

# **Kubernetes Runtime Security** with Falco and Sysdig

Jorge Salamero @bencerillo

#### **About me**



#### Jorge Salamero

- Technical and Product Marketing@ Sysdig
- Used to be a speaker, DevOps and Debian Developer
- Behind many of the Falco integrations and Sysdig content and launches
- GitHub: bencer
- Twitter: @bencerillo



### **About Sysdig**

- The OSS project:
  - 2013 Linux kernel tracing tool
  - Evolution of tcpdump and Wireshark into the system
  - Easy to use (no code required), asynchronous, production performance
  - Container and Kubernetes support
- The company:
  - 2014 Sysdig Monitor
  - 2017 Sysdig Secure
  - Committed to OSS: sysdig, Sysdig Inspect, Falco, eBPF and Prometheus contributor



# Falco, a CNCF project

- Created originally by Sysdig the company
- Currently run independently by the Falco community
- Under the umbrella of the CNCF since 2018
- Sandbox level project, Incubation proposal WIP <a href="https://github.com/cncf/toc/pull/307">https://github.com/cncf/toc/pull/307</a>







# Why we need runtime security?



## **Runtime security**

CI/CD vulnerability scanning + ID management, is that enough?

- 1. Prevention / enforcement
- 2. Detection / audit
- 3. Blocking
- 4. Incident response and forensics



### Runtime security: prevention

- Who can do what within Kubernetes?
- Kubernetes native controls:
  - Admission controller / OPA
  - RBAC
  - Network policies
  - Pod Security Policies
    - seccomp
    - SELinux / AppArmor



### Runtime security: detection

- What happens when a control fails? Last line of defense!
- What do we count on when a catastrophe strikes?
- How do we tell a story when the unexpected happens?
- Applicable very few times, but critical when we need it.

#### But also:

- How can I validate enforcement works?
- Does it break my applications?



### Runtime security: use cases for detection

- Unpatched vulnerabilities, not public vulnerabilities, 0-day exploits
- Insecure configurations
- Leaked or insecure credentials
- Internal threats
- Compliance



### **Existing approaches**

- LD\_PRELOAD
  - Dependency on glibc
  - Changes your app, possible in unknown ways
- ptrace
  - Single PID, captures every system call
  - Changes your app, possible in unknown ways
- sidecars
  - Shared namespaces (process, network, storage, etc)
  - Instrumentation overhead, complexity, limited scope to the pod
- Kernel based



#### Kernel module vs eBPF

#### Kernel module

- Close to total system visibility
- Doesn't change containers or processes
- Asynchronous tracing, lowest impact
- Requires kernel-headers
- Can potentially crash the kernel

#### eBPF probe

- Close to total system visibility
- Doesn't change containers or processes
- Asynchronous tracing, low impact
- Requires kernel-headers
- Safe to run in eBPF VM

https://sysdig.com/blog/sysdig-and-falco-now-powered-by-ebpf/



# Falco: Kubernetes runtime security



#### What is Falco?

- Kubernetes runtime security tool
- Detection engine for anomalous activity in hosts and containers
- Rules built using tcpdump like syntax
- Leverages libscap and libsinsp
- Kubernetes native support (context, kube-apiserver audit)

