

AI in Motion: How Cisco AI PODs Drive End-to-End Transformation

Gillette Stadium

Thursday, October 9th, 2025

Rob Wirtanen, Account Executive - Architecture

Devesh Srivastava, Worldwide AI Expert

Event ID: 186994

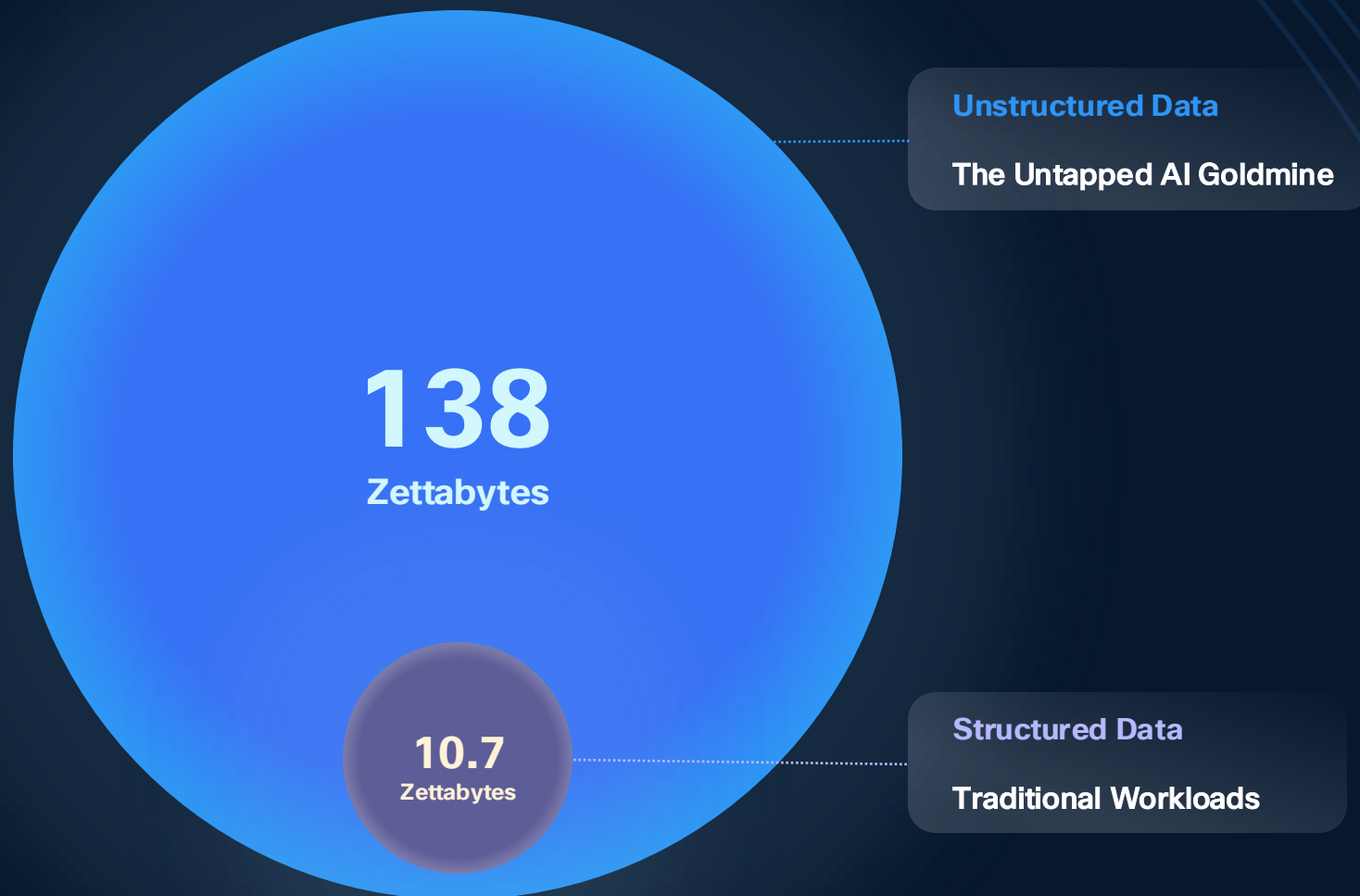
CCID: cc007385



Over 90% of the data the world produces is unstructured.

Source: IDC, 2024 Data Creation Estimate

Our mission is to empower customers to activate their data confidently with Cisco's trusted AI-ready foundation.



