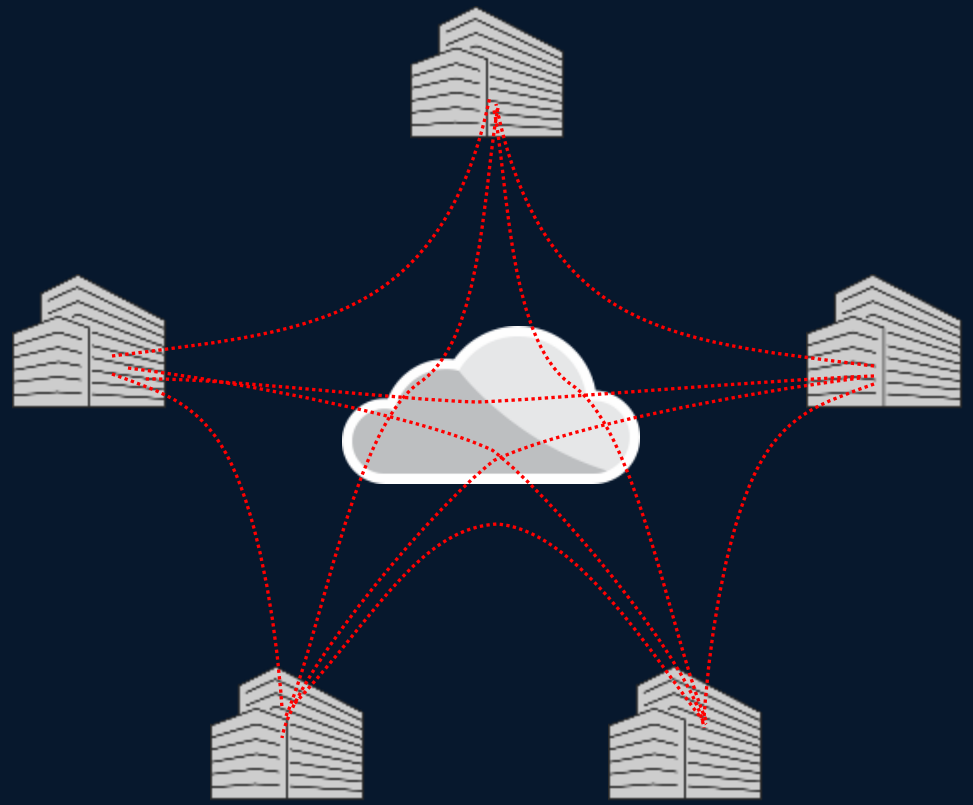
cisco

Demo

MultiSite Support & Roadmap

Overview (Coming Soon)

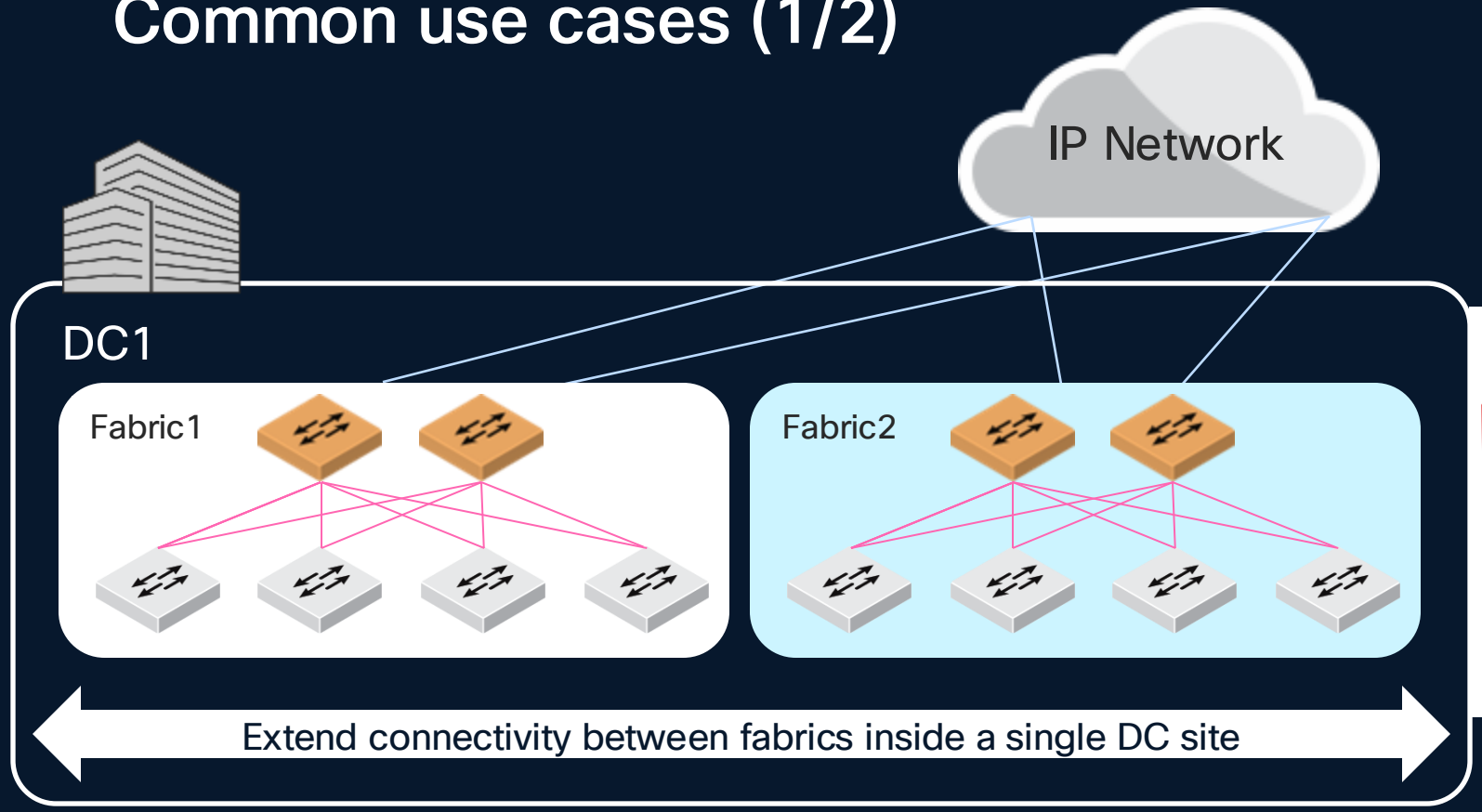
- Hyperfabric Multi-Site (DCI) functionality is a powerful new tool for scaling and connecting EVPN/VXLAN fabrics
- Based on EVPN Border Gateway