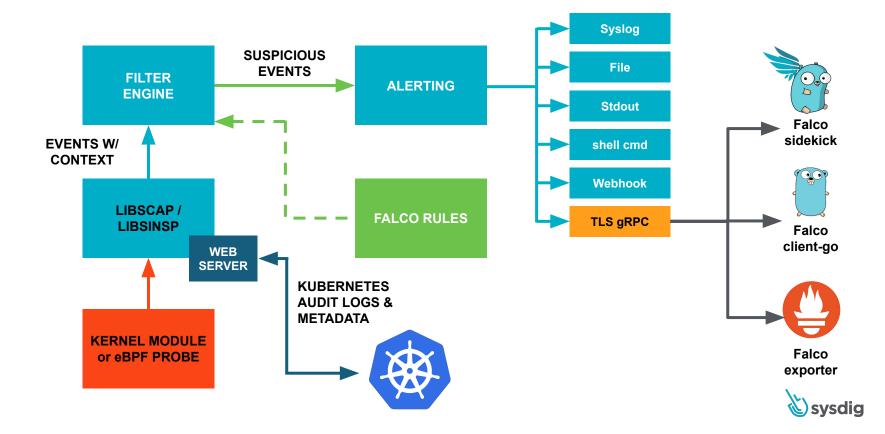


## What kind of problems does it solve?

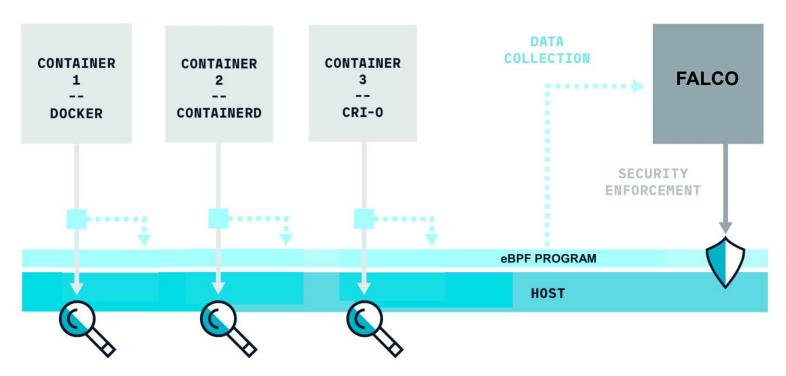
- Are my hosts and containers doing something they shouldn't?
- Spawned processes:
  - Did my PostgreSQL container spawn an unexpected process?
- File system reads, writes:
  - Did someone install a new package or change configuration in a running container?
- Network activity:
  - Did my Nginx container open a new listening port or unexpected outgoing connection?
- User or orchestration activity:
  - Did any K8s user spawn a shell into a privileged container?



#### **Falco architecture**

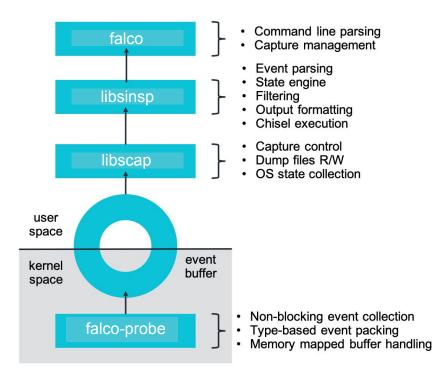


#### Falco low level architecture



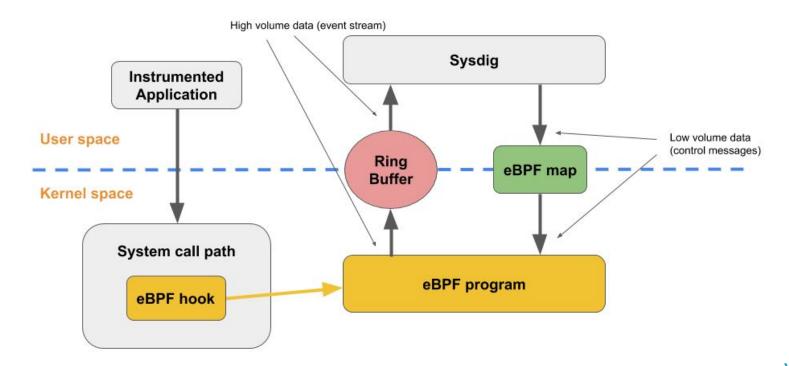


#### Falco lower level architecture





#### Falco lowest level architecture





#### Falco rule: container activity

```
- rule: Node container runs Node binary
 desc: Detect a process that's not node started in a Node container.
  condition: evt.type=execve and k8s.deployment.name=my-node-app and proc.name!=node
 output: Node container started unexpected process
               (user=%user.name command=%proc.cmdline %container.info)
 priority: INFO
 tags: [container, apps]
                                   In a container in
    Something is
                                                                       And the process
                                   my Kubernetes
     executing a
                                                                       name isn't node
                                   deployment for
      program
                                    my-node-app
```



