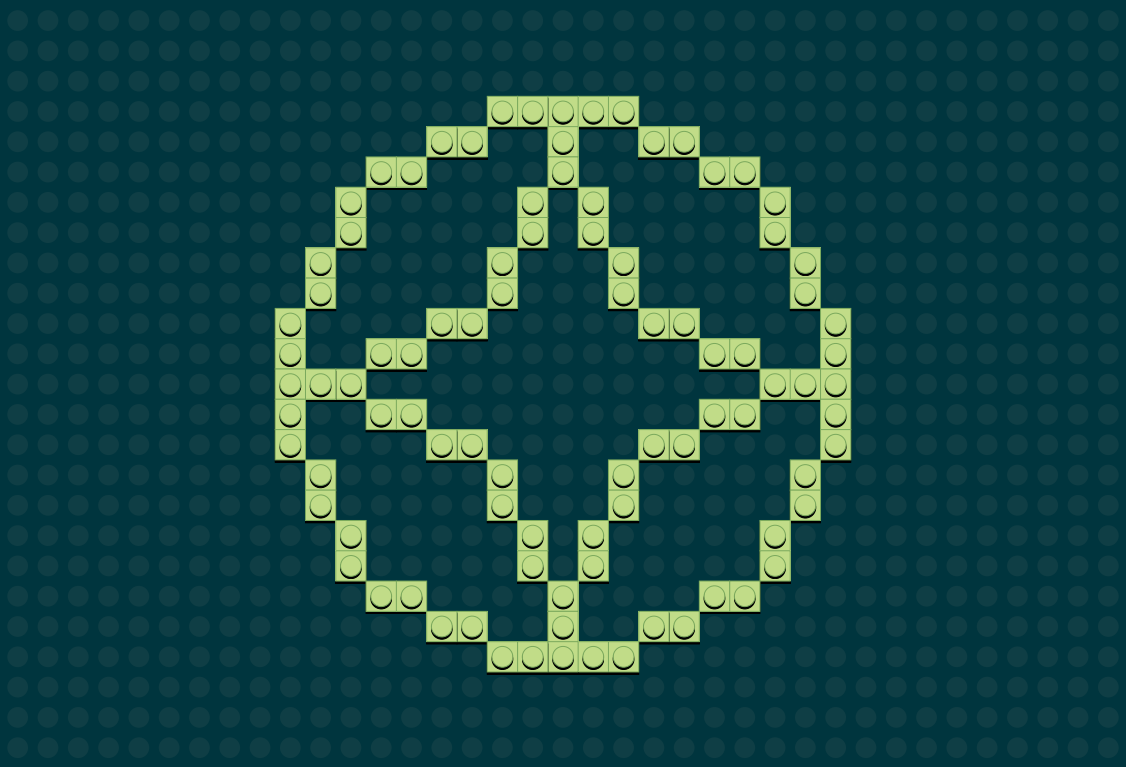
sysdig



Your Blueprint to

Securing AI Workloads, the Right Way

As businesses move workloads to cloud-native environments to support AI initiatives, they encounter a new frontier of security risk. For security managers building cloud security programs, it's crucial to take a step back and ask: Does your current program truly cover the unique risks AI workloads bring? You shouldn't have to choose between innovation and security, and with AI workload security done the right way, you won't have to.

Contents

03
Introduction

05
The changing security paradigm — from castles to carnivals

08
Secure AI workloads the right way: Comprehensive cloud security

16
Assess your security program for AI

19
Conclusion

Introduction: Navigating the new AI attack surface

Organizations worldwide are turning to artificial intelligence (AI) to gain insights, optimize operations, and remain competitive in an increasingly digitized economy. By 2026, it's projected that more than 80% of enterprises will have adopted generative AI, integrating AI-enabled applications and APIs into their operations.[1] Yet as businesses move workloads to cloud-native environments to support these AI initiatives, they encounter a new frontier of security risk. The dynamic, distributed, and ephemeral nature of cloud-native infrastructure combined with the complexity of AI workloads dramatically increases the attack surface.

