Cloud Container Engine

Bulletin

Issue 01

Date 2022-01-17



HUAWEI

HUAWEI TECHNOLOGIES CO., LTD.

Copyright © Huawei Technologies Co., Ltd. 2022. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Contents

1 Latest Announcements....1
2 CCE Security Guide....4
3 Kubernetes Version Support Mechanism....6
4 OS Version Synchronization Mechanism....8
5 Kubernetes Release Notes....11
5.1 CCE Kubernetes 1.21 Release Notes....11
5.2 CCE Kubernetes 1.19 Release Notes....13
5.3 CCE Kubernetes 1.17 Release Notes....16
5.4 CCE Kubernetes 1.15 Release Notes....18
5.5 CCE Kubernetes 1.13 Release Notes....21
5.6 CCE Kubernetes 1.11 Release Notes....22
5.7 Release Notes of CCE Kubernetes 1.9 and Earlier Versions....23
6 Vulnerability Notice....28
6.1 Vulnerability Fixing Policies....28
6.2 Notice on the Vulnerability of Kubernetes subPath Symlink Exchange (CVE-2021-25741)....28
6.3 Notice on runc Vulnerability that Allows a Container Filesystem Breakout via Directory Traversal (CVE-2021-30465)....31
6.4 Notice on the Docker Resource Management Vulnerability (CVE-2021-21285)....32
6.5 Notice on the NVIDIA GPU Driver Vulnerability (CVE-2021-1056)....34
6.6 Notice on the Sudo Buffer Vulnerability (CVE-2021-3156)....36
6.7 Notice on the Kubernetes Security Vulnerability (CVE-2020-8554)....37
6.8 Notice on the Apache containerd Security Vulnerability (CVE-2020-15257)....38
6.9 Notice on the Docker Engine Input Verification Vulnerability (CVE-2020-13401)....39
6.10 Notice on the Kubernetes kube-apiserver Input Verification Vulnerability (CVE-2020-8559)....40
6.11 Notice on the Kubernetes kubelet Resource Management Vulnerability (CVE-2020-8557)....41
6.12 Notice on the Kubernetes kubelet and kube-proxy Authorization Vulnerability (CVE-2020-8558)....43
6.13 Notice on Fixing Kubernetes HTTP/2 Vulnerability....45
6.14 Notice on Fixing Linux Kernel SACK Vulnerabilities....47
6.15 Notice on Fixing the Docker Command Injection Vulnerability (CVE-2019-5736)....48
6.16 Notice on Fixing the Kubernetes Permission and Access Control Vulnerability (CVE-2018-1002105)....50

 Copyright © Huawei Technologies Co., Ltd.

6.17 Notice on Fixing the Kubernetes Dashboard Security Vulnerability (CVE-2018-18264).....52

 Copyright © Huawei Technologies Co., Ltd.

1 Latest Announcements

Notice on the Vulnerability of Kubernetes subPath Symlink Exchange (CVE-2021-25741)

Released: September 17, 2021

A security issue was spotted in Kubernetes where a user may be able to create a container with subPath volume mounted to access files and directories outside of the volume, including those on the host file system.

For details, see **Notice on the Vulnerability of Kubernetes subPath Symlink Exchange (CVE-2021-25741)**.

Notice on the NVIDIA GPU Driver Vulnerability CVE-2021-1056

Released: April 16, 2021

NVIDIA announced the vulnerability (CVE-2021-1056) about the NVIDIA GPU driver. According to the official NVIDIA announcement, if your CCE cluster has a GPU-enabled node (ECS) and uses the recommended NVIDIA GPU driver (Tesla 396.37), your NVIDIA driver is not affected by this vulnerability. If you have installed another driver or updated the driver on your node, this vulnerability may be involved.

For details, see **Notice on the NVIDIA GPU Driver Vulnerability (CVE-2021-1056)**.

Official Release of CCE Turbo Clusters

Released: March 31, 2021

CCE Turbo clusters run on the cloud native 2.0 infrastructure that features software-hardware synergy to support passthrough networking, high security and reliability, and intelligent scheduling.

For details, see **Buying a CCE Turbo Cluster**.

Optimized Key Authentication of Add-on everest

Released: Feb 2, 2021

 Copyright © Huawei Technologies Co., Ltd.

In **version 1.2.0** of the everest add-on, **key authentication** is optimized when OBS is used. After the everest add-on is upgraded from a version earlier than 1.2.0 to 1.2.0 or later, you need to restart all workloads that use OBS in the cluster. Otherwise, workloads may not be able to use OBS.

For details, see **everest**.

Add-on nginx-ingress Available Again

Released: Jan 20, 2021

The sample nginx-ingress chart on the **Charts** page will be taken offline soon. If you have installed nginx-ingress using the sample chart, uninstall the chart release and install the add-on again.

The nginx-ingress add-on in CCE is implemented using the open-source community chart and image. CCE does not maintain the add-on. Therefore, it is not recommended that the nginx-ingress add-on be used commercially.

Constraints: This add-on can be installed only in clusters of v1.15 and later.

For details about the add-on and how to install it, see **nginx-ingress**.

Notice on Upgrading the npd Add-on

Released: Dec 14, 2020

Dear users,

The npd add-on of v1.13.6 is now available for CCE clusters of v1.15.11 and v1.17.9.

Starting from December 30, 2020, CCE will automatically upgrade the npd add-on to the latest version. Containers with this add-on installed on each node will be restarted, but your services will not be affected. You can go to Add-ons on the CCE console to upgrade the add-on now. For details about how to upgrade the add-on, see **npd**.

The upgrade improves the capability and reliability of the npd add-on. If you do not want to upgrade the add-on, you can submit a service ticket to cancel the upgrade and manually upgrade the add-on later.

CVE-2020-8554 Vulnerability Notice

Released: Dec 9, 2020

In the scenario where multiple users in a single cluster share resources, if the creation and update permissions of pods and Services are assigned to untrusted users, the users are susceptible to this vulnerability. You are advised to check all Services that use externalIP and loadBalancerIP to determine whether there are vulnerable Services.

This vulnerability is caused by Kubernetes design defects. The final solution to solve this issue once and for all is still under discussion in the open source community. Currently, you can prevent this issue by restricting the use of externalIP or LoadBalancerIP.

 Copyright © Huawei Technologies Co., Ltd.

For details, see **Notice on the Kubernetes Security Vulnerability (CVE-2020-8554)**.

Removing Support for Creating Clusters of v1.13 and Earlier

Released: Dec 8, 2020

According to the version policy released in **Kubernetes Version Support Mechanism**, from **March 1, 2021**, clusters of Kubernetes version 1.13 and earlier cannot be created in CCE. Existing clusters of v1.13 are still maintained.

Upgrade Notice for Clusters of Kubernetes v1.9

Released: Dec 7, 2020

According to the version policy in **Kubernetes Version Support Mechanism**, CCE will stop maintaining clusters of Kubernetes v1.9 recently. If you are using clusters of v1.9.7 or v1.9.10, you are advised to upgrade your clusters to a later version before **April 30, 2021**. If not, these clusters cannot be upgraded anymore.

For details about how to upgrade your clusters, see **Cluster Upgrade Changelog**.

CVE-2020-15257 Vulnerability Notice

Released: Dec 2, 2020

To avoid this vulnerability, you can prevent containers from running by user **root** (UID is 0) or running on the host network. CCE will provide handling solutions in its versions as soon as possible.

For details, see **Notice on the Apache containerd Security Vulnerability (CVE-2020-15257)**.

npd Add-on Upgrade Notice

Released: Nov 28, 2020

The new npd add-on version 1.13.6 is now available for clusters of v1.15.11 and v1.17.9. You can upgrade the add-on to improve the reliability of your nodes.

For details, see **npd**.

 Copyright © Huawei Technologies Co., Ltd.

2 CCE Security Guide

Node Security Hardening

Cloud Container Engine (CCE) cluster nodes are created from the public images published by Image Management Service (IMS) in the public cloud. CCE performs security hardening on the CCE software installed on nodes. The OS configurations are the same as those of public images by default. You need to perform security hardening based on your own security requirements.

Container Security Suggestions

CCE provides you with an exclusive cluster. You are not advised to share a cluster with multiple tenants. If a cluster must be shared by multiple tenants, perform security hardening by following the suggestions in the following table.

Table 2-1 Container security configuration suggestions

Category	Suggestion

Least privilege	1. Use the sudo whitelist to restrict the Docker commands that can be run in containers by common users.
	2. It is recommended that privileged containers be disabled for Kubernetes pods. To be specific, do not use the following settings:
	`apiVersion: v1`
	`kind: Pod`
	`metadata:`
	`  name: hello-world`
	`spec:`
	`  containers:`
	`  - name: hello-world-container`
	`  # The container definition`
	`  # ...`
	`  securityContext:`
	`  privileged: true`

 Copyright © Huawei Technologies Co., Ltd.

Category	Suggestion

	By default, the token of ServiceAccount **default** is not used.
	• The automatic mounting function is disabled for ServiceAccount **default** by default. If **automountServiceAccountToken** is set to **false**, the created workload will not be mounted to a token by default.
	CAUTION
	Set this field for each namespace as required. After the setting is complete, perform a rolling restart for the pods in the corresponding namespace.
	`apiVersion: v1`
	`kind: ServiceAccount`
	`metadata:`
	`  name: default`
	`automountServiceAccountToken: false`
	• Mount the token of ServiceAccount **default** to specific workloads.
	`...`
	`spec:`
	`  template:`
	`  spec:`
	`  serviceAccountName: default`
	`  automountServiceAccountToken: true`
	`  ...`

Network isolation	**Network isolation between containers in a cluster** Use network policies to implement access control. For details, see **Network Policies**. **Network isolation between containers and hosts outside the cluster** Use security groups in a VPC to implement access control. For details, see **Security Group Configuration Examples**. If network isolation between containers and hosts is required, you are advised to use secure containers of CCE Turbo clusters. **Network isolation between containers and the cluster control plane** CCE clusters are private and workloads would access the cluster API server. Therefore, CCE does not restrict the network communication between containers and the Kubernetes control plane. If you want to ensure the network QoS between Kubernetes components, you can limit the network bandwidth of containers by referring to **Setting the Container Network Bandwidth Limit**.
Mounting of sensitive host directories	• If not necessary, do not mount the **/var/run/docker.sock** file on the host to a container when starting the container. • Do not mount the hostPath directory, such as **/**, **/boot**, **/dev**, **/etc**, **/lib**, **/proc**, **/sys**, and **/usr**, on the host to the container. If necessary, mount the hostPath directory in read-only mode.

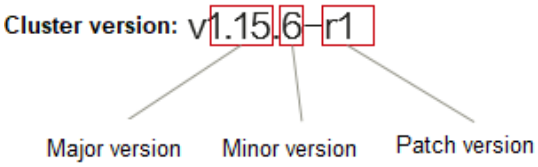
Copyright © Huawei Technologies Co., Ltd.

3 Kubernetes Version Support Mechanism

This section describes the Kubernetes version support mechanism of CCE.

Version Description

Version number: The format is **x.y.z**, where **x.y** is the major version and **z** is the minor version. If the version number is followed by **-r**, the version is a patch version, for example, v1.15.6-r1.



NOTE

Starting from Kubernetes 1.21, CCE only displays the major version number, for example, v1.21.

Term Description

Offline: After a version is brought offline, a cluster of this version cannot be created on the CCE console and no new features will be released for the clusters of this version.

Obsolete: CCE will no longer provide support for this version, including release of new functions, community bug fixes, vulnerability management, and upgrade.

