The evolution of the AI-ready edge

Blueprinting tomorrow's computing with AI environments, not servers

James Leach
Director, Product Management - UCS Compute

Eric Williams
Principal Engineer - UCS Compute

Session ID: BRKCOM-2538

Cisco Webex App

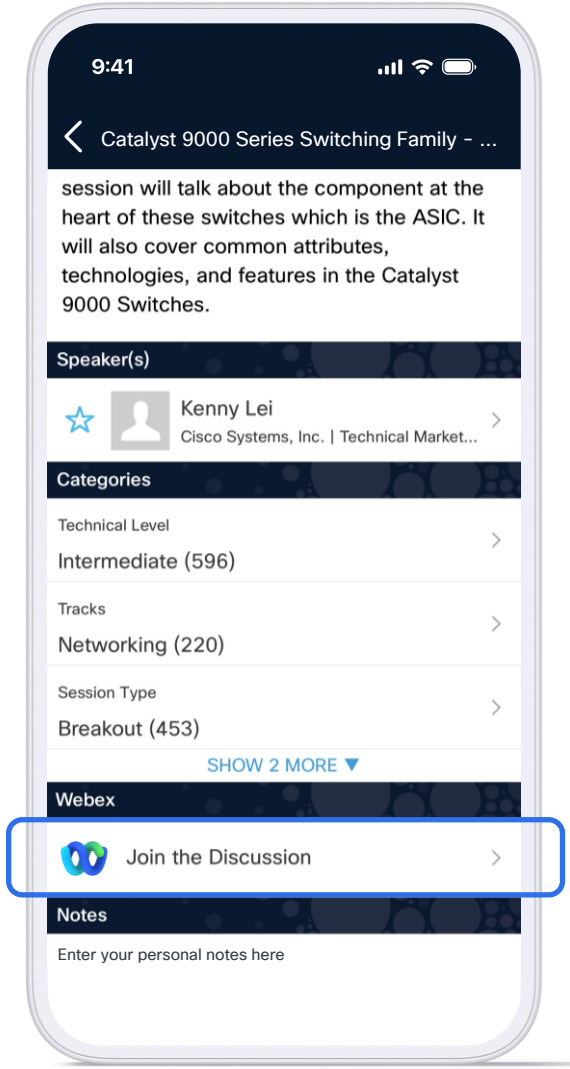
Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until 14 November 2025.



https://ciscolive.ciscoevents.com/
ciscolivebot/**#BRCOM-2538**

© 2025 Cisco and/or its affiliates. All rights reserved.

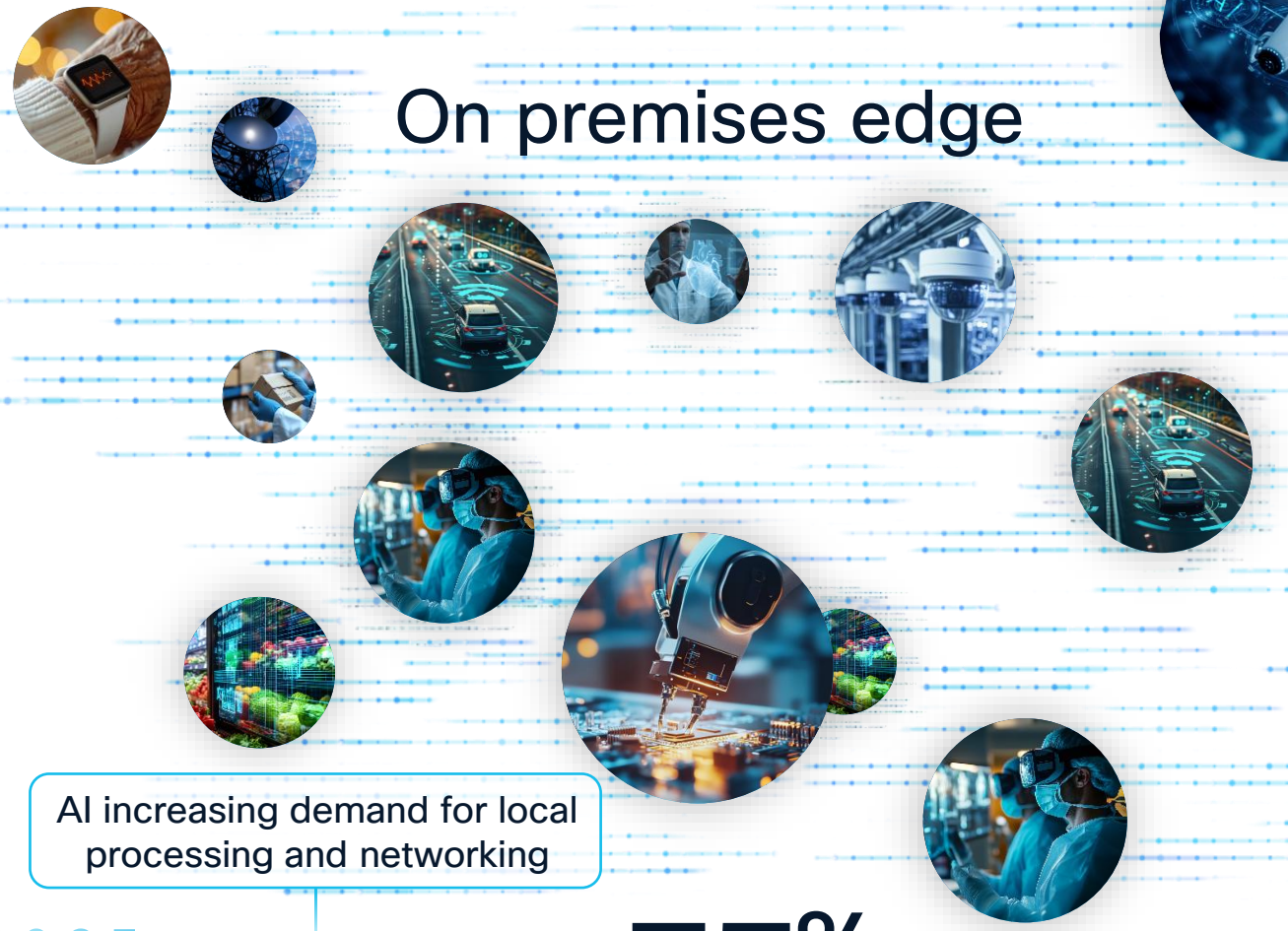
Agenda

1. The edge AI era
2. Cisco’s vision for the edge
3. Cisco Unified Edge
4. Future-ready performance
5. Full-scale simplicity
6. Designed-in security
7. Fleet Management Demo
8. AI-ready platform for the next decade

© 2025 Cisco and/or its affiliates. All rights reserved.

Data center

On premises edge



AI increasing demand for local processing and networking

	2027	
25%	DATA PROCESSING	75%
90%	2021	10%

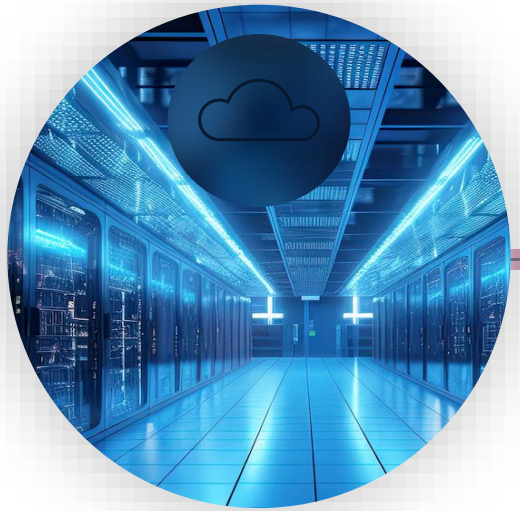
© 2025 Cisco and/or its affiliates. All rights reserved.

2025

77B

AI robotics	Remote patient monitoring
Retail	Worker safety
Smart logistics	Surgical robotics
Autonomous vehicles	Threat detection
...	...

