Cloud Container Engine

Product Bulletin

Issue	01
Date	2024-11-11

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Trademarks and Permissions


and other Huawei trademarks are the property 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 Cloud 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.

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 Latest Notices

CCE has released the latest notices.

No.	Title	Type	Release Date
1	EOM of CentOS	Product Change	2024-10-23
2	Notice of Container Escape Vulnerability in NVIDIA Container Toolkit (CVE-2024-0132)	Vulnerability	2024-10-11
3	Notice of Linux Remote Code Execution Vulnerability in CUPS (CVE-2024-47076, CVE-2024-47175, CVE-2024-47176, and CVE-2024-47177)	Vulnerability	2024-10-11
4	Notice of the NGINX Ingress Controller Vulnerability That Allows Attackers to Bypass Annotation Validation (CVE-2024-7646)	Vulnerability	2024-08-26
5	Notice of Docker Engine Vulnerability That Allows Attackers to Bypass AuthZ (CVE-2024-41110)	Vulnerability	2024-07-31
6	Notice of Linux Kernel Privilege Escalation Vulnerability (CVE-2024-1086)	Vulnerability	2024-07-16
7	Notice of OpenSSH Remote Code Execution Vulnerability (CVE-2024-6387)	Vulnerability	2024-07-03
8	Notice of runC systemd Attribute Injection Vulnerability (CVE-2024-3154)	Vulnerability	2024-04-29
9	Reliability Hardening for Cluster Networks and Storage Functions	Product Change	2024-04-26

No.	Title	Type	Release Date
10	End of Maintenance for Clusters 1.23	Cluster Version	2024-04-25
11	Support for Docker	Product Change	2024-02-19
12	Notice of the Impact of runC Vulnerability (CVE-2024-21626)	Vulnerability	2024-02-01
13	End of Maintenance for Clusters 1.21	Cluster Version	2024-01-22
14	End of Maintenance for Clusters 1.19	Cluster Version	2023-08-02
15	Service Account Token Security Improvement	Product Change	2022-11-24
16	Notice on the Kubernetes Security Vulnerability (CVE-2022-3172)	Vulnerability	2022-09-23
17	Privilege Escalation Vulnerability in Linux Kernel openvswitch Module (CVE-2022-2639)	Vulnerability	2022-09-16

For more historical notices, see **Product Change Notices**, **Cluster Version Release Notes**, and **Vulnerability Notices**.

2 Product Change Notices

2.1 EOM of CentOS

Released: Oct 23, 2024

CentOS has reached its end of maintenance (EOM) date, which means it will no longer receive updates or support. The CentOS public images on CCE are sourced from the official CentOS. As a result, CCE will no longer provide support for CentOS once it is no longer maintained. This section describes the effects of the CentOS EOM and offers solutions to mitigate them.

Background

On December 8, 2020, CentOS officially announced the plan to stop maintaining CentOS Linux and launched CentOS Stream. For more information, see the **CentOS official documentation**.

CentOS 8 ended on December 31, 2021, and CentOS 7 ended on June 30, 2024. CentOS 9 and any future versions are no longer available, and there will be no further software patches or updates provided by the CentOS official. CentOS services may be exposed to risks or even become unavailable, and it will not be possible to restore them in a timely manner.

Impact

According to the official CentOS change plan, there will be the following impacts on users:

- CentOS 7 users cannot receive any software maintenance or support, including bug fixes and feature updates, after June 30, 2024.
- CCE will not remove CentOS 7 public images, and nodes created with CentOS 7 will not be affected, but images will no longer get updated.
- CCE will provide support for CentOS in line with the official CentOS release date.

Solution

- **For a newly created node pool or node**

  When creating a node pool or node, change CentOS to . Huawei Cloud EulerOS is recommended.

- **For an existing node pool**

  Change CentOS to . If you do not need to modify the node configurations, such as the number and type of VPCs and disks, but you want to modify the node OS image and your software is not tightly linked to the original OS, it is recommended that you reset the node to change the OS.

  a. In the navigation pane, choose **Nodes**. On the displayed page, click the **Node Pools** tab.

  b. Select the node pool to be updated, click **Update**, and change CentOS to a supported OS. Huawei Cloud EulerOS is recommended.

  c. In the node list of the target node pool, select a node and choose **More** > **Reset Node** in the **Operation** column. (After a node is reset, the node OS will be reinstalled. Before resetting a node, drain the node to gracefully evict the pods running on the node to other available nodes. Perform this operation during off-peak hours.)