Version Requirements

CCE releases only odd major Kubernetes versions, such as v1.17, v1.15, and v1.13. The specific version support policies in different scenarios are as follows:

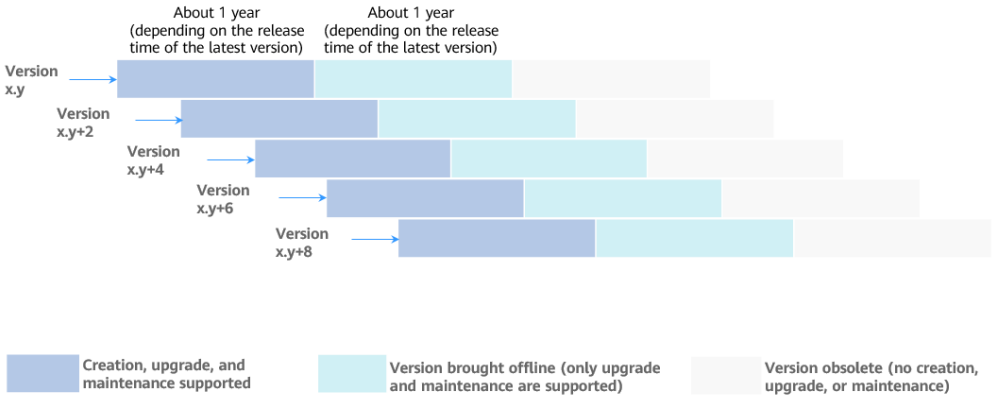
- Cluster creation

  CCE allows you to create clusters of two latest major Kubernetes versions, for example, v1.17 and v1.15. Earlier versions (such as v1.13) will go offline, and you will not be able to create clusters of these versions on the CCE console.

 Copyright © Huawei Technologies Co., Ltd.

- Cluster maintenance

  CCE maintains clusters of four major Kubernetes versions at most, such as v1.11, v1.13, v1.15, and v1.17. For example, after v1.17 is put into commercial use, earlier versions (such as v1.9) will be obsolete.



- Cluster upgrade

  CCE allows you to upgrade clusters of **3 major versions** at the same time. Each version is maintained for one year. For example, after v1.17 is available, you can upgrade your clusters of v1.11, v1.13, and v1.15, but not earlier versions (such as v1.9). Therefore, you must upgrade your Kubernetes cluster before the maintenance period ends.

  - Cluster version upgrade: After clusters of the latest major version (for example, v1.17) go online, CCE allows you to upgrade clusters from an earlier version to the last stable version of the second-latest major version, for example, v1.15.11. For details, see **Changelog for Upgrading Clusters**.
  - Cluster patch upgrade: For existing clusters running on the live network, if there are major Kubernetes issues or vulnerabilities, CCE will perform the patch upgrade on these clusters in the background. Users are unaware of the patch upgrade. If the patch upgrade has adverse impact on user services, CCE will release a notice one week in advance.

Version Release Cycle

Kubernetes releases a major version in about three months. CCE follows the same frequency as Kubernetes to release major versions. To be specific, a new CCE version is released about three months after a new Kubernetes version is released in the community. For example, Kubernetes v1.13 was released in December 2018 and CCE v1.13 was released in February 2019.

Version Constraints

After a cluster is upgraded, it cannot be rolled back to the source version.

 Copyright © Huawei Technologies Co., Ltd.

4 OS Version Synchronization Mechanism

Version Support Mechanism

The cluster versions released by CCE are bound to the OS versions of nodes. Currently, three OSs are supported: EulerOS, CentOS, and Ubuntu.

Synchronization Mechanism

OS synchronization mechanism: OS versions are released or upgraded with cluster versions.

Major OS vulnerability fixing: The policy for fixing major OS vulnerabilities is consistent with the cluster patch upgrade policy.

For details, see **Cluster upgrade**.

Mappings Between Cluster Versions and OS Versions

Table 4-1 lists the mappings between released cluster versions and OS versions.

Table 4-1 Mappings between VM node OS versions and cluster versions

OS Version	Cluster Version	Kernel Information
CentOS Linux release 7.6	v1.21	3.10.0-1160.25.1.el7.x86_64
	v1.19.10	3.10.0-1160.25.1.el7.x86_64
	v1.19.8	3.10.0-1160.15.2.el7.x86_64
	v1.17.17	3.10.0-1160.15.2.el7.x86_64
	v1.17.9	3.10.0-1062.12.1.el7.x86_64
	v1.15.11	3.10.0-1062.12.1.el7.x86_64
	v1.15.6-r1	3.10.0-1062.1.1.el7.x86_64
	v1.13.10-r1	3.10.0-957.21.3.el7.x86_64
	v1.13.7-r0	3.10.0-957.21.3.el7.x86_64

 Copyright © Huawei Technologies Co., Ltd.

OS Version	Cluster Version	Kernel Information
EulerOS release 2.9	v1.21	4.18.0-147.5.1.6.h541.eulerosv2r9.x86_64
	v1.19.10	4.18.0-147.5.1.6.h541.eulerosv2r9.x86_64
EulerOS release 2.8 (ARM)	v1.19.10	4.19.36-vhulk1907.1.0.h962.eulerosv2r8.aarch64
	v1.17.17	4.19.36-vhulk1907.1.0.h962.eulerosv2r8.aarch64
	v1.15.11	4.19.36-vhulk1907.1.0.h702.eulerosv2r8.aarch64
EulerOS release 2.5	v1.21	3.10.0-862.14.1.5.h520.eulerosv2r7.x86_64
	v1.19.10	3.10.0-862.14.1.5.h520.eulerosv2r7.x86_64
	v1.19.8	3.10.0-862.14.1.5.h520.eulerosv2r7.x86_64
	v1.17.17	3.10.0-862.14.1.5.h470.eulerosv2r7.x86_64
	v1.17.9	3.10.0-862.14.1.5.h428.eulerosv2r7.x86_64
	v1.15.11	3.10.0-862.14.1.5.h428.eulerosv2r7.x86_64
	v1.15.6-r1	3.10.0-862.14.1.5.h328.eulerosv2r7.x86_64
	v1.13.10-r1	3.10.0-862.14.1.2.h249.eulerosv2r7.x86_64
	v1.13.7-r0	3.10.0-862.14.1.0.h197.eulerosv2r7.x86_64
Ubuntu 18.04 server 64bit	v1.21	4.15.0-136-generic
	v1.19.10	4.15.0-136-generic
	v1.19.8	4.15.0-136-generic
	v1.17.17	4.15.0-136-generic

 Copyright © Huawei Technologies Co., Ltd.

Table 4-2 Mappings between BMS node OS versions and cluster versions

OS Version	Cluster Version	Kernel Information
EulerOS release 2.10 (elastic bare-metal server)	v1.21	4.18.0-147.5.2.1.h579.eulerosv2r10.x86_64
	v1.19.10	4.18.0-147.5.2.1.h579.eulerosv2r10.x86_64
EulerOS release 2.9 (elastic bare-metal server)	v1.21	4.18.0-147.5.1.2.h340.eulerosv2r9.x86_64
	v1.19.10 v1.19.8	4.18.0-147.5.1.2.h340.eulerosv2r9.x86_64 For details about the instance specifications, see **C6s ECS specifications**.
EulerOS release 2.3 (bare metal server)	v1.19.8 v1.17.9 v1.15.11	3.10.0-514.41.4.28.h62.x86_64 For details about the specifications of the server, see **Instance Family**.

Copyright © Huawei Technologies Co., Ltd.

5 Kubernetes Release Notes

5.1 CCE Kubernetes 1.21 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.21.

Version Upgrades

CCE provides full-link component optimization and upgrade for Kubernetes 1.21.

Table 5-1 Core components and descriptions

Cluster Type	Core Component	Version	Upgrade Precautions
CCE cluster	Kubernetes	1.21	None
	Docker	EulerOS/CentOS: 18.9.0 Ubuntu: 18.09.9	None
	OS	CentOS Linux release 7.6	Starting from CCE 1.21, OverlayFS is used for Docker storage on CentOS 7.6 nodes in a CCE cluster.
		EulerOS release 2.5	None
		Ubuntu 18.04 server 64bit	None
CCE Turbo cluster	Kubernetes	1.21	None
	Docker	CentOS: 18.9.0 Ubuntu: 18.09.9	None
	Containerd	1.4.1	None

 Copyright © Huawei Technologies Co., Ltd.

Cluster Type	Core Component	Version	Upgrade Precautions
	OS (ECS)	CentOS Linux release 7.6 Ubuntu 18.04 server 64bit	None
	OS (BMS)	EulerOS release 2.9	None

Resource Changes and Deprecations

Kubernetes 1.21 Release Notes

- CronJob is now in the stable state, and the version number changes to batch/v1.
- The immutable Secret and ConfigMap have now been upgraded to the stable state. A new immutable field is added to these objects to reject changes. The rejection protects clusters from accidental updates that may cause application outages. As these resources are immutable, kubelet does not monitor or poll for changes. This reduces the load of kube-apiserver and improves scalability and performance of your clusters. For more information, see **Immutable ConfigMaps**.
- Graceful node shutdown has been upgraded to the test state. With this update, kubelet can detect that a node is shut down and gracefully terminate the pods on the node. Prior to this update, when the node was shut down, its pod did not follow the expected termination lifecycle, which caused workload problems. Now kubelet can use systemd to detect the systems that are about to be shut down and notify the running pods to terminate them gracefully.
- For a pod with multiple containers, you can use **kubectl.kubernetes.io/** to pre-select containers.
- PodSecurityPolicy is deprecated. For details, see **https://kubernetes.io/blog/2021/04/06/podsecuritypolicy-deprecation-past-present-and-future/**.

Kubernetes 1.20 Release Notes

- The API priority and fairness have reached the test state and are enabled by default. This allows kube-apiserver to classify incoming requests by priority. For more information, see **API Priority and Fairness**.
- The bug of **exec probe timeouts** is fixed. Before this bug is fixed, the exec probe does not consider the **timeoutSeconds** field. Instead, the probe will run indefinitely, even beyond its configured deadline. It will stop until the result is returned. Now, if no value is specified, the default value is used, that is, one second. If the detection time exceeds one second, the application health check may fail. Update the **timeoutSeconds** field for the applications that use this feature during the upgrade. The repair provided by the newly introduced ExecProbeTimeout feature gating enables the cluster operator to restore the previous behavior, but this behavior will be locked and removed in later versions.

 Copyright © Huawei Technologies Co., Ltd.

- RuntimeClass enters the stable state. RuntimeClass provides a mechanism to support multiple runtimes in a cluster and expose information about the container runtime to the control plane.
- kubectl debugging has reached the test state. kubectl debugging provides support for common debugging workflows.
- Dockershim was marked as deprecated in Kubernetes 1.20. Currently, you can continue to use Docker in the cluster. This change is irrelevant to the container image used by clusters. You can still use Docker to build your images. For more information, see **Dockershim Deprecation FAQ**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.21 and other versions, see the following documents:

- **Kubernetes v1.21 Release Notes**
- **Kubernetes v1.20 Release Notes**

5.2 CCE Kubernetes 1.19 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.19.

Version Upgrades

CCE provides full-link component optimization and upgrade for Kubernetes 1.19.

Table 5-2 Core components and descriptions

Cluster Type	Core Component	Version	Upgrade Precautions
CCE cluster	Kubernetes	1.19.10	Some common API versions are deprecated in Kubernetes 1.19. You are advised to upgrade the deprecated API versions listed in this document before upgrading the cluster.
	Docker	EulerOS/CentOS: 18.09.0.100 Ubuntu: 18.09.9	None
	OS	CentOS Linux release 7.6	None
		EulerOS release 2.5	None
		Ubuntu 18.04 server 64bit	None

 Copyright © Huawei Technologies Co., Ltd.