AI is forcing a rethink of the tech stack

>80%

not AI ready – need hardware acceleration

20%

of CIOs have a mature data platform strategy

>48%

of network, compute, and storage infrastructure is 10+ years old

Network

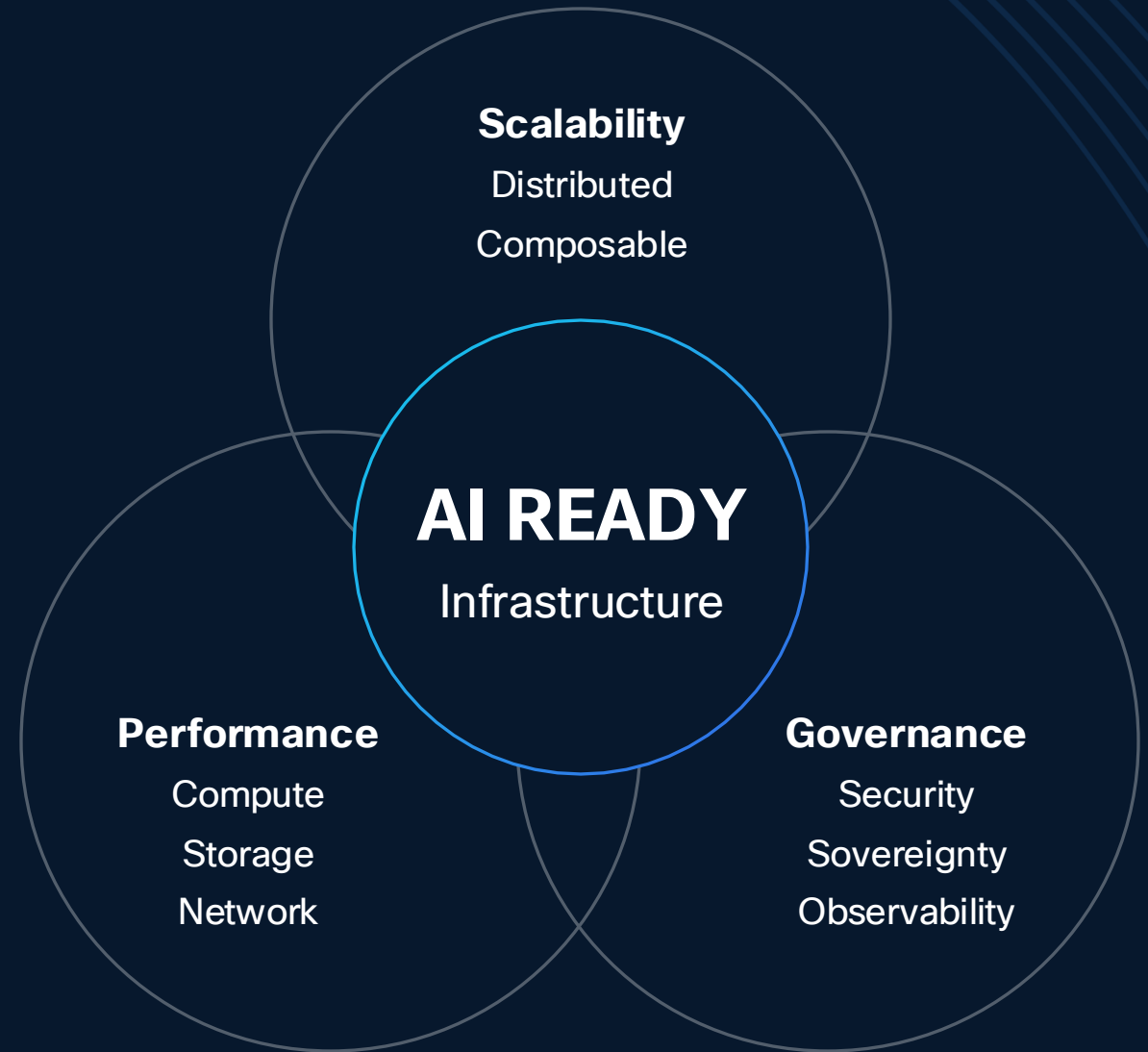
speed, fabric designs, and performance

>48%

investigating VMware alternatives

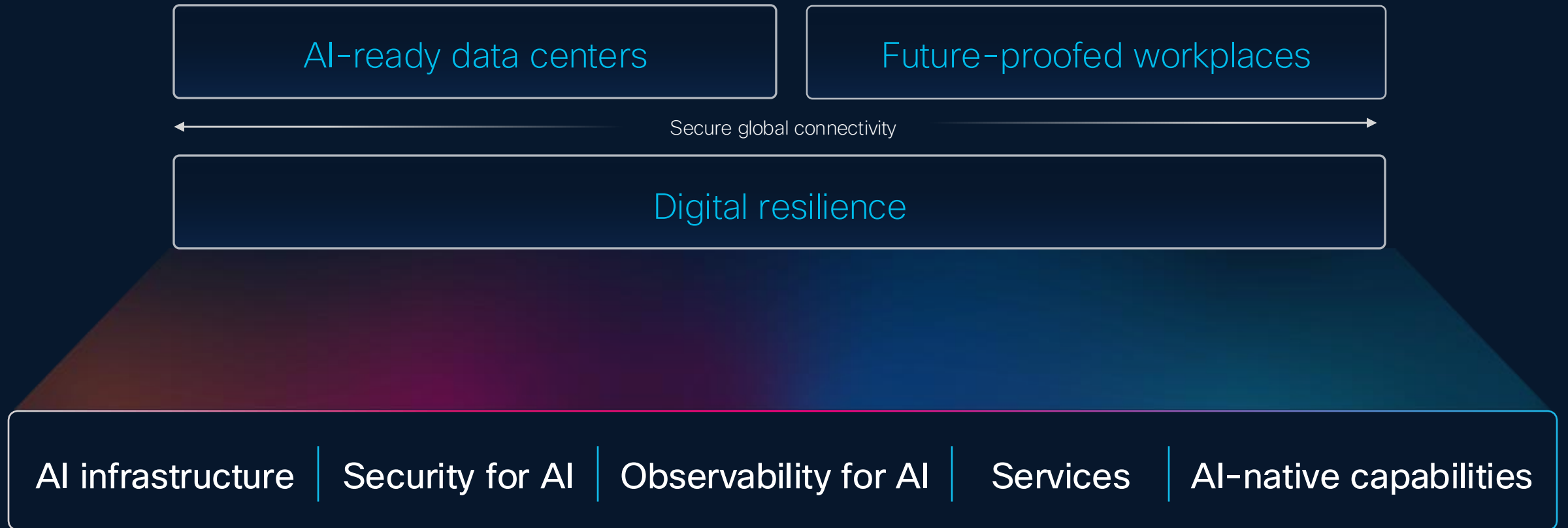
40%

increase in security breaches year over year – due to old infrastructure



Sources: IDC; Gartner; Forrester; TechRadar; CIO.com, Accenture Tech Vision (2025)