Edge devices



Surgical robotics

1 Teraflop

of computational power

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Defining the AI Edge

Central data center/Cloud

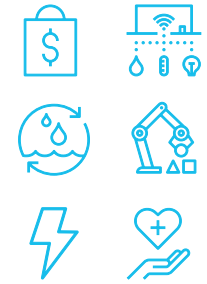
Edge computing

Endpoints

Large data processing

Data warehousing

Full stack infrastructure deployed outside central data centers



IoT endpoints generating data

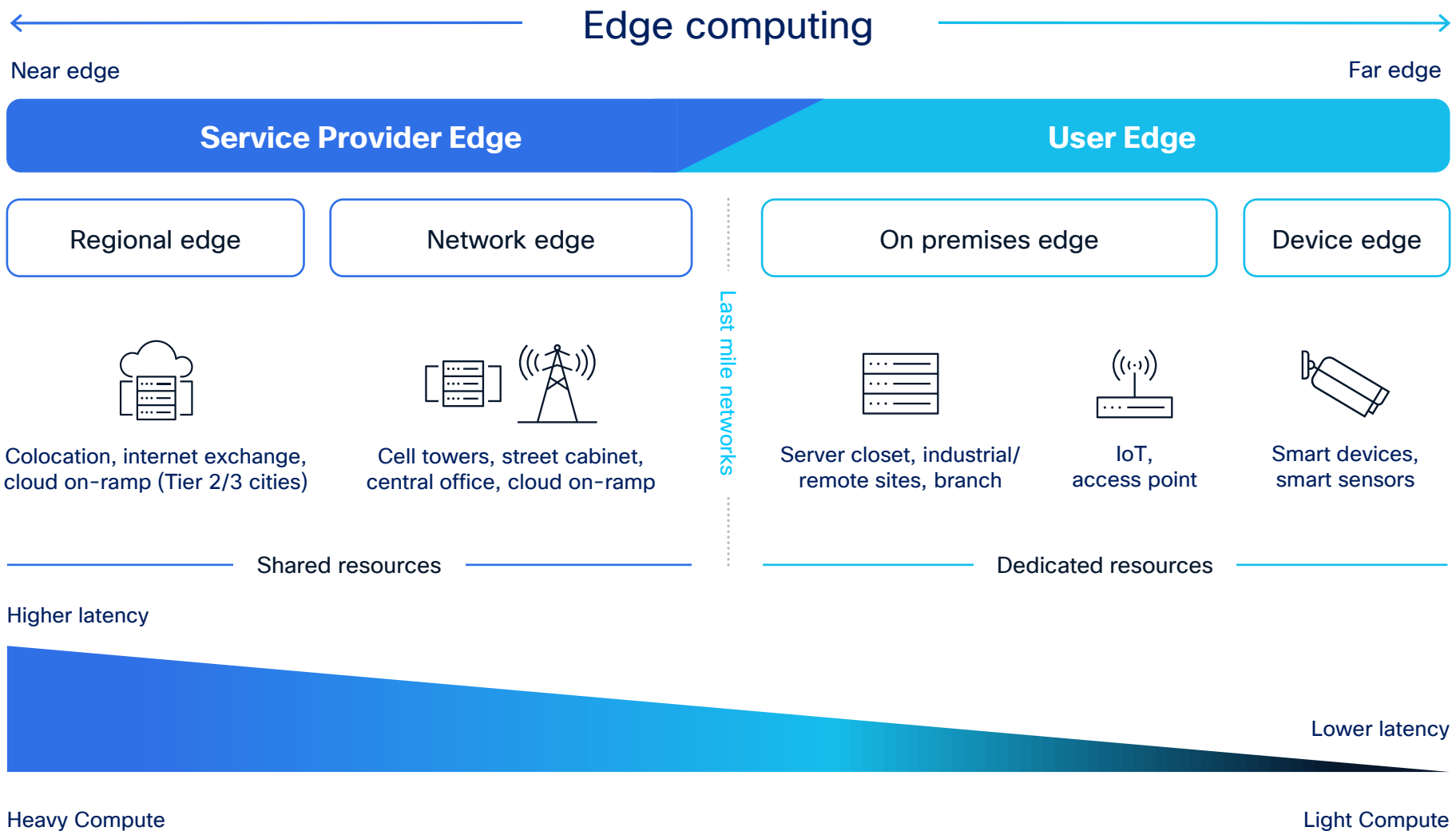
- Distributed computing (server, networking, storage)
- Closer to data, enabling real-time event response
- Reduced latency, minimized bandwidth, cost efficient
- Infrastructure modernization; operational efficiency

Modern infra | Close to endpoints | Low latency | Low bandwidth | Less costly

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

The edge computing spectrum



© 2025 Cisco and/or its affiliates. All rights reserved.

Industry use cases are accelerating the need for AI inferencing and compute at the enterprise edge

Industry-specific use cases and requirements are being evaluated

Retail
Drive thru optimization

Financial
Financial crime/fraud detection

Manufacturing
Asset visibility and control

Healthcare
Augmented diagnosis system

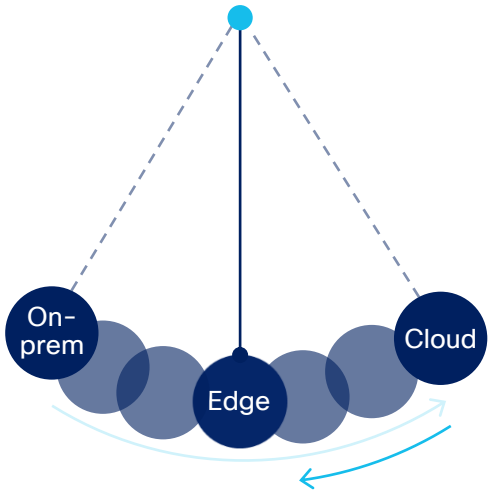
Accelerating the need for AI inference and applications at the edge

Data sovereignty

Latency considerations

Bandwidth needs

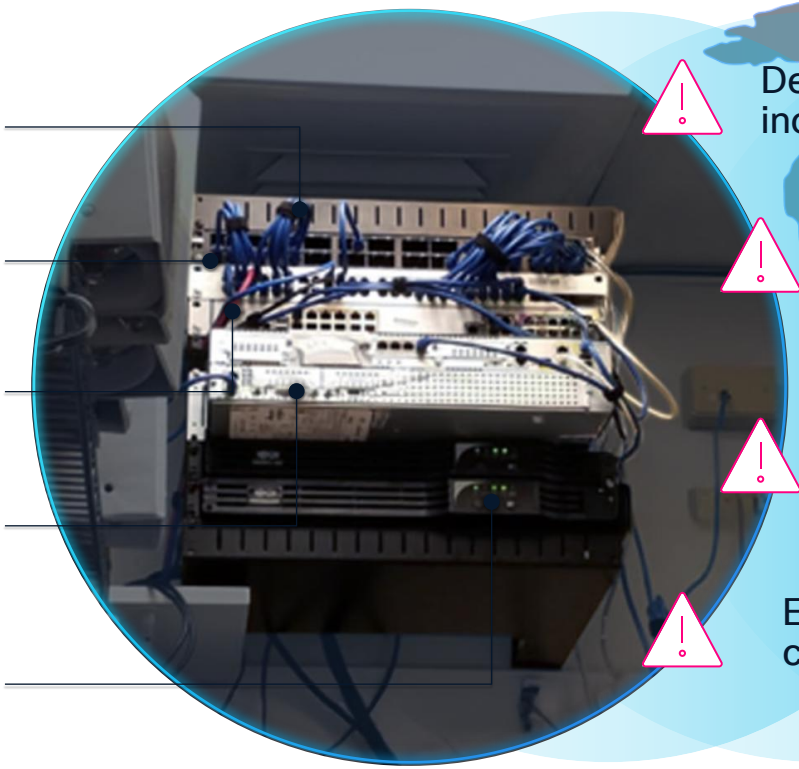
Creating a paradigm shift from the centralized cloud model to Edge AI



© 2025 Cisco and/or its affiliates. All rights reserved.

However, the status quo is not ready for this future

Legacy edge infrastructure



Deployment inconsistency

Product incompatibility

Limited technical expertise

Environmental constraints

Deployed **everywhere**

Operational complexity

Security risks

© 2025 Cisco and/or its affiliates. All rights reserved.

Cisco Confidential

Cisco’s Vision

Blueprinting tomorrow’s computing with AI systems, not servers.