AI workloads often interact with sensitive data, rely on APIs, and span multicloud and hybrid infrastructures. Each of these characteristics introduces potential vulnerabilities that adversaries can exploit. The result? Securing AI workloads calls for a fresh approach to security programs that reflects the complexities of cloud-native environments and the unique challenges AI introduces. While the principles of end-to-end security remain vital, they must now be complemented by frameworks tailored for these cloud and AI security use cases, such as those from MITRE ATT&CK and the Open Web Application Security Project (OWASP).

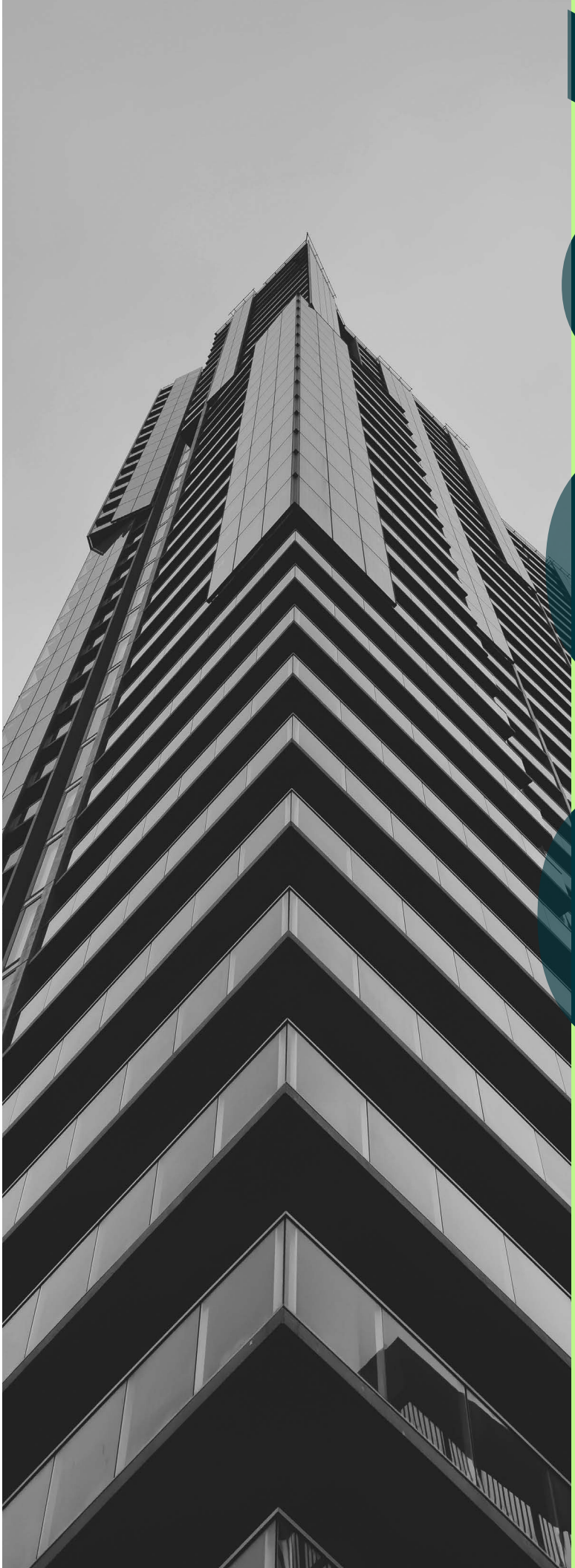
In the face of all of this, it can often seem as if teams have to choose between innovation and security. But AI workload security done the right way means not having to compromise on either. In fact, security the right way can be a strategic advantage, powering innovation and accelerating growth.

What does AI workload security, the right way, look like?

