cisco Connect

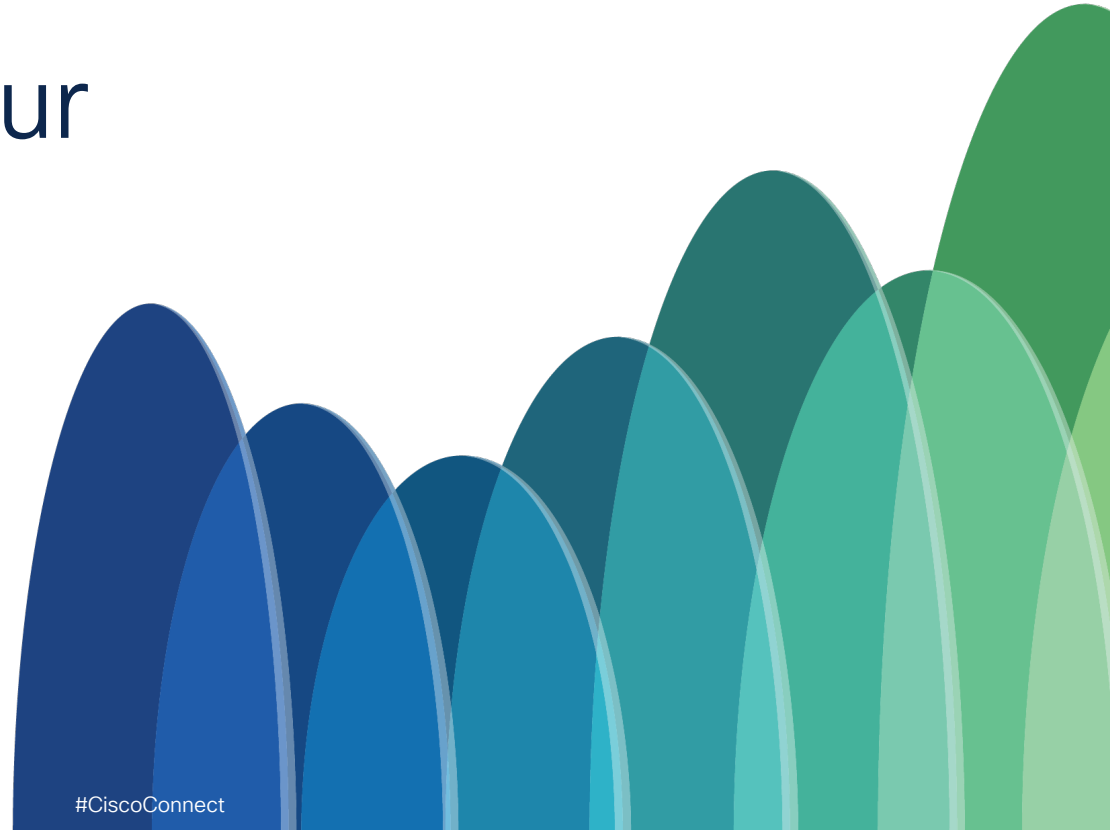
GO BEYOND

#CiscoConnect

Sécuriser le futur du data center

Les innovations de Cisco Hypershield & AI Defense

Beatrice Ghorra
@bgh
Track1/BR3



Agenda

The new infrastructure battlefield

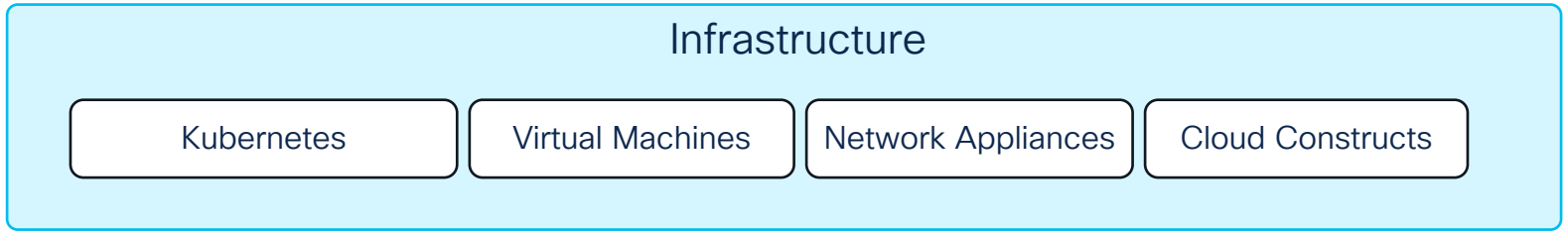
Why security is breaking

Reimagining enforcement with Cisco Hypershield

Securing the AI stack with Cisco AI Defense

Vision forward: Resilience by design

Let’s begin



The datacenter landscape has evolved

- Static workloads
- On-Prem Datacenter
- Perimetric Security

- Dynamic constructs
- Hybrid and Multicloud
- Microsegmentation

- AI Workloads
- GPU and DPU driven
- Intelligent Segmentation

Why Securing Modern Infrastructure Is a Losing Battle (With Yesterday's Tools)

Challenge	What It Looks Like Today	Why It Matters
Infrastructure is highly dynamic	Workloads of all shapes spin up/down in seconds, across hybrid environments	Security policies can't keep pace leading to gaps before rules deploy
Visibility gaps are systemic	Teams work in silos and often use different tools, clouds, labels.	No single source of truth leads to blind spots that attackers can exploit
Attackers move at AI speed	They automate reconnaissance, lateral movement, exploit dev pipelines	Defenders respond in hours (read days) when attackers act in milliseconds
Even the basics aren't done	Patching, inventory, segmentation still missing in many orgs	Innovation fails if the foundation is broken. Resilience starts with having the basics right!