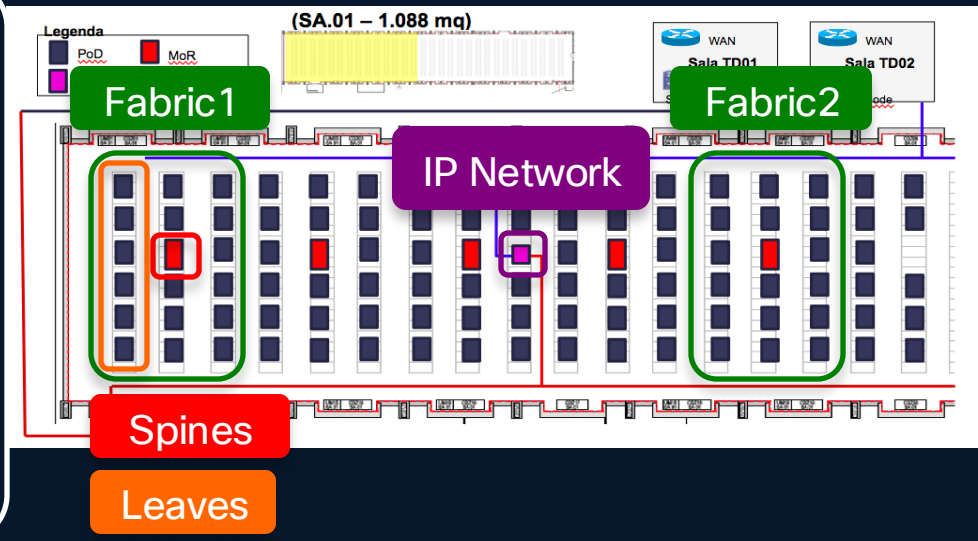


Multi-Site Border Gateway allows:

1. Remote Sites: connect remote fabrics via 3rd party IP networks
2. Scale: hide fabric complexity to peers
3. Native EVPN/VXLAN connectivity
4. Policy control over L2VNI+L3VNI routes
5. Enforce Ownership boundary between fabrics

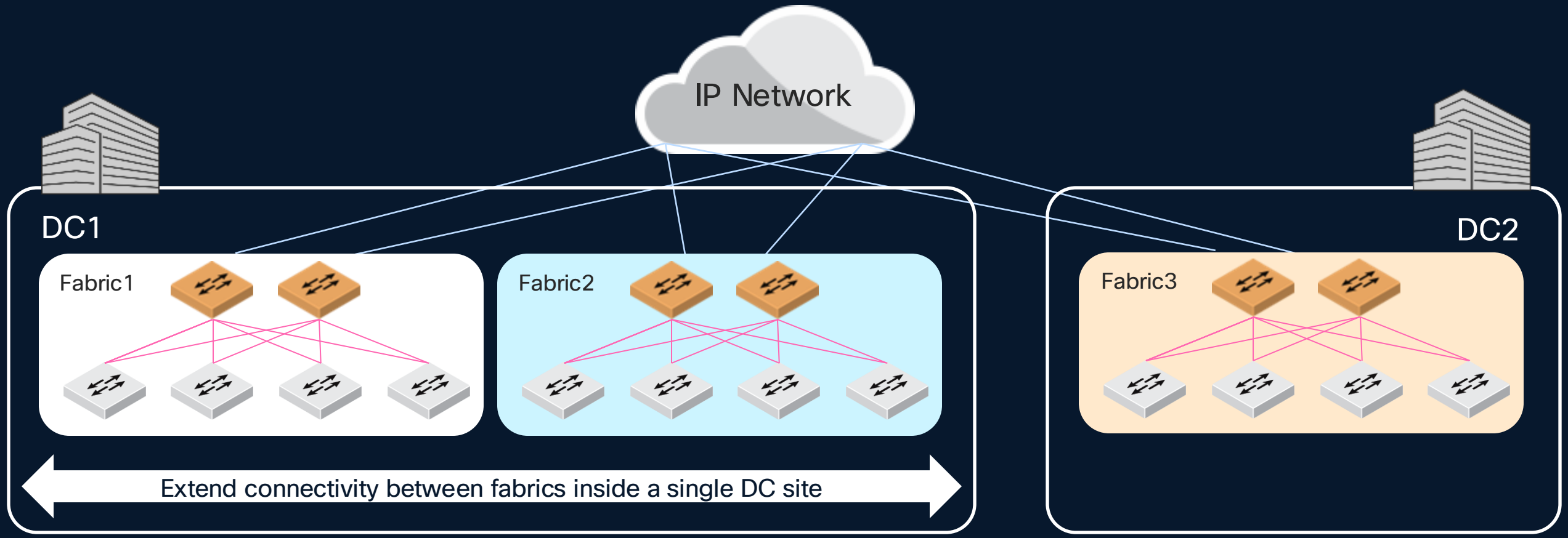
Multi-Site
Common use cases (1/2)





- Scale and IP mobility use cases
- Active/Active fabrics

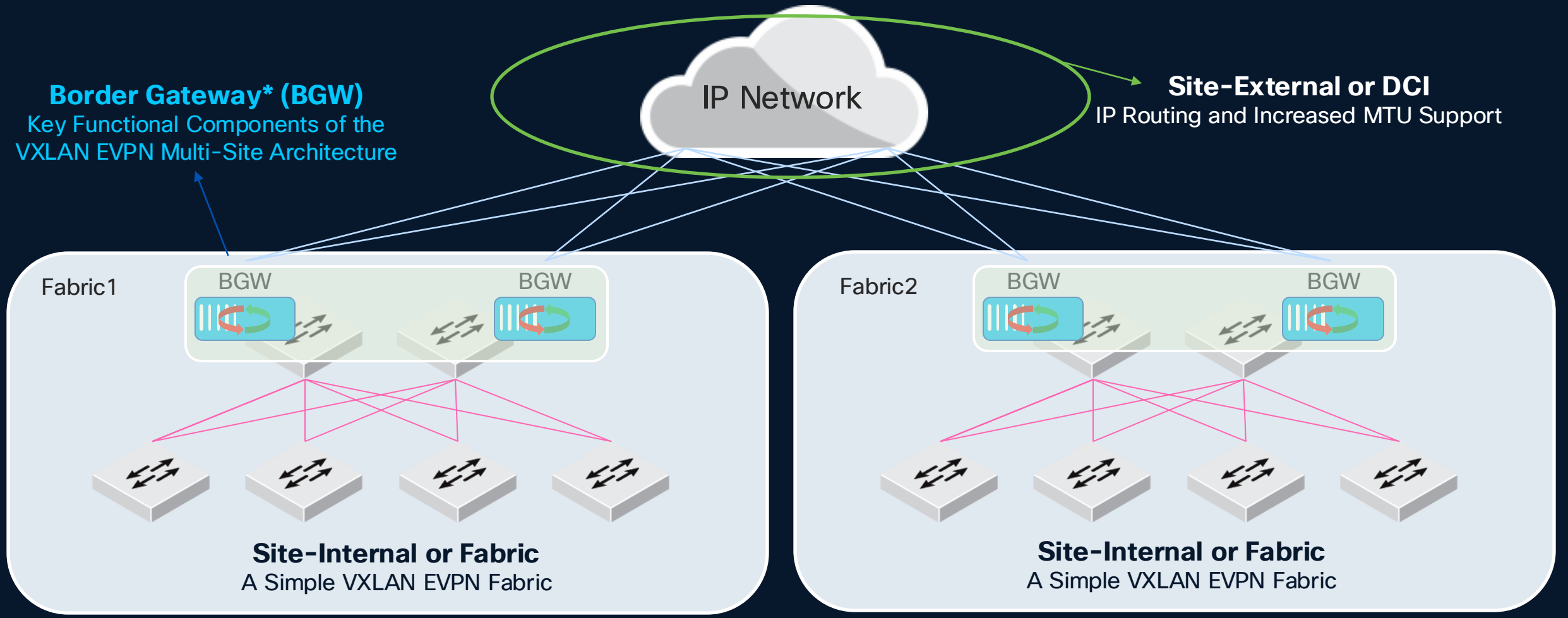
Multi-Site
Common use cases (2/2)