For security managers building cloud security programs, it's crucial to take a step back and ask if your current program truly covers the unique risks AI workloads bring. This means evaluating your practices against proven standards such as vulnerability management, runtime protection, and posture management. Beyond that, it's about equipping your team to perform deeper technical assessments to identify and address potential gaps that may expose AI workloads to risk.

This blueprint explores how AI is changing the cloud security conversation, and provides an assessment to help you evaluate your current security program to understand where it may fall short in addressing the unique needs of AI workloads. With these insights, you will be empowered to advance your security program, reduce risk, and enable your organization to securely leverage AI for innovation and growth — the right way.

1 Forbes, **Adopting AI: From Interest To Action**



80%

By 2026, it's projected that more than 80% of enterprises will have adopted generative AI.

The changing security paradigm — from castles to carnivals

First, it's important to understand how the cloud has changed the security game. In the early days of enterprise IT, organizations built their security strategies around the "castle" model. A central data center functioned as a fortified stronghold with one clearly defined entry and exit point, guarded by firewalls, access controls, and intrusion prevention systems (IPSs). Security efforts focused on defending the network perimeter and monitoring inbound and outbound traffic to prevent unauthorized access, under the assumption that everything inside the castle walls was trusted and relatively safe.

However, the migration to cloud-native environments has dismantled the castle walls. Today's infrastructure resembles a bustling and dynamic carnival. The complexity of modern workloads, such as containerized applications, microservices architectures, and distributed environments, has dramatically transformed the security scope.

Workloads in cloud-native environments are often distributed across multiple data centers, hybrid environments, or even multicloud architectures. The dynamic nature of these environments means that workloads can spin up or down in seconds, making it almost impossible to predict exactly where a particular workload resides at any given moment. In fact, according to Sysdig research, 70% of containers live for less than five minutes,[2] which creates a high level of churn.

This shift dismantles the traditional perimeter, replacing it with a distributed network of constantly evolving, interconnected systems. APIs play a crucial role in this environment, enabling communication between microservices and third-party systems. Each of these APIs is also a potential entry point for attackers, expanding the attack surface and introducing new risks. Sensitive data flows across hybrid and multicloud ecosystems, spanning across services including virtual private clouds (VPCs), Simple Storage Service (S3) buckets, Relational Database Service (RDS) instances, and even third-party systems, such as identity management services. Securing this vast, interconnected landscape requires a shift to a more nuanced, cloud-tailored approach that focuses on securing data, services, and workloads wherever they reside across the entire cloud ecosystem.

2 **Sysdig 2024 Cloud-Native Security and Usage Report**

AI workloads amplify security risks

The introduction of AI workloads has greatly expanded the attack surface, further amplifying the security challenges in cloud-native environments. AI systems operate within a highly interconnected and dynamic infrastructure that brings unique risks. These systems often require vast datasets, many of which contain sensitive or regulated information, such as personally identifiable information (PII), financial data, or proprietary business intelligence. As AI becomes more deeply integrated into business operations, it also becomes an increasingly high-value target for adversaries.

The complexity of AI models further complicates security. Many AI models are considered “black box” systems, where internal decision-making processes are difficult to understand and audit. This opacity creates significant challenges for security teams in detecting anomalous or malicious behavior, which puts organizations at greater risk of operational disruptions and security breaches that may go undetected until it’s too late.

In Kubernetes environments, where workloads are containerized and distributed across clusters, this complexity is multiplied. Each AI workload running within a container may have its own communication channels, privileges, and dependencies, all of which must be secured. While providing orchestration for these workloads, Kubernetes also introduces unique challenges in managing access controls, network policies, and workload isolation. A breach within one container could potentially lead to lateral movement, allowing attackers to compromise additional workloads within the same cluster.

AI systems also rely heavily on APIs, external data sources, and shared infrastructures, further expanding the attack surface. For instance, an AI model that consumes data from external sources or interacts with other cloud services via APIs is only as secure as the weakest link in that chain. If an API or third-party service is compromised, attackers can potentially gain access to the entire AI ecosystem, including sensitive data and intellectual property.