Cluster Type	Core Component	Version	Upgrade Precautions
CCE Turbo cluster	Kubernetes	1.19.10	None
	Docker	CentOS: 18.09.0.100 Ubuntu: 18.09.9	None
	Containerd	1.4.1	None
	OS (ECS)	CentOS Linux release 7.6 Ubuntu 18.04 server 64bit	None
	OS (elastic bare-metal server)	EulerOS release 2.9	None

Resource Changes and Deprecations

Kubernetes 1.19 Release Notes

- vSphere in-tree volumes can be migrated to vSphere CSI drivers. The in-tree vSphere Volume plugin is no longer used and will be deleted in later versions.
- **apiextensions.k8s.io/v1beta1** has been deprecated. You are advised to use **apiextensions.k8s.io/v1**.
- **apiregistration.k8s.io/v1beta1** has been deprecated. You are advised to use **apiregistration.k8s.io/v1**.
- **authentication.k8s.io/v1beta1** and **authorization.k8s.io/v1beta1** have been deprecated and will be removed from Kubernetes 1.22. You are advised to use **authentication.k8s.io/v1** and **authorization.k8s.io/v1**.
- **autoscaling/v2beta1** has been deprecated. You are advised to use **autoscaling/v2beta2**.
- **coordination.k8s.io/v1beta1** has been deprecated in Kubernetes 1.19 and will be removed from version 1.22. You are advised to use **coordination.k8s.io/v1**.
- kube-apiserver: The **componentstatus** API has been deprecated.
- kubeadm: The **kubeadm config view** command has been deprecated and will be deleted in later versions. Use **kubectl get cm -o yaml -n kube-system kubeadm-config** to directly obtain the kubeadm configuration.
- kubeadm: The **kubeadm alpha kubelet config enable-dynamic** command has been deprecated.
- kubeadm: The **--use-api** flag in the **kubeadm alpha certs renew** command has been deprecated.
- Kubernetes no longer supports **hyperkube** image creation.
- The **--export** flag is removed from the **kubectl get** command.
- The alpha feature **ResourceLimitsPriorityFunction** has been deleted.

 Copyright © Huawei Technologies Co., Ltd.

- **storage.k8s.io/v1beta1** has been deprecated. You are advised to use **storage.k8s.io/v1**.

Kubernetes 1.18 Release Notes

- kube-apiserver
  - All resources in the **apps/v1beta1** and **apps/v1beta2** API versions are no longer served. You can use the **apps/v1** API version.
  - DaemonSets, Deployments, and ReplicaSets in the **extensions/v1beta1** API version are no longer served. You can use the **apps/v1** API version.
  - NetworkPolicies in the **extensions/v1beta1** API version are no longer served. You can use the **networking.k8s.io/v1** API version.
  - PodSecurityPolicies in the **extensions/v1beta1** API version are no longer served. Migrate to use the **policy/v1beta1** API version.
- kubelet
  - **--redirect-container-streaming** is not recommended and will be deprecated in v1.20.
  - The resource measurement endpoint **/metrics/resource/v1alpha1** and all measurement standards under this endpoint have been deprecated. Use the measurement standards under the endpoint **/metrics/resource** instead:
    - scrape_error --> scrape_error
    - node_cpu_usage_seconds_total --> node_cpu_usage_seconds
    - node_memory_working_set_bytes --> node_memory_working_set_bytes
    - container_cpu_usage_seconds_total --> container_cpu_usage_seconds
    - container_memory_working_set_bytes --> container_memory_working_set_bytes
    - scrape_error --> scrape_error
  - In future releases, kubelet will no longer create the target directory **CSI NodePublishVolume** according to the CSI specifications. You may need to update the CSI driver accordingly to correctly create and process the target path.
- kube-proxy
  - You are not advised to use the **--healthz-port** and **--metrics-port** flags. Use **--healthz-bind-address** and **--metrics-bind-address** instead.
  - The **EndpointSliceProxying** function option is added to control the use of EndpointSlices in kube-proxy. This function is disabled by default.
- kubeadm
  - The **--kubelet-version** flag of **kubeadm upgrade node** has been deprecated and will be deleted in later versions.
  - The **--use-api** flag in the **kubeadm alpha certs renew** command has been deprecated.
  - kube-dns has been deprecated and will no longer be supported in future versions.

 Copyright © Huawei Technologies Co., Ltd.

  - The ClusterStatus structure in the kubeadm-config ConfigMap has been deprecated and will be deleted in later versions.
- kubectl
  - You are not advised to use boolean and unset values for **--dry-run**. **server|client|none** is used in the new version.
  - **--server-dry-run** has been deprecated for **kubectl apply** and replaced by **--dry-run=server**.
- add-ons

The cluster-monitoring add-on is deleted.

- kube-scheduler
  - The **scheduling_duration_seconds** metric has been deprecated.
  - The **scheduling_algorithm_predicate_evaluation_seconds** and **scheduling_algorithm_priority_evaluation_seconds counters** metrics are no longer used and are replaced by **framework_extension_point_duration_seconds[extension_point="Filter"]** and **framework_extension_point_duration_seconds[extension_point="Score"]**.
  - The scheduler policy AlwaysCheckAllPredictes has been deprecated.
- Other changes
  - The k8s.io/node-api component is no longer updated. Instead, you can use the **RuntimeClass** type in **k8s.io/api** and the generated clients in **k8s.io/client-go**.
  - The **client** label has been deleted from **apiserver_request_total**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.19 and other versions, see the following documents:

- **Kubernetes v1.19.0 Release Notes**
- **Kubernetes v1.18.0 Release Notes**

5.3 CCE Kubernetes 1.17 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.17.

Resource Changes and Deprecations

- All resources in the **apps/v1beta1** and **apps/v1beta2** API versions are no longer served. Migrate to use the **apps/v1** API version.
- DaemonSets, Deployments, and ReplicaSets in the **extensions/v1beta1** API version are no longer served. You can use the **apps/v1** API version.
- NetworkPolicies in the **extensions/v1beta1** API version are no longer served. Migrate to use the **networking.k8s.io/v1** API version.
- PodSecurityPolicies in the **extensions/v1beta1** API version are no longer served. Migrate to use the **policy/v1beta1** API version.

 Copyright © Huawei Technologies Co., Ltd.

- Ingresses in the **extensions/v1beta1** API version will no longer be served in v1.20. Migrate to use the **networking.k8s.io/v1beta1** API version.
- PriorityClass in the **scheduling.k8s.io/v1beta1** and **scheduling.k8s.io/v1alpha1** API versions is no longer served in v1.17. Migrate to use the **scheduling.k8s.io/v1** API version.
- The **event series.state** field in the **events.k8s.io/v1beta1** API version has been deprecated and will be removed from v1.18.
- **CustomResourceDefinition** in the **apiextensions.k8s.io/v1beta1** API version has been deprecated and will no longer be served in v1.19. Use the **apiextensions.k8s.io/v1** API version.
- **MutatingWebhookConfiguration** and **ValidatingWebhookConfiguration** in the **admissionregistration.k8s.io/v1beta1** API version have been deprecated and will no longer be served in v1.19. You can use the **admissionregistration.k8s.io/v1** API version.
- The **rbac.authorization.k8s.io/v1alpha1** and **rbac.authorization.k8s.io/v1beta1** API versions have been deprecated and will no longer be served in v1.20. Use the **rbac.authorization.k8s.io/v1** API version.
- The **CSINode** object of **storage.k8s.io/v1beta1** has been deprecated and will be removed in later versions.

Other Deprecations and Removals

- **OutOfDisk node condition** is removed in favor of **DiskPressure**.
- The **scheduler.alpha.kubernetes.io/critical-pod** annotation is removed in favor of **priorityClassName**.
- **beta.kubernetes.io/os** and **beta.kubernetes.io/arch** have been deprecated in v1.14 and will be removed in v1.18.
- Do not use **--node-labels** to set labels prefixed with **kubernetes.io** and **k8s.io**. The **kubernetes.io/availablezone** label in earlier versions is removed in v1.17 and changed to **failure-domain.beta.kubernetes.io/zone**.
- The **beta.kubernetes.io/instance-type** is deprecated in favor of **node.kubernetes.io/instance-type**.
- Remove the **{kubelet_root_dir}/plugins** path.
- Remove the built-in cluster roles **system:csi-external-provisioner** and **system:csi-external-attacher**.

Container security

 Copyright © Huawei Technologies Co., Ltd.

No.	Feature	Description
1	CCE integrates HUAWEI CLOUD CGS.	CCE integrates Container Guard Service (CGS). CGS scans vulnerabilities and configurations in images, helping enterprises detect the container environment, which cannot be found by the traditional security software. CGS also delivers functions such as process whitelist configuration, read-only file protection, and container escape detection to minimize the security risks for a running container.

References

For more details about the performance comparison and function evolution between Kubernetes 1.17 and other versions, see the following documents:

- **Kubernetes v1.17.0 Release Notes**
- **Kubernetes v1.16.0 Release Notes**

5.4 CCE Kubernetes 1.15 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.15.

To enable interoperability from one Kubernetes installation to the next, you must upgrade your Kubernetes clusters before the maintenance period ends. For details about how to upgrade CCE clusters, see **Cluster Upgrade Changelog**.

Description

CCE provides full-link component optimization and upgrade for Kubernetes v1.15, which includes two minor versions v1.15.11 and v1.15.6-r1.

Resource Changes and Deprecations

- Ingress in the **extensions/v1beta1** API version has been deprecated. It will be no longer served from Kubernetes 1.19. You can use the **networking.k8s.io/v1beta1** API version.
- NetworkPolicy in the **extensions/v1beta1** API version will be officially suspended in 1.16. Migrate to use the **networking.k8s.io/v1** API version.
- PodSecurityPolicy in the **extensions/v1beta1** API version will be officially suspended in 1.16. Migrate to use the **policy/v1beta1** API version.
- DaemonSets, Deployments, and ReplicaSets in the **extensions/v1beta1**, **apps/v1beta1**, and **apps/v1beta2** API versions will not be served in 1.16. You can use the **apps/v1** API version.

 Copyright © Huawei Technologies Co., Ltd.

- PriorityClass is upgraded to **scheduling.k8s.io/v1**, **scheduling.k8s.io/v1beta1**, and **scheduling.k8s.io/v1alpha1**. It will be deprecated in 1.17.
- The **series.state** field in the **events.k8s.io/v1beta1** Event API version has been deprecated and will be removed from 1.18.

Network management

Table 5-3 New features in network management

No.	Feature	Description
1	The path matching of ingress supports the regular expression in the perl format.	The path matching of ingress supports the regular expression in the perl format.
2	CCE supports a SNI certificate with multiple domain names of ELB v3.	If multiple domain names are required for external access and different domain names use independent certificates, you need to enable the SNI function. After SNI is enabled, the client is allowed to submit the requested domain name information when initiating an SSL handshake request.
3	Dedicated load balancers can directly access containers.	The native ENI of the VPC is deeply integrated. The VPC CIDR block is used to allocate container addresses. Access requests can directly reach containers through load balancers, providing optimal networking performance.

Table 5-4 Network management bugfix

No.	Name
1	Fixing the issue that the container domain name resolution is slow.
2	Fix the issue that the load balancing API is slow when load balancing services are being updated and causes an update failure. Fixing the issue that the listener cannot be added after the service type is changed to load balancing.

Kunpeng cluster (Arm)

 Copyright © Huawei Technologies Co., Ltd.