Future-ready performance	Full-scale simplicity	Design-in Security

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Introducing Cisco Unified Edge

Future-ready performance

Compute | Storage | Networking | Software | SaaS Management | Analytics | Security

Nutanix Stack for Edge

Red Hat Edge Stack

VMware Edge Compute Stack

Intel Tiber

Virtual Networking

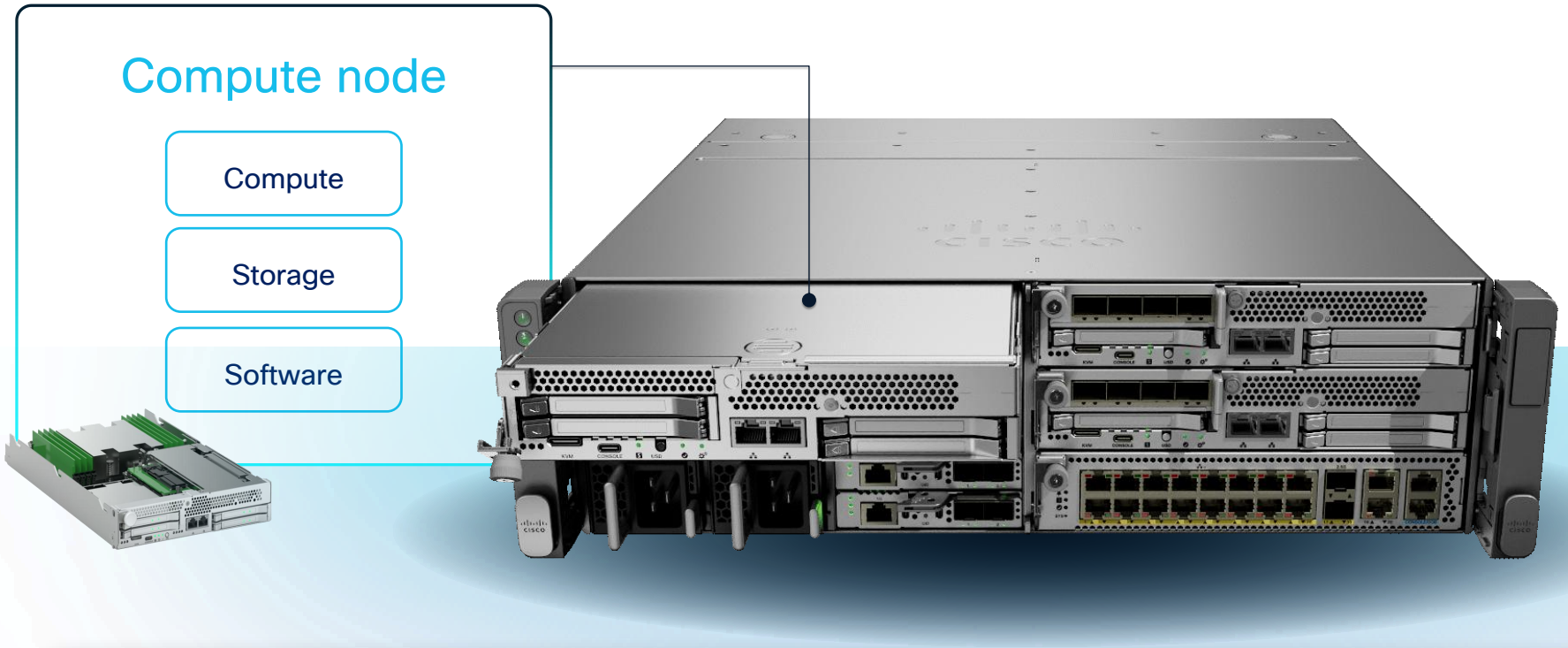
SUSE Edge

© 2025 Cisco and/or its affiliates. All rights reserved.

Cisco Unified Edge: Future-Ready Performance

Future-ready performance

Integrates compute, networking, storage, and security



NUTANIX Nutanix Stack for Edge

Red Hat Edge Stack

vmware by Broadcom VMware Edge Compute Stack

intel Intel Tiber

cisco Virtual Networking

RANCHER BY SUSE SUSE Edge

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

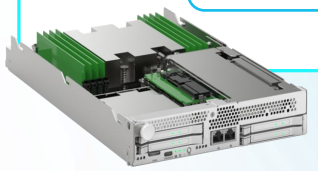
Cisco Unified Edge

Fully validated, full-stack environment that integrates advanced network, compute, storage and security

Future-ready performance

Compute node

- Compute
- Storage
- Software



Networking node

- Network
- Security

Nutanix Stack for Edge

Red Hat Edge Stack

VMware Edge Compute Stack

Intel Tiber

Virtual Networking

SUSE Edge

© 2025 Cisco and/or its affiliates. All rights reserved.

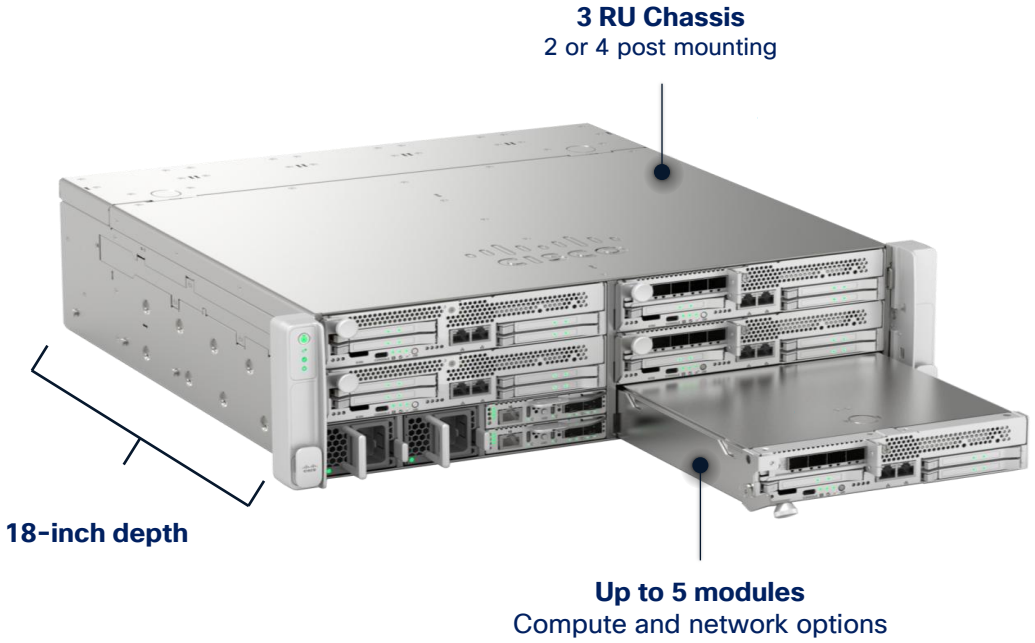
Agile edge system

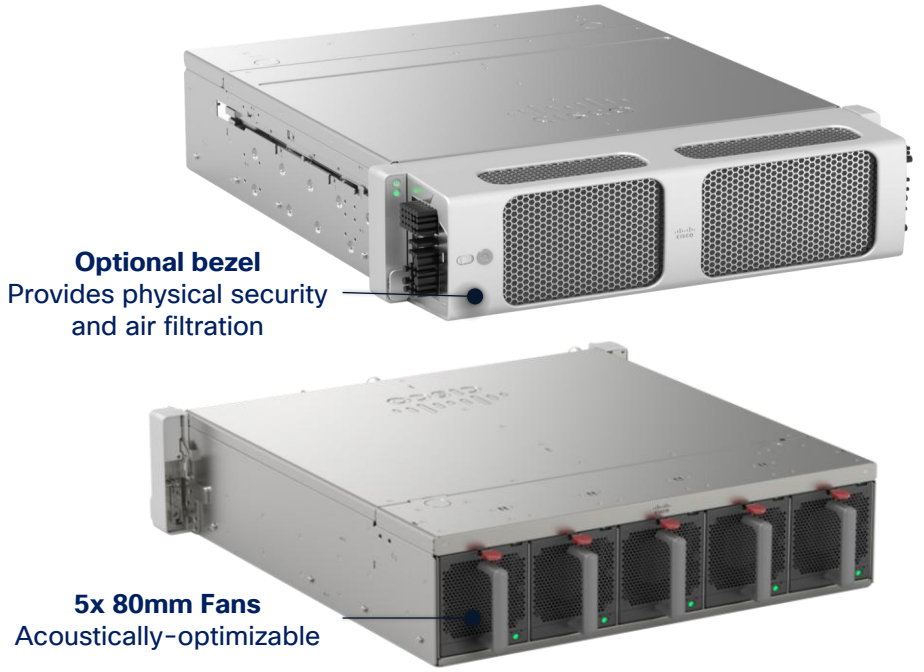
Future-ready performance

Personalized

Serviceable

Modular





Secure by design
(physical and digital)

Investment protection
over multiple generations

Consolidation of power,
cooling and cabling

© 2025 Cisco and/or its affiliates. All rights reserved.