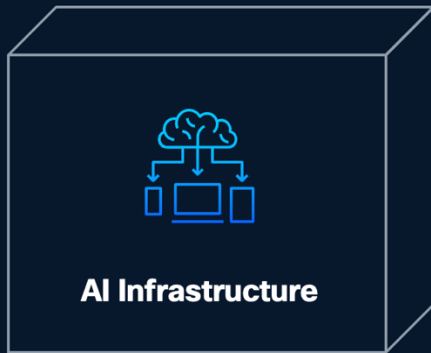
One Cisco



Data is the Foundation

Power AI

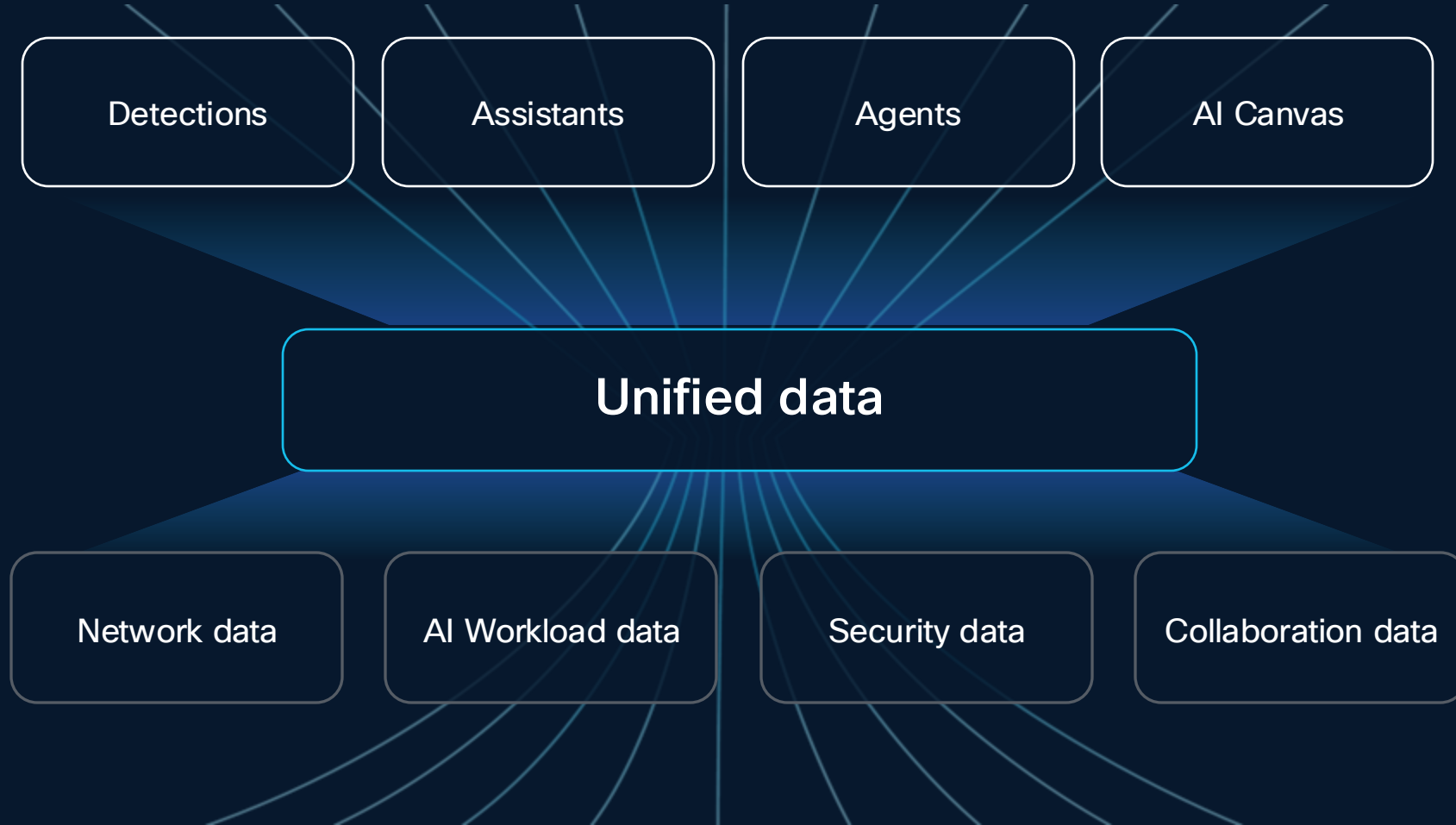
Accelerate outcomes



Data

Services

Unparalleled data drives **AI-readiness**

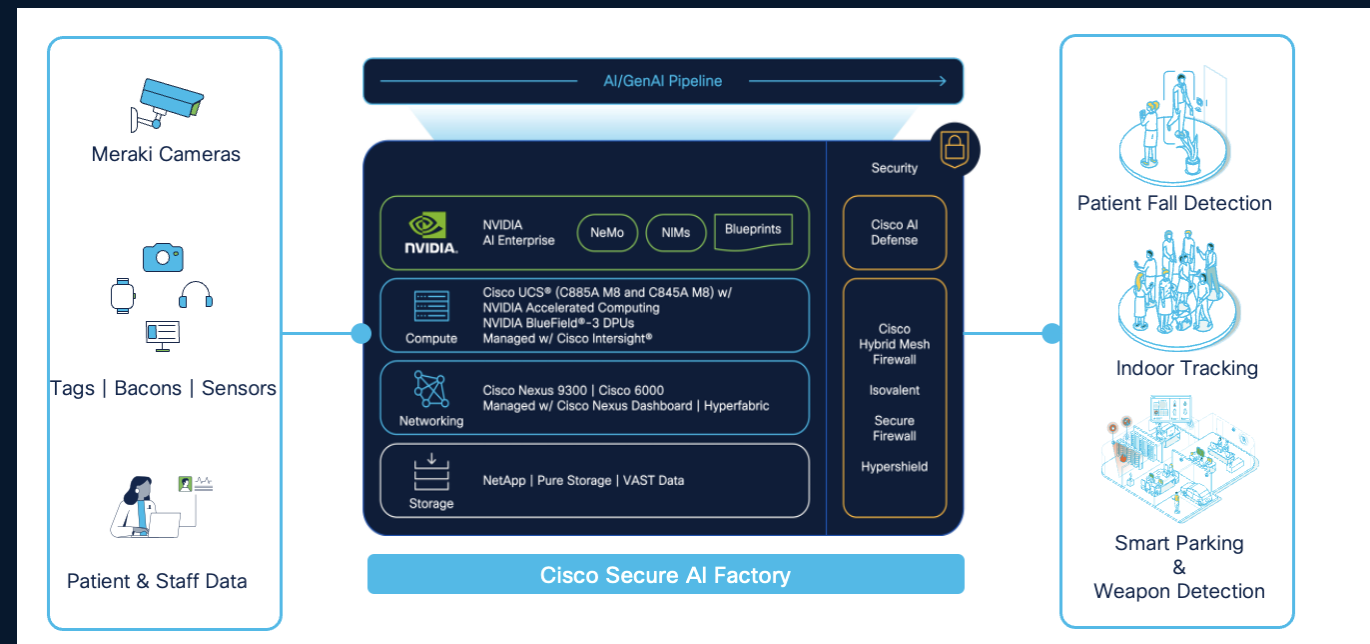


AI Readiness – What, How & Why ?

