Cisco Hypershield

AI Scale | Cloud Native | Highly Distributed

Security for the AI-scale data center

Artificial intelligence (AI) is transformative, driving huge productivity gains across businesses while driving an explosive growth inside data centers — the engine of AI. **Infrastructure** is evolving to advanced computational systems like graphic processing units (GPUs) and data processing units (DPUs), and the **applications** that run on the data center have also evolved from the traditional 3-tier architecture to using multiple microservices running on Virtual Machines (VMs) or containers.

To secure the AI-scale data center, we must reimagine security, as traditional appliances won't do the job.

Cisco Hypershield is the first truly distributed, AI-native security architecture with a clear vision to put security wherever it needs to be: in every software component of every application running on the network, on every server, and in public or private cloud deployments.

The solution is AI-powered to automate security policy lifecycle and security infrastructure upgrades. Imagine a network security solution that can write its own rules, test its own rules, deploy its own rules, and lifecycle manage its own rules. And it can even upgrade itself! At the same time, Hypershield is designed to empower customers to determine the level of autonomy they are comfortable with — using test, record, and report capabilities to earn trust.

Hypershield deeply combines security and networking in a way that only Cisco can, by taking the network security functions that used to come in a box and "melting" them into the network.

More a fabric than a fence, the solution allows security enforcement to be placed everywhere it needs to be. It lets you embed security in VMs or Kubernetes clusters in public clouds. In the private cloud, security can be inserted in VMs. Hypershield's unique and scalable architecture allows it to support various types of enforcement points. For instance, in the future, Cisco Hypershield will be deployable to high-performance servers DPUs and hardware accelerators running on networking devices such as switches, providing security beyond the data center to IoT/OT environments.

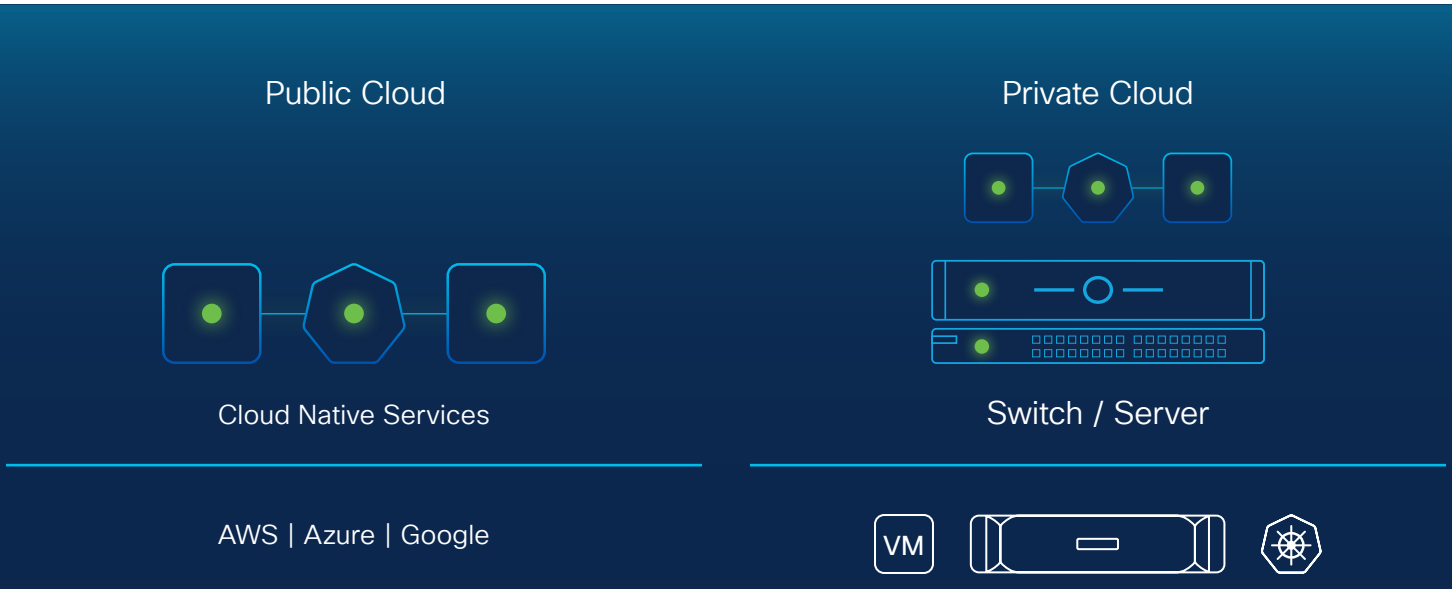

Figure 1. Hypershield puts security wherever you need it

Not the next generation of anything. The first generation of something new.