Personalized

Serviceable

Modular

Status LEDs

Easy-to-read simplified health and status indicators

2x 2400W PSUs

N+N redundancy capable

2x Chassis

Management controllers fully redundant

Tool-less, hot swappable modules

Automated configuration upon replacement

Built-in redundancy and resiliency

Real-time health monitoring

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Personalized
Serviceable
Modular

Storage-optimized configuration

Xeon-6 SoC
P-core processor
12/20/32 Core options

8x Memory DIMMs
Up to 768GB

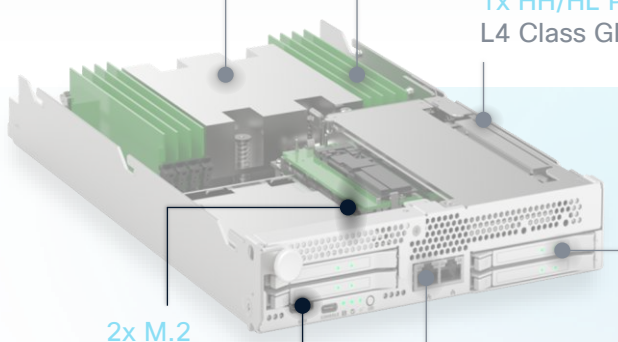
1x HH/HL PCIe G5
L4 Class GPU

4x NVMe
E3.S drives
Up to 120TB

2x M.2
SATA drives
HW RAID

2x 1/10
G NIC ports

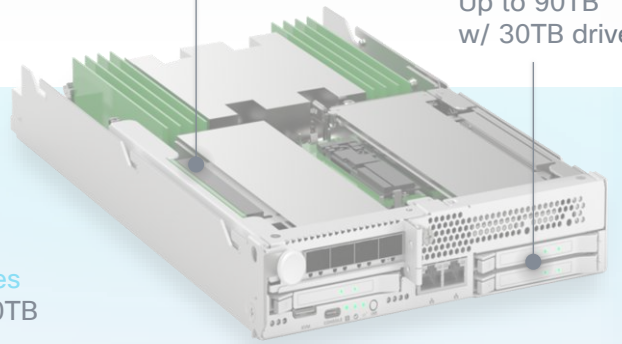
Console / KVM ports
For local access



IO-optimized configuration

1x HH/HL PCIe G5
NIC / DPU

3x NVMe E3.S drives
Up to 90TB
w/ 30TB drives



High-performance networking

25G internal network
Node to node connectivity

Up to 5 modules
Mix and match nodes

Node level personalization

Chassis level personalization

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

High-performance networking

Future-ready performance

Data center fabric at the edge

25G internal network

Node to node connectivity

- In-chassis east-west networking keeps traffic local and fast
- Ideal for node-to-node communication
- Clustered environments

Networking node

- Fourteen 1G ports; Four 2.5mGig ports
  - PoE enabled – fewer cables, faster installs
- Time-Sensitive Networking (TSN) support

© 2025 Cisco and/or its affiliates. All rights reserved.

Validated solutions with ecosystem partners

Future-ready performance

Proven best practices for deploying full-stack solutions

Reduces risk and accelerates time to value

Tested to ensure performance and resiliency

Cisco Intersight

- Zero touch deployment
- Global visualization

Vertical - 3rd party software

- Retail
- Manufacturing
- Healthcare
- AI

Infrastructure software

- Nutanix Stack for Edge
- Red Hat Edge Stack
- VMware Edge Compute Stack
- Intel Tiber
- Virtual Networking
- SUSE Edge

Cisco Unified Edge

- Compute
- Network
- Storage
- Security

© 2025 Cisco and/or its affiliates. All rights reserved.

New Compatible for Cisco Compute Program
Solutions Ecosystem for Unified Edge

	Retail	Manufacturing / Industrial	Healthcare	Cross Vertical AI
Vertical Leaders Solutions	NCR VOYIX Point of Sale Transformation	Rockwell Automation Autonomous AI Factory Transformation	Vertical Target for CY26	NVIDIA METROPOLIS, intel OpenVINO
Target ISVs Solutions at Launch (compatibility tested on beta hardware)	WaitTime Vision AI Customer Traffic Analytics PreciTaste Vision AI Restaurant Operations Optimization Quividi Real-Time Customer Segmentation Targeting	AVEVA Schneider Electric Real-time Production Data Analytics EPIC IO Vision AI Industrial Operations Safety ultralytics Vision AI Parts Inspection AXXONSOFT Vision AI Production Optimization	Medical Informatics Medical Informatics Analytics Platform chooch Vision AI for Medical Equipment Operations	Red Hat AI Inference Server Edge AI Inferencing Server Arcee AI Edge Small Language Models (SLM) and Digital Assistants ONSTAK Cross-Vertical AI Solutions VAIDIO AI-driven Operations Vision Platform

© 2025 Cisco and/or its affiliates. All rights reserved.

Blueprinting tomorrow's computing with AI systems, not servers.

Future-ready performance | **Full-scale simplicity** | Design-in Security

© 2025 Cisco and/or its affiliates. All rights reserved.

What is different about management at the edge?

Deployment scale is horizontal vs. vertical

01 Datacenter: Deploy hundreds to thousands of devices across a few sites

02 Edge: Deploy small number of devices at hundreds to thousands of sites

Deployments are generally very cookie cutter from site to site

01 Manage edge deployments like herds of cattle, not pets

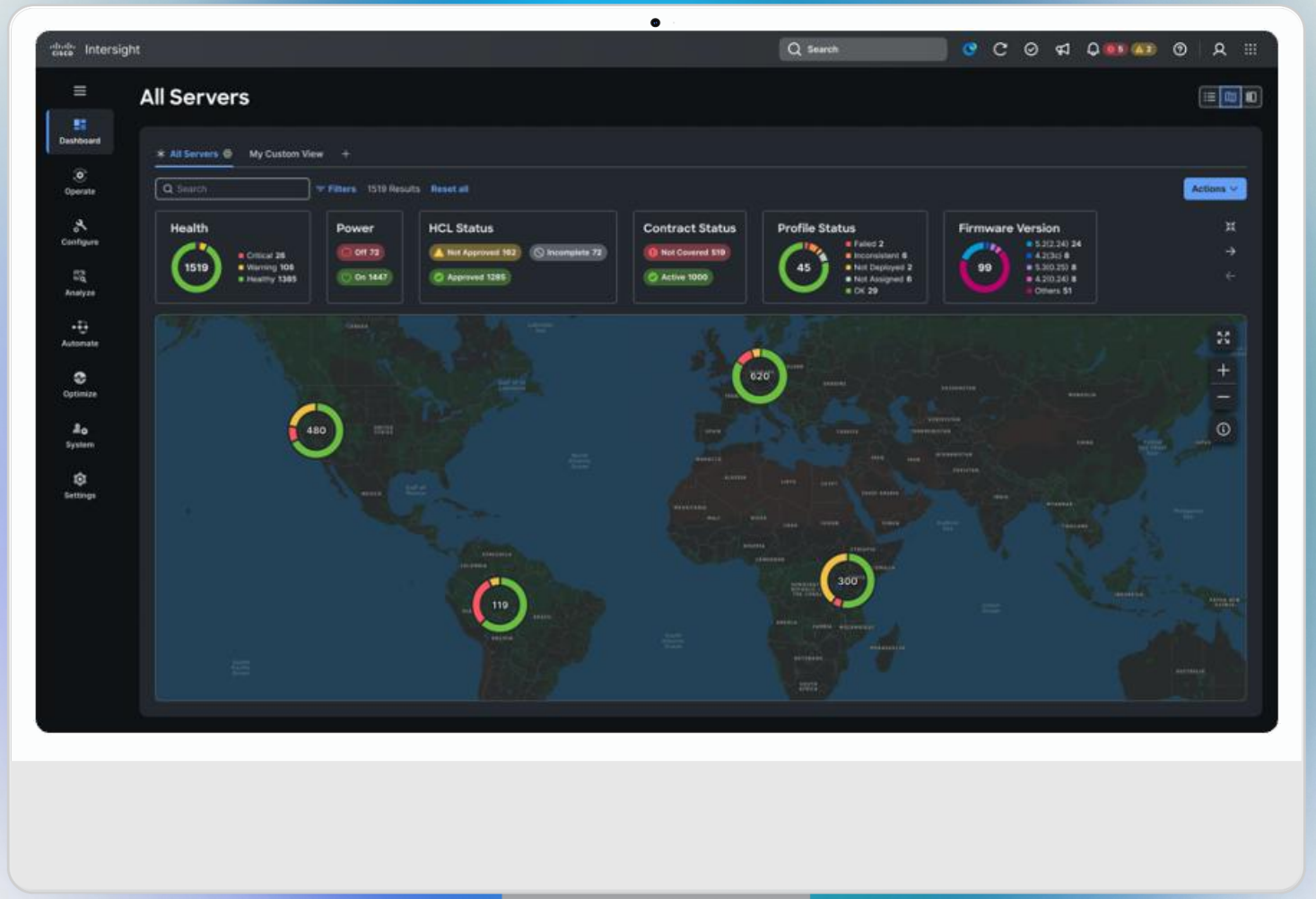
02 Consistency in deployment across sites types is key

03 Replacing hardware/rebuilding configuration needs to be a repeatable process that “just works,” given lack of IT staff on premise to troubleshoot

© 2025 Cisco and/or its affiliates. All rights reserved.

