sysdig

WHITE PAPER

In Cloud Security, Architecture Matters

Comprehensive, real-time cloud security with agentless and agent protection

The effectiveness of cloud security platforms largely depends on the ability to tackle the complexities and diversity of modern cloud environments. An integrated strategy is required that employs a combination of agent and agentless solutions. Optimized designs thoughtfully consider aspects such as performance, scalability, and adaptability for varying workloads and deployment scales. In addition, the ability to seamlessly integrate into existing technology stacks to provide critical insights and actionable outcomes is imperative.

This paper explores the necessity for an advanced, comprehensive solution, adept at merging different data sources and enriching collected data to produce valuable insights in real time.

Table of Contents

03
Spoiler Alert: Cloud Security Requires Both Agentless and Agent

07
Not All Agents are Created Equal

09
Let's Get Techy: A Kernel of Knowledge About Agents

14
Cloud Security Demands Flexible, Scalable Insight

Spoiler Alert: Cloud Security Requires Both Agentless and Agent

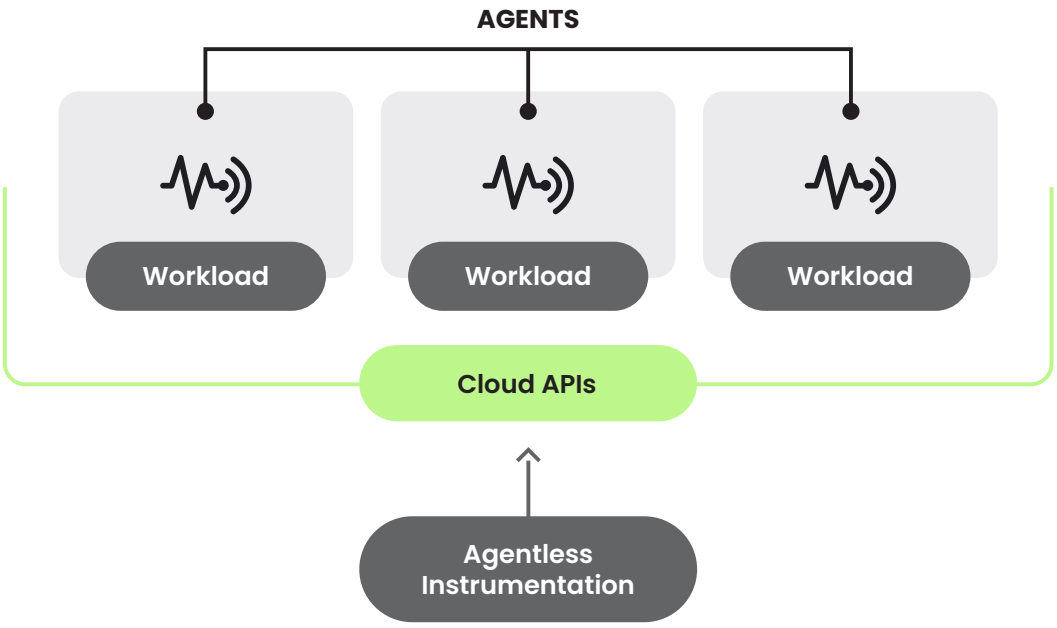
Developing and effectively deploying software in the cloud is a complex challenge, encompassing a multitude of dimensions from managing expensive cloud resources, addressing evolving needs and technologies, and securing them in a perimeter that is often either not clearly defined or doesn't exist at all. Such a complex landscape demands a calculated and strategic approach in securing the cloud, embracing both a proactive "shift left" methodology, to harden and protect in the development process, and to detect and respond tactics to safeguard against potential threats in the operational stage. Merging these disparate philosophies, and ensuring they co-exist with new regulatory requirements focused on cyber-resiliency, requires a shift in an organizational approach to cloud security.