Extend connectivity between DC sites

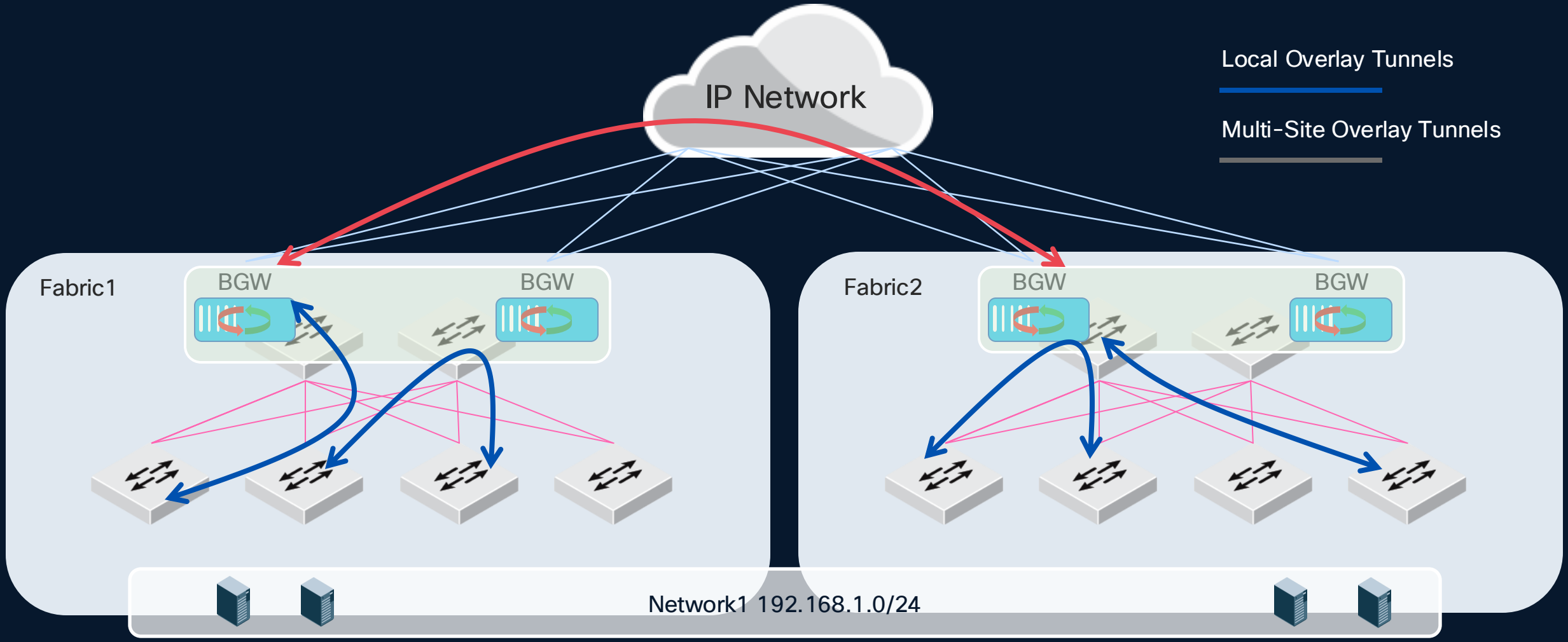
- Disaster Recovery and IP mobility use cases
- Active/Active and Active/Standby

Functional components
Border Gateway (BGW)



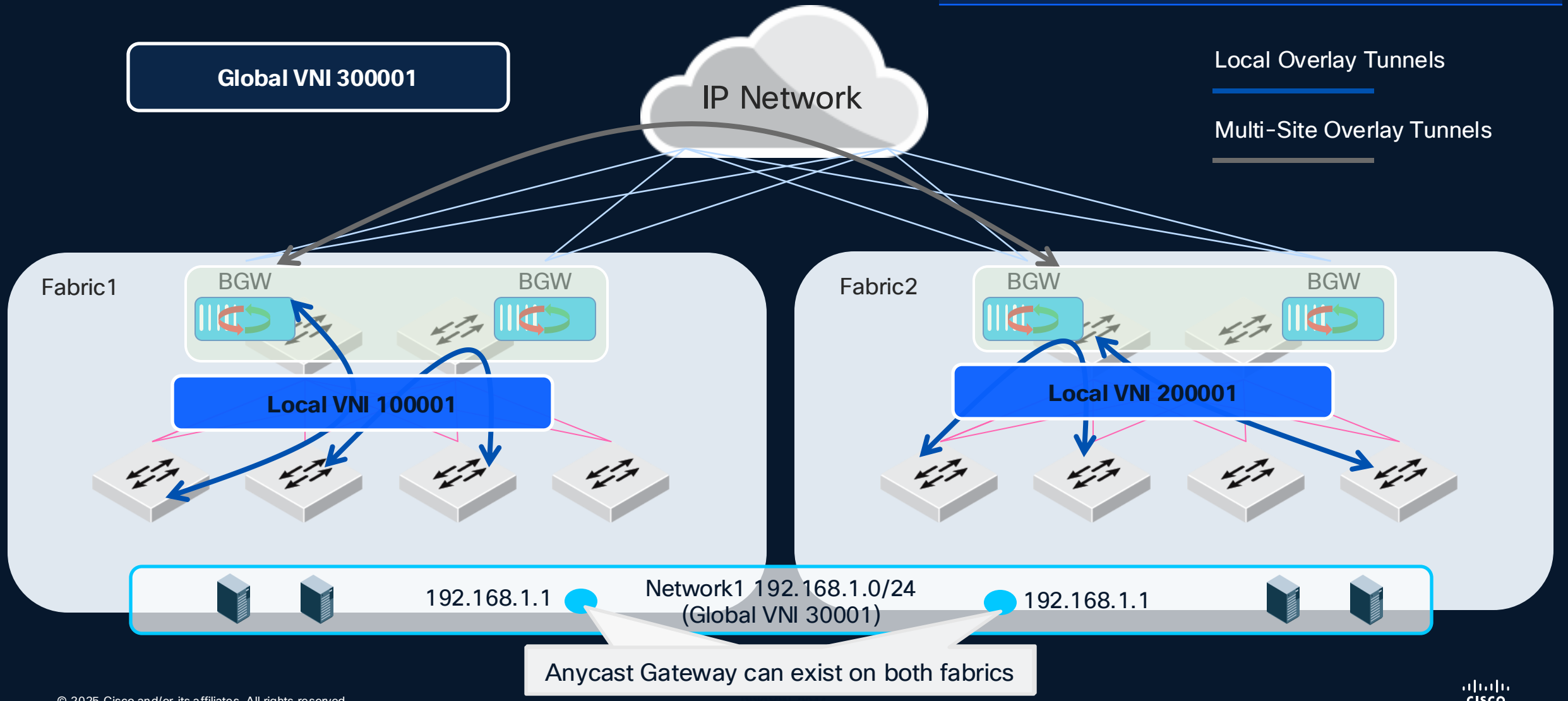
*BGW and spine functions can coexist on the same physical devices