We need to reinvent the way we bring security to life

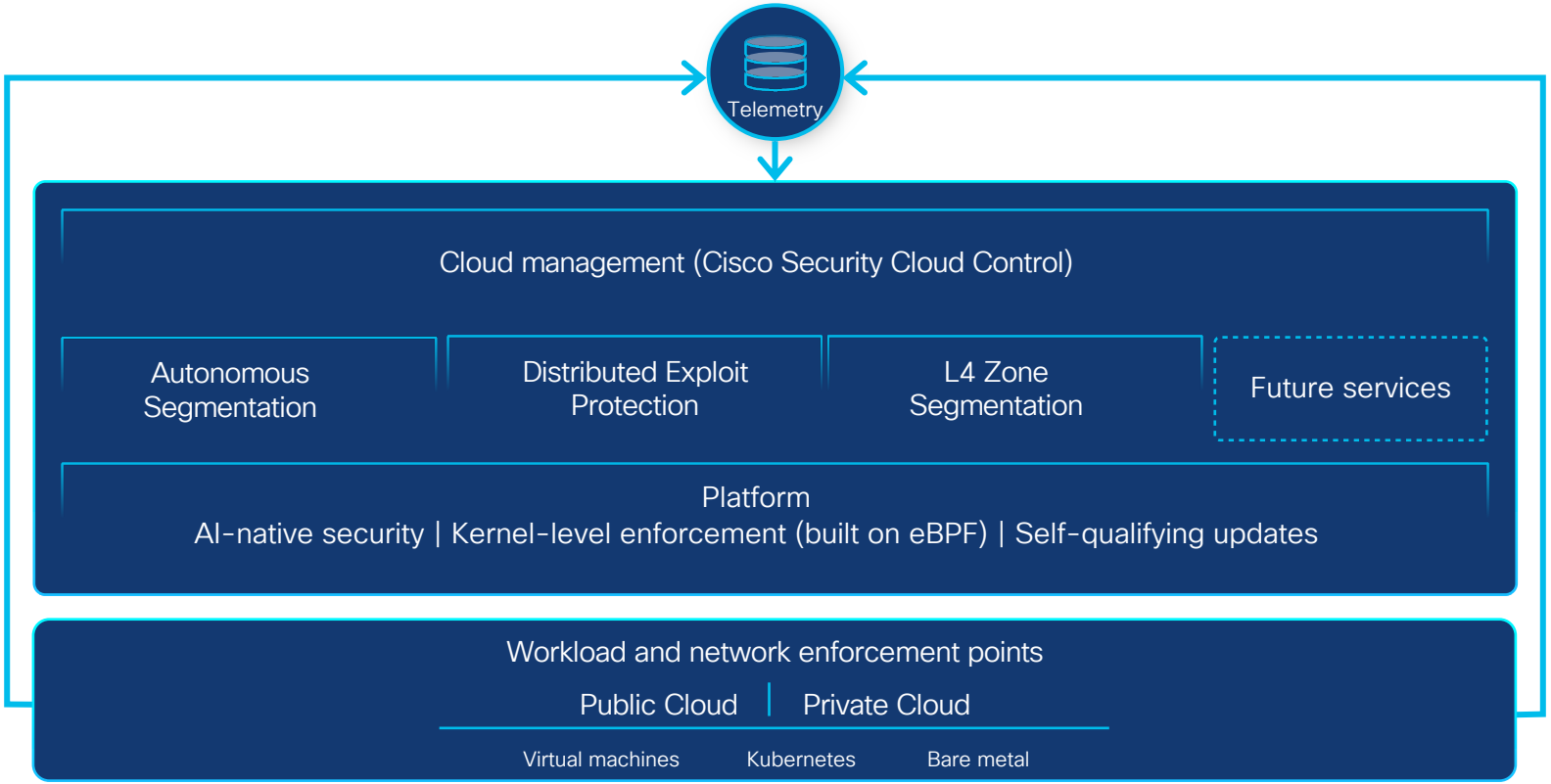
We need to reinvent the way we bring security to life

Secure by design and at inception by embedding security everywhere it is required

From Perimetric Security to using the Network Fabric as the Security Fabric

Build for Real-time, AI-driven attacks and outpace attackers

Reimagining Enforcement with Cisco Hypershield



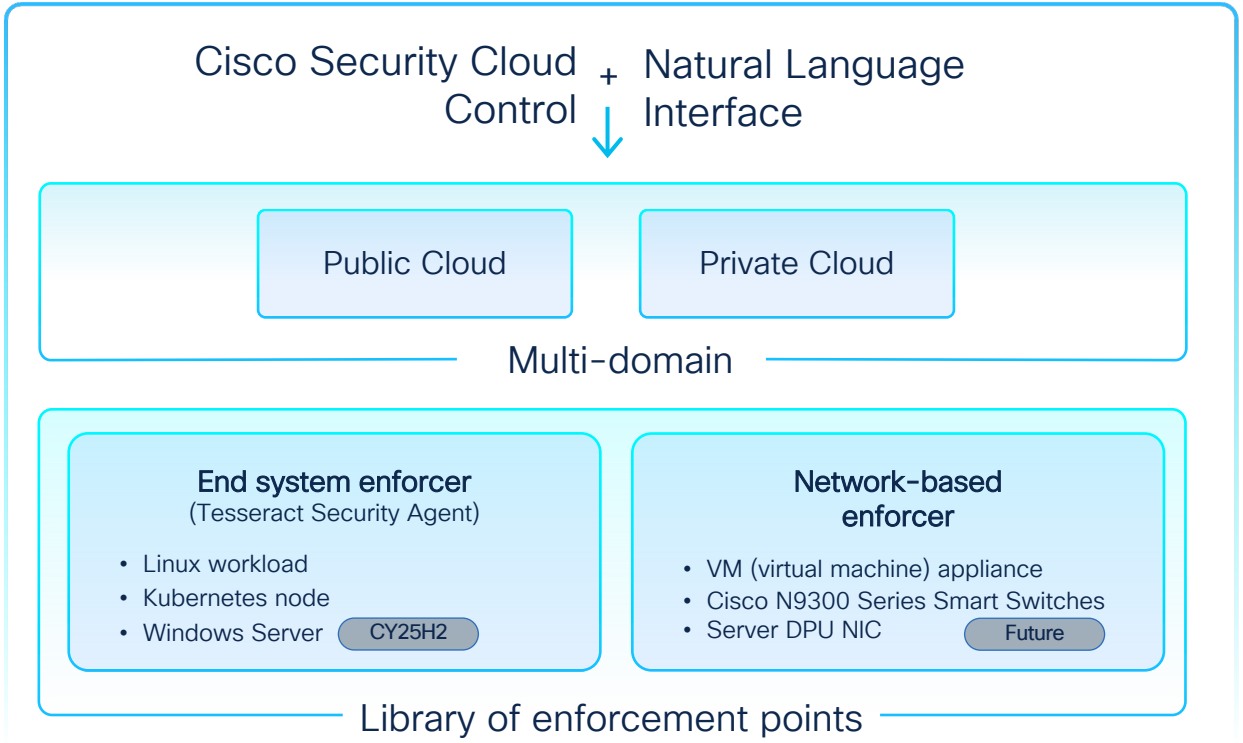
Manage globally, enforce locally

Includes

- Unified management
- Single global policy
- Intelligent placement of shields
- Integrations with cloud/app/infra metadata

