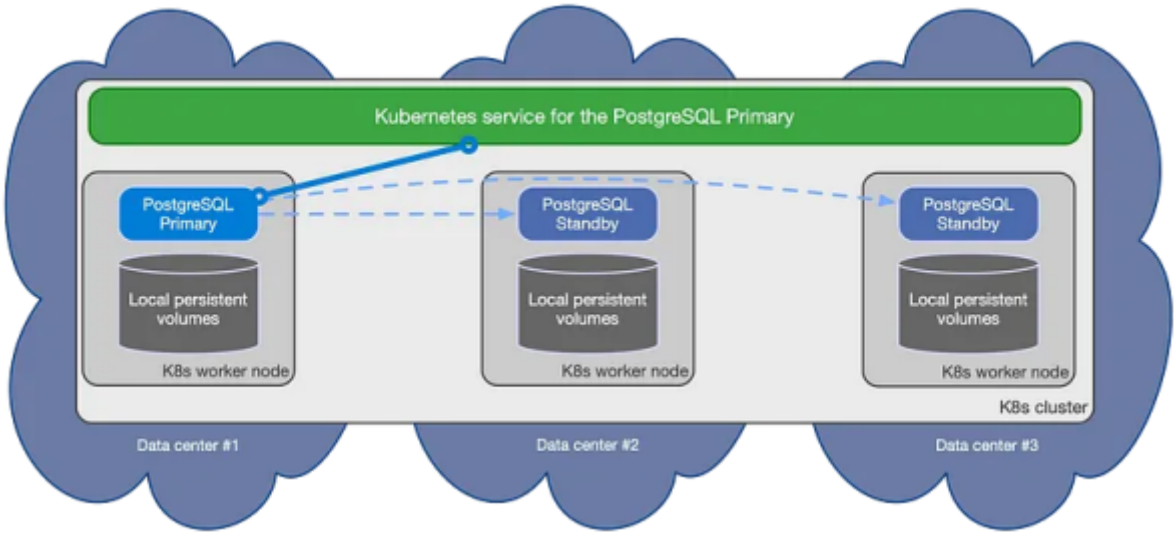
Pod Security Admission

Die Kubernetes Pod Security Admission (PSA) ist eine im Core von Kubernetes integrierte Admission-Controller-Strategie, um die Sicherheit von Pods zu gewährleisten. Sie ersetzt den früheren `PodSecurityPolicy` Mechanismus, der ab Kubernetes v1.25 als veraltet gilt. Mit Pod Security Admission lassen sich für Namespaces einfache, deklarative Sicherheitsstandards festlegen, die direkt im Namespace-Objekt hinterlegt werden.



Funktionsweise

Der Pod Security Admission Controller überprüft beim Erstellen oder Aktualisieren eines Pods, ob dessen Spezifikation den im Namespace konfigurierten Sicherheitsstufen entspricht. Pods, die restriktive Einstellungen unterlaufen, werden abgewiesen oder mit Warnungen versehen.

Es stehen drei Sicherheitsprofile zur Verfügung:

- **privileged**: minimale Einschränkungen (für vertrauenswürdige Workloads)
- **baseline**: Standard-Sicherheitsfunktionen, empfohlen für allgemeine Anwendungen
- **restricted**: strengste Einstellungen, empfohlen für Produktionsumgebungen

Konfiguration

Die Sicherheitsprofile werden auf Namespace-Ebene über Labels definiert. Jedes Profil kann drei Modi haben:

- `enforce`: Verhindert die Erstellung unzulässiger Pods.
- `audit`: Schreibt Verstöße ins Audit-Log, blockiert den Vorgang aber nicht.
- `warn`: Gibt eine Warnung zurück, aber erstellt trotzdem den Pod.

Beispiel: Namespace-Labelling

```
kubectl label namespace <namespace-name> \
    pod-security.kubernetes.io/enforce=baseline \
    pod-security.kubernetes.io/enforce-version=v1.34 \
    pod-security.kubernetes.io/warn=restricted \
    pod-security.kubernetes.io/warn-version=v1.34
```

Mit diesem Beispiel:

- Werden nur Pods akzeptiert, die das Profil `baseline` erfüllen.
- Bei Verstoß gegen `restricted` gibt es eine Warnung.