Cisco Hypershield reimagines security with a unique architecture built specifically for AI workloads.

Hypershield is a composable, **subscription-based solution** that sits on top of existing hardware. **Modules** built on top of the solution's core capabilities deliver specific security use cases – including segmentation and protecting against vulnerability exploits.

- **AI-native security:** By designing Cisco Hypershield from the ground up to leverage the power of AI, it's orders-of-magnitude more autonomous than other security solutions. In fact, since Hypershield was built from the beginning around AI management, we think of it as AI-native, as opposed to an AI layer bolted on top of a traditional product.
- **Kernel-level enforcement:** Hypershield provides deep visibility and enforcement actions at the workload level with its Tesseract Security Agent, built on top of Isovalent's (now part of Cisco) Tetragon and eBPF. eBPF provides a safe way to extend kernel capabilities without modifying the kernel itself or risking system stability. This allows Hypershield to gain deep visibility into workload behavior and implement fine-grained security controls while ensuring applications continue to run as Hypershield recommends, tests and deploys policy updates.
- **Self-qualifying updates:** Hypershield was designed to be self-upgrading and updating. Hypershield's recommended policies and its own software updates are tested in live production traffic, not in a lab or a simulation. The results of these tests are presented as a report that includes the deployment confidence score and the deployment effectiveness score. This helps security admins approve updates with confidence without the risk of disrupting the business, and once approved, the updates are deployed.

Figure 2. Cisco Hypershield architecture

Figure 3. Centralized management for distributed enforcement

Centralized security policy

Highly distributed IT environments require a new way to manage and enforce security policies across multiple domains. The scale of policies within an enterprise is staggering, with hundreds of thousands of rules that go in but never come out, written by folks who might no longer be within the organization. This level of complexity has been difficult to manage – until now. With AI-powered capabilities, Cisco helps security admins implement strong, consistent and dynamic policies at scale.

Hypershield is AI-powered and uniquely architected to implement a truly intent-based policy model that is centralized and easy to manage. No matter the form factor or location of the enforcement point, the policy being enforced is organized at a central location by Hypershield's management console. When a new policy is created or an old one is updated, it is "compiled" and intelligently placed on the appropriate enforcement points. Security admins always have an overview of the deployed policies, no matter the degree of distribution in the enforcement points. Policies can follow workloads as they move, from on premises to the native public cloud.

Solving real customer challenges.

Cisco Hypershield can solve the critical problems organizations are facing today.

1. Autonomous Segmentation

Companies have been using segmentation for decades to specify which workloads and applications can access which parts of the network to prevent lateral movement in case of an attack. But segmentation can be hard and time consuming, taking up to 40 days or more to define rules for a single application. That's way too slow. Current segmentation tools lack a deep understanding of the application and try to baseline application behavior based on a period in time without taking into account app-specific events.

For example, imagine an application that orchestrates the delivery of sheet metal to where it's needed on the factory floor. Traditional security tools will observe the app for a certain period of time like 90 days and formulate segmentation policies based on the behavior observed in that period. But what if the factory runs out of sheet metal on day 91 and this event triggers communications from the app to the various systems to place a new order? This legitimate behavior will come across as random when it's not. That's why you need a continuous and intimate understanding of the application.

The AI-native Hypershield looks beyond the network flows that other solutions focus on to get trained. The full scope of observed behaviors is informed by what's happening across all the environments it's protecting; what threat intelligence teaches it about behaviors that should never happen; the latest attack vectors, techniques, and vulnerabilities; what the system has learned and observed based on best practices that model how the customer modifies recommended policies; and what the customer does when they step in when under attack.

The system begins with an understanding of the applications in the environment and how they communicate and work. Based on that, macro guardrails are created to capture governance requirements and keep the business safe. As Hypershield learns it tightens policies further. This is a continuous and dynamic process, so if the application changes or moves, the segmentation policies will relax, and as the system learns about the new behavior, it can tighten them back again.