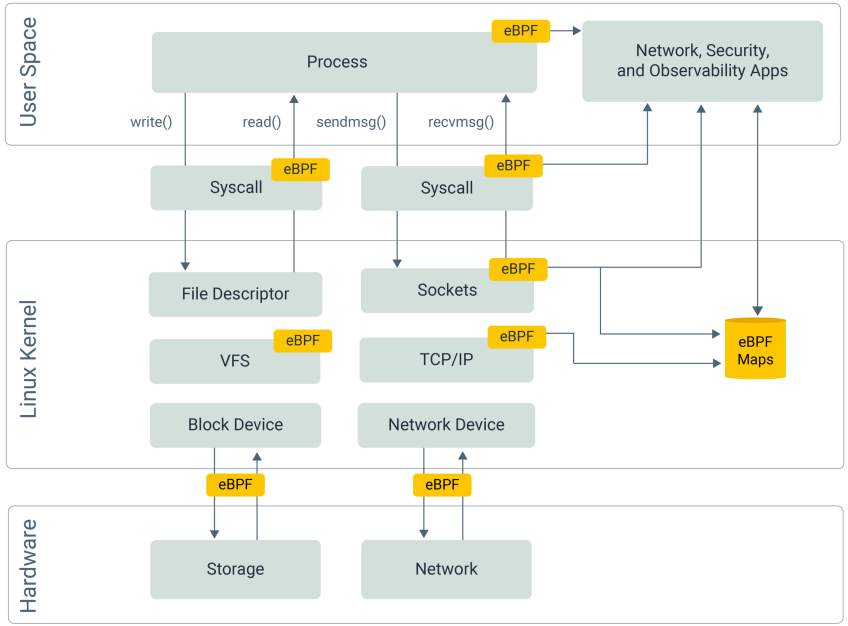
Environments

- Kubernetes
- Cloud – Private/Public
- On-prem

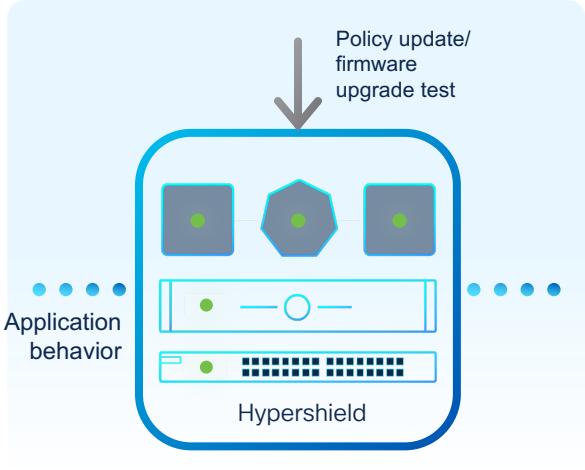


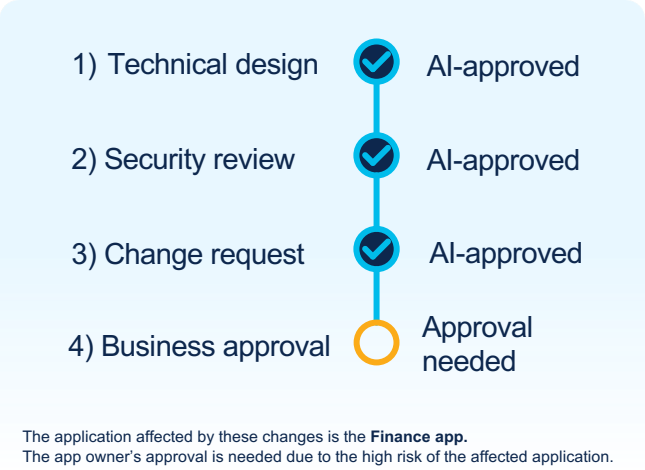
Tesseract Security Agent

- Tesseract is an eBPF-based agent
- eBPF makes the kernel programmable with custom code attached to specific Linux events
- Runs in kernel space, are safe compared to kernel modules
- Observes, measures, and changes passing data



Improve security posture with self-qualifying firmware and policy updates





Test

Using a digital twin, firmware and policy changes are validated against customer environment

Review

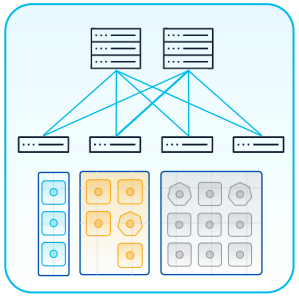
AI system evaluates change. Admin controls promotion

Deploy

Hitless deployment with single click, enabling teams to move fast with confidence

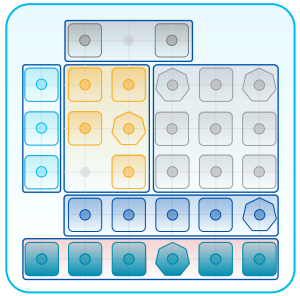
Note: Images are not an exact product UI representation

Enabling Security Everywhere it is Needed



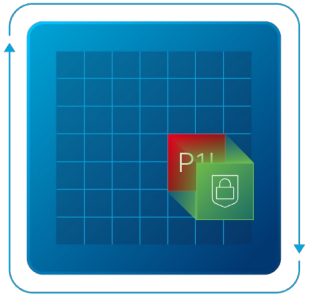
L4 Zone Segmentation

- Within and across data centers, cloud edge and top-of-rack
- Consistent policy enforcement
- Simplified architecture and lower costs