Table 5-5 New Features of Arm

No.	Feature	Description
1	CCE supports creating Kunpeng clusters (Arm-based) of v1.15.	Kunpeng container cluster (Arm-based instruction set) enables containers to run on cloud servers that use Arm architecture and Kunpeng processors. Kunpeng-accelerated cloud servers are easy to deploy and provide comparable scaling and scheduling performance as x86-based cloud servers at only a fraction of what x86-based cloud servers would cost.
2	Some add-ons can be installed in Kunpeng clusters (Arm-based).	The autoscaler, coredns, everest, and dashboard add-ons on the **Add-on Marketplace** page can be installed in Kunpeng clusters (Arm-based) and support the dual-architecture to adapt to the x86/Arm hybrid scenario.

Container security

No.	Feature	Description
1	1	CCE integrates Container Guard Service (CGS). CGS scans vulnerabilities and configurations in images, helping enterprises detect the container environment, which cannot be found by the traditional security software. CGS also delivers functions such as process whitelist configuration, read-only file protection, and container escape detection to minimize the security risks for a running container.

References

Changelog from v1.13 to v1.15

- Changelog from v1.14 to v1.15:

  https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.15.md
- Changelog from v1.13 to v1.14:

 Copyright © Huawei Technologies Co., Ltd.

https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.14.md

5.5 CCE Kubernetes 1.13 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.13.

Table 5-6 v1.13 cluster description

Kubernetes (CCE Enhanced Version)	Description
v1.13.10-r0	**Highlights**: • Arm nodes can be added to a CCE cluster. • The load balancer name is configurable. • Layer-4 load balancing supports health check, and layer-7 load balancing supports health check, allocation policy, and sticky session. • BMS nodes can be created in a CCE cluster (when the tunnel network model is used). • Ascend-accelerated nodes (powered by HiSilicon Ascend 310 AI processors) apply to scenarios such as image recognition, video processing, inference computing, and machine learning. • The docker baseSize is configurable. • Namespace affinity scheduling is supported. • User space can be partitioned in node data disks. • Cluster CPU management policies can be configured. • Nodes in a cluster can be configured across subnets (when the tunnel network mode is used).
v1.13.7-r0	**Highlights**: • Features of Kubernetes v1.13.7 are incorporated. • The network attachment definition is supported.

References

Changelog from v1.11 to v1.13

- Changelog from v1.12 to v1.13:

  https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.13.md
- Changelog from v1.11 to v1.12:

  https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.12.md

 Copyright © Huawei Technologies Co., Ltd.

5.6 CCE Kubernetes 1.11 Release Notes

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.11.

Table 5-7 v1.11 cluster description

Kubernetes (CCE Enhanced Version)	Description
v1.11.7-r2	**Highlights**: • Support for GPU V100. • Support for permission management.
v1.11.7-r0	**Highlights**: • Features of Kubernetes v1.11.7 are incorporated. • Node pools, VMs, and Kunpeng clusters can be created. • BMS nodes can be created in a CCE cluster (when the VPC network model is used), and hybrid deployment of BMSs and VMs is supported. • Support for GPU V100. • AOM notifies users when alarms are generated for container clusters of v1.11. • Access type switching is supported for Services. • Service network segments can be configured. • The number of IP addresses allocated to a node in a cluster can be customized.
v1.11.3-r2	**Highlights**: • Clusters support IPv6 dual stack. • ELB load balancing algorithms: source IP hash and sticky sessions with backend servers.
v1.11.3-r1	**Highlights**: • Perl regular expressions can be used for matching ingress URLs.
v1.11.3-r0	**Highlights**: • Features of Kubernetes v1.11.3 are incorporated. • Master nodes of a cluster can be deployed across multiple AZs. • CCE works with SFS Turbo to provide container storage.

 Copyright © Huawei Technologies Co., Ltd.

References

Changelog from v1.9 to v1.11

- Changelog from v1.10 to v1.11:
  https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.11.md
- Changelog from v1.9 to v1.10:
  https://github.com/kubernetes/kubernetes/blob/master/CHANGELOG/CHANGELOG-1.10.md

5.7 Release Notes of CCE Kubernetes 1.9 and Earlier Versions

HUAWEI CLOUD CCE strictly complies with the community consistency certification. This section describes the updates in CCE Kubernetes 1.9 and earlier versions.

Table 5-8 Description of v1.9 and earlier version clusters

Kubernetes (CCE Enhanced Version)	Description
v1.9.10-r2	**Highlights**: • ELB load balancing algorithms: source IP hash and sticky sessions with backend servers.
v1.9.10-r1	**Highlights**: • CCE works with SFS. • Enhanced ELBs can be automatically created for Services. • Transparent transmission of source IP addresses is supported for enhanced ELBs on the public network. • The maximum number of pods on a node can be set.
v1.9.10-r0	**Highlights**: • Use of ELB/ingress for Kubernetes clusters; new traffic control mechanism. • Features of Kubernetes v1.9.10 are incorporated. • Kubernetes RBAC capability authorization is supported. **Fault rectification**: • Occasional memory leak on nodes, which is caused by kernel cgroup bugs.

Copyright © Huawei Technologies Co., Ltd.

Kubernetes (CCE Enhanced Version)	Description
v1.9.7-r1	**Highlights**: ● The mechanism for reporting PVC and PV events is enhanced. Events can be viewed on the PVC details page. ● CCE works with a third-party authentication system. ● Physical machines that use EulerOS 2.3 can be managed. ● Data disk allocation can be user-defined. ● Elastic Volume Service (EVS) disks are supported for BMSs. ● InfiniBand NICs are supported for BMSs. ● Nodes can be created using the CM-v3 API in BMS scenarios.
v1.9.7-r0	**Highlights**: ● The Docker version of new clusters is upgraded to v17.06. ● DNS cascading is supported. ● Add-ons can be managed. ● Features of Kubernetes v1.9.7 are incorporated. ● The HTTPS of layer-7 ingress is supported. ● StatefulSets can be migrated, scheduled, updated, and upgraded.
v1.9.2-r3	**Highlights**: ● Cluster nodes that use CentOS 7.4 can be created or managed. ● DNAT Services are supported. ● NetworkPolicy APIs are provided. ● Multiple ports can be configured for a Kubernetes Service that uses an ELB. **Fault rectification**: ● Incomplete pod resource recycling caused by a disconnection with kube-apiserver ● Data inaccuracy during auto node scaling

 Copyright © Huawei Technologies Co., Ltd.

Kubernetes (CCE Enhanced Version)	Description
v1.9.2-r2	**Highlights**: ● Custom health check ports can be configured for classic load balancers. ● Performance of classic load balancers is enhanced. ● Kubernetes Service ports can be configured for layer-4 load balancing. **Fault rectification**: ● Bugs in network add-ons, which cause deadlocks in health checks. ● A limited number of HAProxy connections in an HA cluster.
v1.9.2-r1	**Highlights**: ● Features of Kubernetes v1.9.2 are incorporated. ● Cluster nodes support CentOS 7.1. ● GPU nodes are supported and GPU resource use can be restricted. ● The web-terminal add-on is supported.
v1.7.3-r13	**Highlights**: ● The Docker version of new clusters is upgraded to v17.06. ● DNS cascading is supported. ● Add-ons can be managed. ● The mechanism for reporting PVC and PV events is enhanced. ● OBS is supported for BMS clusters.
v1.7.3-r12	**Highlights**: ● Cluster nodes that use CentOS 7.4 can be created or managed. ● DNAT Services are supported. ● NetworkPolicy APIs are provided. ● Multiple ports can be configured for a Kubernetes Service that uses an ELB. **Fault rectification**: ● Incomplete pod resource recycling caused by a disconnection with kube-apiserver ● Data inaccuracy during auto node scaling ● The event aging period prompt is modified. The cluster aging period is 1 hour.

 Copyright © Huawei Technologies Co., Ltd.

Kubernetes (CCE Enhanced Version)	Description
v1.7.3-r11	**Highlights**: ● Custom health check ports can be configured for classic load balancers. ● Performance of classic load balancers is enhanced. ● Kubernetes Service ports can be configured for layer-4 load balancing. ● Namespaces can be deleted. ● EVS disks can be unbound. ● Migration policies can be configured. **Fault rectification**: ● Bugs in network add-ons, which cause deadlocks in health checks. ● A limited number of HAProxy connections in an HA cluster.
v1.7.3-r10	**Highlights**: ● Overlay L2 container networks are supported. ● Cluster nodes can be GPU-accelerated VMs. ● Cluster nodes support CentOS 7.1 and the operating system can be selected. ● Windows clusters support ELBs. ● CCE allows exporting data from SFS. ● BMS clusters support SFS and OBS.
v1.7.3-r9	**Highlights**: ● Cross-AZ deployment is supported for workloads. ● Containers support OBS. ● Layer-7 load balancing is supported. ● Windows clusters support EVS. ● Device mapper in direct-lvm mode is supported in BMS scenarios.
v1.7.3-r8	**Highlights**: ● Auto scaling is supported for cluster nodes. ● Arm nodes can be managed.

 Copyright © Huawei Technologies Co., Ltd.

Kubernetes (CCE Enhanced Version)	Description
v1.7.3-r7	**Highlights**: ● SUSE 12 sp2 nodes can be managed in the container clusters (in the tunnel network mode). ● Docker supports the device mapper in direct-lvm mode. ● Clusters support the dashboard add-on. ● Windows clusters can be created.
v1.7.3-r6	**Highlights**: ● Native EVS APIs are supported for clusters.
v1.7.3-r5	**Highlights**: ● HA clusters can be created. **Fault rectification**: ● Container network disconnection after a node restart
v1.7.3-r4	**Highlights**: ● Cluster performance is enhanced. ● Interconnection with ELB is allowed in BMS scenarios.
v1.7.3-r3	**Highlights**: ● Storage can be attached to kernel-based virtual machines (KVMs).
v1.7.3-r2	**Highlights**: ● SFS is supported to provide container storage. ● Custom logs can be configured for workloads. ● Graceful scaling-in is supported for workloads. **Fault rectification**: ● Expiration of Access Key ID/Secret Access Key (AK/SK) of container storage volumes.
v1.7.3-r1	**Highlights**: ● External domain names can be resolved by kube-dns.
v1.7.3-r0	**Highlights**: ● Features of Kubernetes v1.7.3 are incorporated. ● Elastic Load Balance (ELB) is supported. ● Storage can be attached to Xen VMs. ● EVS is supported to provide container storage.

 Copyright © Huawei Technologies Co., Ltd.

6 Vulnerability Notice

6.1 Vulnerability Fixing Policies

Cluster Vulnerability Fixing Cycle

- High-risk vulnerabilities:
  - CCE fixes vulnerabilities within one month after the Kubernetes community detects them and releases fixing solutions. The fixing policies are the same as those of the community.
  - Emergency vulnerabilities of the operating system are released according to the HUAWEI CLOUD operating system fixing policies and procedure. Generally, a fixing solution is provided within one month, and users need to fix the vulnerabilities by themselves.
- Other vulnerabilities:

  Other vulnerabilities can be fixed through a normal upgrade.

Fixing Statement

To prevent customers from being exposed to unexpected risks, CCE does not provide other information about the vulnerability except the vulnerability background, details, technical analysis, affected functions/versions/scenarios, solutions, and reference information.

In addition, CCE provides the same information for all customers to protect all customers equally. CCE will not notify individual customers in advance.

CCE does not develop or release exploitable intrusive code (or code for verification) using the vulnerabilities in the product.

6.2 Notice on the Vulnerability of Kubernetes subPath Symlink Exchange (CVE-2021-25741)

A security issue was spotted in Kubernetes where a user may be able to create a container with subPath volume mounted to access files and directories outside of the volume, including those on the host file system.

 Copyright © Huawei Technologies Co., Ltd.