The result is higher-confidence, data-backed recommendations based not on what might have happened in the past but on what is happening now. The recommendations are automatically tested against live traffic and presented to the user with results on effectiveness and any performance impact of the policies. Once the user is satisfied and accepts the recommendation, only then it is deployed. So, the system is autonomous, but it earns trust and is continuously learning.

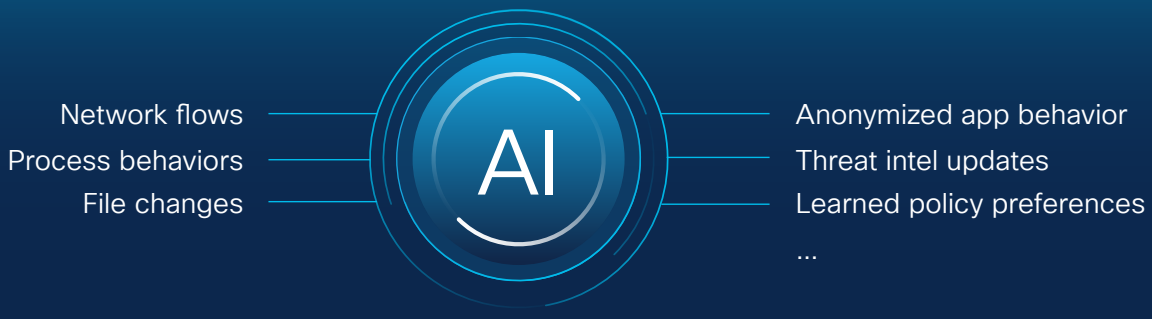

Figure 4 . Hypershield will use comprehensive set of inputs for segmentation policy creation

2. Distributed Exploit Protection

Attackers have gotten very good at either stealing a credential or compromising the services of an application so that they can move through legitimate application pathways. You need to understand the services that make up an application and its vulnerabilities. But even if the vulnerabilities are identified, patching them is hard; it takes time (weeks or months) and, in some cases, is not possible. Whereas attackers begin exploits within a few hours of a vulnerability being published.

Hypershield draws a picture of vulnerable assets across your entire environment. However, with around 500–1000 CVEs being published every week, it is difficult to prioritize and fix those vulnerabilities.

Hypershield's AI capabilities and deep understanding of the application will help prioritize the most critical vulnerabilities specific to the organization's environment. These prioritizations are based on three key questions:

- Is the vulnerable code module running in memory?
- Is the vulnerability theoretical, or is it being exploited in the wild?
- Is this vulnerability impacting a high-value asset?

While the application team takes the time to qualify the patch, Hypershield applies a surgical mitigating control or a mitigation shield in the path of the application to prevent an exploit. Once the patch is applied, the mitigating control gets removed automatically.