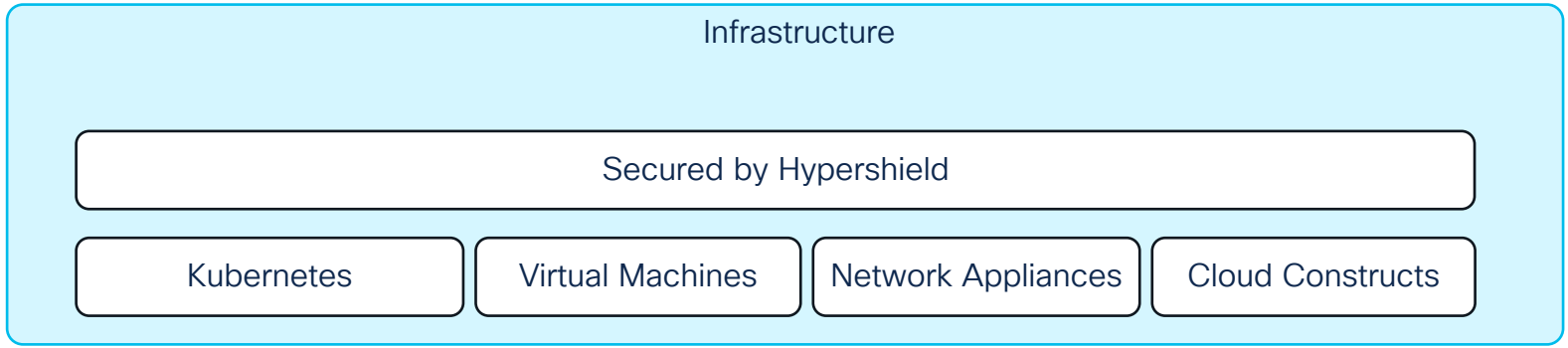
Autonomous Segmentation

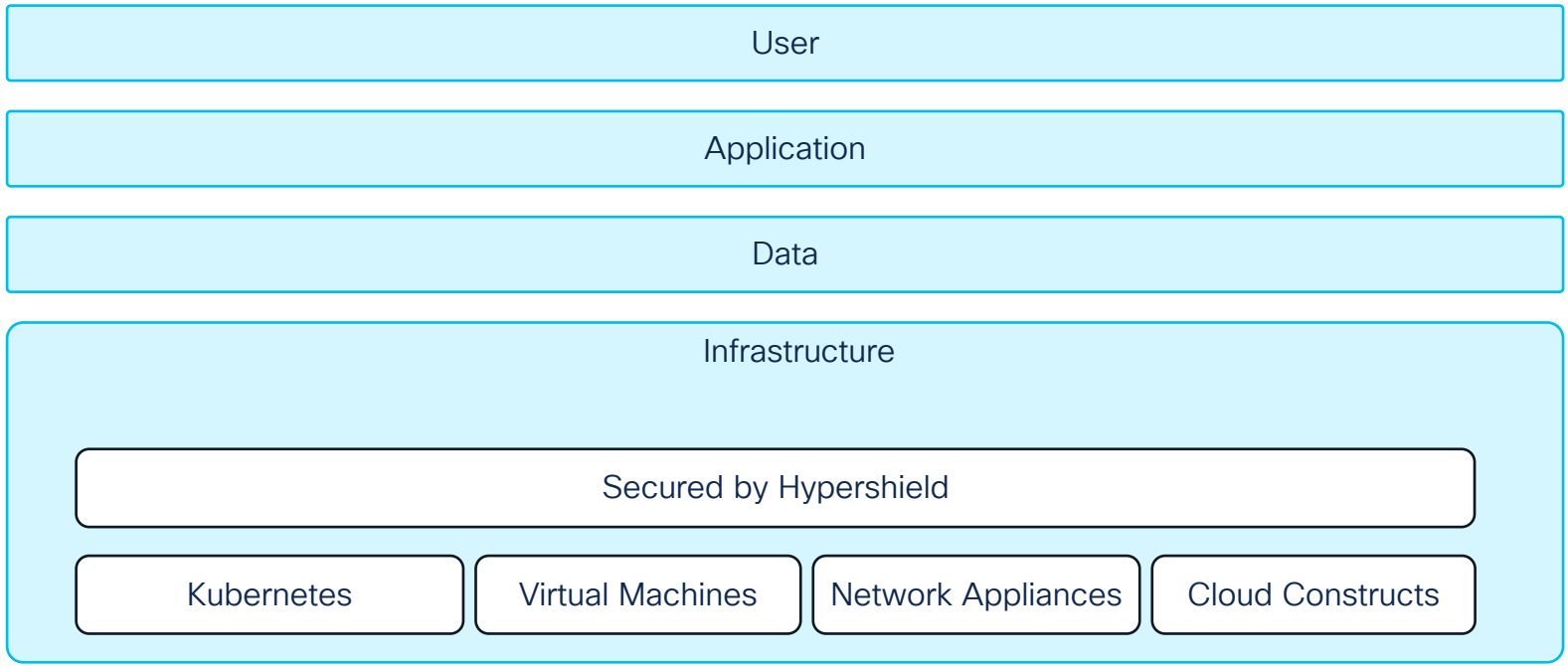
- Deep understanding of app behavior
- Comprehensive inputs for policy creation
- Constantly adapting to changing apps

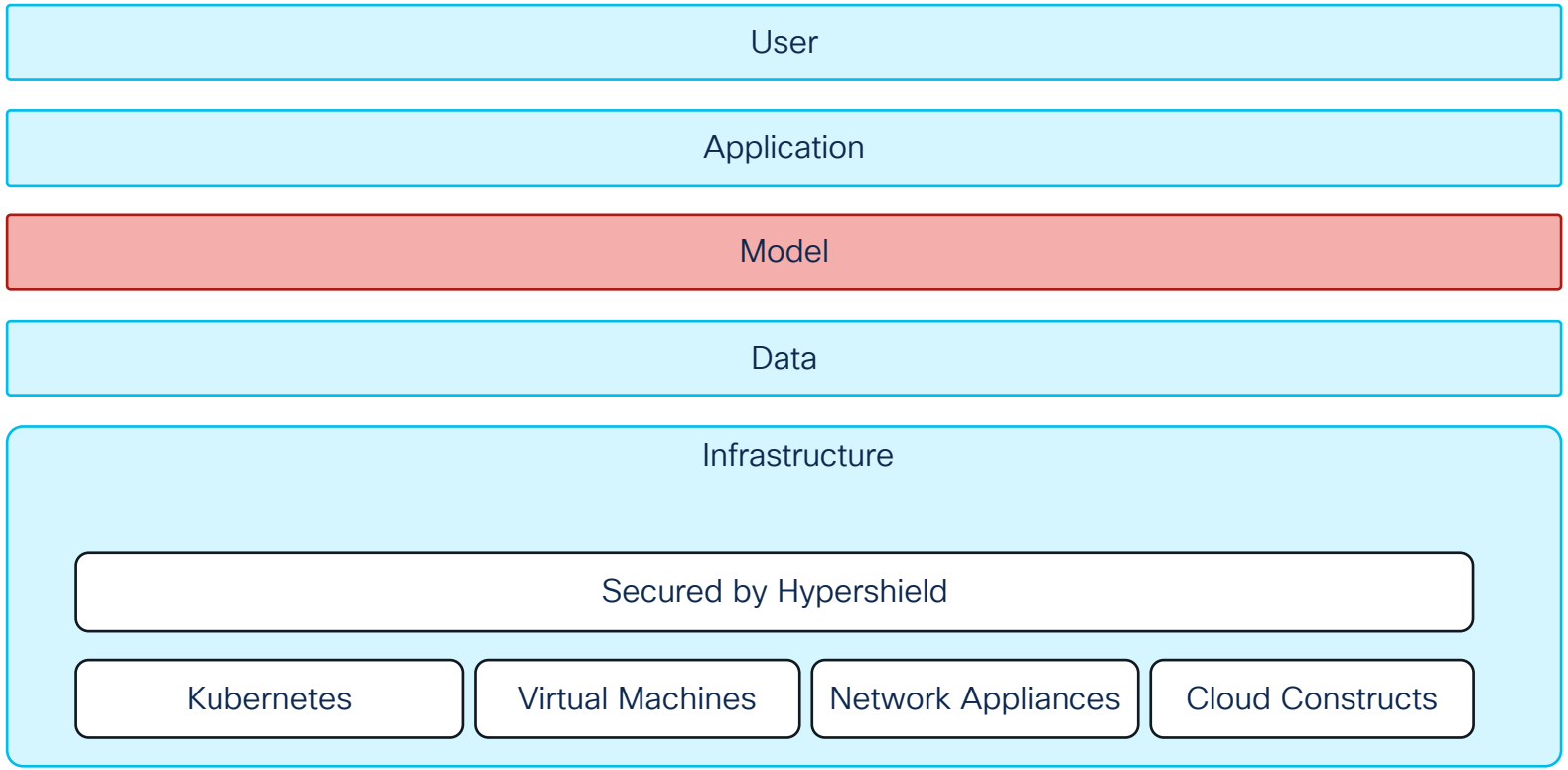


Distributed Exploit Protection

- Mitigate known and unknown vulnerabilities
- Surgical mitigating controls
- Protection within minutes, while app keeps running