When beginning a cloud security journey, security teams sometimes start with an agentless approach due to ease of deployment. Agentless security solutions leverage cloud logs, APIs, and volume snapshotting to provide runtime visibility by monitoring cloud security controls, detecting configuration changes, identifying misconfigurations, and preventing drift across multiple cloud accounts. They only require an account configured with the proper Identity and Access Management (IAM) roles that are easy to create with the click of a button, simplifying the initial setup process and allowing for a quicker implementation of basic security measures.

Developing and effectively deploying software in the cloud is a **complex challenge**, encompassing a multitude of dimensions

As organizations progress further in cloud security, the desire for stronger controls and workload visibility dictates the need for augmenting agentless solutions with an agent-based approach. Agentless solutions offer a simple and quick approach to achieve basic posture and vulnerability management, and they may also provide runtime visibility. However, compared to agent-based solutions, they lack granular visibility into system-level activities and the context of what is happening in real time. Agentless only solutions are reliant on a wide variety of cloud mechanisms, such as periodic scanning to report changes on a schedule, typically 3-6 hours apart, cloud log parsing, or passing cloud logs to a SIEM for parsing, analysis, and detection.

Deploying software agents on cloud compute workloads allows organizations to gain more comprehensive insights into processes, user, file activity, network connections, and other system-specific details. This enables more effective cloud threat detection and response capabilities, including advanced techniques such as behavioral analysis and machine learning algorithms.



By incorporating both agentless and agent-based approaches, Sysdig partners with organizations to ease their initial deployment and provide deep visibility and stronger security as customers progress in their cloud security journey. This flexible and comprehensive approach allows for greater protection and enables organizations to effectively adapt to evolving security challenges as they mature their cloud presence. This is also confirmed by the market trends; in the "**Market Guide for Cloud-Native Application Protection Platforms**" published earlier this year, Gartner highlights how the most advanced CNAPP solutions used agentless inspection to augment the data gathered by agents.

Agentless Should Be Real Time, Not Rearview Mirror

The powerful management APIs and snapshotting services exposed by cloud service providers enable an agentless approach to cloud security. Leveraging these services and the additional context they provide makes it possible to gather meaningful, high-level information on the deployed resources and their configuration. This can satisfy use cases like vulnerability assessments, asset discovery, or "static" cloud security posture management, but does not address active cloud risk. Active cloud risks are real-time changes or events that may pose a threat and are captured with cloud event logs that allow security teams to gather insights about services where agent-based telemetry is not available, making it possible to implement active risk and threat detection for an ecosystem of third-party services.

Agentless solutions are generally deployed in three different ways: snapshots, APIs, and logs/events.

Snapshots enable point-in-time images to be shipped to remote storage and analyzed without interfering with the workload execution. This enables users to scan the disk contents for vulnerable packages, malware traces, secrets, or other indicators of compromise (IoCs). Once the security tool has been granted the appropriate permissions to access the cloud resources, security scanning can occur without affecting workload performance or requiring additional maintenance. Scanning snapshots enables lower overhead data discovery and categorization looking for sensitive or regulated data, such as PII or HIPAA. Snapshot-based scanning, however, does have important drawbacks:

- The scans are limited to disk contents, so there's no runtime context such as open network connections, running processes, etc.
- They take a significant amount of time and are only feasible with reduced frequency (typically every 12 or 24 hours), so they are not suitable for threat detection use cases where the detection time needs to be as low as possible.
- They require resources on which to run the scanner software. These resources could be hosted either in the customer organization, increasing its own cloud spending, or in the security tool provider cloud organization, requiring shipment of these snapshots to a trusted third-party and increasing spending for the user.

API-based solutions take advantage of the APIs exposed by cloud providers to collect, correlate, and consolidate data on the cloud infrastructure. By pulling metadata on the cloud resources and their configurations, it is possible to reliably map the infrastructure and identify potential vulnerabilities, such as outdated software, exposed endpoints, and overly permissive roles. This type of detection is subject to the limits imposed by the cloud provider APIs, both in terms of visibility and call or bandwidth quotas.