This interconnectedness means that protecting AI workloads requires more than just securing the models themselves. It involves securing the entire ecosystem of services, data flows, and communication channels.



20%

of businesses have suffered an attack on their AI models in the past 12 months.[3]

3 Forbes, **AI Models Under Attack: Protecting Your Business From AI Cyberthreats**

How to

Secure AI workloads the right way: Comprehensive cloud security

To effectively secure this chaotic new breed of cloud environments, businesses need comprehensive cloud security capabilities that cover everything from data and applications to infrastructure and user access. This approach goes beyond simple defenses, ensuring that the entire cloud ecosystem is protected against evolving cyberthreats. By combining strong controls with continuous monitoring, comprehensive cloud security upholds data confidentiality, integrity, and availability, minimizing the risk of breaches and unauthorized access.

This strategy should span the entire cloud stack, from posture management to real-time runtime protection. This is where cloud-native application protection platform (CNAPP) solutions play an important role by covering the span of prevention to defense capabilities. By adopting comprehensive security across the entire cloud ecosystem, businesses can better defend against sophisticated attacks, adapt to evolving risks, and ensure that AI-driven innovations remain secure.

Truly comprehensive AI workload security should cover three critical pillars: visibility, prevention, and detection and response.



> Truly comprehensive AI workload security should cover three critical pillars: visibility, prevention, and detection and response.

Pillar

01

Visibility into all AI in your environment

AI has quickly become a cornerstone of modern business, and ignoring its role in your security strategy risks leaving security gaps unaddressed. As security leaders, the question you need to ask is: How can you support the organization's goal to leverage AI while ensuring an acceptable level of risk?

To ensure that your security program is aligned with modern needs, start by evaluating your AI usage. This requires asking the following questions:

Where are models being used?

To effectively protect AI models, you must first know where they exist. This requires building a comprehensive AI inventory. You can't protect what you can't see, so identifying every instance where an AI model is deployed is critical to understanding its potential exposure and risks. Whether models are used internally or externally significantly influences their security posture and how best to monitor them.

What data do these models access, and how is that data protected?

AI models rely on data, and knowing what data they access is fundamental to protecting sensitive information. You'll need to ensure that all data, especially proprietary or personally identifiable data, is encrypted and protected with the right access controls to avoid unauthorized exposure.

Are they internal or external?

Understanding whether AI models are deployed within internal systems or exposed to external environments is essential. Publicly accessible models increase the attack surface, as malicious actors can target them. In contrast, internal-only models may require a different level of access control and monitoring to ensure that only authorized users have the necessary permissions to interact with them.

What permissions are granted to the models, and who is consuming them?

Tracking who has permission to use AI models and what level of access they have is essential. It's easy to inadvertently grant broad permissions with the cloud, so it's crucial to restrict access to only those who truly need it, and then regularly audit these permissions to prevent privilege creep.

Answering these questions will provide clarity on your organization's current AI usage and security posture. Once you have a full understanding of where your AI models are operating, what data they access, and who has permission to use them, you can assess whether your current security program is sufficient or if it needs to be updated.

Pillar

02

Prevention of threats with a hardened security posture

Once you have visibility into your AI environment, the next step is to shore up weaknesses in your security posture before they become active threats. AI workloads process vast amounts of sensitive data, depend on fast-changing open source tools, and often interact with other systems through exposed APIs and machine identities. These dynamics make prevention a moving target. Securing AI effectively requires a layered defense that enforces strict access controls, reduces software vulnerabilities, and continuously maintains a secure configuration baseline across your cloud footprint.

Identity and access management is the foundation of preventative security in AI environments. AI systems introduce a broader web of machine-to-machine interactions between models, data pipelines, APIs, and orchestrators like Kubernetes. Applying the principle of least privilege across these relationships is critical to reducing risk, ensuring that users, models, and services have only the access they need and nothing more. Enforcing granular permissions, tightening identity governance, and regularly rotating credentials are all critical to reduce opportunities for privilege escalation or unauthorized model interaction.