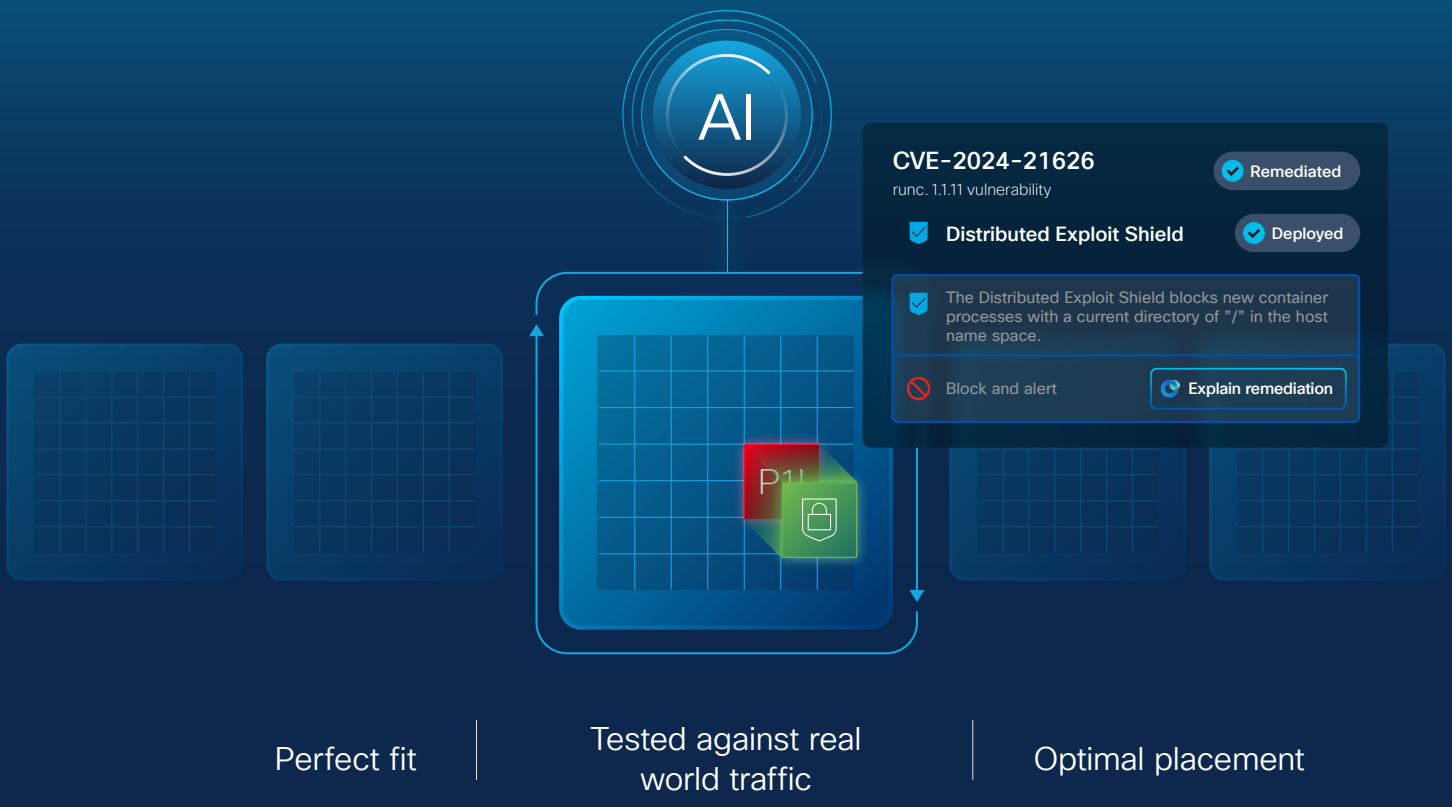

Figure 5 . Hypershield applies mitigation shields to protect against vulnerability exploits