Log-based detection leverages the various audit logs exposed by cloud providers, hosted operating systems, and applications to identify potentially significant security events. Sysdig enriches static risk findings and overlays active risk information for prioritization, investigation, and remediation. The riskiest combinations of static and active risks are surfaced to the top with attack path visualization to speed investigation. Guided remediation is integrated in the workflow to help security teams fix issues fast when every second counts. Moreover, logs might be the only event source available for third-party systems, such as external IAM providers.

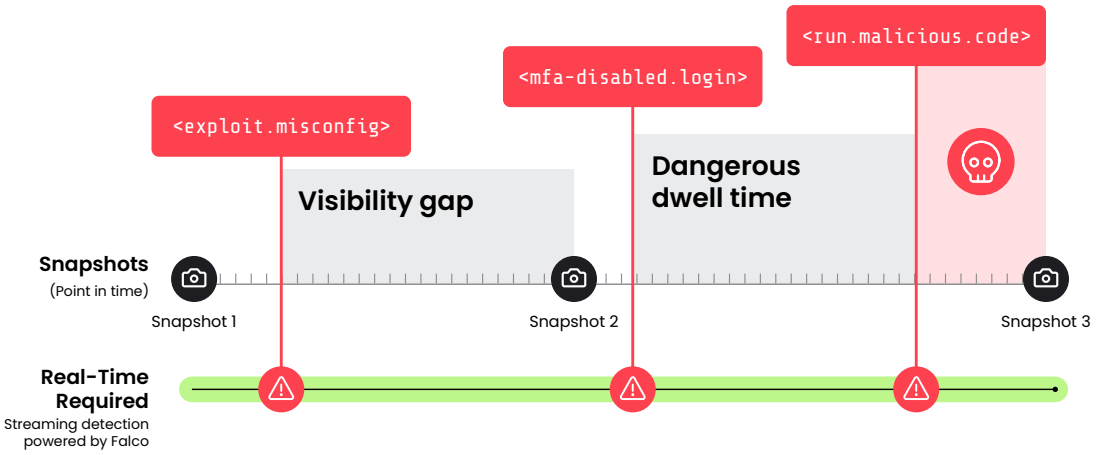
Sysdig has developed a unique agentless implementation that can analyze and process event sources in real-time, enhancing the security of resources beyond compute workloads, and identify threats as early as possible on the attack path. Based on Falco, this unique streaming processing allows for real-time detections across disparate cloud and on-premises environments versus waiting potentially hours for periodic scan-based approaches or post-processed log analysis to detect and identify active malicious threats.

Cloud Risks are Active.
Why is Your CSPM Static?

Point in time assessments leaves visibility gaps of hours or more

Static checks miss active movement and dynamic changes

Static checks miss runtime context and active imminent risks



Streaming detection detects malicious activity in real time.

Not All Agents are Created Equal

Performance, scalability, accuracy, and maintainability considerations

Performance is a primary consideration when discussing agent based solutions. An agent needs to be tested in real-world scenarios, under various system load situations, and protecting critical workloads rather than competing with them for resources. While what constitutes “acceptable performance” is by nature highly subjective, any agent-based solution should never hinder the proper execution of the traced processes.

Scalability in complex cloud environments requires a solution designed to scale from a few processes to thousands, in a single machine or in a cluster with hundreds or thousands of nodes, all without posing unreasonable limits on the detection rules or performance and without sacrificing efficacy. That is why scalability reinforces the importance of a streaming approach — events need to be evaluated and consumed as soon as possible. Batch or snapshot approaches inevitably result in missing time-sensitive events, especially in large scale deployments.

Any solution charged with safeguarding such a dynamic, ethereal, complex, and varied “infrastructure” as the cloud needs to be able to accurately reduce the noise of false positives and efficiently highlight the highest risk alerts. As use of cloud infrastructure continues to grow, implementing a single security posture across the organization becomes more complex, and creates potential gaps that can be exploited by malicious actors. Simplifying the security architecture and leveraging tighter, more seamless integrations between technology partners creates the foundation for a more consistent security posture.