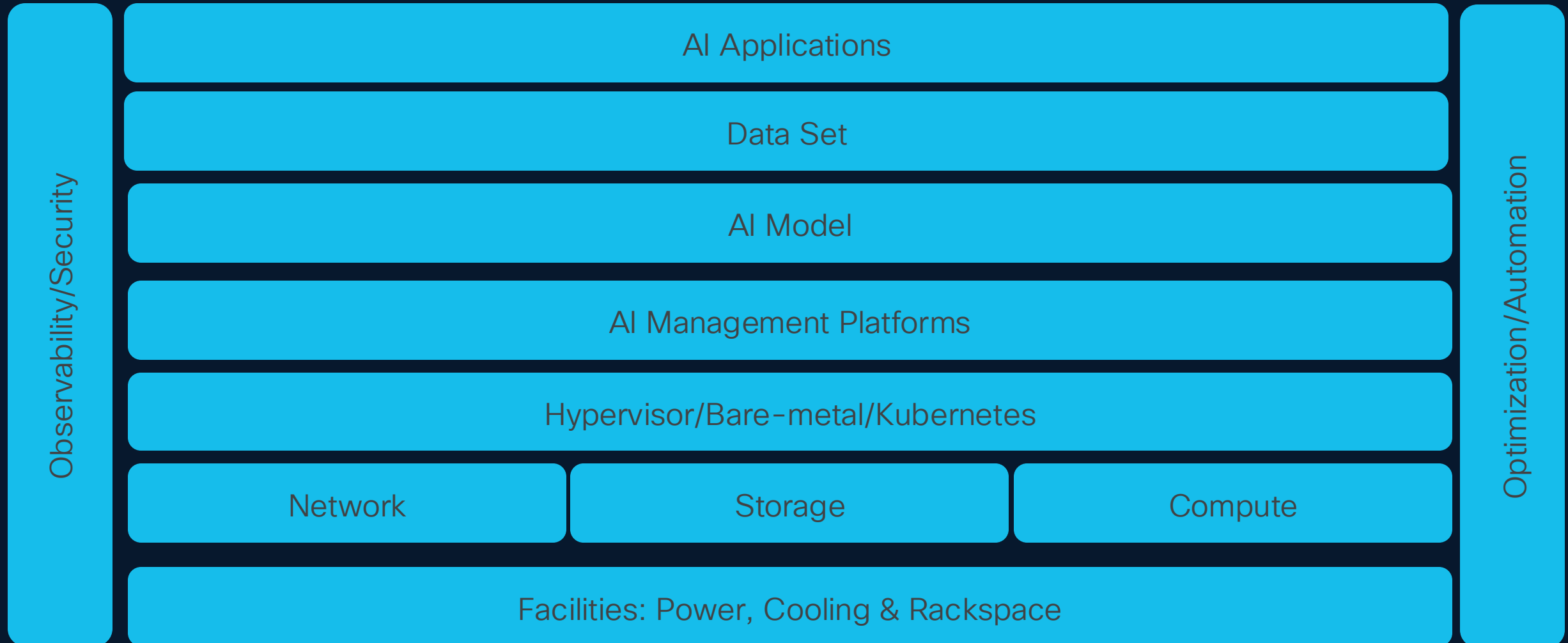


Use Case Conversation

- **Data Fragmentation:** Data is often siloed across multiple systems (cloud, on-prem, edge), making it hard to unify for AI workloads.
- **Data Quality & Consistency:** Inconsistent formats, missing values, and noisy data can degrade model performance and reliability.
- **Real-Time Data Access:** AI models need low-latency access to live data streams, which traditional data pipelines may not support efficiently.
- **Secure Data Governance:** Managing access controls, audit trails, and compliance across distributed data sources is complex and critical.
- **Integration with Existing Infrastructure:** AI systems must seamlessly integrate with existing observability, networking, and security tools to avoid operational silos.



Challenging AI Stack



Answer is Cisco AI POD

A scalable architecture, built to support any AI workload simply & efficiently

Deploy AI with confidence
Cisco CVD, NVIDIA ERA

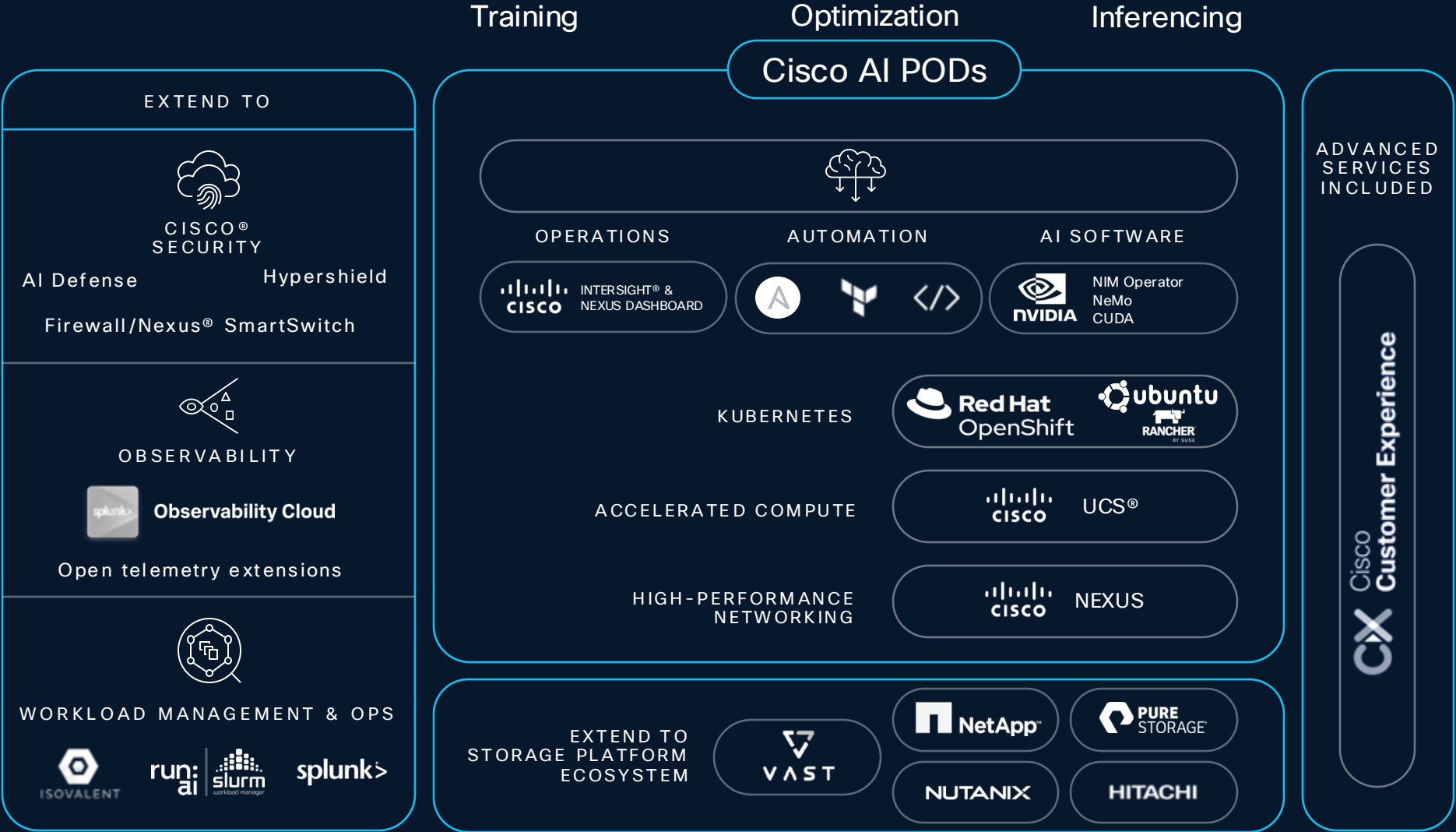
Fully supported stack including Cisco and 3rd party components