#### **Kubernetes filters**



```
About Features Download Docs Blog 中文 Chinese
 k8s Field Class
Fields to filter on Kubernetes metadata. Allows rules to apply to particular namespaces ( k8s.ns.name ) or a resource's labels.
                 Kubernetes pod name.
  k8s.pod.name
  k8s.pod.id
                  Kubernetes pod id.
  k8s.pod.label Kubernetes pod label. E.g. 'k8s.pod.label.foo'.
  k8s.pod.labels Kubernetes pod comma-separated key/value labels, E.g. 'foo1:bar
                  1,foo2:bar2'.
                  Kubernetes replication controller name.
  k8s.rc.name
                  Kubernetes replication controller id.
  k8s.rc.id
  k8s.rc.label
                 Kubernetes replication controller label. E.g. 'k8s.rc.label.foo
  k8s.rc.labels Kubernetes replication controller comma-separated key/value lab
                  els. E.g. 'foo1:bar1,foo2:bar2'.
  k8s.svc.name
                 Kubernetes service name (can return more than one value, concat
  k8s.svc.id
                  Kubernetes service id (can return more than one value, concaten
  k8s.svc.label Kubernetes service label. E.g. 'k8s.svc.label.foo' (can return
                  more than one value, concatenated).
  k8s.svc.labels Kubernetes service comma-separated key/value labels, E.g. 'foo1
                  :bar1,foo2:bar2'.
  k8s.ns.name
                  Kubernetes namespace name.
  k8s.ns.id
                  Kubernetes namespace id.
  k8s.ns.label
                 Kubernetes namespace label. E.g. 'k8s.ns.label.foo'.
  k8s.ns.labels Kubernetes namespace comma-separated key/value labels. E.g. 'fo
                  o1:bar1.foo2:bar2'.
  k8s.rs.name
                  Kubernetes replica set name.
  k8s.rs.id
                  Kubernetes replica set id.
  k8s.rs.label
                 Kubernetes replica set label. E.g. 'k8s.rs.label.foo'.
  k8s.rs.labels Kubernetes replica set comma-separated key/value labels. E.g. '
                  foo1:bar1.foo2:bar2'.
  k8s.deployment.name
                  Kubernetes deployment name.
  k8s.deployment.id
                  Kubernetes deployment id.
  k8s.deployment.label
```



CLOUD NATIVE

#### Falco rule: Kubernetes activity

```
- macro: contains private credentials
 condition: >
   (ka.req.configmap.obj contains "aws access key id" or
   ka.reg.configmap.obj contains "aws s3 access key id" or
   ka.reg.configmap.obi contains "password")
- macro: configmap
 condition: ka.target.resource=configmaps
- macro: modify
 condition: (ka.verb in (create, update, patch))
- rule: Create/modify Configmap with private credentials
  desc: Detect creating/modifying a configmap containing a private credential
     (aws key, password, etc.)
  condition: configmap and modify and contains private credentials
  output: K8s configmap with private credential (user=%ka.user.name
           verb=%ka.verb name=%ka.reg.configmap.name
           configmap=%ka.req.configmap.name config=%ka.req.configmap.obj)
  priority: WARNING
  source: k8s audit
  tags: [k8s]
```



## Top runtime security violations

2019 Container Usage Report

#### Top runtime policy violations

We looked at policy violations as measured by the volume of alerts customers are receiving. This indicates the types of runtime security risks that container users are uncovering most frequently. Each of the following

violations are detected by Falco security policies that are enabled by default in Sysdig Secure. Below, we provide the top 10 violations in order of frequency, along with a description of each to explain the possible threat.