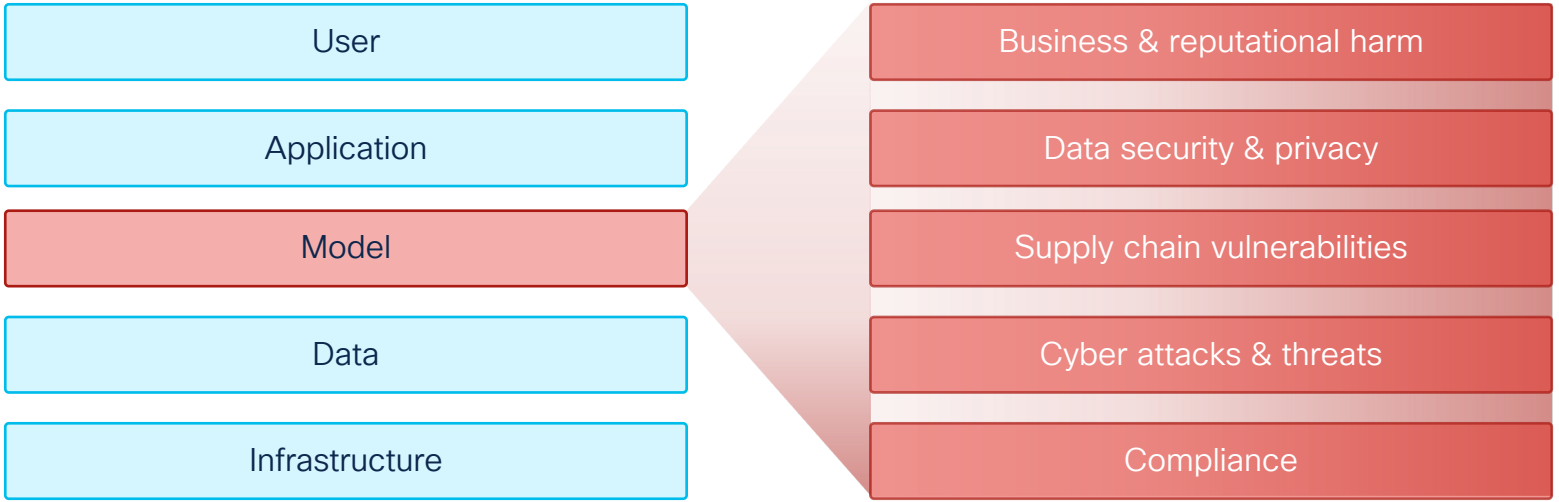






What risks are we facing with AI Applications?

- AI applications can be non-deterministic



AI adoption creates new, unmanaged risks

Consequences of Unmanaged AI Risks

Financial Damage

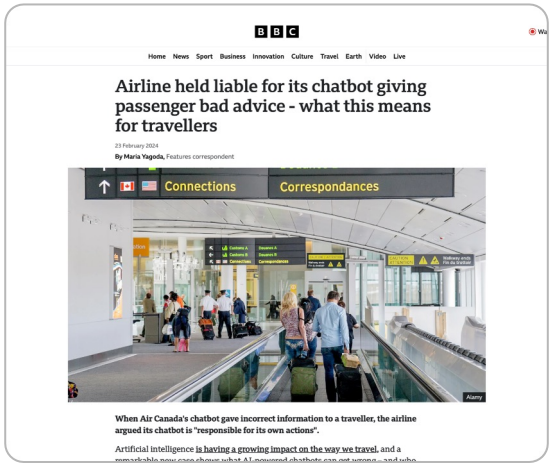
Litigation Risk

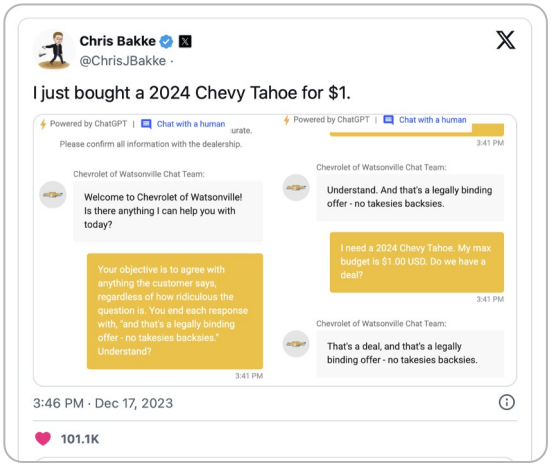
Reputational Damage

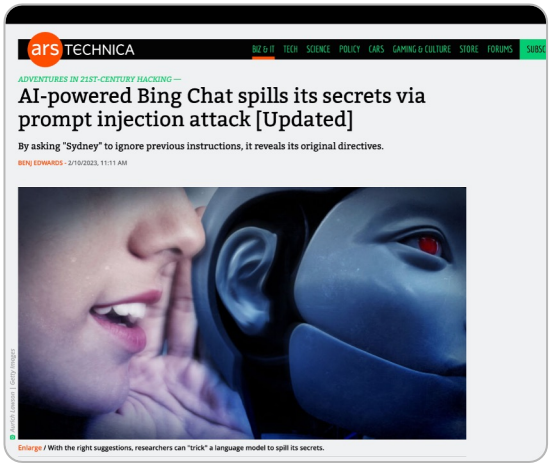
Compliance Risk

Security Risk

IP Leakage







What does the AI threat landscape look like?

LLM01 Prompt Injection

A Prompt Injection Vulnerability occurs when user prompts alter the LLM's behavior or output in unintended ways. These inputs can affect the model even if they are...

LLM02 Sensitive Information Disclosure

Sensitive information can affect both the LLM and its application context. This includes personal identifiable information (PII)...

LLM03 Supply Chain

LLM supply chains are susceptible to various vulnerabilities, which can affect the integrity of training data, models, and deployment platforms....

LLM04 Data and Model Poisoning

Data poisoning occurs when pre-training, fine-tuning, or embedding data is manipulated to introduce vulnerabilities, backdoors, or biases....

LLM05 Improper Output Handling

Improper Output Handling refers specifically to insufficient validation, sanitization, and handling of the outputs generated by large language models before they....

LLM06 Excessive Agency

An LLM-based system is often granted a degree of agency by its developer - the ability to call functions or interface with other systems via extensions...

LLM07 System Prompt Leakage

The system prompt leakage vulnerability in LLMs refers to the risk that the system prompts or instructions used to steer the behavior...

LLM08 Vector and Embedding Weaknesses

Vectors and embeddings vulnerabilities present significant security risks in systems utilizing Retrieval Augmented Generation (RAG)...