Functional components
Local and Multi-site overlays



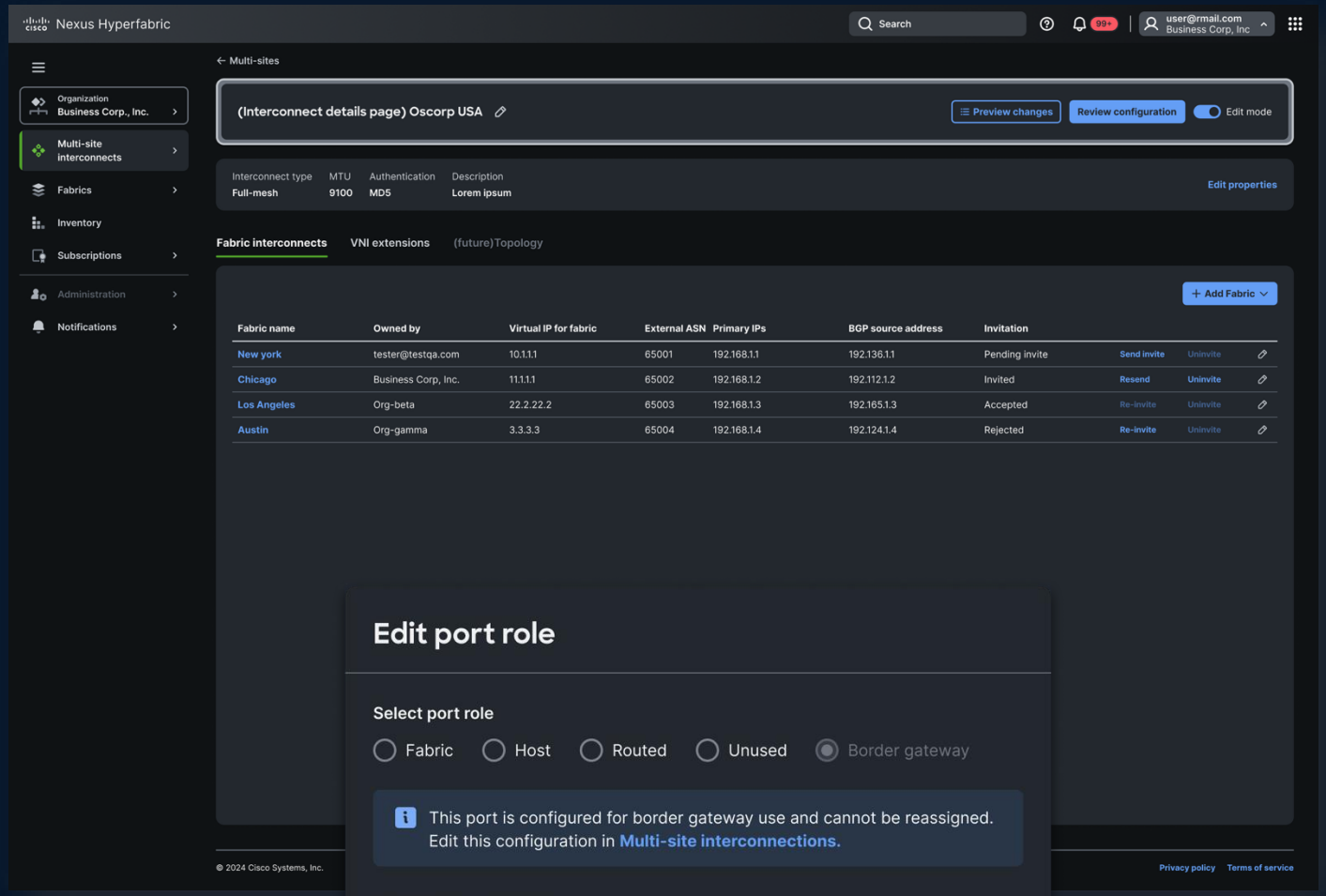
Functional components
VNID translation

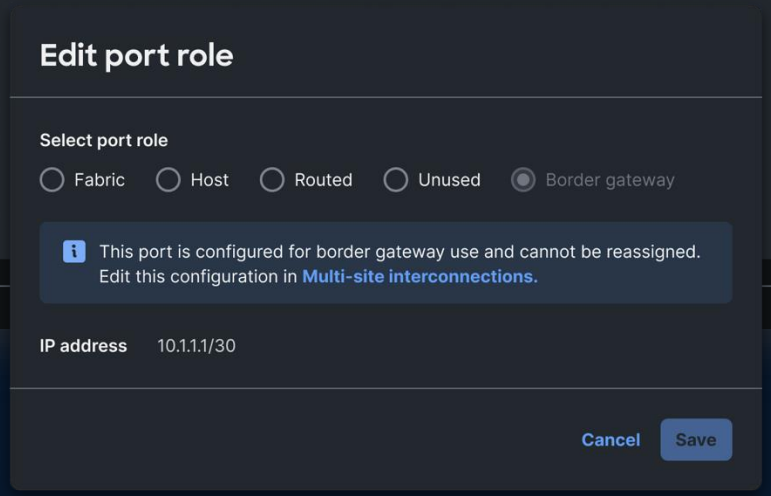
Global VNI	Fabric1 local VNI	Fabric2 local VNI
300001	100001	200001