Violation	What it is	Why it's a security threat
Write below etc	Attempt to write to any file below the /etc directory	Adding or altering files in /etc, could be an attempt to change the application behavior.
Write below root	Attempt to write to any file directly below / or /root	Modifying data in these directories could be an unauthorized attempt to install software on the container.
Launch privileged container	Starting a privileged container	Privileged containers can interact with host system devices, cause harm to the host OS, and gain access to other containers.
Change thread namespace	Attempt to change a program/ thread's namespace by calling setns	Could indicate a privilege escalation and an attempt to gain access to other containers.
Launch sensitive mount container	Starting a container that has a file system mount from a sensitive host directory	Indicates the container has to access to data volumes that might contain sensitive files.
Non sudo setuid	Attempt to change users by calling setuid	Could indicate an attempt by a process to elevate its privileges.
Write below binary dir	Attempt to write to any file below a set of binary directories	Could indicate a malicious attempt to install unauthorized software like backdoors.
Run shell untrusted	Attempt to spawn a shell below a non-shell application	Enables an attacker to manipulate the system, download malware, or initiate other malicious activity.
System procs network activity	Network activity performed by system binaries that are not expected to send or receive network traffic	Binaries that are should not have network activity have network activity, indicating that the binary has been compromised.
Terminal shell in container	A shell was used as the entrypoint/ exec point into a container with an attached terminal	Enables an attacker to manipulate the system, download malware, or initiate other malicious activity

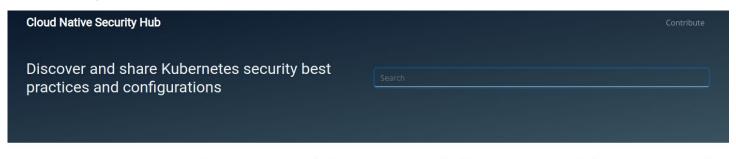


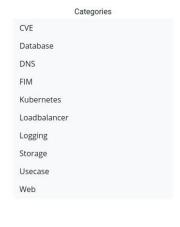
# Popular Falco detection rules

Best practices	Compliance	Vulnerabilities	Cloud Native Stack
Update packages	FIM	CVE-2019-11246	K8s control plane
Modify /bin /usr	Privileged pod	kubectl cp	Nginx
Write below /etc	ConfigMap creds		Elasticsearch
Read sensitive file	kubectl exec/attach	CVE-2019-5736	Redis
DB spawned proc	Role changes audit	runc breakout	HAproxy
Change namespace	PCI		Rook
Privileged container	NIST	CVE-2019-14287	MongoDB
Sensitive mount		sudo bypass	PostgreSQL
Terminal shell			