Intersight fleet management

- Simplified onboarding & Zero-touch provisioning
- Automated lifecycle management
- Global visualization

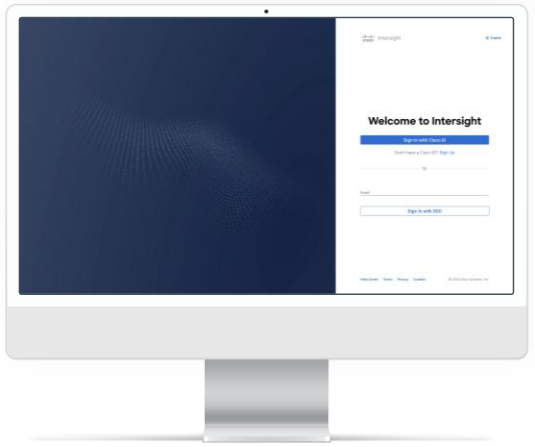


© 2025 Cisco and/or its affiliates. All rights reserved.

Day 0: Zero touch provisioning

Full-scale simplicity

Consistent, repeatable AI-ready infrastructure deployments across multiple sites



Cisco Intersight

Cisco Edge blueprints
(Retail, manufacturing, healthcare)

- Application
- Cluster
- O S
- Network config
- Server configuration
- GPU configuration
- Storage configuration
- Security configuration

Golden configurations based on Cisco validated designs

Faster deployments with fewer errors

Supported by Cisco TAC

Available in a marketplace

© 2025 Cisco and/or its affiliates. All rights reserved.

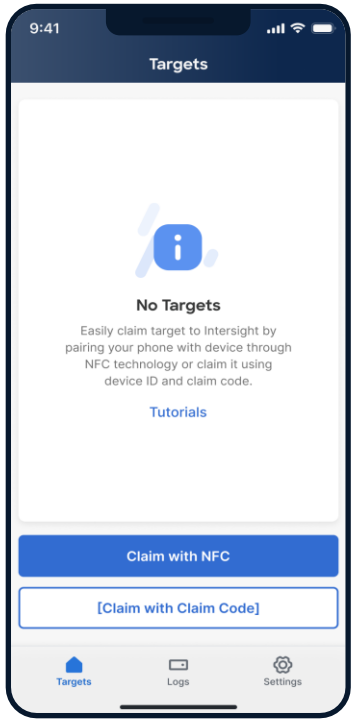
Day 1: NFC-based claim via mobile app

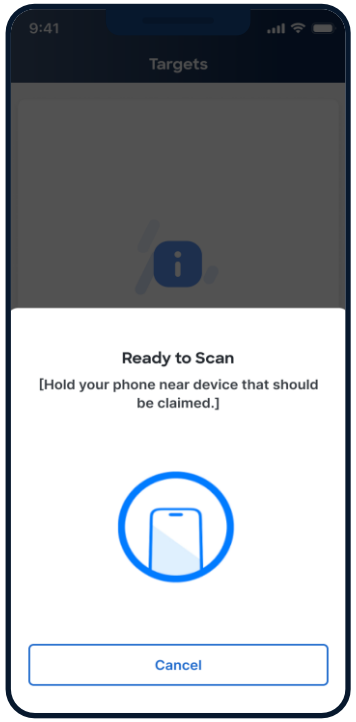
Full-scale simplicity

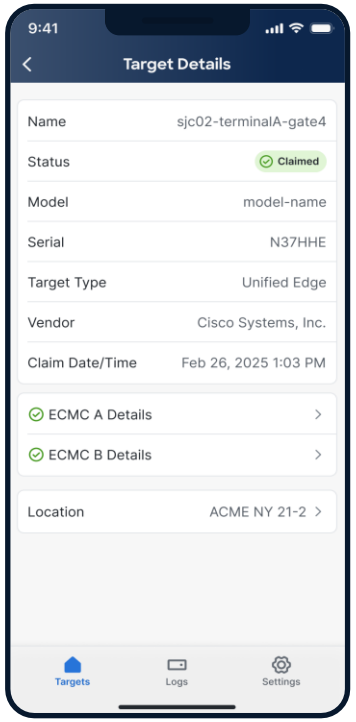
NFC tag is embedded within the slam latch of the chassis to allow communication to the Intersight mobile application

01 **Passive mode**—Programmed with same details physically present on the chassis (PID, serial number, revision, etc .) when chassis is powered off

02 **Active mode**—Intersight mobile app communicates to the device connectors in the eCMC via NFC tag when chassis is powered on







Fewer truck rolls – no need to onboard systems centrally

No technical expertise needed to claim device

Pre-claimed capabilities accessible to authorized on-site staff

© 2025 Cisco and/or its affiliates. All rights reserved.