Solutions that require complex redeployments when updating detection rules, for example, will inevitably end up outdated for most of their lifecycle, severely limiting the security they are supposed to provide. Maintaining, updating, and troubleshooting in complex environments needs to provide granular, relevant information, and also be efficient. Time spent maintaining security software opens a window to a potential breach, making maintenance an often overlooked but important aspect to security. Security solutions to protect cloud architecture must adapt as quickly as the programs they protect, otherwise security teams will forever be playing catch up.

The cloud landscape: additional challenges and needs

The cloud introduces new challenges and needs for protecting critical resources, both in terms of the scale and flexibility required to adapt to the different cloud resources and the limited visibility they provide. These are some of the different use cases that need to be addressed by modern CNAPP solutions:

- CaaS (Container as a Service) solutions, such as AWS Fargate, often referred to as “serverless”: These services completely abstract container execution from the underlying hosts, and thus require novel solutions to maintain kernel-level visibility on the workloads.
- Cloud provider logs: Every major cloud provider offers access to audit logs and event sources, granting some visibility within the cloud environment, including user authentication and authorization, resource provisioning and configuration changes, network traffic, and more.
- IAM services: Identity services are the cloud perimeter, and thus it is crucial to properly monitor activity in IAM services and intercept anomalous activity, coming both from real users and machine accounts.
- IaC security: Scanning IaC manifests to identify misconfigurations and security risks before deployment while preventing drift.
- Vulnerability management / Supply chain security: Identifying, prioritizing, and fixing vulnerabilities across your software supply chain (SCM, CI/CD, registry, and runtime environments).
- Configuration and access management: Hardening posture by managing misconfigurations and excessive permissions across cloud environments (cloud resources, users, and even ephemeral services like Lambda).
- Threat detection and response across cloud workloads, users, and services: Multi-layered detection approach that combines rules and ML-based policies, enhanced with threat intelligence, along with a detailed audit trail for forensics/IR.
- Compliance: Meeting compliance standards for dynamic cloud/container environments against PCI, NIST, HIPAA, etc.
- Third-party services: Complex cloud environments often rely on third-party tools and services, such as secret management services, multi-cloud orchestrators, etc. A comprehensive cloud security approach must integrate and monitor these tools to provide a holistic view of the security posture.

The final and most critical challenge is integrating and consolidating all these events and detections, coming from diverse and heterogeneous resources. This is where most solutions fall short, since they are often the result of separate designs tacked together from smaller, niche solutions built with a singular, limited focus.

Sysdig has taken a comprehensive end-to-end approach to protecting the cloud. We built a single, unified platform, powered by Falco, that uses multi-tier enhancement to correlate insights coming from agent and agentless instrumentation across cloud services, workloads, identities, and third-party tools. Our agent is optimized to deliver industry-leading performance and is proven at enterprise scale. These adaptive optimizations make it possible to achieve performance levels similar to kernel-level solutions and maintain excellent visibility into the workloads.

Let's Get Techy: A Kernel of Knowledge About Agents

Agent instrumentation

Inspecting activity on the host must be as unobtrusive and lightweight as possible in order to not compromise the workloads, while at the same time provide the highest degree of visibility on the system.

This visibility requires a privileged point of view. The kernel is the core of the operating system; it manages hardware resources and provides basic services to applications through system calls (abbreviated as syscalls). Syscalls are effectively the ultimate source of truth about system activity, and monitoring them is at the basis of modern security solutions.

There are three primary techniques to inspect processes to collect syscall events, which differ in terms of accuracy and overhead.

Too hot

One approach is using ptrace and other similar system level introspection techniques to pause and inspect a process, as done by most debugging tools. These techniques have great accuracy, since they are based on functionalities exported by the kernel itself. However, being based on user space APIs, they require multiple context switches to collect the event data, and thus incur a significant performance penalty. Reducing the amount of data collected using these techniques improves performance but reduces accuracy. There is also the risk associated with how tools such as ptrace modify the active memory of running applications and workloads, which can lead to instability.