# SecurityHub.dev









fluentd

Falco rules for securing

FluentD



Falco rule 🗽

elasticsearch

Falco rules for securing

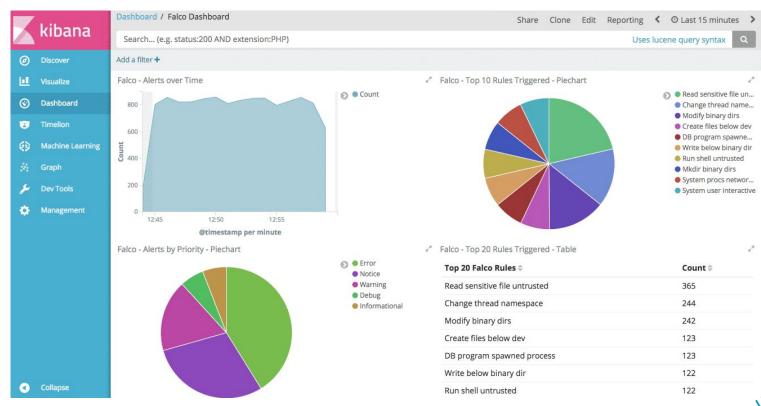
ElasticSearch







#### SIEM

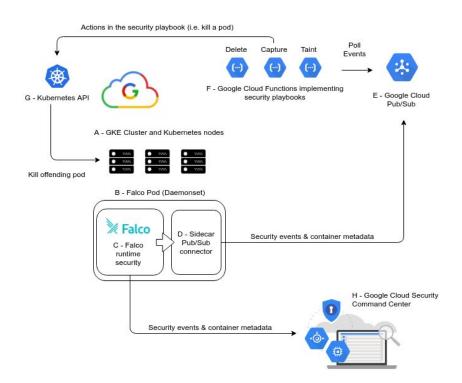




### Response Engine

Trigger automated reactions to events
Blocking component of runtime security
Security playbooks executed as FaaS

- Taint a node NoSchedule
- Isolate pod via Network Policy
- Delete offending pod
- Scale down deployment to 0 pods
- Trigger a Sysdig capture
- Send notifications





#### **ADOPTERS.md**

Booz Allen Hamilton Shopify

Frame.io Sight Machine

League Sumo Logic

Preferral Sysdig

And what about you? Using Falco in production? Talk about it! <a href="https://github.com/falcosecurity/falco/blob/dev/ADOPTERS.md">https://github.com/falcosecurity/falco/blob/dev/ADOPTERS.md</a>



# **Getting involved**

Website

https://falco.org/

Blogs

https://falco.org/blog/

https://sysdig.com/blog/

Github

https://github.com/falcosecurity/

Slack

http://slack.sysdig.com/

Docs

https://falco.org/docs/

SecurityHub

https://securityhub.dev/



# Sysdig Secure extends Falco functionality



### **Extending Falco Across the Lifecycle**



Runtime detection

Build	Run	Respond
<ul><li>Image Scanning</li><li>Configuration</li><li>Validation</li></ul>	<ul> <li>Runtime prevention</li> <li>Automated policy creation</li> <li>Policy editor and rules library</li> <li>Threat blocking</li> </ul>	<ul><li>Incident Response</li><li>Forensics</li><li>Audit</li></ul>
Conti	nuous Compliance (PCI, NIST, CI	S, etc.) ———

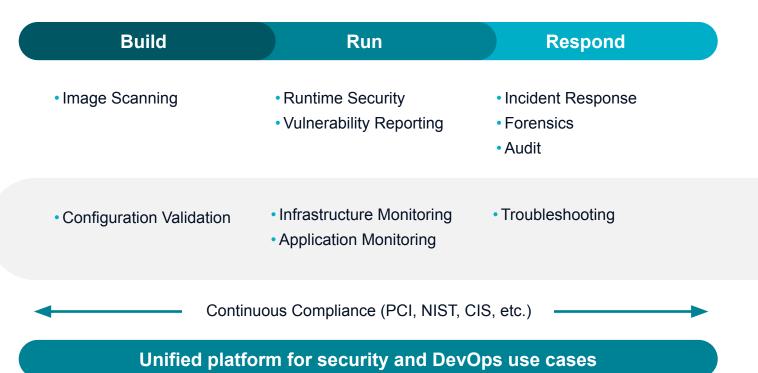


## Platform Built on an Open Foundation

Build Run Respond Sysdig Secure DevOps Platform Adds scale, workflow, K8s, and cloud context **%** Falco sysdig **Prometheus** Image scanning Monitoring **Runtime security** Forensics/troubleshooting Infrastructure and **Detection rules** Vulnerability analysis Deep visibility into application metrics and alerts container activity