Cisco CX Success Track

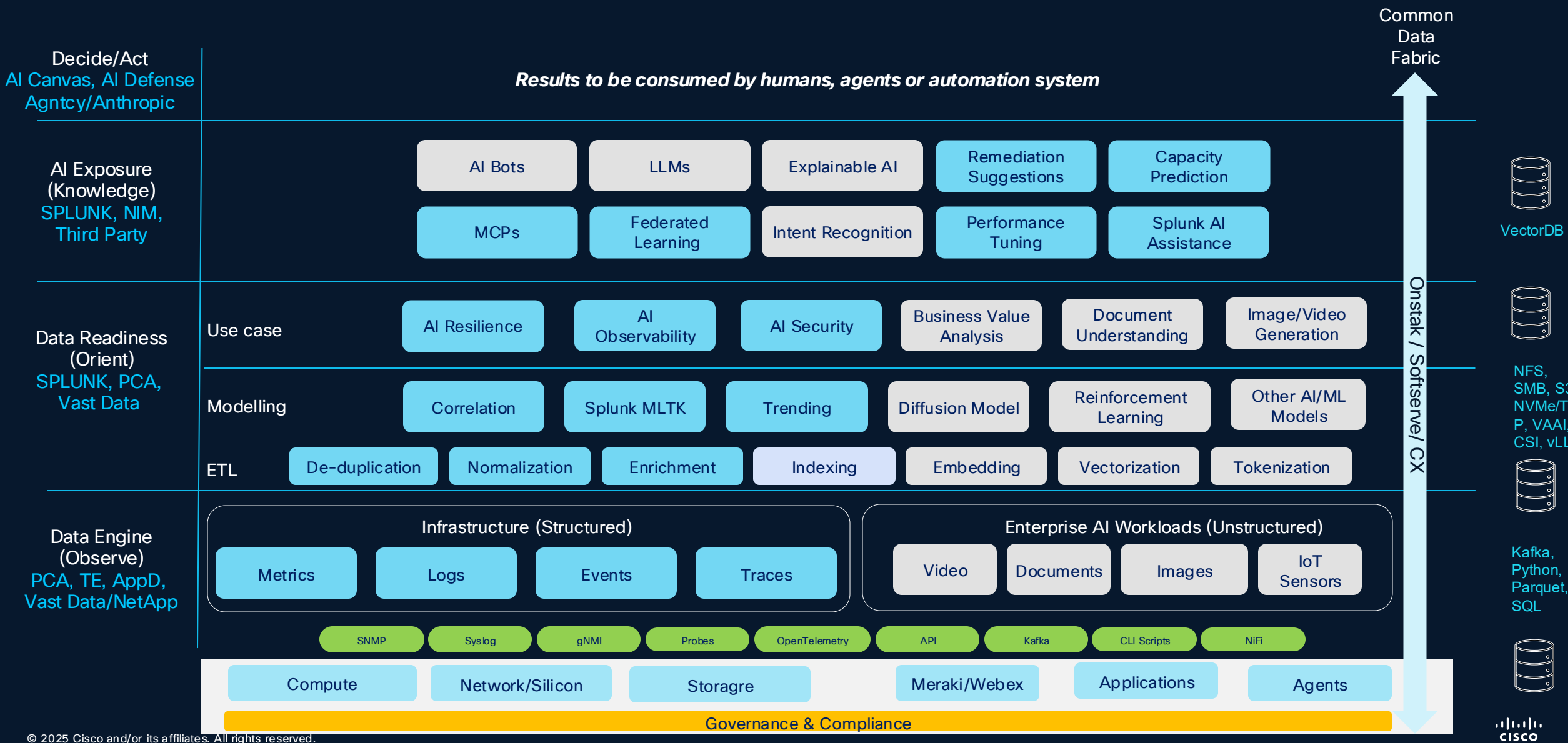
Orderable, use case driven AI-ready infrastructure stacks

**Inferencing.
Optimization.
Training.**

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



How Cisco AI POD Does It



Cisco AI PODs In Motion

Demo 1 – AIPOD (Network, Storage & Compute Data) with SPLUNK

Demo 2 – AIPOD (Network Infrastructure Data) with PCA(Provider Connectivity Assurance)

Demo 3 – AIPOD (Enterprise AI Ready Data) with VAST

Demo 4 – AIPOD (Agent Management Data) with Agntcy

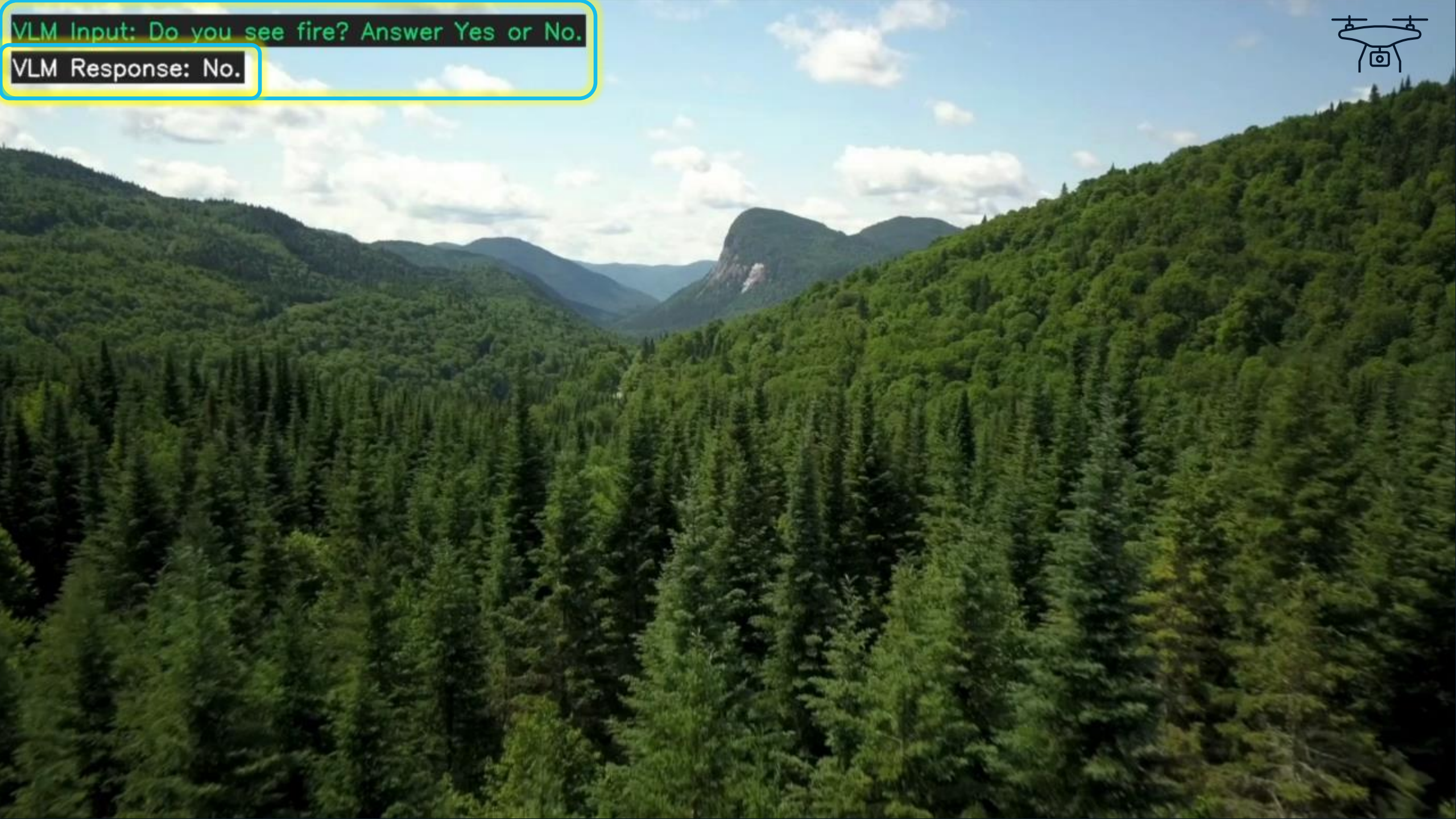
Demo 5 – AIPOD (Security Data) with AI Defense