Vulnerability management reinforces the integrity of the software and infrastructure that power AI. Continuous scanning of container images, open source libraries, and model-serving dependencies uncovers weaknesses before they can be exploited. Because AI pipelines evolve rapidly with frequent retraining, updates, and integration of open source components, prioritization of risk based on business context and automated remediation workflows are vital. Addressing the vulnerabilities that could expose sensitive data, corrupt models, or disrupt AI services keeps the AI stack resilient and trustworthy.

Complementing this, **cloud security posture management (CSPM)** provides the oversight needed to maintain a secure configuration baseline across distributed AI environments. By continuously monitoring for misconfigurations and other posture risks, CSPM helps teams identify and remediate systemic weaknesses before they become entry points for attacks.

As organizations mature in their AI adoption, some are beginning to explore AI security posture management (AI-SPM) as the next step of their cloud security strategy. Building on the principles of CSPM, AI-SPM extends visibility and governance to AI models, data pipelines, and services. Though still emerging, AI-SPM underscores the growing need for organizations to identify and manage risks that are unique to AI workloads.

While identity and access management, vulnerability management, and posture management are foundational to any modern security program, they take on new urgency and complexity in AI environments. AI workloads operate at a faster cadence, rely heavily on open source frameworks, and continuously exchange data. Misconfigurations or unpatched vulnerabilities can not only expose sensitive information but also alter model behavior or corrupt training data. Strengthening these core preventative controls with an AI-aware lens ensures that traditional defenses evolve to protect the dynamic, data-driven nature of AI workloads.

Pillar

03

A real-time detection and response program

In the cloud, businesses are navigating an ecosystem that is fast-paced, distributed, and inherently interconnected. With the modern cloud operating like a carnival, things are constantly changing: Workloads spin up and down in seconds, APIs multiply, and sensitive data flows freely across hybrid and multicloud environments. When adding AI workloads to this equation, bringing their own reliance on sensitive data, APIs, and external integrations, the complexity and risks multiply.

Even organizations that have updated their security practices for the cloud must reevaluate their strategies to address the unique challenges posed by AI. At the core of this shift is adopting a security strategy that operates under the assumption of a breach. Why? Because in the cloud, malicious access isn't a distant possibility; it's an inevitability.

Why assuming a breach should be the default

In today's accelerated cloud environments, everything moves at rapid speed. A misconfigured API can be exploited within seconds, allowing attackers to move laterally or escalate privileges across systems in minutes. This pace makes the traditional "keep the attackers out" approach obsolete.

Cloud security isn't just about stopping threats before they happen; it's also about understanding that attacks that get past prevention controls are inevitable. By working from this understanding, organizations can ensure that they're ready to detect, contain, and mitigate threats before they escalate into significant incidents.

Of course, preventative controls are foundational, like your helmet when you're skiing down the cybersecurity slopes. You wouldn't hit the slopes without it, just like you wouldn't deploy an application without making sure that basic security hygiene is in place. But just like a helmet won't stop you from falling, preventative controls won't stop every threat. You have to assume that at some point, an attack will get past prevention controls. And that's where detection capabilities step in to ensure that you can detect an attack the moment it happens.

Preventative controls provide a sense of safety, but they're inherently backward-looking. They focus on stopping what could happen, not what is happening right now. And with cloud attacks happening in seconds, the only way to stay ahead and know where to focus security investigations is with comprehensive cloud security that includes real-time threat detection capabilities.

Uncover threats in real time with runtime protection

Effective detection starts with runtime visibility. Continuous monitoring and analysis of workload behavior, model execution, and API interactions allows teams to spot anomalous behavior that signals compromise. This could be unexpected network connections, unusual data access patterns, or unauthorized actions.

Equally important are threat correlation and investigation. By unifying telemetry from across containers, cloud services, and AI systems, security teams can connect the dots between seemingly isolated events to reveal coordinated attacks. Correlating indicators of compromise across multiple layers, from model runtime to underlying infrastructure, accelerates investigation and response time.