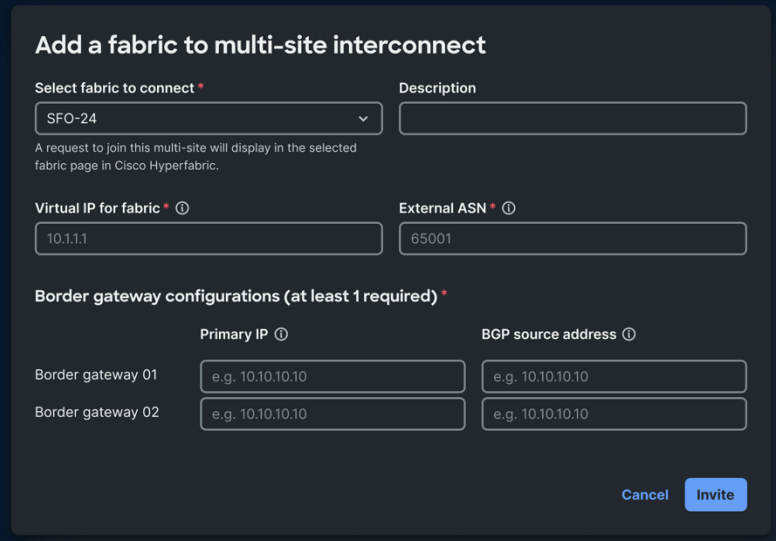


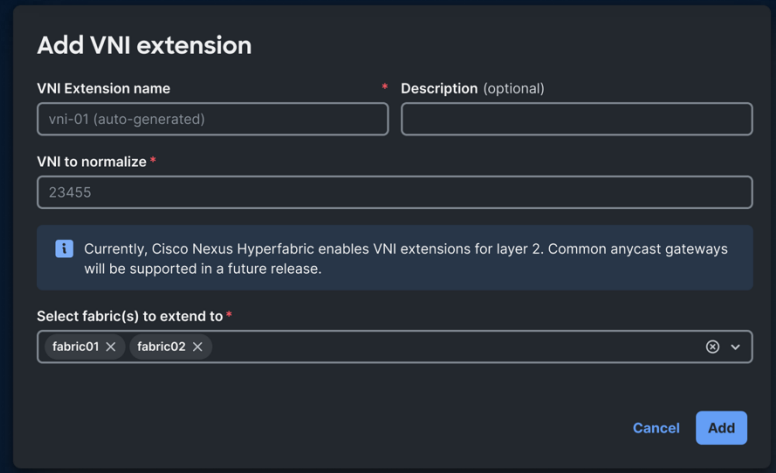
Multi-Site Config (Sneak Preview)

Note: User interface example subject to change

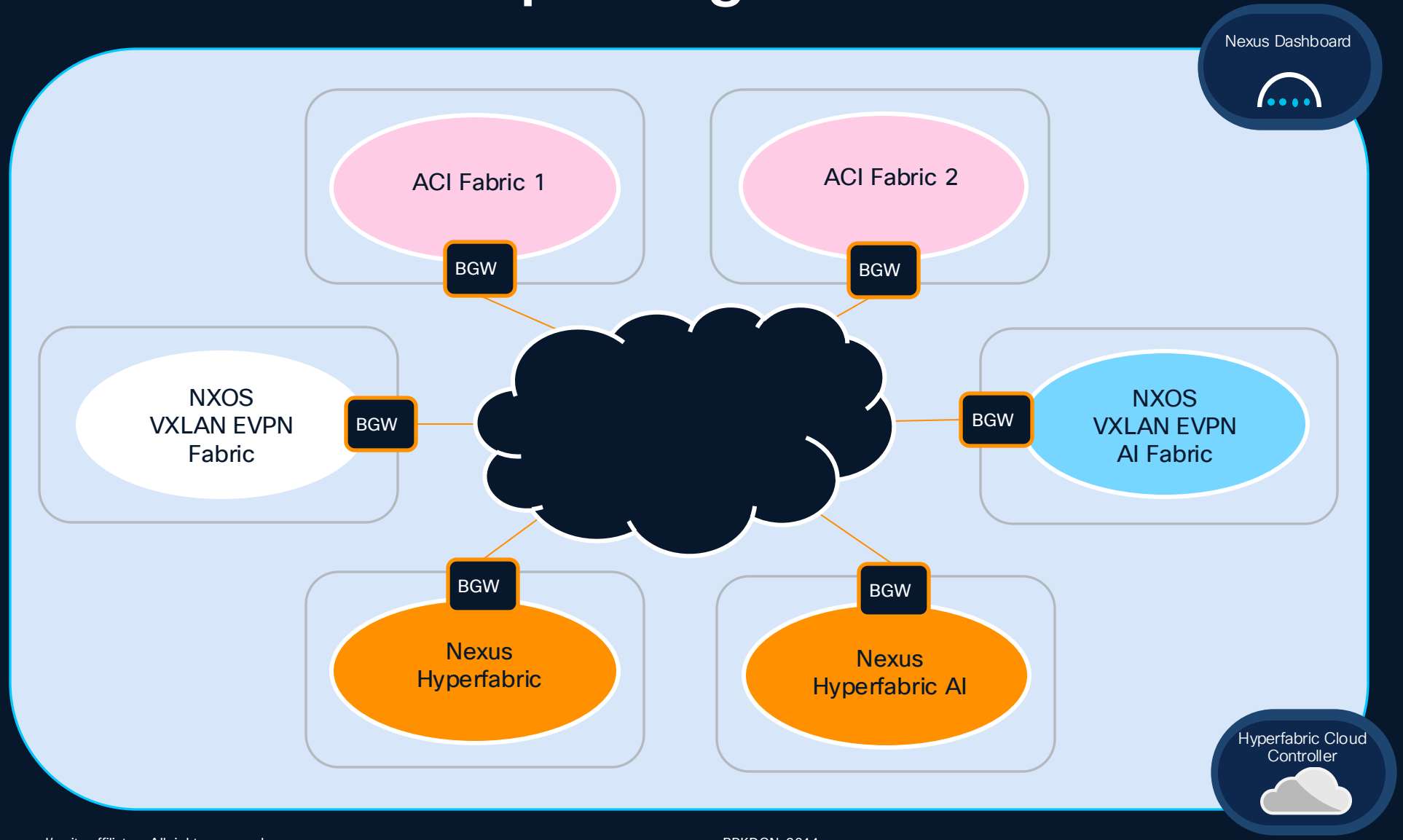








Multi-Site as native peering model for Nexus One



Hyperfabric roadmap

Major releases

3QC25	4QC25	1QC26	2QC26	3QC26+
	Hyperfabric AI HF6100-64ED (800G)	Nexus 9300 Switch (1/10G Cu, FX3) UCS C845A M8 (RTX) Hyperfabric EU sovereign cloud		Hyperfabric Enterprise DC • Advantage subscription package • Super-spine, 10Ks host ports • Security / Policy • Multi-Site via Nexus Dashboard with ACI & NXOS