Too cold

Another technique is leveraging dynamic linking of libraries, replacing a system library (through LD_PRELOAD) with an instrumented version to trace syscalls for a process. This is relatively efficient but has low accuracy, since it will only work for programs using dynamically-linked libraries and can be easily circumvented. It also doesn't support statically-linked binaries, such as those compiled by languages like Go. While this approach has merits, agents using LD_PRELOAD tend to introduce instability into the systems they are monitoring.

Just right ... when done right!

Kernel-level inspection stays within the kernel context to collect data. This approach provides:

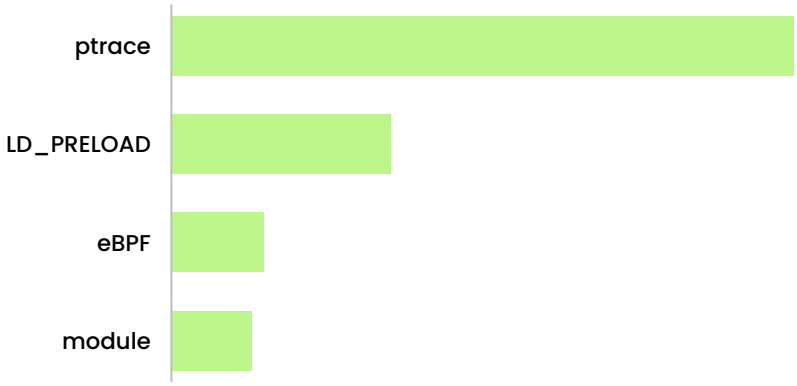
1. Enhanced visibility: Kernel-level instrumentation provides the most comprehensive visibility into system operations. It enables monitoring and captures a wide range of events, including process creation, file system activities, network traffic, and more. This level of visibility allows for effective threat detection, performance analysis, and system monitoring.

2. Lowest overhead: Kernel-level instrumentation incurs lower overhead compared to user-space techniques. By operating at a lower level of the system, it can intercept and process events more efficiently, reducing performance impacts on the overall system.
3. Greater capabilities: Kernel-based agents can enforce access control policies, detect and prevent malicious behavior, and protect against kernel-level exploits and vulnerabilities. By residing in the kernel, the agents have more privileged access and can actively monitor and respond to security threats.

There are two main techniques at the kernel level: using kernel modules (the traditional way) or using eBPF (extended Berkeley Packet Filter) programs. Kernel modules are more invasive but offer the lowest overhead and the most versatility, including working on older Linux distributions. eBPF programs are safer than kernel modules, since they run in a virtual machine at kernel level and are verified to be safe before execution, but offer the same great accuracy and performance. They grant performance nearly as good as those of kernel modules, thanks to just-in-time (JIT) compilation; this efficiency enables real-time monitoring and analysis with minimal impact on system performance.

Thanks to its powerful and flexible programming framework, eBPF has gained significant popularity and has an active community. Its increasing support and constant evolution is driving the development of novel tools and libraries, especially in performance critical areas.

Kernel instrumentation performance overhead



Designing an agent to operate at the kernel-level, specifically leveraging eBPF, offers enhanced visibility, lower overhead, increased security, safety, performance efficiency, and programmability. These factors make it an excellent choice for building robust and efficient agents for monitoring, security, and performance analysis purposes.

Sysdig has its roots in Falco, the open source solution for runtime security. Falco uses state-of-the-art kernel-level instrumentation, leveraging both kernel modules and modern eBPF to give excellent visibility on syscall events with industry-leading performance and support for new and old kernels.