Ein weiteres Beispiel als Manifestfile

```
apiVersion: v1
kind: Namespace
metadata:
  name: my-baseline-namespace
  labels:
    pod-security.kubernetes.io/enforce: baseline
    pod-security.kubernetes.io/enforce-version: v1.34

    # We are setting these to our _desired_ `enforce` level.
    pod-security.kubernetes.io/audit: restricted
    pod-security.kubernetes.io/audit-version: v1.34
    pod-security.kubernetes.io/warn: restricted
    pod-security.kubernetes.io/warn-version: v1.34
```

Es besteht auch die Möglichkeit statt einer eindeutigen Version einfach `latest` zu nehmen.

Profile im Detail

1. **Privileged**
   - Kaum Einschränkungen.
   - Erlaubt z. B. privilegierte Container, Host-Networking, HostPath Volumes.
2. **Baseline**
   - Verbietet die meisten eskalierenden Features.
   - Erlaubt, was für viele Anwendungen benötigt wird, wie z. B. einige Capabilities.
3. **Restricted**
   - Strikte Isolation des Pods.
   - Kein privilegierter Zugriff, keine systemkritischen Capabilities.

Mehr zu den genauen Profile-Einstellungen in der offiziellen Dokumentation.

Referenzen

- Kubernetes Documentation: Pod Security Admission
- Security Standards: Baseline & Restricted

Hinweis: Für eine produktive Umgebung sollte die Konfiguration regelmäßig geprüft und mit weiteren Maßnahmen (z. B. Role-based Access Control, Image Policies) kombiniert werden.

Anwendung von Pod Security Standards verstehen

`--dry-run=server` dient dazu, zu verstehen was passiert, wenn verschiedene Pod-Sicherheitsstandards angewendet werden:

1. **Privileged**

```
kubectl label --dry-run=server --overwrite ns --all \
pod-security.kubernetes.io/enforce=privileged
```

Die Ausgabe sieht in etwa so aus:

```
namespace/default labeled
namespace/kube-node-lease labeled
namespace/kube-public labeled
namespace/kube-system labeled
namespace/local-path-storage labeled
```

2. **Baseline**

```
kubectl label --dry-run=server --overwrite ns --all \
pod-security.kubernetes.io/enforce=baseline
```

Die Ausgabe sieht in etwa so aus:

```
namespace/default labeled
namespace/kube-node-lease labeled
namespace/kube-public labeled
Warning: existing pods in namespace "kube-system" violate the new
PodSecurity enforce level "baseline:latest"
Warning: etcd-psa-wo-cluster-pss-control-plane (and 3 other pods): host
namespaces, hostPath volumes
Warning: kindnet-vzj42: non-default capabilities, host namespaces,
hostPath volumes
Warning: kube-proxy-m6hwf: host namespaces, hostPath volumes,
privileged
namespace/kube-system labeled
namespace/local-path-storage labeled
```

3. **Restricted**

```
kubectl label --dry-run=server --overwrite ns --all \
pod-security.kubernetes.io/enforce=restricted
```

Die Ausgabe sieht in etwa so aus:

```
namespace/default labeled
```

```
namespace/kube-node-lease labeled
namespace/kube-public labeled
Warning: existing pods in namespace "kube-system" violate the new
PodSecurity enforce level "restricted:latest"
Warning: coredns-7bb9c7b568-hsptc (and 1 other pod): unrestricted
capabilities, runAsNonRoot != true, seccompProfile
Warning: etcd-psa-wo-cluster-pss-control-plane (and 3 other pods): host
namespaces, hostPath volumes, allowPrivilegeEscalation != false,
unrestricted capabilities, restricted volume types, runAsNonRoot !=
true
Warning: kindnet-vzj42: non-default capabilities, host namespaces,
hostPath volumes, allowPrivilegeEscalation != false, unrestricted
capabilities, restricted volume types, runAsNonRoot != true,
seccompProfile
Warning: kube-proxy-m6hwf: host namespaces, hostPath volumes,
privileged, allowPrivilegeEscalation != false, unrestricted
capabilities, restricted volume types, runAsNonRoot != true,
seccompProfile
namespace/kube-system labeled
Warning: existing pods in namespace "local-path-storage" violate the
new PodSecurity enforce level "restricted:latest"
Warning: local-path-provisioner-d6d9f7ffc-lw9lh:
allowPrivilegeEscalation != false, unrestricted capabilities,
runAsNonRoot != true, seccompProfile
namespace/local-path-storage labeled
```

Aus der vorherigen Ausgabe geht hervor, dass die Anwendung des privileged Pod-Sicherheitsstandards keine Warnungen für irgendwelche Namespaces ausgibt. Die Standards `baseline` und `restricted` hingegen weisen beide Warnungen auf, insbesondere im `kube-system` Namespace.