Essentials subscription features

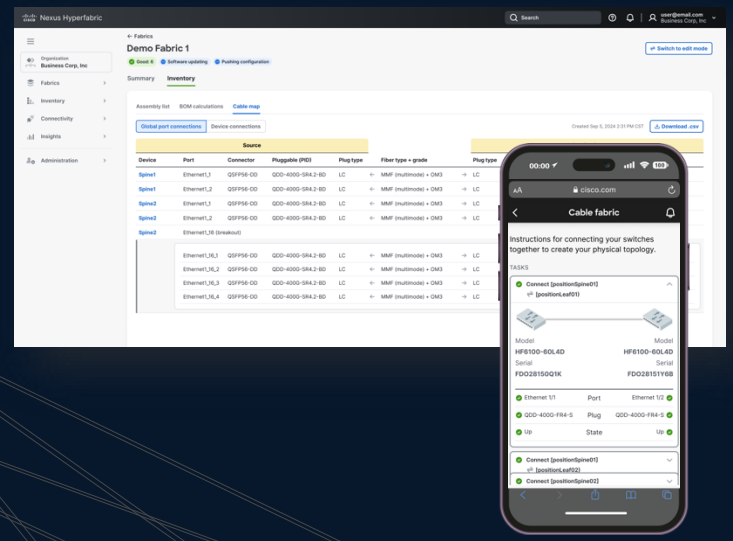
	• Multi-Site between Hyperfabrics • External notifications • SAML authenticaiton federation • Predefined CoS and QoS profiles	• Southbound BGP • Splunk Integration • Cisco AI POD support • IPSLA tracked static routes	**Future roadmap features** • Policy-based routing • Per-port VLAN mapping • Advanced loop prevention • ERSPAN + flow export • IPM probing and visibility

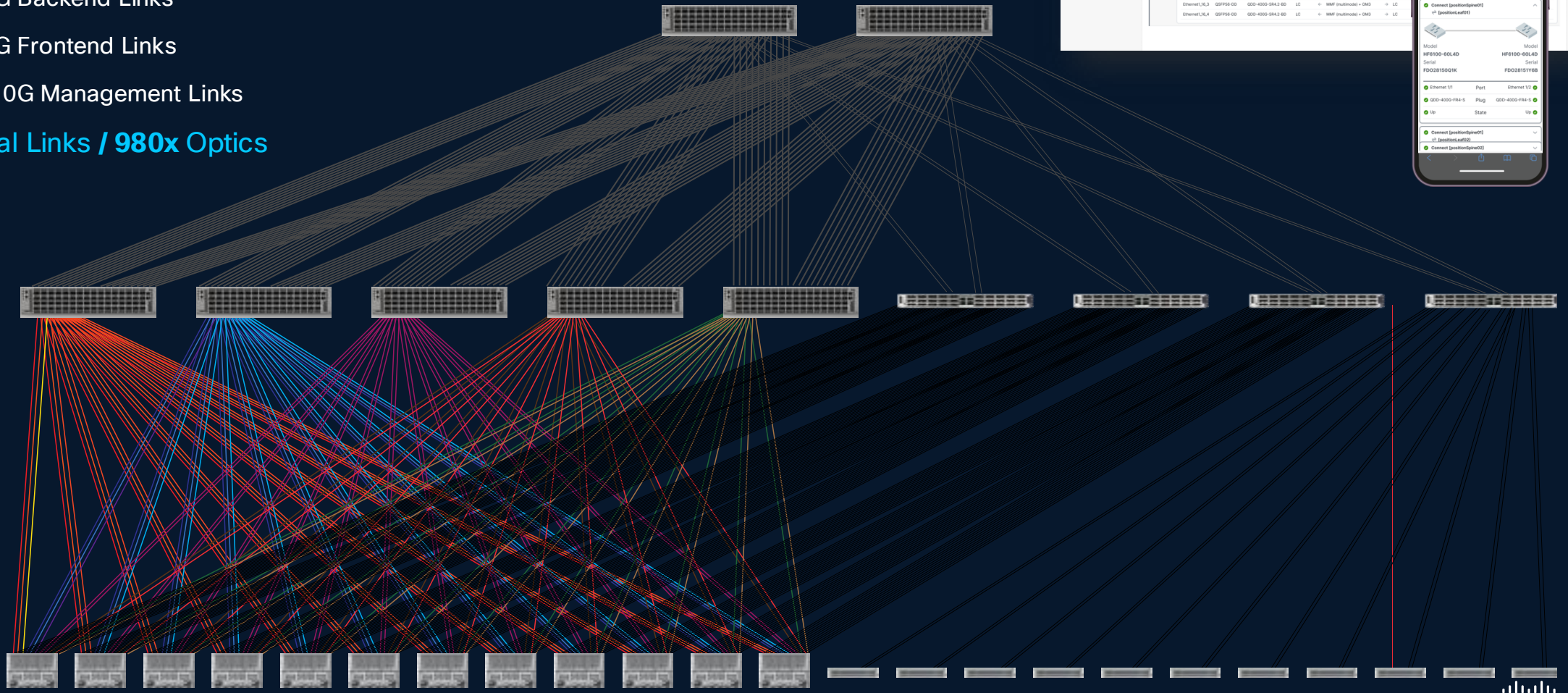
Subject to change

Hyperfabric AI

A “small” $5M AI Training Cluster

- **12 UCS GPU Servers (96x GPUs) + Storage**
  - 112x 800/400G Fabric Links
  - 96x 400G Backend Links
  - 48x 400G Frontend Links
  - 189x 1/10G Management Links
- **= 445** Total Links **/ 980x** Optics