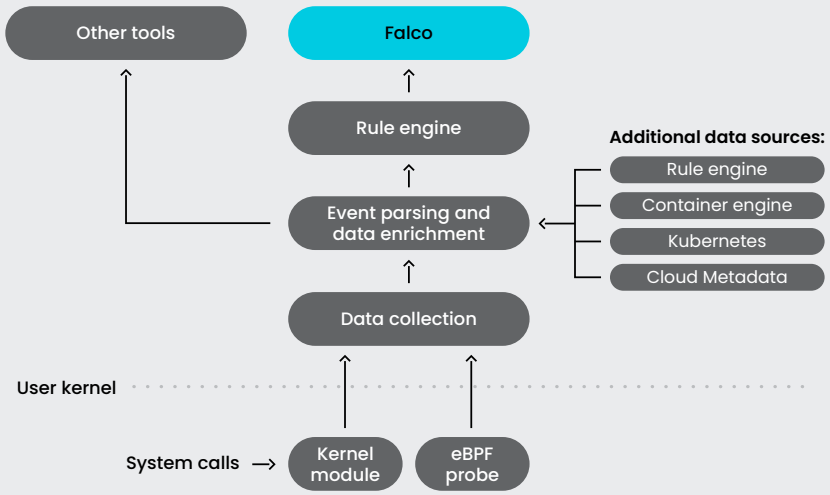
Data collection

Collecting all the data that security tools provide is a challenge, especially at cloud scale. Moreover, the data collected also needs to be enriched, categorized, and consolidated in a way that can be consumed by detection engines and humans.

Collecting and making data available to consume is also a time-sensitive operation, especially when threat detection is a primary use case. The scale of cloud deployments is an additional challenge in achieving these goals.

Security tools should be able to collect data both at kernel level and at user level. The granular low-level details gathered from the syscall events need to be enriched with details on the process, user, container, Kubernetes namespace, and all other available correlated metadata, and the enrichment needs to happen before the evaluation of the event by the detection engine. Moreover, the data collection needs to be comprehensive and powerful enough to support additional data sources. Being able to ingest and enrich data coming from third parties (e.g., plugins for novel services) future-proofs solutions, and ensures they are ready to address the challenges coming from technological evolutions.

Sysdig and Falco perform data collection and enrichment in an optimized way. The syscall events are augmented with relevant context, including running processes and threads, the files they have open, the containers, and the associated Kubernetes objects. All of this context is available to the rules engine and the output. This makes it possible to write meaningful and precise detectors and rules, and it helps incident responders utilize real-time insights to identify and prioritize high-risk events quickly and accurately. Since Falco is open source, this powerful capability can be extended by the community of users and contributors. Since the data is collected and evaluated on the same host of the agent, horizontal scalability can be easily achieved, even for larger cloud environments.



Multi-tier enrichment

Collecting data solely from system calls provides valuable insights into the behavior of individual processes and the underlying operating system. However, raw syscall data is useless without the proper context and information. What makes an event relevant for security is the context in which it happens.

To provide a comprehensive and contextual understanding of the cloud environment, it is crucial to enrich this data with additional metadata and context from various sources, such as the cloud infrastructure, Kubernetes, container runtimes, and more. It is crucial this enrichment happens as quickly and efficiently as possible, keeping the detection time and the agent footprint at a minimum.

These are some contexts where multi-domain correlation is essential:

1. Local enrichment: Adding contextual information from the system gives meaning to all the data collected through the syscalls. Mere ID numbers are translated into process names, users, file paths, connections, container names, etc., making it possible to write human readable detection rules and identify the meaningful security events.

It is crucial to enrich data with additional metadata and context from various sources

2. Cloud environment context: The cloud environment introduces a dynamic and complex ecosystem with various services, virtualized resources, and network configurations. By incorporating metadata from the cloud infrastructure, such as virtual machines, storage, networking, and identity services, agents can contextualize system call data within the broader cloud context. This allows for better correlation, identification of dependencies, and detection of threats and anomalous behaviors specific to the cloud environment.
3. Kubernetes orchestration: In containerized environments managed by Kubernetes, enriching syscall data with Kubernetes-specific metadata is crucial. This includes information about pods, containers, deployments, services, and labels. By understanding the relationships and configurations orchestrated by Kubernetes, agents can provide deeper insights into container behavior, workload distribution, resource utilization, and security events specific to the Kubernetes environment.
4. Container runtimes: Containers rely on specific runtime engines, such as Docker, containerd, or CRI-O. Enriching syscall data with metadata from these container runtimes enables better visibility into container lifecycle events, image details, container network namespaces, and resource utilization. This context allows for more accurate monitoring, security analysis, and performance optimization within containerized environments.
5. Identity and access context: Incorporating identity and access-related metadata, such as user information, roles, permissions, and authentication mechanisms, helps attribute system call activities to specific users or entities. This contextual information aids in auditing, compliance, and detecting potential security threats associated with user behavior.