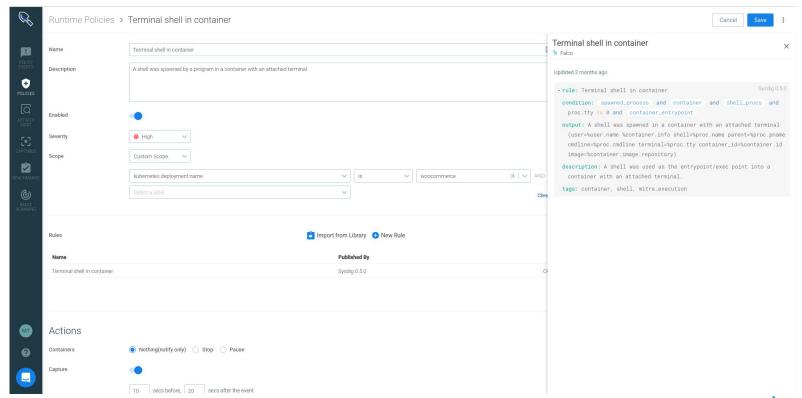
# Secure DevOps Across Cloud-Native Lifecycle





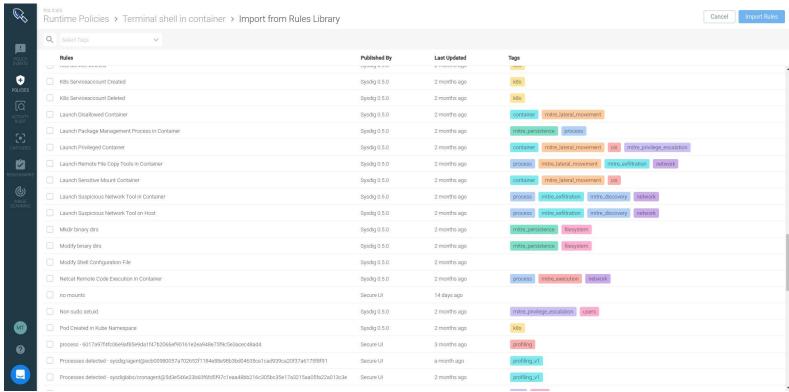


## Inside Sysdig Secure: Falco editor





# Inside Sysdig Secure: Falco library





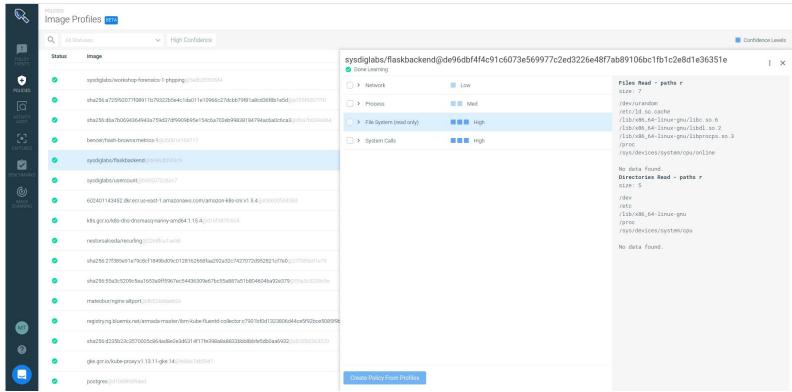
## Inside Sysdig Secure: Falco tuning

#### #policy-event-onprem & (i) & ☆ | & 6 | 육 0 | Secure policy tuning for individual customer Thursday, November 7th 9 events generated in 5 minutes (Burst Ratio: 0.02) across [O containers | 9 hosts | O images | 1 processes | 1 file descriptors] Processes: [sed 9]] Process + FD: [{sed /etc/pam.d/ 9}] Sysdig Secure Nov 7th Suggesting changes in 6 falco rules 1. Read sensitive file untrusted - macro: host sed access files condition: (proc.name=sed and (fd.name startswith "/etc/pam.d/")) 2. Write below root - macro: host\_exe\_access\_files condition: (proc.name=exe and (fd.name startswith "/hosts" or fd.name startswith "/test.vml")) - macro: host runc:[1:CHILD] access files condition: (proc.name=runc:[1:CHILD] and (fd.name startswith "/exec.fifo")) 3. Write below etc - macro: host insights-client access files condition: (proc.name=insights-client and (fd.name startswith "/etc/insights-client/")) 4. Write below binary dir - macro: host\_exe\_access\_files condition: (proc.name=exe and (fd.name startswith "/usr/sbin/" or fd.name startswith "/usr/bin/")) 5. Write below rpm database - macro: host exe access files condition: (proc.name=exe and (fd.name startswith "/var/lib/")) 6. Launch Privileged Container - macro: filebeat image condition: (container.image.repository endswith beats/filebeat) Sysdig Secure Nov 7th