Vulnerability Details

Table 6-1 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Resource management flaw	CVE-2021-25741	2021-09-15

Threat Severity

Medium

Impact and Risks

- If a malicious user creates a container with a subPath volume mounted, the user can access files and directories outside the volume, including those on the host file system.
- Clusters for which the cluster administrator has restricted the ability to create hostPath mounts are most severely affected. An attacker can exploit this vulnerability to perform access similar to hostPath without using the hostPath function, thereby bypassing the restriction.
- In the default Kubernetes environment, vulnerability exploitation can be used to mask the abuse of granted privileges.

Trigger Conditions

The VolumeSubpath feature is enabled by default.

Root Cause

When a container uses subPath to mount some files or directories, attackers may use Symlink Exchange to access directories other than the mount directory or files on the host, causing unauthorized operations.

Identification Method

This vulnerability has been fixed for clusters of version 1.19.10 in CN South-Guangzhou, CN North-Beijing4, CN Southwest-Guiyang1, CN East-Shanghai2, CN East-Shanghai1, and CN North-Beijing1. Clusters of v1.19.10 in these regions are not affected.

Log in to the node and run the following command to check BuidDate. If BuidDate is later than August 20, 2021, the vulnerability has been fixed and the system is not affected by the vulnerability.

```
[root@prometheus-38892-wsb84 ~]# kubelet --version=raw
version.Info{Major:"1", Minor:"19+", GitVersion:"v1.19.10-r1.0.0-source-121-gb9675686c54267", GitCommit:"b9675686
c54267276a35579d4921c91be3d226f2", GitTreeState:"clean", BuildDate:"2021-09-03T09:35:06Z", GoVersion:"go1.15.7",
Compiler:"gc", Platform:"linux/amd64"}
```

 Copyright © Huawei Technologies Co., Ltd.

Workarounds and Mitigation Measures

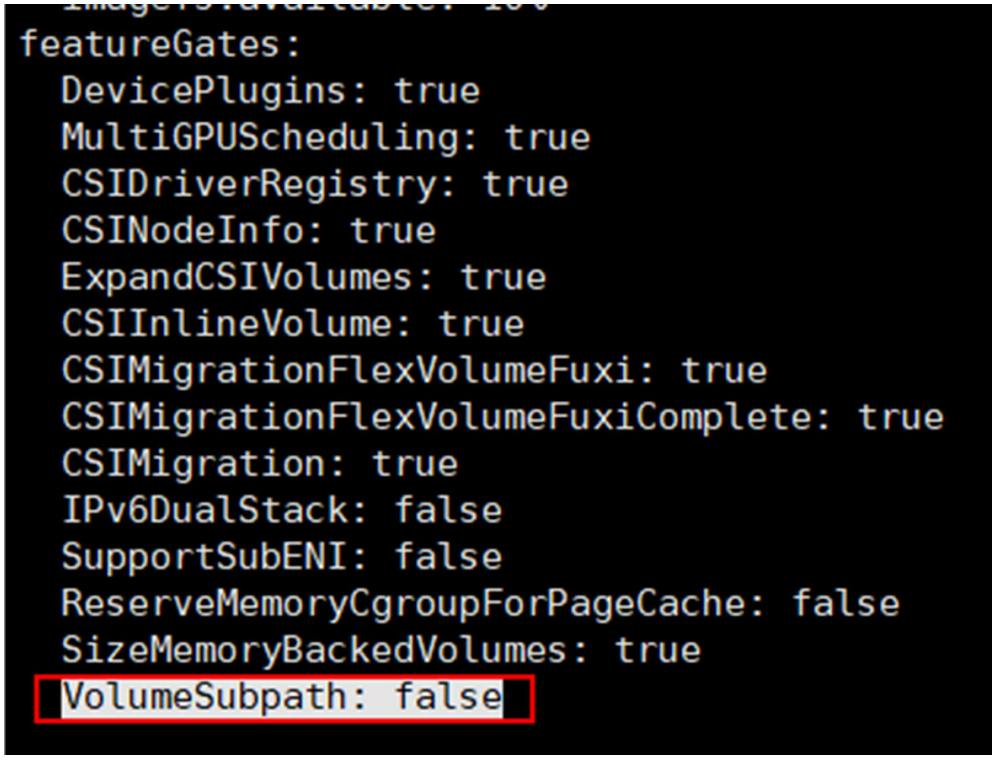
You can disable VolumeSubpath feature gate on kubelet and delete any existing pods that use the subPath function.

Step 1 Log in to each CCE node as user **root**.

Step 2 Modify the kubelet configuration parameter to disable the VolumeSubpath feature.

vi /opt/cloud/cce/kubernetes/kubelet/kubelet_config.yaml

Add the **VolumeSubpath: false** field.



Step 3 Restart kubelet.

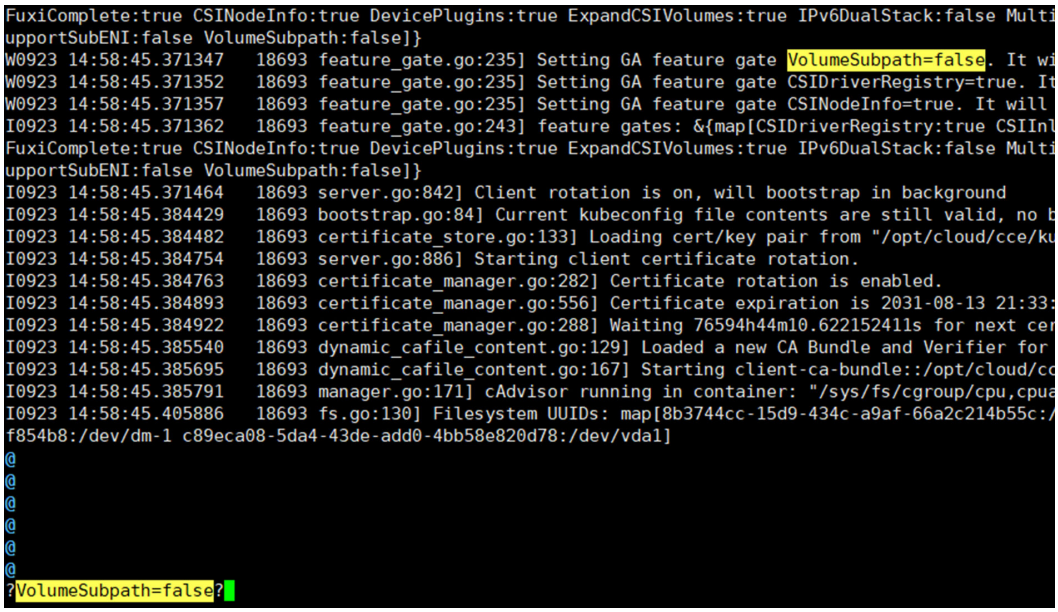
systemctl restart kubelet

Step 4 Ensure that the new kubelet process is started and VolumeSubpath is disabled.

vi /var/paas/sys/log/kubernetes/kubelet.log

Search for **VolumeSubpath=false**. If it can be found, the function is successfully disabled.

 Copyright © Huawei Technologies Co., Ltd.



Step 5 Delete any pod that uses the subPath function.

----End

Enabling or Rolling Back the VolumeSubpath Feature

Step 1 Modify the kubelet configuration file and delete the **VolumeSubpath** field.

vi /opt/cloud/cce/kubernetes/kubelet/kubelet_config.yaml

Step 2 Restart kubelet.

systemctl restart kubelet

Step 3 Check that the new kubelet process is started and the **kubelet.log** file does not contain **VolumeSubpath=false**.

----End

Reference

https://github.com/kubernetes/kubernetes/issues/104980

6.3 Notice on runc Vulnerability that Allows a Container Filesystem Breakout via Directory Traversal (CVE-2021-30465)

runc is vulnerable to a symlink exchange attack whereby an attacker can request a seemingly-innocuous pod configuration that actually results in the host filesystem being bind-mounted into the container (allowing for a container escape). CVE-2021-30465 has been assigned for this vulnerability. Currently, the details and POC of this vulnerability have been disclosed and the risk is high.

 Copyright © Huawei Technologies Co., Ltd.

Vulnerability Details

Table 6-2 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Container escape	CVE-2021-30465	2021-05-31

Threat Severity

Critical

(Severity: low, moderate, important, and critical)

Impact and Risks

An attacker can create a malicious pod, mount the host directory to the container, and exploit a runc symlink and race condition vulnerability, allowing container escape and host filesystem access.

Impact Scope

All versions of runc on CCE cluster nodes are affected by this vulnerability.

Workarounds and Mitigation Measures

- Restrict untrusted users from creating workloads, especially configuring volume mounting parameters.
- Restrict the permissions of the container.
  - Use a non-root user.
  - Use capabilities to restrict the privileges of containers, such as CAP_DAC_OVERRIDE, CAP_DAC_READ_SEARCH, and CAP_SYS_ADMIN.
  - Use seccomp to restrict the attacker's system call permissions on the host kernel. For details, see **Restrict a Container's Syscalls with Seccomp**.

Solution

CCE will release patches for this vulnerability soon. Please follow the official announcement.

References

https://github.com/opencontainers/runc/security/advisories/GHSA-c3xm-pvg7-gh7r

6.4 Notice on the Docker Resource Management Vulnerability (CVE-2021-21285)

Docker is an open-source container engine. It allows you to create a container (a lightweight VM) on the Linux system and achieve automatic installation,

 Copyright © Huawei Technologies Co., Ltd.

deployment, running, and upgrade through configuration files. Docker before versions 19.03.15 and 20.10.3 have a resource management vulnerability that may be exploited by attackers to crash the Docker daemon (dockerd).

Vulnerability Details

Table 6-3 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Resource management flaw	CVE-2021-21285	2021-02-02

Threat Severity

Medium

Impact and Risks

- If an image is maliciously damaged, pulling it may crash the dockerd daemon.
- If you use HUAWEI CLOUD SWR and your images are obtained from SWR, digest verification will be performed on the image uploaded to the image repository, and the Docker daemon will not be affected.
- This vulnerability does not affect the running containers.

Trigger Conditions

1. Manually run **docker pull** on a node in the cluster to pull a maliciously damaged image.
2. kubelet automatically pulls a maliciously damaged image defined in the workload template during workload deployment.

Root Cause

The Docker daemon does not verify the digest at the image layer during image pull.

Identification Method

1. For EulerOS or CentOS nodes, run the following command to check the security package version:
   ```
   rpm -qa |grep docker
   ```
2. For a node running on EulerOS or CentOS, if the Docker version is earlier than **18.09.0.100.51.h10.51.h3-1.h15.eulerosv2r7**, the Docker package will be affected by this vulnerability.
3. For nodes that use other OSs, such as Ubuntu, you can run the **docker version** command to view the Docker version. If the version is earlier than 19.03.15 and 20.10.3, this vulnerability is involved.

 Copyright © Huawei Technologies Co., Ltd.

Workarounds and Mitigation Measures

Do not use images from unknown sources. You are advised to use HUAWEI CLOUD **SoftWare Repository for Container (SWR)**.

Helpful Links

Go to **GitHub** for the upgrade patch provided by the vendor.

6.5 Notice on the NVIDIA GPU Driver Vulnerability (CVE-2021-1056)

Recently, NVIDIA has announced the CVE-2021-1056 vulnerability about the NVIDIA GPU driver. According to the official NVIDIA announcement, if your CCE cluster has a GPU (ECS) node and uses the recommended NVIDIA GPU driver (Tesla 396.37), your NVIDIA driver is not affected by this vulnerability. If you have installed or updated the NVIDIA GPU driver on your node, this vulnerability may be involved.

This document describes the background, impact, and solution of this vulnerability.