Pod-Sicherheitsstandards auf Clusterebene

Es werden die folgenden Pod-Sicherheitsstandards auf die Version latest angewendet:

- baseline standard in enforce mode
- restricted standard in warn and audit mode

Der baselinePod Security Standard bietet einen praktischen Mittelweg, der es ermöglicht, die Ausnahmeliste kurz zu halten und bekannte Rechteausweitungen zu verhindern. Um außerdem zu verhindern, dass Pods in

- kube-system
- calico-system

fehlschlagen, werden diese Namespaces von der Anwendung der Pod-Sicherheitsstandards ausgenommen.

Bei der Implementierung von Pod Security Admission auf Cluster Ebene sollte Folgendes beachtet werden:

- Abhängig von der Risikobewertung eines Clusters könnte ein strengerer Pod-Sicherheitsstandard die bessere Wahl sein. Beispielsweise `restricted`

Für die Clusterweite Anwendung wird der Admission Controller konfiguriert

```
apiVersion: apiserver.config.k8s.io/v1
kind: AdmissionConfiguration
plugins:
- name: PodSecurity
  configuration:
    apiVersion: pod-security.admission.config.k8s.io/v1 # see compatibility note
    kind: PodSecurityConfiguration
    # Defaults applied when a mode label is not set.
    #
    # Level label values must be one of:
    # - "privileged" (default)
    # - "baseline"
    # - "restricted"
    #
    # Version label values must be one of:
    # - "latest" (default)
    # - specific version like "v1.33"
    defaults:
      enforce: "baseline"
      enforce-version: "latest"
      audit: "restricted"
      audit-version: "latest"
      warn: "restricted"
      warn-version: "latest"
    exemptions:
      # Array of authenticated usernames to exempt.
      usernames: []
      # Array of runtime class names to exempt.
      runtimeClasses: []
      # Array of namespaces to exempt.
      namespaces: [kube-system,calico-system]
```

--admission-control-config-file
Das obige Manifest muss über den kube-apiserver angegeben werden.

1. Bereiten Sie das Konfigurationsverzeichnis vor:

```
mkdir -p /etc/kubernetes/psa
cp psa-config.yaml /etc/kubernetes/psa/
```

2. Ändern Sie die API-Serverkonfiguration:

```
apiServer:
  extraArgs:
    admission-control-config-file:  /etc/kubernetes/psa/psa-config.yaml
  extraVolumes:
    -  name:  psa-config
      hostPath:  /etc/kubernetes/psa
      mountPath:  /etc/kubernetes/psa
      readOnly:  true
```

3. Starten Sie den API-Server neu:
   Nach der Aktualisierung der Konfiguration muss der API-Server neu gestartet werden, um die Änderungen anzuwenden.

Container und Container Image Security

trivy

Trivy has scanners that look for security issues, and targets where it can find those issues.

Targets (what Trivy can scan):

- Container Image
- Filesystem
- Git Repository (remote)
- Virtual Machine Image
- Kubernetes
- AWS

Scanners (what Trivy can find there):

- OS packages and software dependencies in use (SBOM)
- Known vulnerabilities (CVEs)
- IaC issues and misconfigurations
- Sensitive information and secrets
- Software licenses

kube-bench

kube-bench is a tool that checks whether Kubernetes is deployed securely by running the checks documented in the CIS Kubernetes Benchmark.

CIS Scanning as part of Trivy and the Trivy Operator

Trivy, the all in one cloud native security scanner, can be deployed as a Kubernetes Operator inside a cluster. Both, the Trivy CLI, and the Trivy Operator support CIS Kubernetes Benchmark scanning among several other features.