When threats are identified, automated response and containment capabilities become critical. The ability to isolate a compromised container, revoke credentials, or block a malicious API call within seconds can mean the difference between a contained incident and a full-scale breach. Real-time detections must seamlessly feed into workflows for security operations centers (SOCs) and incident response teams to support rapid investigation, forensics, and remediation.

Ultimately, the goal of this pillar is resilience. By assuming that some threats will always slip past preventative controls, organizations can design their programs to detect faster, respond smarter, and recover sooner, ensuring AI innovation continues securely even in the face of evolving attacks.

Assess your security program for AI

To determine whether your organization's security program is equipped to handle the unique risks posed by AI, start with this self-assessment. These questions are designed to help you evaluate your organization's current AI security readiness and identify areas that may require immediate attention.

For each question, answer **yes** or **no** based on your organization's practices.

01

Can you identify all AI models in your environment?

02

Are permissions for accessing your AI applications strictly limited to what's necessary?

03

Do you know which data sources your AI models can access? Are they protected?

04

Are you actively monitoring for runtime anomalies?

05

Do you know if your AI workloads are publicly exposed?

06

Does your cloud service provider have AI monitoring enabled by default?

07

Can you validate the security processes for pre-trained and third-party models prior to implementation?

08

Is access to underlying AI systems and training data restricted?

09

Are associated API keys securely managed and regularly rotated?

10

Do you have an incident response plan specifically for AI-related incidents?

Scoring

For each “no” answer, assign 10 points. For each “yes” answer, assign 0 points.

Now tally your points to determine your organization’s current level of AI security maturity. The more “yes” answers you have, the more aligned your security posture is with the demands of AI.

0-10

Advanced program

Your AI security program is mature.

11-20

Moderate maturity

Your program is on the right track, but you probably should update your security strategy to stay ahead of AI security risks.

30+

High risk

You should take immediate action to implement foundational AI security practices.

By completing this self-assessment, you’ll gain insights into where your organization stands and where improvements are necessary. If your results indicate a need for stronger AI security practices, our guide on adopting an AI security framework can help you get started on the path to protecting your AI-driven future.

Conclusion

AI workloads are driving innovation, but they also introduce growing risks. In cloud-native environments, where the pace of change is constant, securing these workloads requires a paradigm shift.

To ensure your security program aligns with modern needs, start by evaluating your AI usage. Ask critical questions about your AI footprint: where models are deployed, what data they access, and what permissions they have. This self-assessment will help you identify any gaps and determine if your current security strategy needs to evolve to address the demands of AI.

If your self-assessment reveals the need for an AI security program, it's time to take action. By combining strong preventative controls with real-time detective capabilities, organizations can build security programs that support strategic goals while ensuring that AI drives business outcomes — without introducing unacceptable levels of risk.

The stakes are high, but the rewards are greater. For security leaders, it should never be a question of whether to choose speed or safety. By securing innovation and AI in the cloud the right way, organizations can move forward without compromise.

See how Sysdig secures AI workloads, the right way.

REQUEST A DEMO →

Sysdig helps security and development teams prevent, detect, and respond to threats instantly. Founded by the creators of Falco and Wireshark, and built on agentic AI, Sysdig delivers real-time defense grounded in the uncompromising truth of runtime. No guesswork. No black boxes. Just cloud security, the right way.

ABOUT THE AUTHOR



Matt Kim,
Senior Product Marketing Manager at Sysdig

Matt is a product marketer with a focus on cybersecurity, particularly in container security and vulnerability management. With over a decade of experience, he brings keen insight into translating complex security topics into straightforward, actionable insights that empower teams to tackle today's toughest security challenges.

sysdig

BLUEPRINT

COPYRIGHT © 2025 SYSDIG, INC.
ALL RIGHTS RESERVED.

WP-014 REV. B 11/25