Background

Recently, NVIDIA announced that the NVIDIA GPU driver contains a vulnerability related to device-level isolation. When a container is started in the non-privileged mode, an attacker can exploit this vulnerability to create a special character device file in the container to obtain the access permission of all GPU devices on the host machine.

For more information about this vulnerability, see **CVE-2021-1056**.

Vulnerability Details

Table 6-4 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Privilege escalation	CVE-2021-1056	2021-01-07

Querying the NVIDIA Driver Version of a GPU Node

Log in to your GPU node and run the following command to view the driver version.

NOTE

For details about how to log in to a GPU node, see **Linux ECS Login Overview**.

```
[root@XXX36 bin]# ./nvidia-smi
Fri Apr 16 10:28:28 2021
+-----------------------------------------------------------------------------+
| NVIDIA-SMI 460.32.03    Driver Version: 460.32.03    CUDA Version: 11.2     |
|-------------------------------+----------------------+----------------------+
```

 Copyright © Huawei Technologies Co., Ltd.

```
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|                               |                      |               MIG M. |
|===============================+======================+======================|
|   0  Tesla T4            Off  | 00000000:21:01.0 Off |                    0 |
| N/A   68C    P0    31W /  70W |      0MiB / 15109MiB |      0%      Default |
|                               |                      |                  N/A |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Processes:                                                                  |
|  GPU   GI   CI        PID   Type   Process name                  GPU Memory |
|        ID   ID                                                   Usage      |
|=============================================================================|
|  No running processes found                                                 |
+-----------------------------------------------------------------------------+
```

The preceding command output indicates that the GPU driver version of the node is 460.32.03.

Impact Scope

According to the vulnerability notice provided by NVIDIA, the affected NVIDIA GPU driver versions are as follows:

CVE IDs Addressed	Software Product	Operating System	Driver Branch	Affected Versions	Updated Driver Version
CVE-2021-1052 CVE-2021-1053	GeForce	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
	NVIDIA RTX/Quadro, NVS	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
	Tesla	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
CVE-2021-1056	GeForce	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
	NVIDIA RTX/Quadro, NVS	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
			R390	All version prior to 390.141	390.141
	Tesla	Linux	R460	All versions prior to 460.32.03	460.32.03
			R450	All versions prior to 450.102.04	450.102.04
			R418	All versions prior to 418.181.07	418.181.07

For more information, visit the **official NVIDIA website**.

Impact:

- The NVIDIA GPU driver version recommended for CCE clusters and the gpu-beta add-on has not yet been listed in the affected versions disclosed on the NVIDIA official website. If there are official updates, you will be notified and provided possible solutions to fix this vulnerability.
- If you have selected a custom NVIDIA GPU driver version or updated the GPU driver on the node, check whether your GPU driver is affected by this vulnerability by referring to the preceding table.

Solution

According to the official NVIDIA vulnerability notice, upgrade the node to the corresponding driver version to fix the vulnerability.

NOTE

After upgrading your NVIDIA GPU driver, you need to restart the GPU node, which will temporarily affect your services.

- If your node driver version belongs to 418 series, upgrade it to 418.181.07.

 Copyright © Huawei Technologies Co., Ltd.

- If your node driver version belongs to 450 series, upgrade it to 450.102.04.
- If your node driver version belongs to 460 series, upgrade it to 460.32.03.

If you upgrade the GPU driver of a CCE cluster node, upgrade or reinstall the gpu-beta add-on, and enter the download address of the repaired NVIDIA GPU driver when installing the add-on. For details, see **gpu-beta**.

Helpful Links

- NVIDIA security notice: **https://nvidia.custhelp.com/app/answers/detail/a_id/5142**
- Ubuntu security notice: **https://ubuntu.com/security/CVE-2021-1056**
- CVE: **https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-1056**
- NVD: **https://nvd.nist.gov/vuln/detail/CVE-2021-1056**
- CVE PoC: **https://github.com/pokerfaceSad/CVE-2021-1056**
- GPUMounter: **https://github.com/pokerfaceSad/GPUMounter**

6.6 Notice on the Sudo Buffer Vulnerability (CVE-2021-3156)

A security team has officially disclosed the heap-based buffer overflow vulnerability in sudo (CVE-2021-3156), a near-ubiquitous utility available on major Unix-like operating systems. All legacy versions from 1.8.2 to 1.8.31p2 and all stable versions from 1.9.0 to 1.9.5p1 are affected. Any unprivileged user can gain root privileges on a vulnerable host using a default sudo configuration by exploiting this vulnerability.

Sudo is a powerful utility included in most if not all Unix- and Linux-based OSs. It allows users to run programs with the security privileges of another user.

Vulnerability Details

Table 6-5 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Privilege escalation	CVE-2021-3156	2021-01-26

Threat Severity

Critical

Affected Versions

- All legacy versions from 1.8.2 to 1.8.31p2 (default configuration)
- All stable versions from 1.9.0 to 1.9.5p1 (default configuration)

 Copyright © Huawei Technologies Co., Ltd.

Vulnerability Scanning

1. Log in to the system as a non-root user.
2. Run the **sudoedit -s /** command to scan the vulnerability.
   - If the system is vulnerable, it will respond with an error that starts with **sudoedit:**.
   - If the system is patched, it will respond with an error that starts with **usage:**.

Solution

Upgrade sudo to a secure version and perform a self-check before the upgrade.

- For CentOS: sudo 1.9.5p2 or later

  Click **here** to select a sudo version to download.
- For EulerOS:
  - EulerOS 2.2: **https://mirrors.huaweicloud.com/euler/2.2/os/x86_64/updates/sudo-1.8.6p7-23.h9.x86_64.rpm**
  - EulerOS 2.5: **https://mirrors.huaweicloud.com/euler/2.5/os/x86_64/updates/sudo-1.8.19p2-14.h9.eulerosv2r7.x86_64.rpm**
  - EulerOS 2.8: **https://mirrors.huaweicloud.com/euler/2.8/os/aarch64/updates/sudo-1.8.23-3.h18.eulerosv2r8.aarch64.rpm**

Helpful Links

https://blog.qualys.com/vulnerabilities-research/2021/01/26/cve-2021-3156-heap-based-buffer-overflow-in-sudo-baron-samedit

6.7 Notice on the Kubernetes Security Vulnerability (CVE-2020-8554)

CVE-2020-8554 is a man-in-the-middle (MITM) vulnerability that exists in every version of Kubernetes with the most significant impact on multi-tenant clusters. A potential attacker who has the permissions to create and update Services and pods is able to intercept traffic from other pods or nodes in the cluster. By setting the **spec.externalIPs** field of a Service, a potential attacker can intercept the traffic of other pods or nodes that access this externalIP (for example, a well-known public IP address) and forward the traffic to a malicious pod created by the attacker, causing a man-in-the-middle attack. For Services, attackers can initiate MITM attacks by modifying the **status.loadBalancer.ingress.ip** field.

Vulnerability Details

Table 6-6 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Traffic interception	CVE-2020-8554	2020-12-07

 Copyright © Huawei Technologies Co., Ltd.

Affected Versions

Multi-tenant clusters;

Clusters of all Kubernetes versions

Preventive Measures

You are advised to check all Services that use externalIP and loadBalancerIP to determine whether there are vulnerable Services.

This vulnerability is caused by Kubernetes design defects. The final solution to solve this issue once and for all is still under discussion in the open source community. Currently, you can take the following measures to prevent this vulnerability:

- **Restrict the use of externalIP.**
  - Method 1: Use the Admission Webhook container (k8s.gcr.io/multitenancy/externalip-webhook:v1.0.0). The source code and deployment description are released at **https://github.com/kubernetes-sigs/externalip-webhook**.
  - Method 2: Use the open-source OPA Gatekeeper. The example constraint template and constraints are released at **https://github.com/open-policy-agent/gatekeeper-library/tree/master/library/general/externalip**.
- **Restrict the use of loadBalancerIP.**

  The Kubernetes community does not recommend that the cluster administrator assign the patch permission of the Service and status objects to users in the cluster. Therefore, the community does not provide preventive measures for loadBalancerIP. If you need to restrict the use of loadBalancerIP, you can refer to the preventive measures for externalIP.

Helpful Links

https://github.com/kubernetes/kubernetes/issues/97076

6.8 Notice on the Apache containerd Security Vulnerability (CVE-2020-15257)

CVE-2020-15257 is a Docker container escape vulnerability officially released by containerd. containerd is a container runtime underpinning Docker and common Kubernetes configurations. It handles abstractions related to containerization and provides APIs to manage container lifecycles. Attackers, under certain circumstances, can access the containerd-shim API to implement Docker container escape.

 Copyright © Huawei Technologies Co., Ltd.

Vulnerability Details

Table 6-7 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Docker container escape	CVE-2020-15257	2020-11-30

Affected Versions

CCE clusters from v1.9 to v1.17.9.

If the host network is not used and the processes in a container are not run by user **root** (UID is 0), this vulnerability is not involved.

Preventive Measures

You are advised to run containers with least privilege and impose the following restrictions on untrusted containers:

1. Host network cannot be used.
2. Processes in a container cannot be run by user **root**.

Helpful Links

containerd-shim API exposed to host network containers

Please follow the latest news about and official patches for this vulnerability as CCE will do.

6.9 Notice on the Docker Engine Input Verification Vulnerability (CVE-2020-13401)

IPv6 address dynamic allocation can be implemented through Dynamic Host Configuration Protocol (DHCP) or Router Advertisement. This causes the CVE-2020-13401 vulnerability. Router Advertisement allows the router to periodically notify nodes of the network status, including routing records. The client configures the network through Neighbor Discovery Protocol (NDP). This section describes the impacts of the vulnerability.

Vulnerability Details

Table 6-8 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Input validation flaw	CVE-2020-13401	2020-06-01

 Copyright © Huawei Technologies Co., Ltd.

Affected Versions

Nodes on which IPv6 is enabled and the Container Network Interface (CNI) plug-in version is earlier than v0.8.6

Impact

A malicious attacker can tamper with the IPv6 routing records of other containers on the host or the host itself to initiate a man-in-the-middle attack. Even if there was no IPv6 traffic before, if the DNS returns A (IPv4) and AAAA (IPv6) records, many HTTP libraries will try to use the IPv6 record for connections first then fall back to the IPv4 record, giving an opportunity to the attacker to respond. This vulnerability received a CVSS rating of 6.0 (Medium).

Kubernetes is not affected by this vulnerability. However, the CNI plug-in (see **containernetworking / plugins#484** for details) used by Kubernetes is affected. The following kubelet versions contain the affected CNI plug-in:

- kubelet v1.18.0–v1.18.3
- kubelet v1.17.0–v1.17.6
- All kubelet versions earlier than v1.16.11

Preventive Measures

- Change the value of the host kernel parameter **net.ipv6.conf.all.accept_ra** to **0** to reject IPv6 route advertisements.
- Use service containers together with TLS and proper certificate verification to prevent man-in-the-middle spoofing.
- Do not set the CAP_NET_RAW capability in pods to prevent malicious containers from tampering with IPv6 routes.

```
securityContext:
    capabilities:
      drop: ["NET_RAW"]
```

6.10 Notice on the Kubernetes kube-apiserver Input Verification Vulnerability (CVE-2020-8559)

Kubernetes recently disclosed a security vulnerability in the kube-apiserver component. An attacker can intercept certain upgrade requests sent to kubelet of a node and forward the requests to other target nodes using the original access credentials in the requests. This can lead permission escalation. This section describes the affected versions, impact, and preventive measures of the vulnerability.

Vulnerability Details

Table 6-9 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Others	CVE-2020-8559	2020-07-15

 Copyright © Huawei Technologies Co., Ltd.