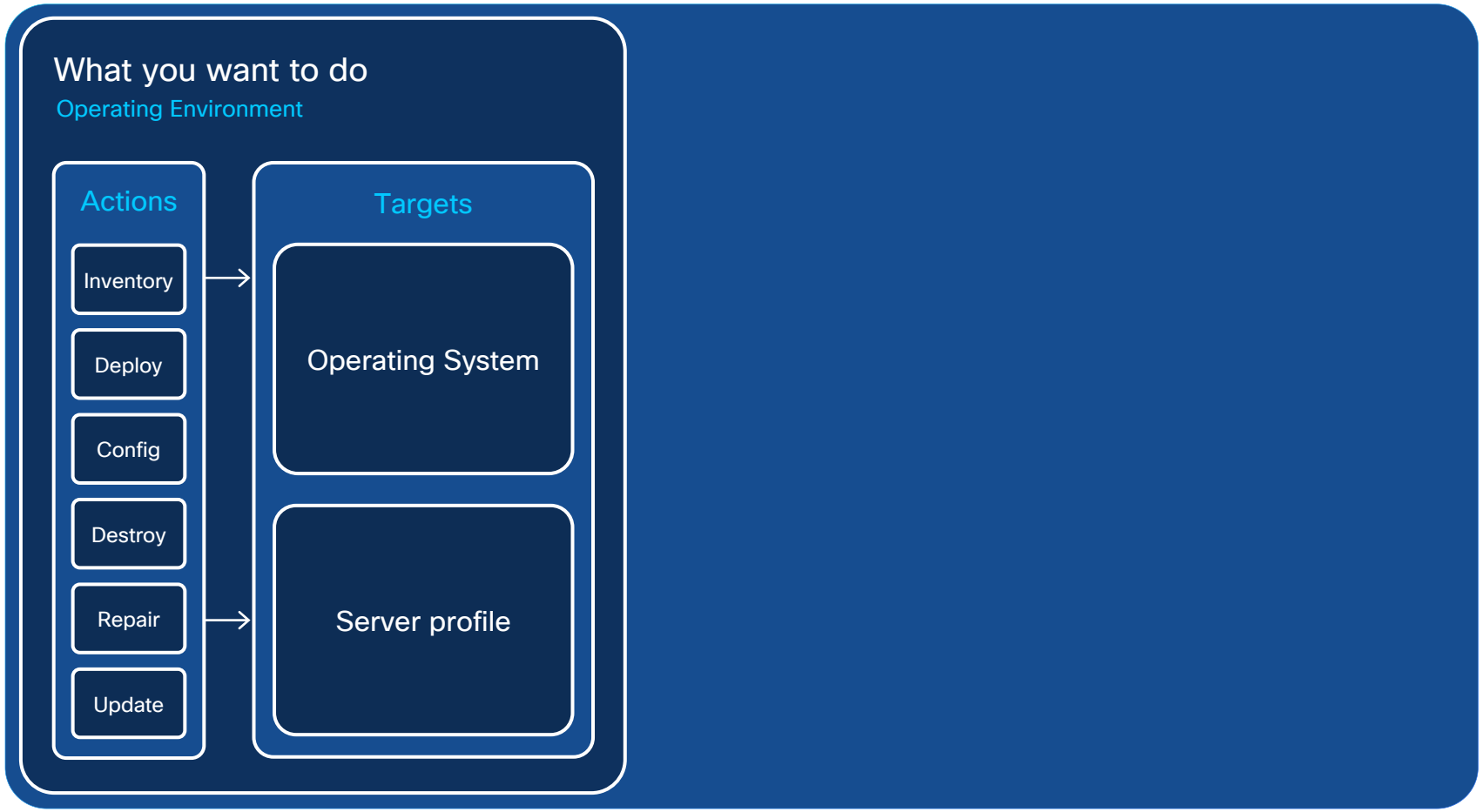
Blueprints

Compute Node

Blueprints are curated, tested, and authored by Cisco

Available in a marketplace as options for deployment

Maintained and fully supported by Cisco engineering and TAC



Supported Operating Systems

Windows Server | ORACLE | Rocky Linux™ | vmware® | SUSE | Canonical Ubuntu | Red Hat

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

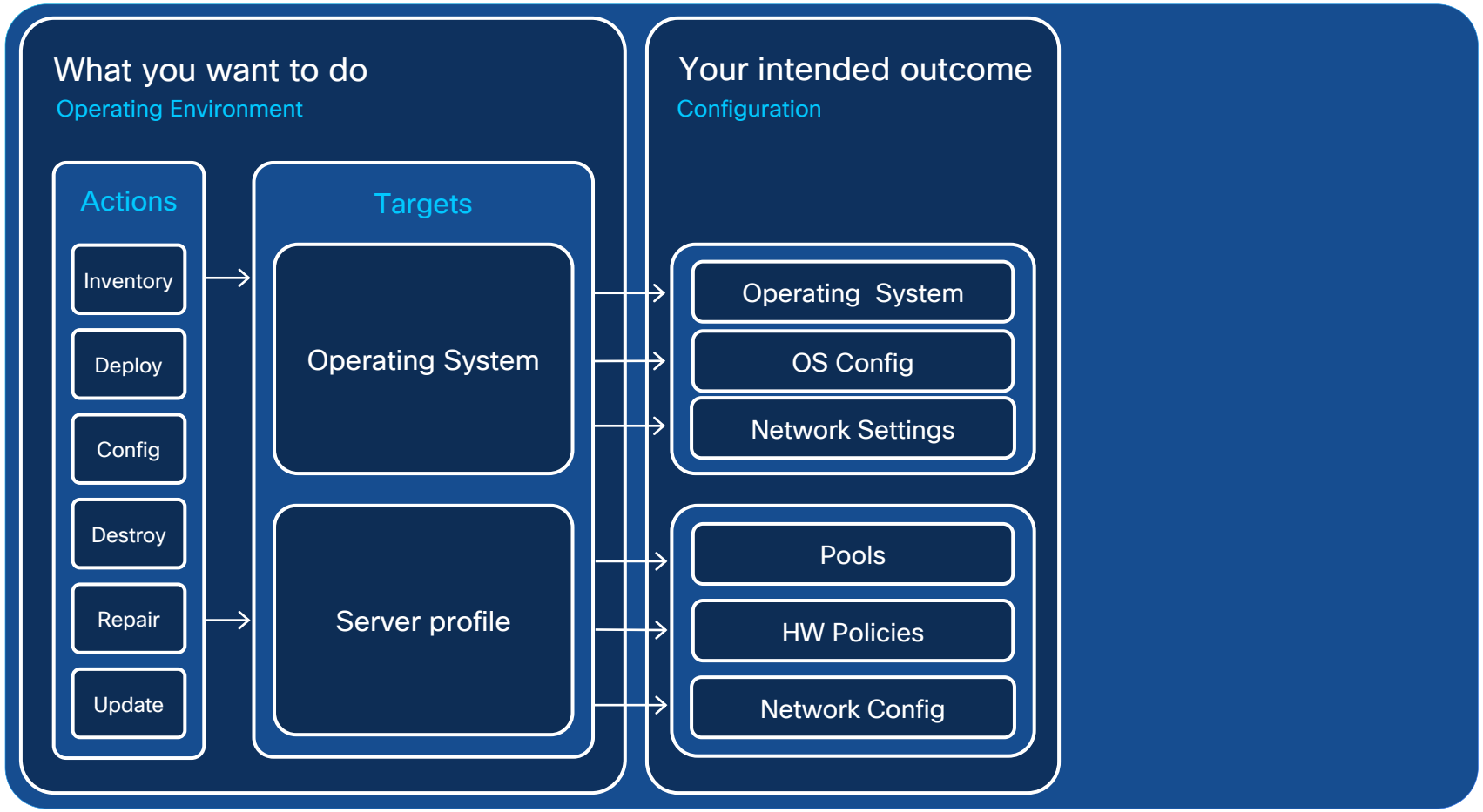
Blueprints

Compute Node

Blueprints are curated, tested, and authored by Cisco

Available in a marketplace as options for deployment

Maintained and fully supported by Cisco engineering and TAC



Supported Operating Systems

Windows Server | ORACLE | Rocky Linux™ | vmware® | SUSE | Canonical Ubuntu | Red Hat