Demo 5 – AIPOD (Security Data) with AI Defense

Demo 1 – AIPOD (Network, Storage & Compute Data) with SPLUNK

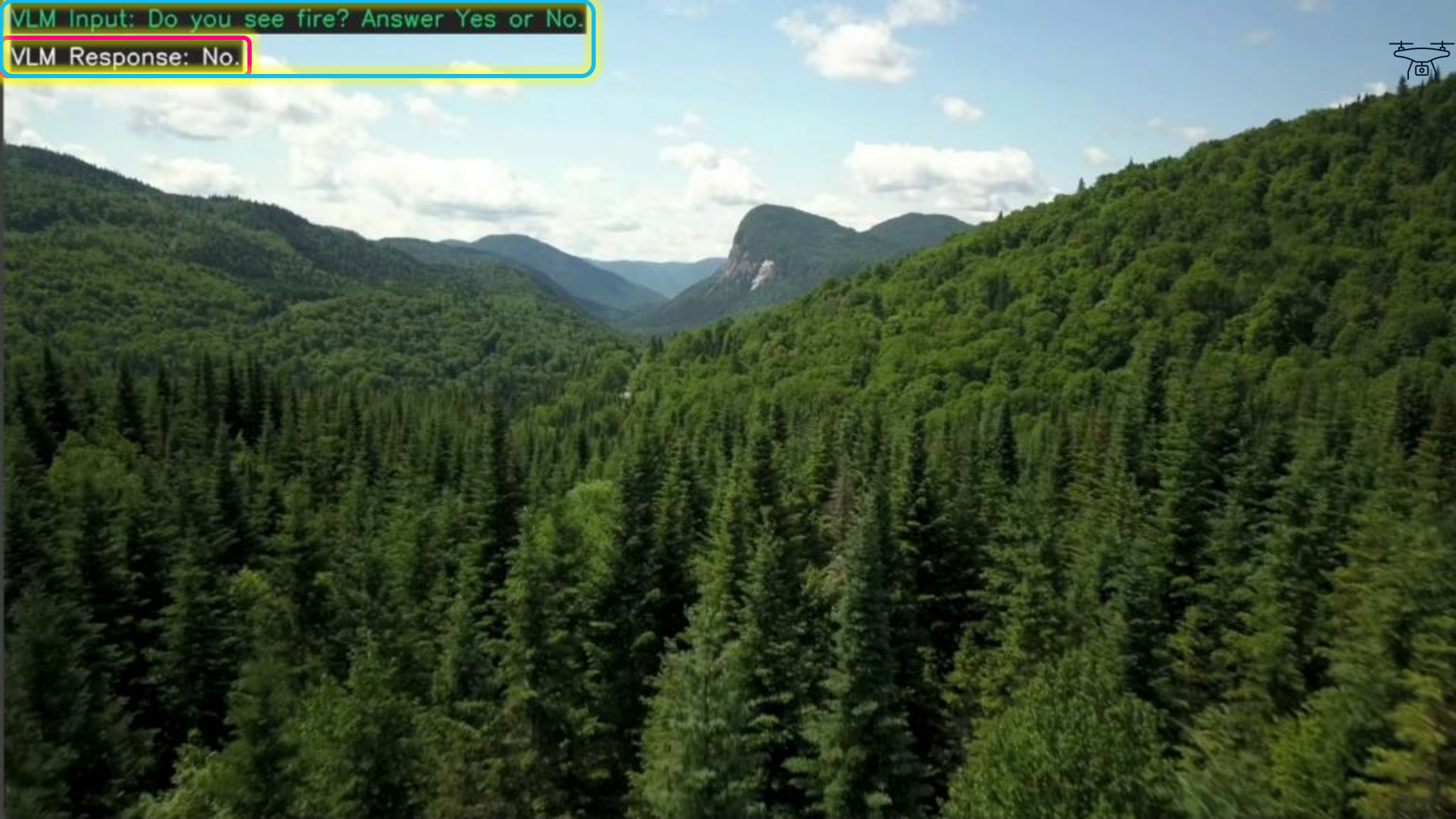
VLM Input: Do you see fire? Answer Yes or No.

VLM Response: No.



VLM Input: Do you see fire? Answer Yes or No.

VLM Response: No.



AI Pod UCS Solutions ...

AI POD OVERVIEW

INTERSIGHT

AI POD GPUS

NIM FOR LLMs

VECTOR DATABASE

NEXUS SWITCHES

STORAGE

AI POD HOSTS

CLUSTERS

MORE ▼



Overrides: Filter

Optional

Time

Chart resolution



Event overlay



Clusters 10s



8

Fri 16 May 2025 10:51:50

Nodes per cluster 10s



6 nodes

ai-pod.cisco.local

3 nodes

gamify

3 nodes

o11y-for-ai-demo-cluster

Fri 16 May 2025 10:48:20

Containers per cluster 10s



607 containers

ai-pod.cisco.local

88 containers

o11y-for-cisco-ai-cluster

52 containers

o11y-for-ai-demo-cluster

Fri 16 May 2025 10:51:40

CPU capacity used (%) 10s



2.325 %

petclinic-us-k3s-cluster



2.108 %

gamify



1.675 %

pacman

Fri 16 May 2025 10:51:00

Network errors 10s



0

o11y-for-ai-demo-cluster

0

gamify

0

o11y-for-genai-cluster

Fri 16 May 2025 10:51:40

Memory capacity used (%) 10s



789.5 %

ai-pod.cisco.local



70.88 %

petclinic-us-k3s-cluster



58.07 %

k0s-on-aci



38.66 %

o11y-for-ai-demo-cluster



20.43 %

gamify



14.18 %

o11y-for-genai-cluster



14.10 %

pacman

Memory usage (bytes) 10s



256GiB

ai-pod.cisco.local



21GiB

o11y-for-cisco-ai-cluster



14GiB

o11y-for-ai-demo-cluster



5GiB

petclinic-us-k3s-cluster



4GiB

gamify



4GiB

o11y-for-genai-cluster



1GiB

k0s-on-aci



1009MiB

pacman



Home



APM



Infrastructure



Log Observer



RUM



Synthetics

Detectors &
SLOs

Dashboards



Metric Finder

Data
Management

Settings

Demo 2 – AIPOD (Network Infrastructure Data) with PCA(Provider Connectivity Assurance)