Affected Versions

All kube-apiserver components from v1.6.0 to the following fixed versions are affected by this vulnerability:

- kube-apiserver v1.18.6
- kube-apiserver v1.17.9
- kube-apiserver v1.16.13

The following application scenarios are also affected by this vulnerability:

- A cluster is shared by multiple tenants and nodes are used as security boundaries for tenant isolation.
- Clusters share certificate authorities (CAs) and authentication credentials.

Impact

The kube-apiserver component allows the proxied backends to send upgrade requests back to the original client. An attacker can intercept certain upgrade requests sent to kubelet of a node and forward the requests to other target nodes using the original access credentials in the requests. This can lead permission escalation. This vulnerability received a CVSS rating of 6.4 (Medium).

If multiple clusters share the same CA and authentication credential, this vulnerability may allow an attacker to attack other clusters. In this case, this vulnerability should be considered **High** severity.

In the cross-cluster scenarios, each CCE cluster uses an independently issued CA and authentication credentials of different clusters are isolated from each other. The cross-cluster scenarios are not affected by this vulnerability.

Preventive Measures

You are advised to take the following security measures to prevent cross-node attacks in a cluster:

- Keep authentication credentials secure.
- Follow the principle of the least privilege when granting permissions to IAM users. Use RBAC policies to restrict the access to the pods/exec, pods/attach, pods/portforward, and proxy resources.

6.11 Notice on the Kubernetes kubelet Resource Management Vulnerability (CVE-2020-8557)

The eviction manager of kubelet does not manage the temporary storage usage of the **/etc/hosts** file mounted to pods. For this vulnerability, if a pod writes a large amount of data to its mounted **/etc/hosts** file to occupy the storage space of a node, a denial of service occurs on the node. This section describes the affected versions, impact, and preventive measures of the vulnerability.

 Copyright © Huawei Technologies Co., Ltd.

Vulnerability Details

Table 6-10 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Resource management flaw	CVE-2020-8557	2020-07-15

Affected Versions

The following kubelet versions are affected by this vulnerability:

- kubelet v1.18.0–v1.18.5
- kubelet v1.17.0–v1.17.8
- All kubelet versions earlier than v1.16.13

Impact

The eviction manager of kubelet does not manage the temporary storage usage of the **/etc/hosts** file mounted to pods. For this vulnerability, if a pod writes a large amount of data to its mounted **/etc/hosts** file to occupy the storage space of a node, a denial of service occurs on the node. This vulnerability received a CVSS rating of 5.5 (Medium).

Clusters running pods with sufficient privileges to write to their own **/etc/hosts** files are affected. The following pods are included:

- Containers running with CAP_DAC_OVERRIDE (which is granted by default)
- Containers running as the **root** user (with **UID** set to **0**)
- Containers running with security context that have the flag **allowPrivilegeEscalation** set to **true** (which is the default behavior when **Privileged Container** is **On** or the pods have the CAP_SYS_ADMIN permission)

Preventive Measures

You are advised to take the following security measures:

- Set the cluster pod security policy or the admission mechanism to force pods to delete the CAP_DAC_OVERRIDE system permission.
```
securityContext:
    capabilities:
      drop: ["DAC_OVERRIDE"]
```
- Set the cluster pod security policy or other admission mechanisms to prevent the **root** user from starting containers, or set the **allowPrivilegeEscalation** parameter to **false**.
```
securityContext:
    allowPrivilegeEscalation: false
```
- Run the following command to monitor the **/etc/hosts** file in containers. If the file size is abnormal, enable the system to report an alarm or take corresponding container isolation measures.

 Copyright © Huawei Technologies Co., Ltd.

```
find /var/lib/kubelet/pods/*/etc-hosts -size +1M
```

6.12 Notice on the Kubernetes kubelet and kube-proxy Authorization Vulnerability (CVE-2020-8558)

Kubernetes officially disclosed that kube-proxy, a core component, has a localhost boundary bypass vulnerability (CVE-2020-8558). With this vulnerability, attackers, through containers in the same LAN, can reach TCP and UDP services bound to 127.0.0.1 running on the node or in the node's network namespace, to obtain interface information. If a service on the port requires no additional authentication, the service is vulnerable to attacks.

Therefore, we kindly remind kube-proxy users to arrange self-check and implement timely security hardening.

For details, visit **https://github.com/kubernetes/kubernetes/issues/92315**.

Vulnerability Details

A security vulnerability was detected in kube-proxy, which allows adjacent hosts (hosts running in the same LAN or layer 2 domain) to access the TCP and UDP services bound to 127.0.0.1 running on the node. For example, if a cluster administrator runs a TCP service on a node that listens on 127.0.0.1:1234, because of this security vulnerability, the TCP service may be accessed by other hosts in the same LAN as the node or by containers running on the same node as the service. If the TCP service on port 1234 did not require additional authentication (because it assumed that only other localhost processes could reach it), the service could be vulnerable to attacks that use this security vulnerability.

Table 6-11 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
Code injection	CVE-2020-8558	2020-07-08

Threat Severity

Important: when the API Server enables insecure port (8080 by default)

Moderate: when the API Server disables the insecure port

NOTE

There are four threat severities: low, moderate, important, and critical.

Affected Versions

The following kube-proxy versions are affected by this vulnerability:

- kube-proxy v1.18.0–v1.18.3
- kube-proxy v1.17.0–v1.17.6

 Copyright © Huawei Technologies Co., Ltd.

- All kube-proxy versions earlier than v1.16.10

The CCE cluster control plane is protected by security groups, and CCE clusters can be accessed from tenant nodes or adjacent nodes through secure ports.

System components on cluster nodes listen on the port mapping to 127.0.0.1. This port is only used for health check and monitoring information query, which will not cause information leakage.

In conclusion, this vulnerability has little impact on CCE clusters.

Impact

If an attacker can configure the host network or runs containers with CAP_NET_RAW, the attacker can obtain the socket information of the service that listens on 127.0.0.1 on the target host. If the target host runs an exposed service that can be accessed from 127.0.0.1 without any further authentication, the service information can be obtained by the attacker. For details, see **Placeholder issue**.

Possible attackers can be:

- Other pods sharing a host in the same switch
- Running container of the local host

Vulnerability Handling

Secure versions have been provided with this vulnerability fixed. If your service version falls into the affected range, upgrade it to a secure version. For details, see the official documentation:

https://kubernetes.io/docs/tasks/administer-cluster/cluster-management/#upgrading-a-cluster

Preventive Measures

You are advised to take the following security measures:

- If your service container needs to use the host network mode and listen on an insecure port, you can manually add an iptables rule on nodes.

  Run the following command to configure an iptables rule in clusters to reject traffic to 127.0.0.1 which does not originate on the nodes.

  ```
  iptables -I INPUT --dst 127.0.0.0/8 ! --src 127.0.0.0/8 -m conntrack ! --ctstate RELATED,ESTABLISHED,DNAT -j DROP
  ```

  If your cluster needs not to enable the API Server insecure port, add the **--insecure-port=0** flag to your Kubernetes API Server command line to disable the insecure port.

- If your cluster runs an untrusted container, run the following command to disable CAP_NET_RAW in the manifest file:

  ```
  securityContext:
      capabilities:
        drop: ["NET_RAW"]
  ```

 Copyright © Huawei Technologies Co., Ltd.

> **CAUTION**
>
> Before fixing vulnerabilities, use **CSBS** to back up your files and conduct a thorough test.

6.13 Notice on Fixing Kubernetes HTTP/2 Vulnerability

Kubernetes recently announced a security advisory against CVE-2019-9512 and CVE-2019-9514 vulnerabilities. We strongly recommend that you select Kubernetes v1.13.10 when creating clusters, because this version will incorporate the security patches. Security patches will be automatically applied to clusters with Kubernetes version v1.13 by Sep 30, 2019. For clusters with Kubernetes version earlier than v1.13, please upgrade them to Kubernetes version v1.13.10.

Vulnerability Details

Recently, the Kubernetes community has released Go-related vulnerabilities: CVE-2019-9512 and CVE-2019-9514. The security issue has been found in the net/http library of the Go language that affects all versions and all components of Kubernetes. These vulnerabilities may cause DoS attacks to all processes that process HTTP or HTTPS Listener.

Go has released versions Go 1.12.9 and Go 1.11.13.

Kubernetes has released v1.13.10 - go1.11.13 using patched versions of Go.

Table 6-12 Vulnerability information

Vulnerability Type	CVE-ID	Discovered
DoS	CVE-2019-9512	2019-08-13
Resource management flaw	CVE-2019-9514	2019-08-13

Impact Scope

Default clusters are protected by VPCs and security groups and therefore not vulnerable.

If cluster APIs are exposed to Internet users, the cluster control plane may be vulnerable.

Solution

- The latest Kubernetes 1.13.10 has been released to fix the vulnerability.
- Security patches will be automatically applied to clusters with Kubernetes version v1.13 by Sep 30, 2019.
- For clusters with Kubernetes version earlier than v1.13, please upgrade them to Kubernetes version v1.13.10.

 Copyright © Huawei Technologies Co., Ltd.

References

Netflix:

https://github.com/Netflix/security-bulletins/blob/master/advisories/third-party/2019-002.md

Bug fixes for Go:

https://golang.org/doc/devel/release.html#go1.12

PRs in the Kubernetes community:

https://github.com/kubernetes/kubernetes/pull/81520

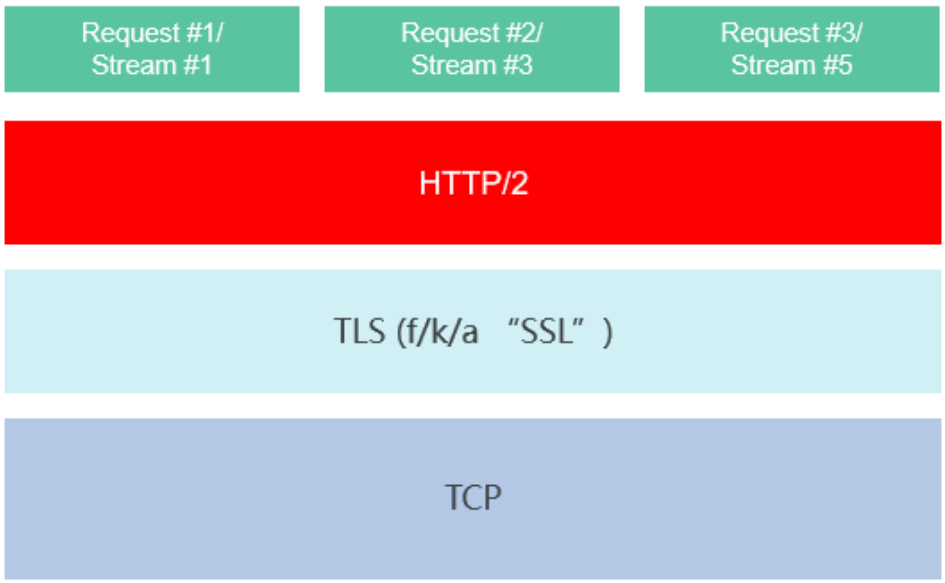
https://github.com/kubernetes/kubernetes/pull/81522

Upgrading a Kubernetes cluster:

https://kubernetes.io/docs/tasks/administer-cluster/cluster-management/#upgrading-a-cluster

Technical Details

Most of these attacks occur at the HTTP/2 layer between request streams and TLS transmission. In fact, many attacks involve zero or one request.



Since the early hypertext transfer protocol, middleware services are request-oriented: logs are separated by requests instead of connections; rate limiting occurs at the request level; throttling is triggered when the number of requests reaches a specified limit.

Few tools can perform logging, rate limiting, and rate modification based on the client behavior at the HTTP/2 layer. Without tools, middleware services may find it even more difficult to detect and block malicious HTTP/2 connections.