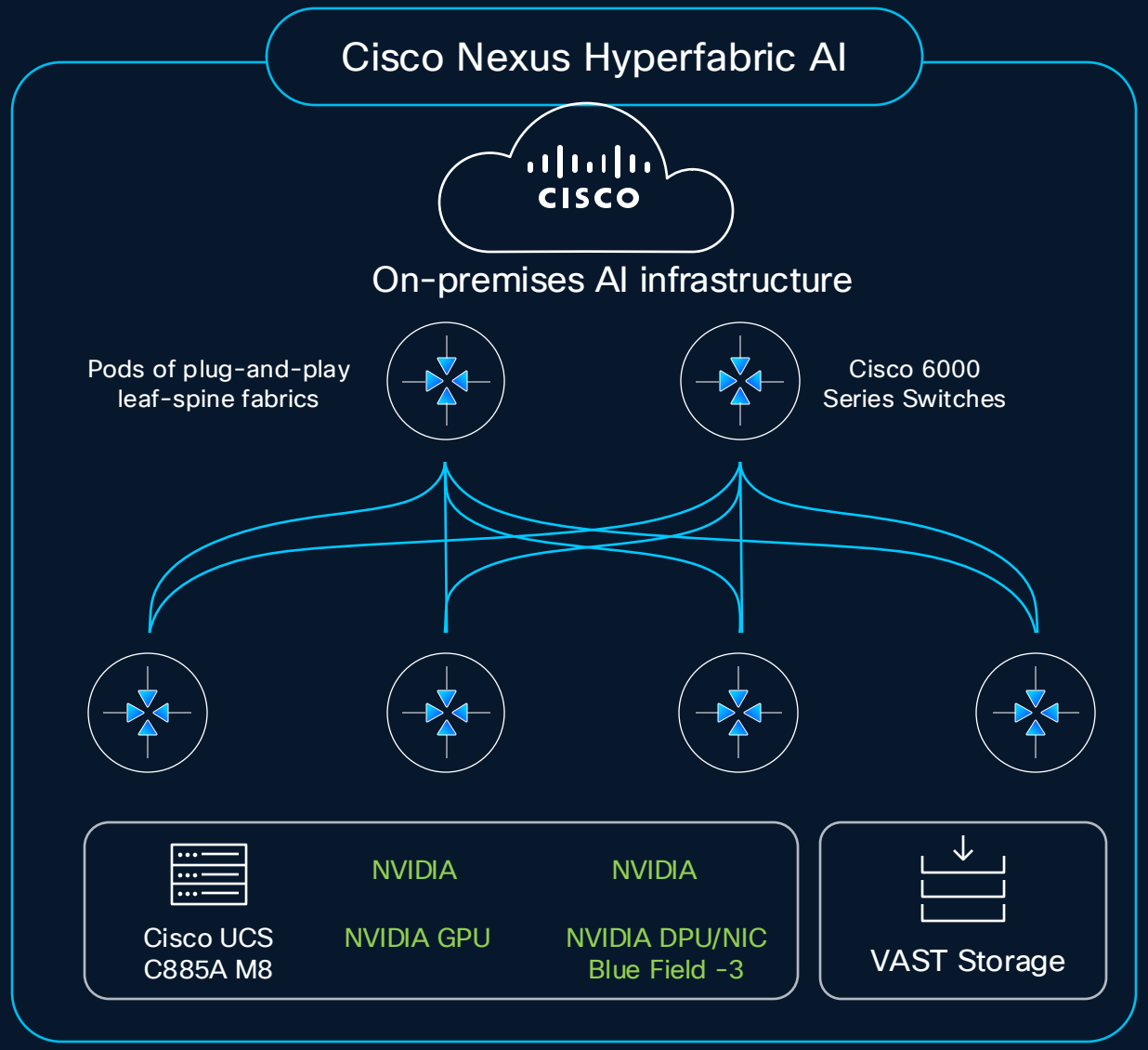




AVAILABLE NOW

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



How are we partnering with NVIDIA?

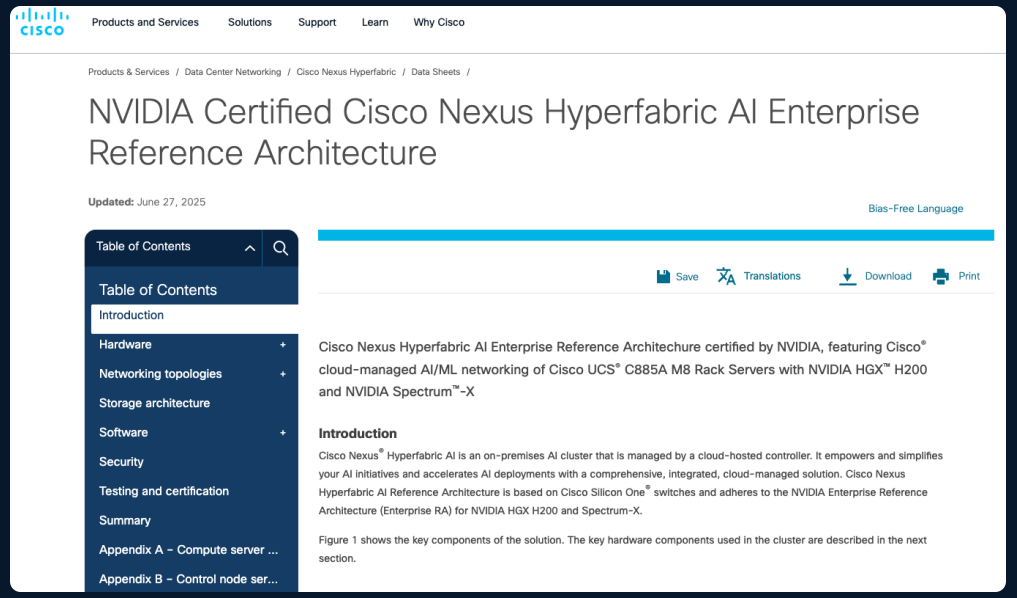
Cisco published Enterprise Reference Architecture with validation

Deliver AI Factory with Cisco Silicon One

NVIDIA Certified Cisco Nexus Hyperfabric AI Enterprise Reference Architecture (ERA):

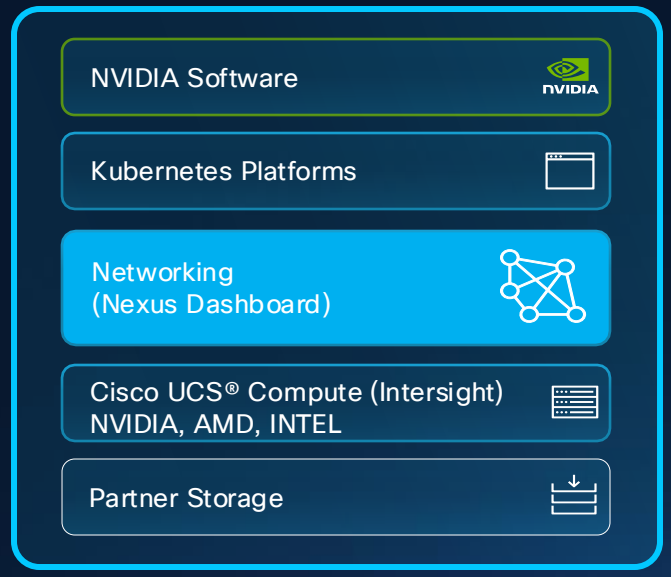
Cisco Nexus Hyperfabric AI Enterprise Reference Architecture certified by NVIDIA, featuring Cisco cloud-managed AI/ML networking of Cisco UCS C885A M8 Rack Servers with NVIDIA HGX H200 and NVIDIA Spectrum-X