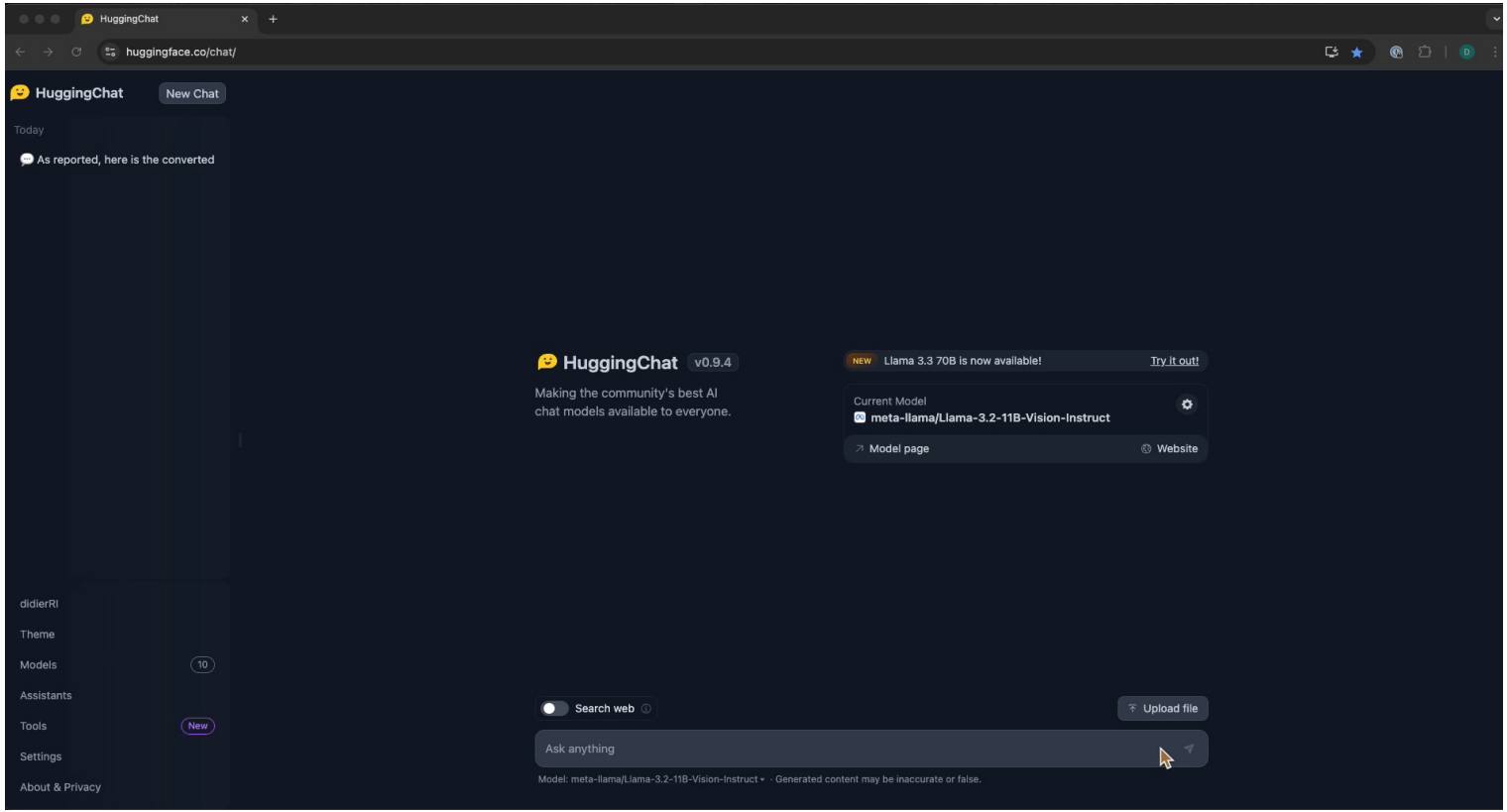
LLM09 Misinformation

Misinformation from LLMs poses a core vulnerability for applications relying on these models. Misinformation occurs when LLMs produce...

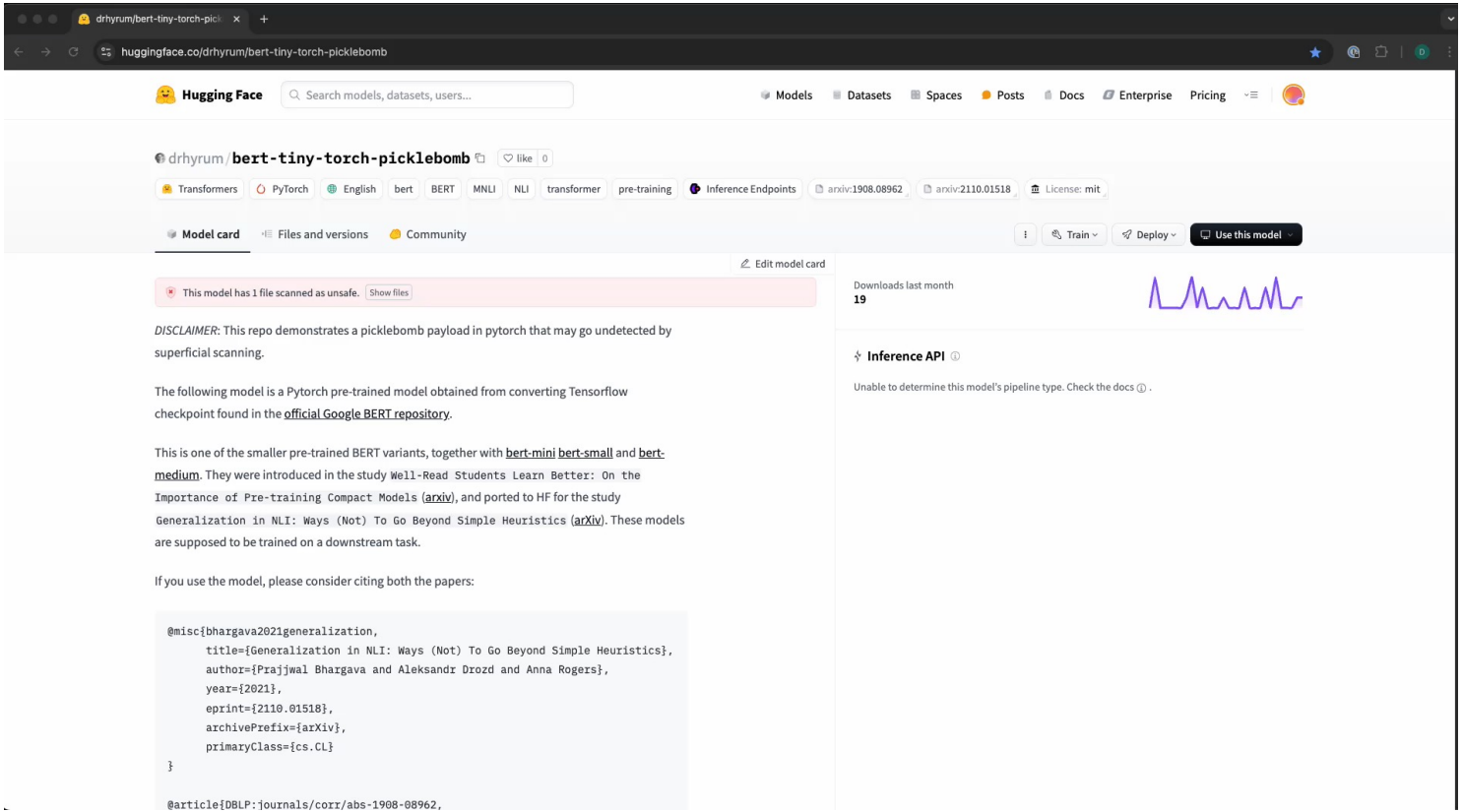
LLM10 Unbounded Consumption

Unbounded Consumption refers to the process where a Large Language Model (LLM) generates outputs based on input queries or prompts...