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

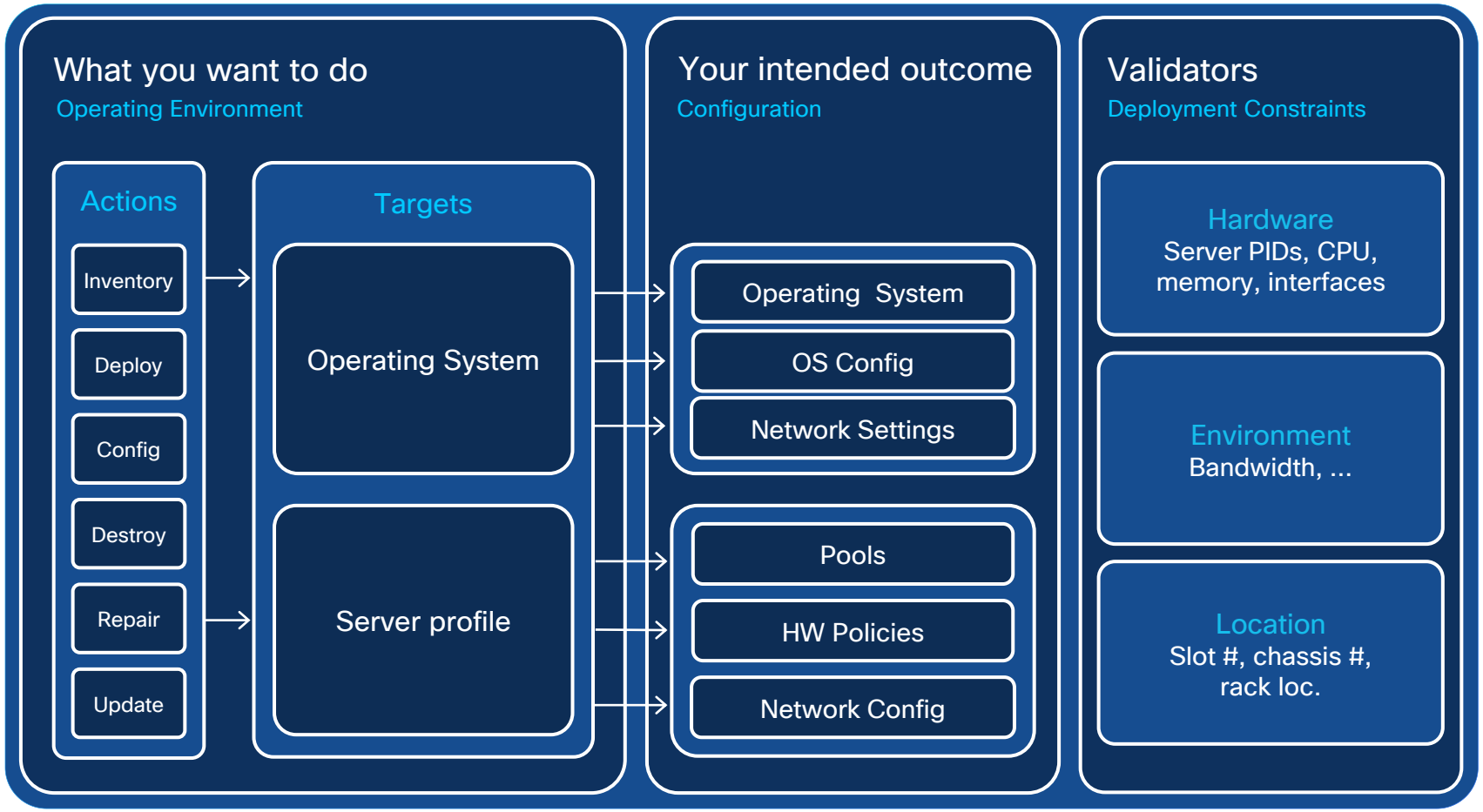
Blueprints

Compute Node

Blueprints are curated, tested, and authored by Cisco

Available in a marketplace as options for deployment

Maintained and fully supported by Cisco engineering and TAC



Supported Operating Systems

Windows Server | ORACLE | Rocky Linux™ | vmware® | SUSE | Canonical Ubuntu | Red Hat

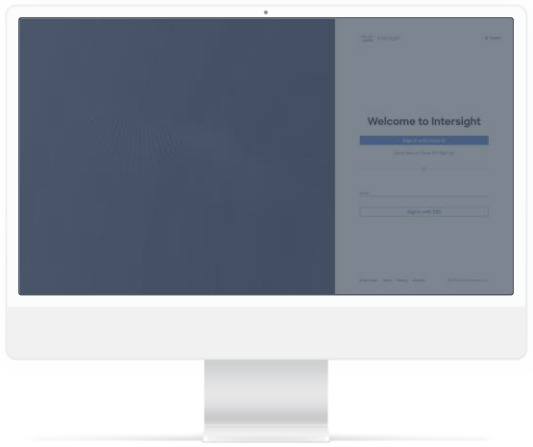
© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Day 2: Deployment at scale

Full-scale simplicity

Consistent, repeatable AI-ready infrastructure deployments across multiple sites

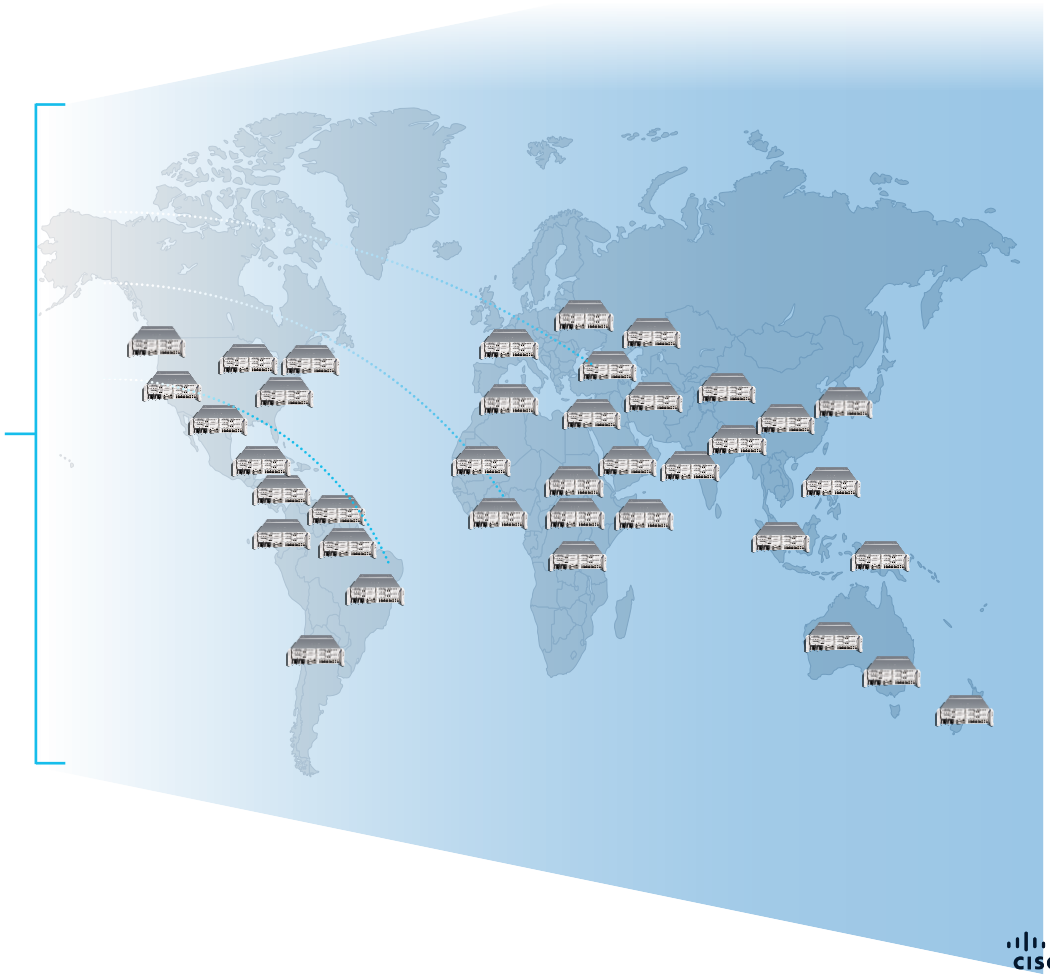


Cisco Intersight

Cisco Edge blueprints

(Retail, manufacturing, healthcare)

- Application
- Cluster
- O S
- Network config
- Server configuration
- GPU configuration
- Storage configuration
- Security configuration



© 2025 Cisco and/or its affiliates. All rights reserved.