The vulnerabilities allow remote attackers to consume excess system resources. Some attacks are very efficient, allowing a single terminal system to cause severe impacts on multiple servers. These impacts include server shutdown, crash of core

 Copyright © Huawei Technologies Co., Ltd.

processes, and suspension. Attacks that are less efficient may cause lead to challenging issues. They only slow down servers and the slowdown may occur intermittently, making it more difficult to detect and prevent attacks.

6.14 Notice on Fixing Linux Kernel SACK Vulnerabilities

The Linux Kernel SACK vulnerabilities have been fixed for HUAWEI CLOUD CCE. This section describes the solution to these vulnerabilities.

Vulnerability Details

On June 18, 2019, Red Hat released a security notice, stating that three security vulnerabilities (CVE-2019-11477, CVE-2019-11478, and CVE-2019-11479) were found on the TCP SACK module of the Linux kernel. These vulnerabilities are related to the maximum segment size (MSS) and TCP selective acknowledgment (SACK) packets. Remote attackers can exploit these vulnerabilities to trigger a denial of service (DoS), resulting in server unavailability or breakdown.

Helpful links:

https://www.suse.com/support/kb/doc/?id=7023928

https://access.redhat.com/security/vulnerabilities/tcpsack

https://www.debian.org/lts/security/2019/dla-1823

https://wiki.ubuntu.com/SecurityTeam/KnowledgeBase/SACKPanic?

https://lists.centos.org/pipermail/centos-announce/2019-June/023332.html

https://github.com/Netflix/security-bulletins/blob/master/advisories/third-party/2019-001.md

Table 6-13 Vulnerability information

Vulnerability Type	CVE-ID	Discovered	Fixed
Input validation flaw	CVE-2019-11477	Jun 17, 2019	Jul 6, 2019
Resource management flaw	CVE-2019-11478	Jun 17, 2019	Jul 6, 2019
Resource management flaw	CVE-2019-11479	Jun 17, 2019	Jul 6, 2019

Impact Scope

Linux kernel version 2.6.29 and later

Solution

Upgrade the OS kernel. For details, see **Upgrading the OS Kernel**.

 Copyright © Huawei Technologies Co., Ltd.

Introduction to TCP SACK

TCP is a connection-oriented protocol. When two parties wish to communicate over a TCP connection, they establish a connection by exchanging certain information such as requesting to initiate (SYN) a connection, initial sequence number, acknowledgement number, maximum segment size (MSS) to use over this connection, and permission to send and process Selective Acknowledgments (SACKs). This connection establishment process is known as 3-way handshake.

TCP sends and receives user data by a unit called Segment. A TCP segment consists of TCP Header, Options and user data. Each TCP segment has a sequence number (SEQ) and an acknowledgement number (ACK).

These SEQ and ACK numbers are used to track which segments are successfully received by the receiver. An ACK number indicates the next segment expected by the receiver.

Example:



In this example, user A sends 1 KB data through 13 segments. Each segment has a header of 20 bytes and contains 100 bytes data in total. On the receiving end, user B receives segments 1, 2, 4, 6, and 8–13. Segments 3, 5, and 7 are lost.

By using ACK numbers, user B will indicate that it is expecting segment 3, which user A reads as none of the segments after 2 were received by user B. Then user A will retransmit all the segments from 3 onwards, even though segments 4, 6, and 8–13 were successfully received by user B. This leads to low performance due to repeated transmissions.

6.15 Notice on Fixing the Docker Command Injection Vulnerability (CVE-2019-5736)

The runC vulnerability CVE-2019-5736 has been fixed in HUAWEI CLOUD CCE. This section describes the impact scope of this vulnerability and the solution.

Vulnerability Details

Runtimes such as Docker and containerd that sit on top of runC have a security vulnerability. This vulnerability allows attackers to obtain the file descriptor handled in runC of the host and overwrite the host runC binary by leveraging the ability to execute a command as root within a new container with a specific image or an existing container that can be attached with docker exec.

Copyright © Huawei Technologies Co., Ltd.

Table 6-14 Vulnerability information

Vulnerability Type	CVE-ID	Discovered	Fixed by HUAWEI CLOUD
Code execution	CVE-2019-5736	2019-02-11	2019-02-12

For details about CVE-2019-5736, see **https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-5736**.

Technical Details

Attacker-controlled images implant malicious functions into a malicious dynamic library such as **libseccomp.so.2** and let execution commands point to **/proc/self/exe**.

When runC is performing a dynamic compilation, it loads the dynamic link library from the attacker-controlled container images and consequently the malicious dynamic library is loaded. Running the **/proc/self/exe** (runC) will execute the malicious program in the malicious dynamic link library. Because the malicious program inherits the file handle opened by runC, runC on the host may be replaced by the file handle.

Then, executing runC-related commands allows for container escape.

Impact Scope

- The runC is the core component of Docker containers and this vulnerability in runC affects most containers. The impact of this vulnerability is often observed in multi-tenant clusters. If multiple users share nodes, any of the users may exploit this vulnerability to control the nodes and attack the entire cluster through penetration.
- **Cloud Container Engine (CCE)**

  The Kubernetes clusters created by CCE are tenant-specific and cannot be shared by multiple tenants. Therefore, this vulnerability has little impact on Kubernetes clusters.

  CCE uses Huawei iSula Docker containers, which are free from this vulnerability because the runC uses static compilation. To ensure security and stability of container service, a non-disruptive upgrade will be performed on the Docker containers running on CCE.
- **Cloud Container Instance (CCI)**

  CCI uses Huawei iSula Kata container engine to ensure that multiple containers on a single node are hypervisor isolated. CCI does not use runC containers and therefore this vulnerability does not affect CCI.

Solution

- CCE

  A non-disruptive upgrade will be performed on the Docker containers running in CCE. Historical Docker versions (such as v1.11.2 and v17.06) will have a

 Copyright © Huawei Technologies Co., Ltd.

security patch installed and need not be upgraded to the latest version 18.09. We will send an upgrade notice prior to the upgrade.

- On-premises Kubernetes or open-source container engine
  - **Upgrade Docker to version 18.09.2**. If the current Docker version is an open-source version earlier than v17.06, the upgrade may interrupt container services. This is because significant changes (including architectural decoupling and restructuring) were made to open-source Docker versions later than 17.06. To minimize the container service downtime, intensively verify the upgrade plan before starting the upgrade and perform a rolling upgrade node by node.
  - **Upgrade only the runC**. For Docker versions 17.06 and earlier, upgrading runC will not interrupt services. Currently, runC has no vulnerability-fixing version. If you want to upgrade runC separately, you can compile it by yourself.
  - The official Docker patch uses the system call provided by Linux kernel v3.17 or later. The patch may not work with certain old versions of Linux kernel. If the patch does not work, upgrade Linux kernel to v3.17 or later. The security patch provided by HUAWEI CLOUD CCE evolves out of the official Docker patch and has been verified to work well on multiple versions of Linux kernel.

6.16 Notice on Fixing the Kubernetes Permission and Access Control Vulnerability (CVE-2018-1002105)

The security vulnerability CVE-2018-1002105 was reported in the Kubernetes community. By forging requests, Kubernetes users can access the backend over established connections through the Kubernetes API server. The HUAWEI CLOUD CCE has securely fixed this vulnerability in a timely manner.

Table 6-15 Vulnerability information

Vulnerability Type	CVE-ID	Discovered	Fixed by HUAWEI CLOUD
Privilege escalation	CVE-2018-1002105	2018-12-05	2018-12-05

For details about the vulnerability, see **https://github.com/kubernetes/kubernetes/issues/71411**.

Affected Versions

- Kubernetes v1.0.x-1.9.x
- Kubernetes v1.10.0-1.10.10 (fixed in v1.10.11)
- Kubernetes v1.11.0-1.11.4 (fixed in v1.11.5)
- Kubernetes v1.12.0-1.12.2 (fixed in v1.12.3)

 Copyright © Huawei Technologies Co., Ltd.

Affected Configurations

Clusters that run aggregated API servers directly accessible from the Kubernetes API server's network;

Clusters visible to attackers, that is, attackers can access the kube-apiserver APIs. If your clusters are deployed on a secure private network, the clusters are not affected.

Clusters that assign **pod exec**/**attach**/**portforward** permissions to users who are not expected to have full access to kubelet APIs

Impact Scenario

If a cluster uses aggregated APIs, the attacker can exploit this vulnerability to send any API request to the aggregated API server, as long as the kube-apiserver is directly connected to the aggregated API server network.

If the access permission of the cluster is granted to anonymous users, anonymous users can also exploit this vulnerability. Unfortunately, Kubernetes allows anonymous access by default, that is, "-- anonymous-auth=true" in the startup parameter settings of kube-apiserver.

Users granted the exec/attach/portforward permission for pods can exploit this vulnerability to make themselves the cluster administrator. Then they can perform destructive operations on any pods.

For more discussion about the vulnerability, see **https://github.com/kubernetes/kubernetes/issues/71411**.

Measures and Suggestions

You do not need to worry about this vulnerability because:

- By default, anonymous access is disabled for clusters created by CCE.
- Clusters created by CCE do not use aggregation APIs.

The HUAWEI CLOUD CCE has completed online patch installation for all Kubernetes clusters of v1.11 and later versions. The Kubernetes community does not provide solutions to fix the vulnerability for clusters of earlier versions. Therefore, the CCE has provided a dedicated patch version for them. Pay attention to the upgrade notices, and install the patch version in time to fix the vulnerability.

NOTE

If you set up Kubernetes clusters without using CCE, you are advised to disable the anonymous access permissions to improve the cluster security.

Upgrade to the vulnerability fixing version provided in the community as soon as possible. When configuring RBAC policies, ensure that the pod exec/attach/portforward permission is granted only to trusted users.

If the Kubernetes version of your clusters is earlier than v1.10, which is not supported by the Kubernetes community, you are advised to add the patch code provided in **https://github.com/kubernetes/kubernetes/pull/71412**.

 Copyright © Huawei Technologies Co., Ltd.

6.17 Notice on Fixing the Kubernetes Dashboard Security Vulnerability (CVE-2018-18264)

The dashboard add-on provided by HUAWEI CLOUD CCE has been upgraded to v1.10.1 and is free of the Kubernetes Dashboard vulnerability CVE-2018-18264. This section describes the version affected by the vulnerability and the solution.

Background

The Kubernetes community has discovered the security vulnerability CVE-2018-18264 in Kubernetes Dashboard v1.10 and earlier versions. This vulnerability allows a user to skip the authentication and query resources that the dashboard service account has access to, such as the private key.

The dashboard add-on provided by HUAWEI CLOUD CCE has been upgraded to v1.10.1 and is free of the Kubernetes Dashboard vulnerability CVE-2018-18264.

Table 6-16 Vulnerability information

Vulnerability Type	CVE-ID	Discovered	Fixed by HUAWEI CLOUD
Access validation error	CVE-2018-18264	2019-01-03	2019-01-05

For details about the security vulnerability CVE-2018-18264, see the following:

- **https://github.com/kubernetes/dashboard/pull/3289**
- **https://github.com/kubernetes/dashboard/pull/3400**
- **https://github.com/kubernetes/dashboard/releases/tag/v1.10.1**

Affected Versions

Kubernetes Dashboard v1.10 or an earlier version (v1.7.0-v1.10.0) that is independently deployed in your Kubernetes clusters, has a login functionality, and uses a custom certificate

Solution

The dashboard add-on provided by HUAWEI CLOUD CCE has been upgraded to v1.10.1 and is free of the Kubernetes Dashboard vulnerability CVE-2018-18264.

 Copyright © Huawei Technologies Co., Ltd.