- Makes suggestions on what to change on the Falco rules (default rules or your own rules).
- You decide what you merge or not (easy with Policy Editor!).
- This Falco rule triggers all the time, these are false positives, it's very noisy creating alert fatigue, how we can fix that at scale for all policies?

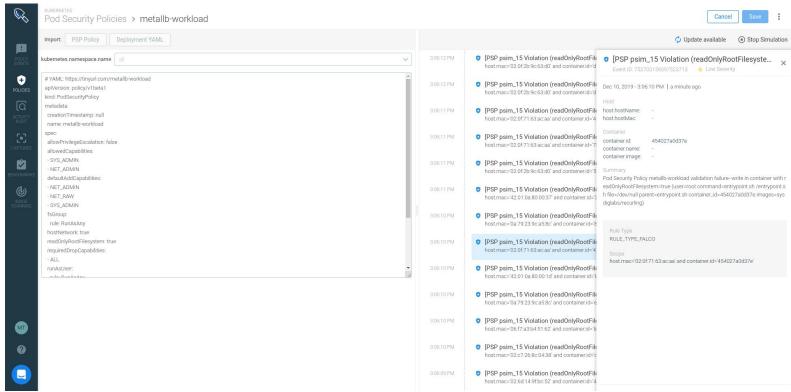


# Inside Sysdig Secure: Profiling



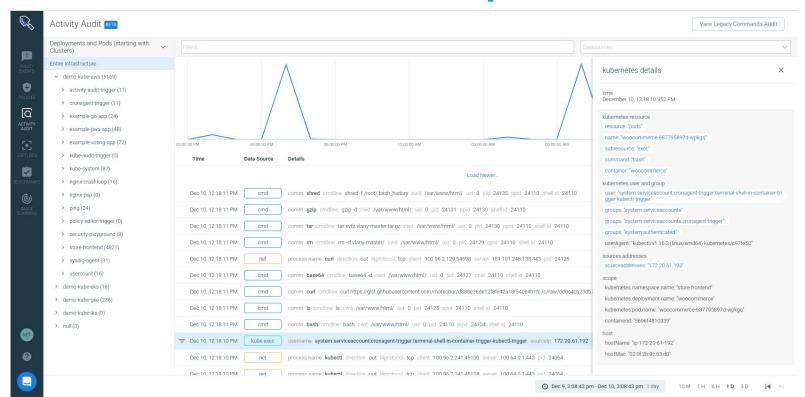


# Inside Sysdig Secure: Policy Advisor





#### Cloud-native Incident Response / Forensics





# **Upcoming events**

• Join us Dec. 19 for **Star Wars** premieres:

https://go.sysdig.com/starwars2019

• In 2020

RSA ~ February 24-28 in San Francisco

KubeCon ~ March 31-April 2 in Amsterdam

Red Hat Summit ~ April 27-29 in San Francisco



#### Learn more

Falco

https://falco.org/

Container Usage Report and webinar

https://sysdig.com/resources/papers/2019-container-usage-report/

Sysdig blog

https://sysdig.com/blog/

What's new in Kubernetes 1.17

https://sysdig.com/blog/whats-new-kubernetes-1-17/





Dig deeper