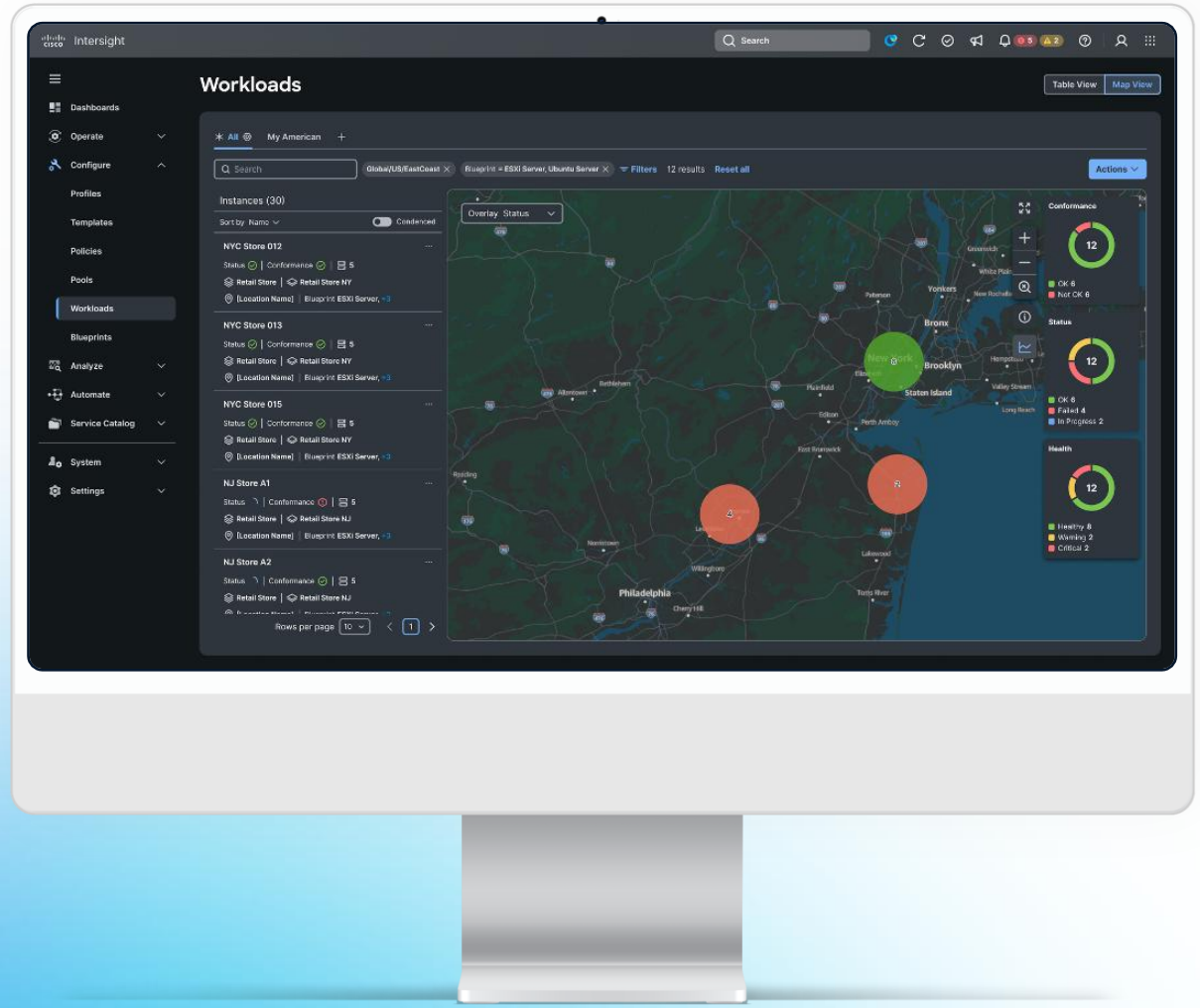
Cisco Confidential

Day 2: Location based global dashboards

Easily observe and remediate issues across a scale of thousands of systems

Overlay options of health, conformance & deployment cycle to switch context as needed

Cross-domain visibility into upstream networking health



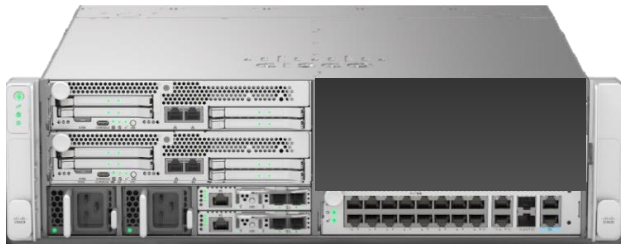
© 2025 Cisco and/or its affiliates. All rights reserved.

Day N: Seamless, scalable operations

Fleet management at global scale

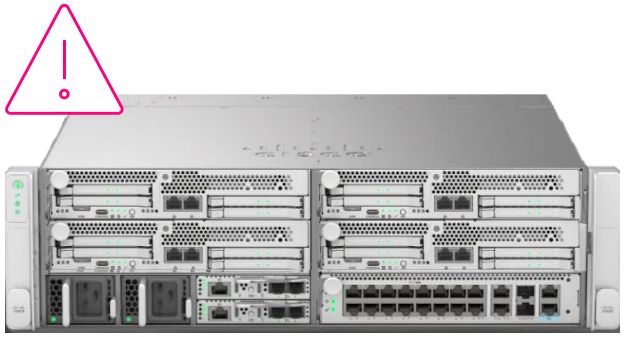
Full-scale simplicity

On-boarding a new service



1. New AI application requires GPU
2. Install a new compute node with GPU
3. No forklift upgrade

Avoiding unscheduled down-time



1. Intersight identifies predictive maintenance item
2. Replacement node sent out automatically
3. Easy on-site install

End to end visibility & observability

Fleet management

\+

splunk>

Observability

\+

ThousandEyes

Security

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Fleet Management Demo

Blueprinting tomorrow's computing with AI systems, not servers.

Future-ready performance	Full-scale simplicity	**Design-in Security**

© 2025 Cisco and/or its affiliates. All rights reserved.

Comprehensive security for the edge

Design-in Security

System protection

- Locking bezel with intrusion detection
- Configuration control via policy-based templates
- Firmware and hardware roots of trust

AI models & apps protection

- AI Defense
- Isovalent (K8s)
- CPU and GPU confidential computing

Network protection

- Firewall, DNS, IPS/IDS
- AMP, segmentation
- Umbrella, ZTNA, TrustSec

© 2025 Cisco and/or its affiliates. All rights reserved.

Centralized management, consistent and distributed security enforcement

Design-in Security

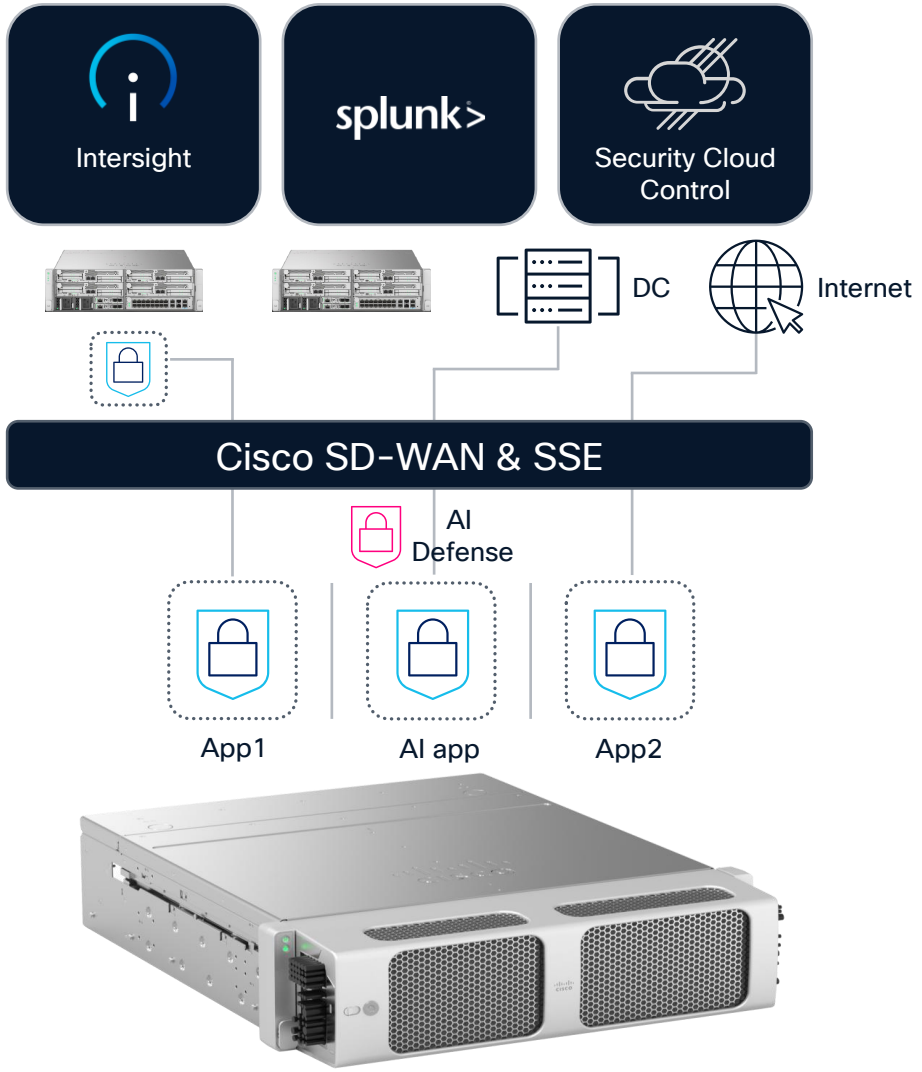
5 Observability & SOC alignment

4 Network security fabric

3 Workload & model security

2 Secure operations at scale

1 Hardware & platform integrity



© 2025 Cisco and/or its affiliates. All rights reserved.

Cisco Unified Edge AI-ready platform for the next decade

Future-ready performance	Full-scale simplicity	Design-in Security

© 2025 Cisco and/or its affiliates. All rights reserved.

cisco

Only Cisco can converge network, compute, storage, and security from the edge up.

Deep networking expertise combined with compute innovation.

Comprehensive SaaS management for global fleet operations.

Built-in, multi-layered security from device to data center.

Proven track record in enterprise infrastructure.

© 2025 Cisco and/or its affiliates. All rights reserved.

Complete your session evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to claim a Cisco Live T-Shirt.

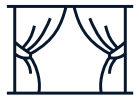
Earn up to 800 points by completing all surveys and climb the Cisco Live Challenge leaderboard.

Level up and earn exclusive prizes!

Complete your surveys in the Cisco Live Events app.

© 2025 Cisco and/or its affiliates. All rights reserved.

Continue your education



Visit the Cisco Stand for related demos

Book your one-on-one Meet the Expert meeting

Attend the interactive education with Capture the Flag, and Walk-in Labs

Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

© 2025 Cisco and/or its affiliates. All rights reserved.

Thank you

CISCO