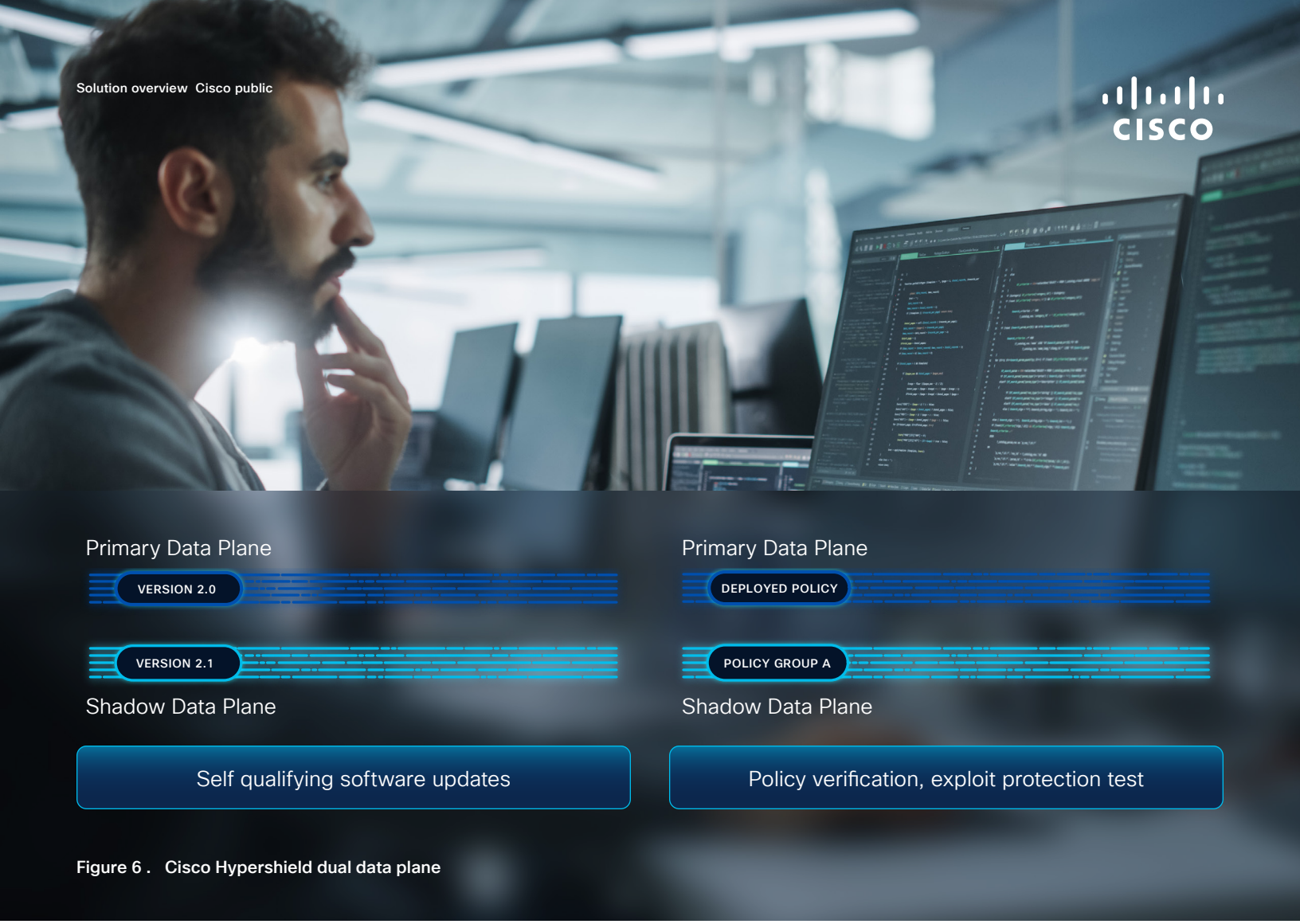

Figure 6 . Cisco Hypershield dual data plane

3. Self-qualifying updates

The nature of security controls is such that they tend to get outdated quickly. Sometimes, this happens because a new software update has been released. Other times, new applications and business processes force a change in security policy. Traditionally, neither scenario has been accommodated well by enforcement points – both acts can be disruptive to the IT infrastructure and present a business risk that few security admins want to undertake. A mechanism that makes software and policy updates seamless and non-disruptive is necessary.

As mentioned earlier, Hypershield is able to self-qualify updates. This reduces risk related to policy updates as well as its own software upgrade to ensure an up-to-date security posture and policy lifecycle management at scale.

The network-based enforcer performs self-qualifying updates using the **dual data plane.** It supports two data paths: a primary data plane and a shadow data plane. Live, real-world traffic is replicated between the primary and the shadow data plane – essentially a digital twin running in every enforcement point within the environment, not in a lab or a simulation. Software updates are first applied to the shadow data plane, and when fully vetted and accepted by the admin, the roles of the primary and shadow data planes are switched. Similarly, new security policies can be applied first to the shadow data plane, and when everything looks good, the shadow becomes the primary.

In addition to the dual data plane, there are similar (but different) methodologies on the end system enforcer (powered by Tesseract Security Agent) to self-qualify updates.

Conclusion

Cisco Hypershield fundamentally reimagines security for the AI-scale data center. The AI-powered capabilities, the deep visibility and enforcement right down to the kernel level and the self-qualifying updates, make Cisco Hypershield a powerful network security solution.

At the same time, customers can set the dial for autonomy within the AI too, increasing it as the system earns your trust with its ability to test, record and report everything.

Additional resources

- Solution web page
- Video: Unveiling a New Era of AI-native Security – Cisco Hypershield launch
- Blog: Cisco Hypershield: Security Reimagined - Hyper-Distributed Security for the AI-Scale Data Center by Jeetu Patel, EVP and GM, Cisco Security and Collaboration Business Group
- Blog: A New Era of Distributed, AI-Native Security by Tom Gillis, SVP Cisco Security Business Group
- Blog: Reimagining Security by Craig Connors, VP and CTO Cisco Security Business Group
- Blog: Our Vision to Combat Unknown Vulnerabilities by Craig Connors
- To stay up-to-date on product availability, demos and other news, please sign up here.