Link: [NVIDIA Certified Hyperfabric AI ERA]



AI PODs: Flexible Operating Models

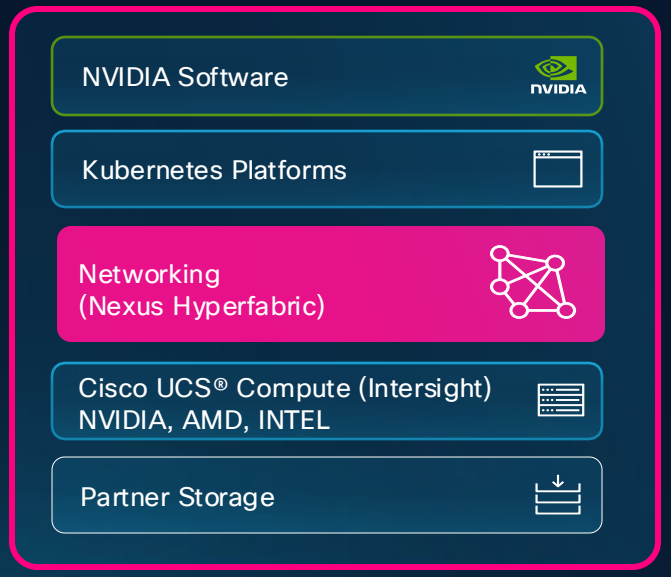
AI POD w/ On-prem management



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 full-stack turn-key infrastructure

AI PODs Solution Options

Choose Based on AI Use-Case and Operational Model

Nexus Dashboard

On-premises managed network

Compute
Cisco UCS
NVIDIA HGX, MGX or
RTX PRO 6000 Servers

Networking
Cisco Nexus 9300
w/ Silicon One
w/ Cloud Scale
w/ NVIDIA Spectrum-X

Storage
NetApp | Pure Storage |
VAST Data | Hitachi Vantara

Training, fine-tuning, inference

Nexus 9300 managed with Nexus Dashboard
NVIDIA aligned Enterprise Reference Architecture
Flexible design consistent performance at any scale
Dedicated AI or multi-purpose DC

Nexus Hyperfabric

Cisco-managed network cloud controller

Compute
Cisco UCS
NVIDIA HGX, MGX or
RTX PRO 6000 Servers

Networking
Cisco Nexus Hyperfabric
w/ Silicon One

Storage
NetApp | Pure Storage |
VAST Data | Hitachi Vantara

Inference

Cloud-managed fabric-as-a-service
Deploy anywhere, easy to scale
BYO servers, GPUs, storage, software stack
Multi-purpose DC with AI & non-AI services

Nexus Hyperfabric AI

Cisco-managed network cloud controller

Compute
Cisco UCS
NVIDIA HGX (885a with 8x H200)
Future: RTX PRO Server

Networking
Cisco Nexus Hyperfabric
w/ Silicon One

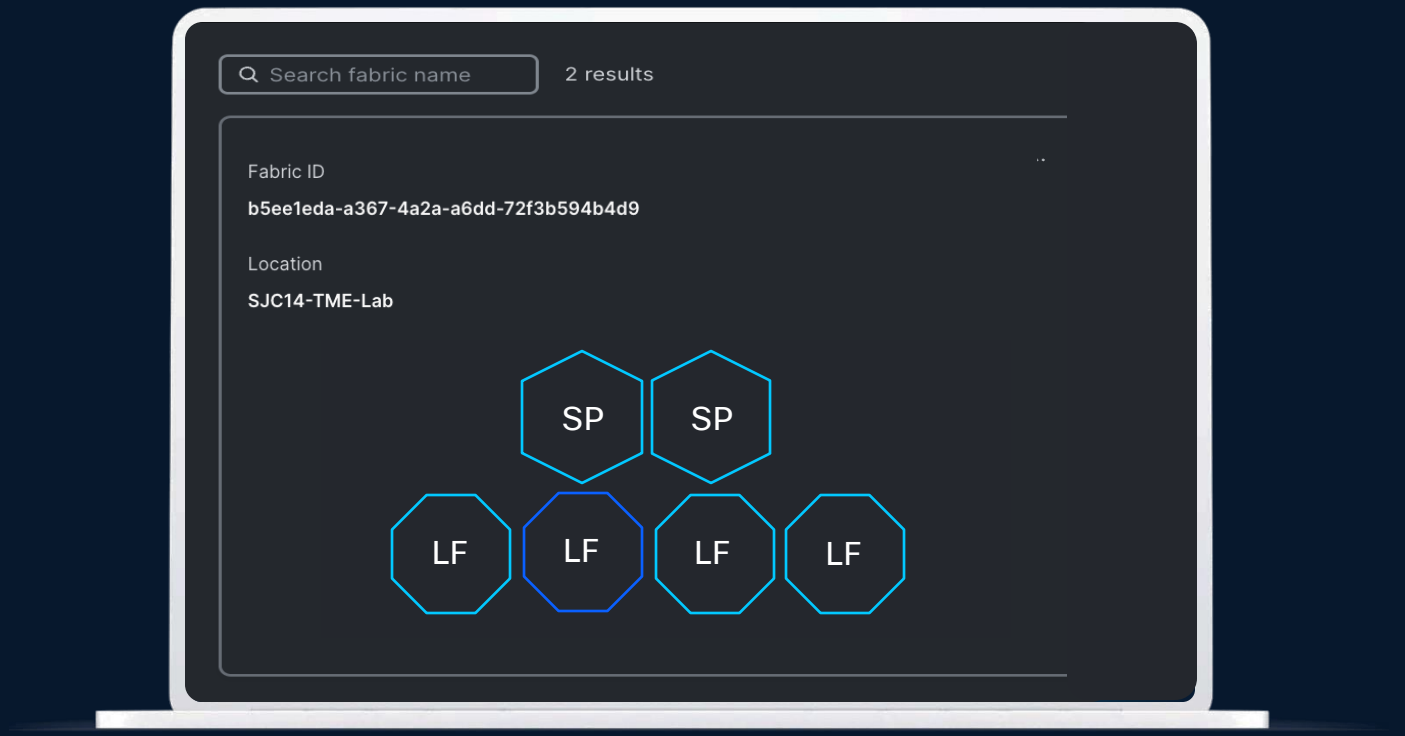
Storage
VAST Data

Training, fine-tuning, inference

Full stack: network, servers, GPU, storage
NVIDIA Enterprise Reference Architecture compliant
Blueprints encompass the entire cluster
Monitoring spans network to compute NIC

Customizable ⟷ *Prescriptive*

Try it yourself



Design a fabric using your Cisco.com ID at
Hyperfabric.Cisco.com

Additional resources

Cisco.com - Cisco Nexus Hyperfabric

Thank you


CISCO