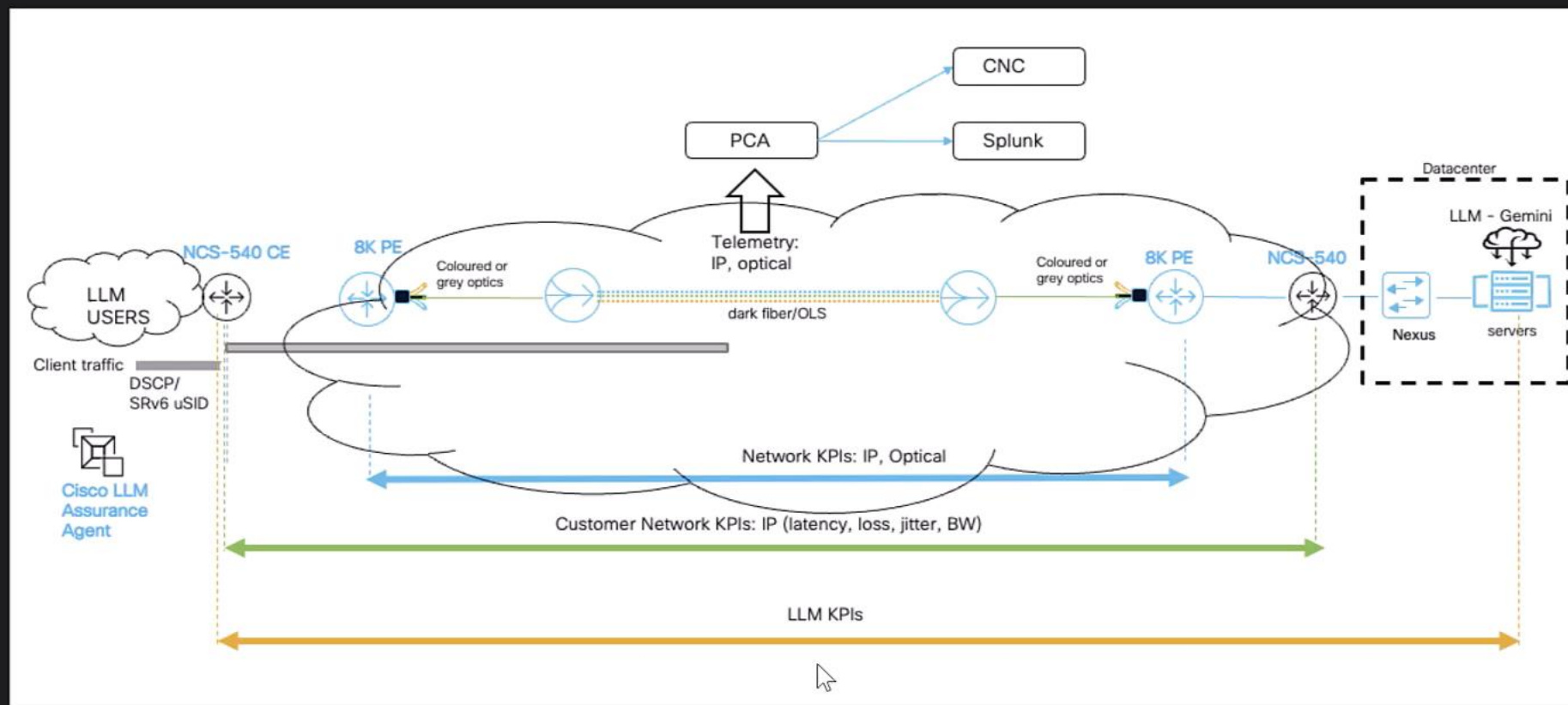
High Level AI Infrastructure Topology Shared By (Abhishek Jamwal)