This additional, multi-domain context helps security teams in these key ways:

- Scoping/partitioning: The context allows security events to be categorized based on any view, both physical (e.g., hosts) and logical (e.g., application centric views).
- Filtering: Events can be filtered out: for example, excluding events for lab environments, reducing alert noise.
- Prioritizing: Precise context facilitates prioritizing the alerts in the most critical environments: for example, internet-facing clusters or environments subject to strict regulations.
- Applying flexible policies: The additional context enables specific policies to be written against the same object in different environments: for example, regulating a container image between a development cluster (loose policy) and a production environment (restrictive policy).
- Assigning ownership: Precisely knowing not only the environment in which an event happens, but also the level of abstraction in which it happens, helps assign the security issues to the relevant work group.

Thanks to the flexible plugin architecture, Sysdig Secure can tap into other data sources alongside system calls. It can monitor AWS CloudTrail logs in real time and alert when there is suspicious user activity on a cloud resource. Having full visibility to the workloads, identities, and cloud services allows correlation across sources and follows the events to precisely trace the attacker's actions. Furthermore, immediate response can be taken on the workload side when a Falco rule is triggered, stopping malicious behavior in real time with precision.

SUMMARY

Cloud Security Demands Flexible, Scalable Insight

The cloud-native landscape continues to rapidly evolve, and security has become an increasingly important concern. As the complexity and scale of cloud-native applications and infrastructure continues to grow and security teams consolidate cloud security, it has become clear that a combination of agent and agentless approaches is needed to ensure effective security. Security teams must carefully evaluate the architectural approach their vendors are deploying, as security requires state-of-the-art instrumentation, not just “check the box” implementations.

The need for both agentless and agent-based approaches is particularly important in the context of cloud-native environments, where complex and dynamic systems require flexible and adaptable security measures. Another development that reinforces the need for both agent and agentless approaches is the rapid adoption of new types of abstractions, such as Lambda’s FaaS services. These specialized environments are designed with security in mind and require a combination of agent and agentless approaches to provide effective protection.

In addition, new application kernels and sandboxes are coming into the mainstream, and these heavily sandboxed environments require a different approach to securing them. These sandboxes increase security by restricting the low-level capabilities available to applications, so novel techniques are needed to effectively monitor and protect these environments. Fortunately, solutions like Sysdig and Falco are designed to be flexible enough to support new technologies and integrate with other projects. Falco’s open design makes it easy to integrate with the latest projects, and its modular architecture ensures that it can adapt to new technologies as they emerge, leveraging plugins for third-party services.

Overall, effective security in the cloud requires the highest levels of flexibility, performance, efficacy, and scalability. Sysdig Secure offers state-of-the-art solutions that satisfy all three requirements and have been battle tested in the biggest production cloud deployments on the planet for many years. Sysdig’s integrated agent and agentless approach is leading cloud security into a new era. By offering a comprehensive and adaptable security platform that effortlessly scales to match the demands of modern cloud environments, Sysdig Secure redefines what is possible in consolidated cloud security, delivering unparalleled resilience and ensuring robust protection for today’s dynamic digital landscapes.

Effective security in the cloud requires the highest levels of flexibility, performance, efficacy, and scalability.

About Sysdig

In the cloud, every second counts. Attacks move at warp speed, and security teams must protect the business without slowing it down. Sysdig stops cloud attacks in real time, instantly detecting changes in risk with runtime insights and open source Falco. We correlate signals across cloud workloads, identities, and services to uncover hidden attack paths and prioritize real risk. From prevention to defense, Sysdig helps enterprises focus on what matters: innovation.

To learn more about Sysdig, visit sysdig.com

REQUEST DEMO →

WP-007 REV. C 03/24