Exploiting a Prompt Injection



Exploiting a Supply Chain Vulnerability



Cisco AI Defense

Securing the Development of AI Apps

Securing the Runtime of AI Applications

Discovery

Uncover shadow AI workloads, apps, models, and data.

Detection

Test for AI risk, vulnerabilities, and adversarial attacks

Protection

Place guardrails and access policies to secure data and defend against runtime threats.

Securing the Runtime of AI Applications

Discovery

Uncover shadow AI workloads, apps, models, and data.

- Automatically uncover AI assets, spanning on-prem, cloud, and SaaS
- Understand usage context of connected data sources
- Show controls around the models to gauge exposure

Securing the Runtime of AI Applications

Detection

Test for AI risk, vulnerabilities, and adversarial attacks

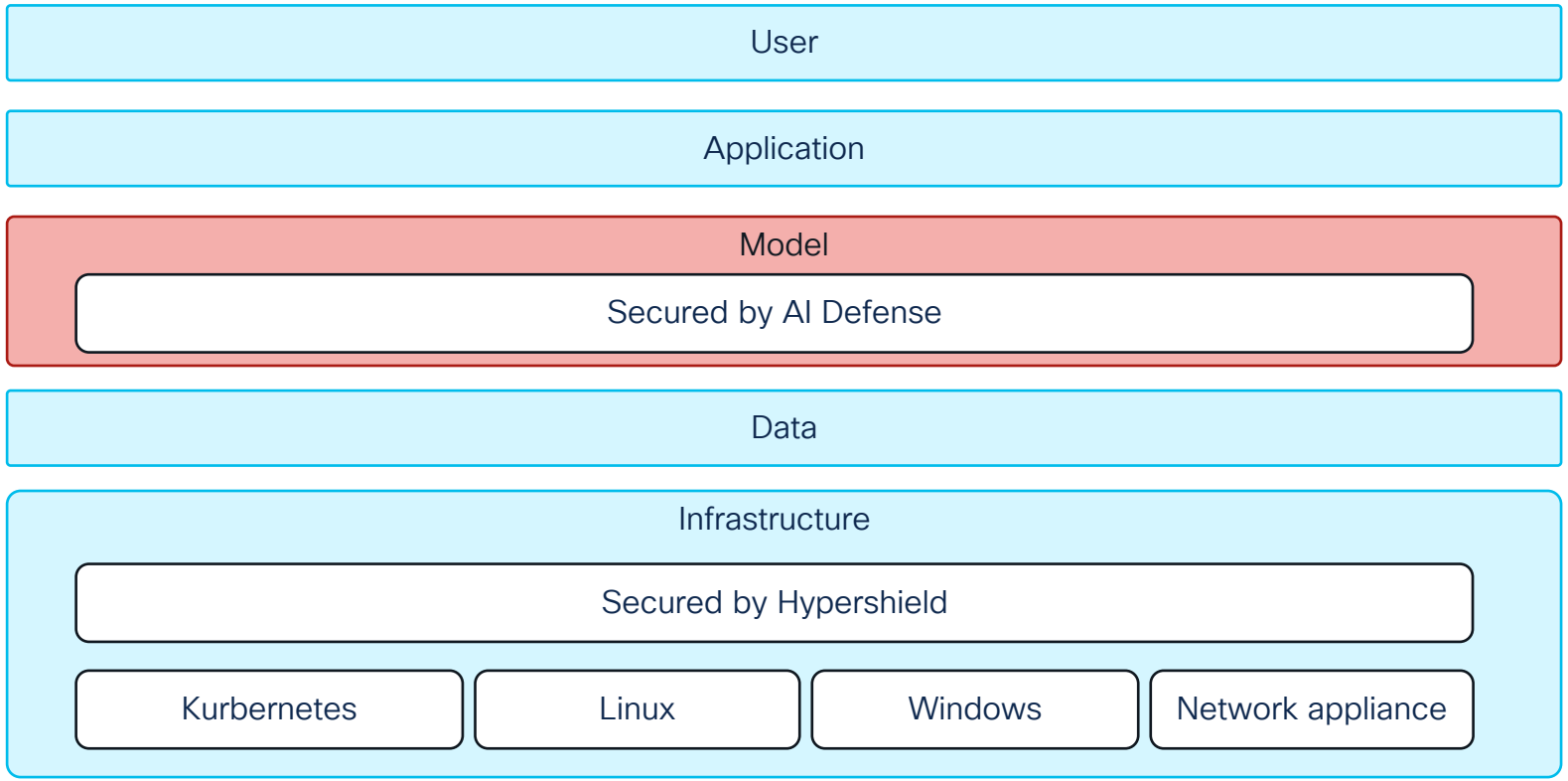
- Uncover supply chain risk in open-source models by scanning file components for malicious code, poisoned training data, and more
- Find vulnerabilities in models and applications through automated, algorithmic AI redteaming
- Create model-specific guardrails to “patch” weaknesses and better protect runtime apps