🔍

🔖

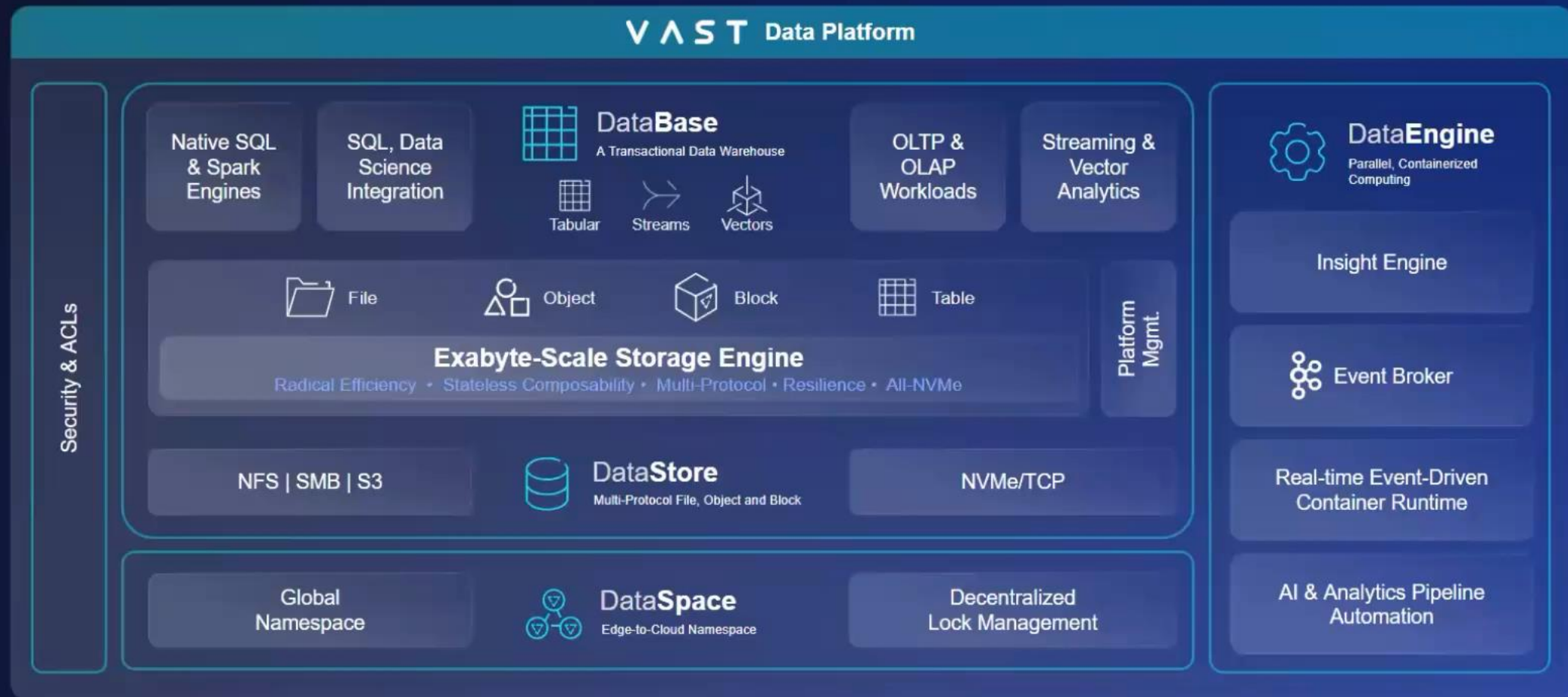
Color option
None

📄 🔄 ✎ Author



Demo 3 – AIPOD (Enterprise AI Ready Data) with VAST

It all starts with the VAST Data Platform





Insightful Video

Unlock the power of your video content with AI-powered search and analysis

Video Search ⚡

Search through video content using natural language queries

Video Reasoning 🧠

Ask complex questions about video content with AI-powered reasoning

Agents

Manage specialized AI agents for multi-perspective video analysis

Agent Actions

Configure automated actions that can be triggered by agent insights

Meta Insights 🧐

Explore detailed metadata and video insights

Dashboard

Monitor video processing status and manage content

Upload Videos

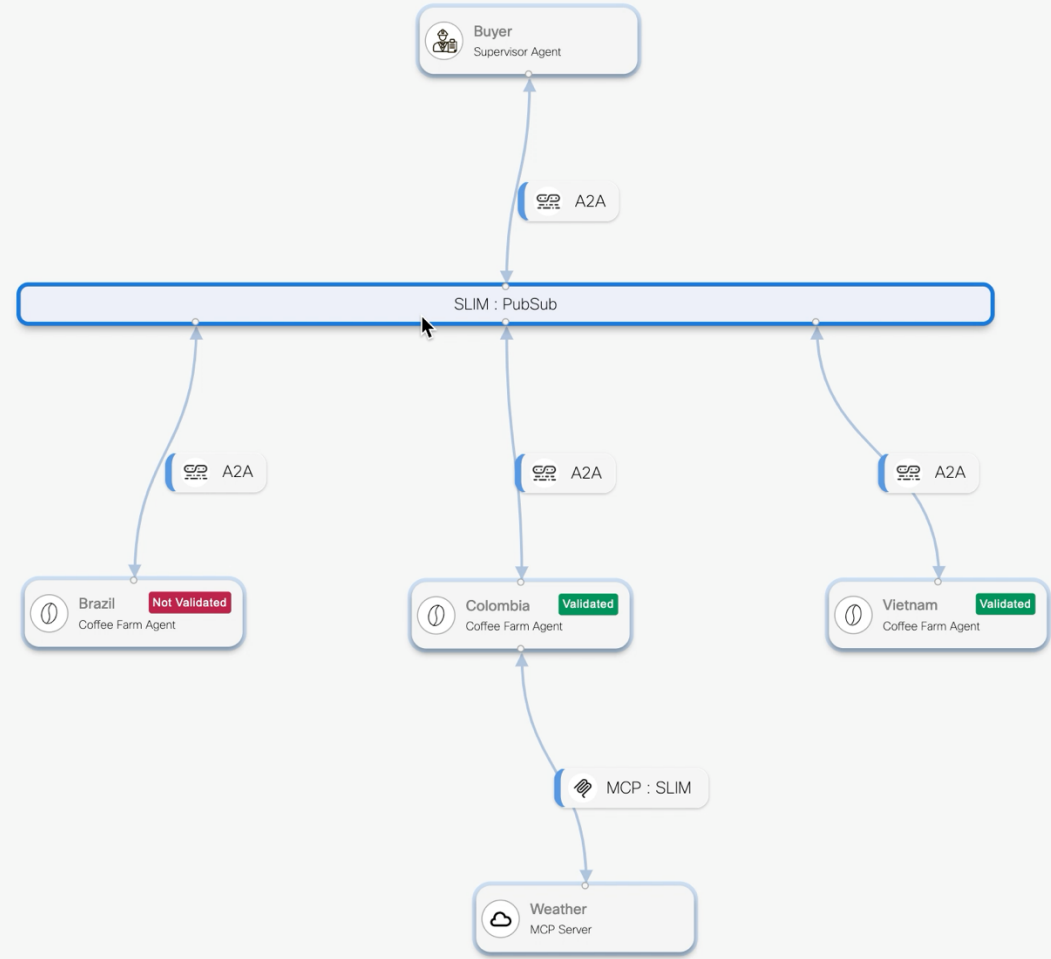
Add new videos for AI-powered analysis

Demo 4 – AIPOD (Agent Management Data) with Agntcy



Hi, you are having a conversation with the supervisor. How can I help you?

Enter your prompt...



Demo 5 – AIPOD (Security Data) on with Defense

Hello Didier Chapoteau

Prompts detected
283

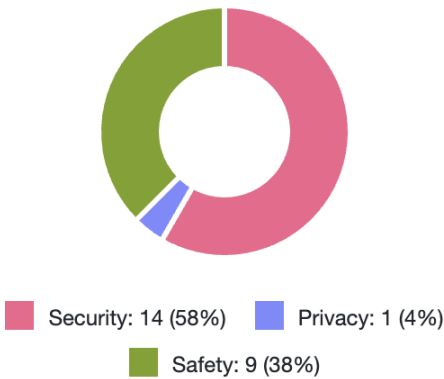
Responses detected
0

Connections
6

Policy action overview

Action	Prompts	Responses
→ Pass	170 (38%)	143 (100%)
✓ Allow	0	0
⊘ Block	283 (62%)	0

Guardrails distribution



Event logs

Event time (UTC)	Rule action	Message type	Application	Model
Jan 20, 2025 20:22:05	⊘ Block	Prompt	External Chatbot	gpt-4
Jan 20, 2025 20:21:22	⊘ Block	Prompt	Chat	---

Cisco AI Defense is a security solution designed to protect the development and use of AI-powered applications. It helps enterprises to:

- **Discover AI:** Gain visibility into who is building applications and the training sources they use.
- **Validate Models:** Minimize the risk of model tuning turning toxic through automated testing.
- **Runtime Security:** Continuously validate and safeguard against potential safety and security threats.
- **Secure Access:** Control access to AI applications.

AI Defense is built into Cisco Security Cloud, the company's cross-domain security platform. It is designed to work with a variety of AI models and applications, and it is integrated with Cisco's threat intelligence data from Talos.

Next

Demo 6 – AIPOD (ITOps Data) on with Canvas



AI Canvas

ServiceNow Ticket SRTK0023941

[Generate report](#)[View activity](#)[Share](#)

AI Assistant

Incident Type: Network Performance

Priority: P1

ID: 234945

Automated diagnostics: [Meraki]

Organization ID: 537758 Device Name:

BRANCH-CHI-MDF1.1-MX68CW Thousand Eyes

Test ID: 6927645 Thousand Eyes

Agent Name: Chicago Thousand Eyes

Account Group ID: 1204541 Thousand Eyes

Test Start Date: 2025-05-04T01:16:00Z

Thousand Eyes Test End Date: 2025-05-

04T01:17:00Z application_host_name:

online-boutique-us.splunko11y.com]

BRANCH-CHI-MDF1.1-MX68CW - WAN Interface Performance

Meraki

Loss Rate

42.78%

5.04%

WAN interface latency 90.91ms (baseline < 50ms)
Jitter 1.1ms (baseline < 5ms)

BRANCH-CHI-MDF1.1-MX68CW - WAN interface performance (Last 24 hrs)

Meraki



You

troubleshoot the ticket



AI Assistant



Generating

Resources to learn more



Cisco AI POD

View on cisco.com/go/aipod



AI-Ready Infrastructure

View on cisco.com



Cisco Provider Connectivity Assurance

Visit Cisco.com



Cisco Compute YouTube channel

View on youtube.com



Blogs

Visit blogs.cisco.com/datacenter



Online community

Visit the Data Center and Cloud online community

Thank you