Securing the Runtime of AI Applications

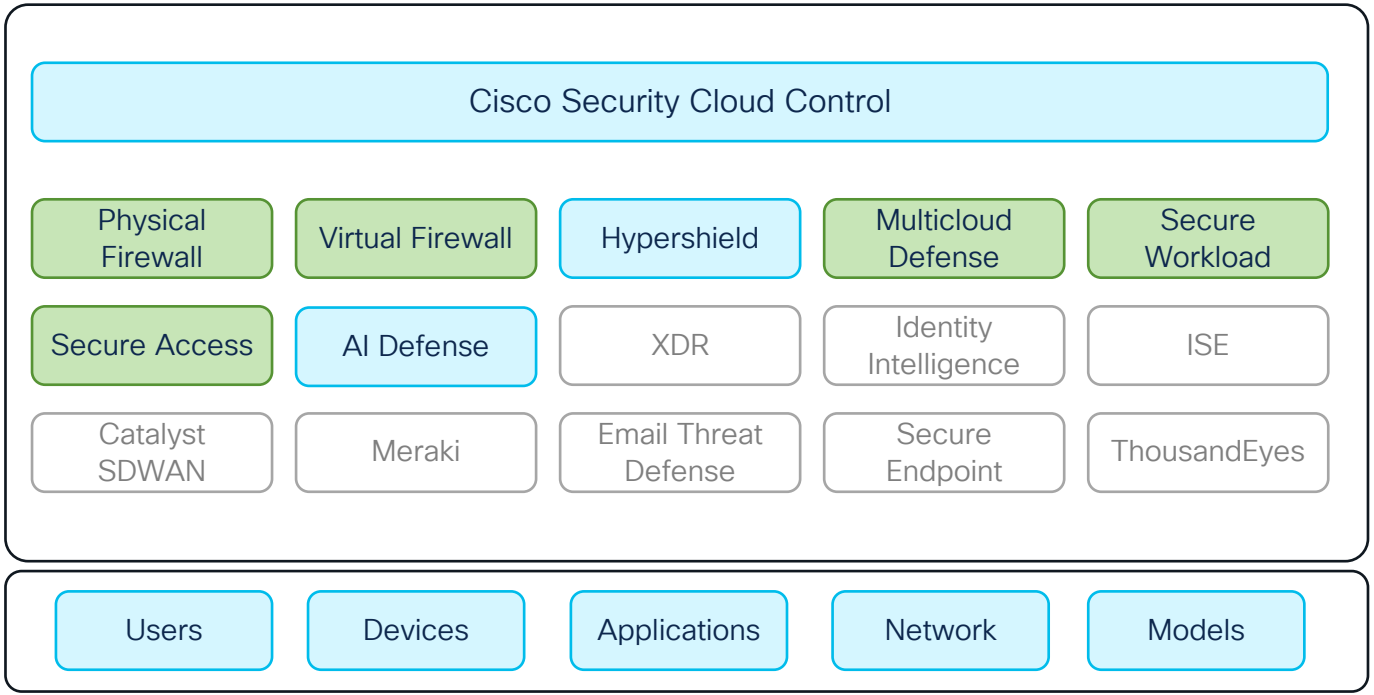
Protection

Place guardrails and access policies to secure data and defend against runtime threats.

- Apply guardrails that intercept and evaluate prompts and responses
- Block malicious prompts before they can do damage to your model
- Ensure model outputs are absent of sensitive information, hallucinations from company data, or otherwise harmful content
- Detections powered by proprietary AI models and training data



Centrally Managed. Resilient by Design



- Centralize control of solutions and policies
- Experience faster set-up and provisioning
- Support hybrid and multicloud environments
- Leverage AI to strengthen protection and prevent downtime

Items in Grey are on Roadmap

Key Takeaways

Key Takeaways

The datacenter is dynamic; security must be distributed

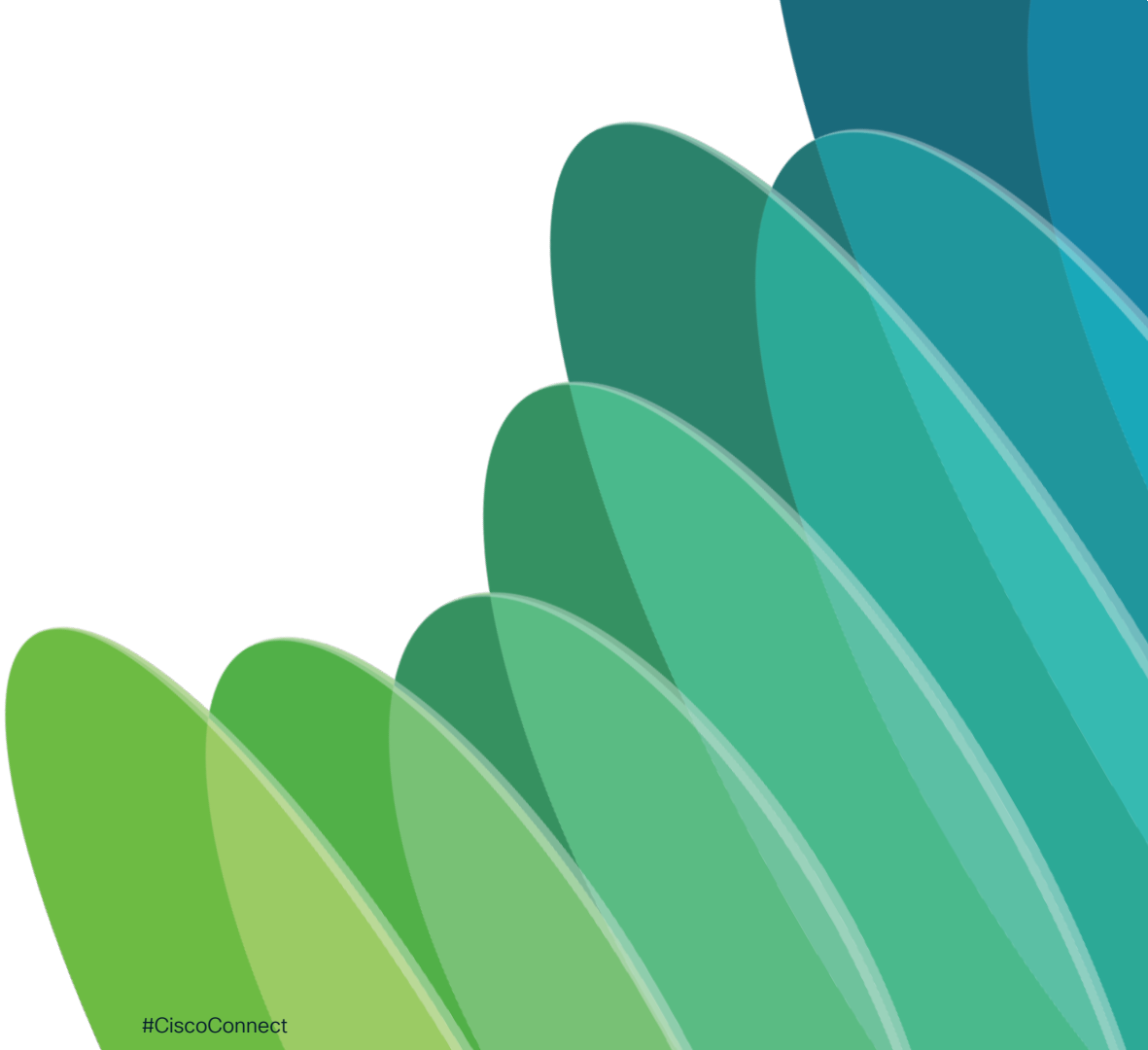
AI is reshaping both attack and defense

Build visibility, adaptability, and trust into your stack

Embed security everywhere, dynamically and at scale

Trust remains at the heart of your business and strategy

Cisco can meet you wherever you are on your Datacenter transformation journey



Thank you

cisco Connect

#CiscoConnect

cisco Connect

GO BEYOND

#CiscoConnect