Supported OS	Description	Intended Audience
Huawei Cloud EulerOS	Huawei Cloud EulerOS is a cloud OS built on openEuler. It offers cloud native, high-performing, secure, and easy-to-migrate capabilities. This accelerates service migration to the cloud and promotes application innovation. You can use it to replace OSs such as CentOS and EulerOS.	Individuals or enterprises who prefer free images and want to keep using images from the open-source communities
Ubuntu	Ubuntu is a Linux distribution. Different OSs reflect different usage habits and application compatibilities.	Individuals or enterprises who can handle the cost of switching OSs

NOTICE

To upgrade a node's OS, replace the system disks in batches. It is important to avoid storing any important data in the system disks or back up the data beforehand. The upgrade process will not affect the data disks.

2.2 Reliability Hardening for Cluster Networks and Storage Functions

Released: Apr 26, 2024

If there is a regional fault in the IAM service, it is likely that an authentication exception will occur. This can impact certain functions within a cluster, such as workload storage volume mounting and load balancing interconnection. The newest version of the CCE clusters has been improved and strengthened to handle this type of fault. To maintain stable service performance, we recommend upgrading your account's clusters to the target version as soon as possible.

Trigger Condition

All the following conditions are met:

1. Cluster version ranges:
   - EOS versions: 1.19 and earlier versions
   - Version 1.21 (end of maintenance on April 30, 2024, 00:00 GMT+08:00): v1.21.1-r0 to v1.21.11-r40
   - Version 1.23: v1.23.1-r0 to v1.23.10-r20
   - Version 1.25: v1.25.1-r0 to v1.25.5-r20
   - Version 1.27: v1.27.1-r0 to v1.27.2-r20
2. During a regional fault in the IAM service, the temporary IAM access key in a cluster expires.
3. When creating or updating a workload in a cluster, pod startup is necessary, and functions like storage volume mounting and load balancing are required.

Solution

Keep an eye on the patch version release records and upgrade your clusters to the target version in a timely manner. For clusters that have reached EOS, upgrade them to versions under maintenance.

Target cluster versions:

- Version 1.21 (end of maintenance on April 30, 2024, 00:00 GMT+08:00): v1.21.12-r0 or later
- Version 1.23: v1.23.11-r0 or later
- Version 1.25: v1.25.6-r0 or later
- Version 1.27: v1.27.3-r0 or later
- Version 1.28: v1.28.1.0 or later

2.3 Support for Docker

Released: Feb 19, 2024

Kubernetes community has removed dockershim from clusters 1.24, in which Docker runtime is no longer supported. Considering that some users still use Docker, CCE will continue to support the creation of Docker containers.

The more lightweight and secure containerd runtimes are recommended. You can migrate the runtimes of existing containers to the containerd runtimes. For details, see **Migrating Nodes from Docker to containerd**.

For details about the differences between containerd and Docker, see **Container Engine**.

2.4 Service Account Token Security Improvement

Released: Nov 24, 2022

In Kubernetes clusters v1.21 or later, pods will not automatically mount permanent tokens. You can obtain tokens using **TokenRequest** API and mount them to pods using the projected volume.

Such tokens are valid for a fixed period (one hour by default). Before expiration, kubelet refreshes the tokens to ensure that the pods always use valid tokens. This feature is enabled by default in Kubernetes clusters v1.21 and later. If you use a Kubernetes client of a to-be-outdated version, the certificate reloading may fail.

2.5 Upgrade of Helm v2 to Helm v3

Released: Aug 30, 2022

The open source Helm on which the charts on CCE depend has upgraded from v2 to v3. From now on, CCE will automatically convert Helm v2 releases in your clusters to Helm v3 ones. Some Helm v2 functions have better implementations on Helm v3, but may be incompatible with the original ones. You need to check the differences between Helm v3 and Helm v2 and perform adaptation verification as described in **Differences Between Helm v2 and Helm v3 and Adaptation Solutions**.

If switching to Helm v3 is hard for now, you can manage and deploy Helm v2 releases through the Helm client in the background. For details, see **Deploying an Application Through the Helm v2 Client**. To better run your services deployed in CCE with sufficient O&M support, you are advised to switch to the Helm v3 before **December 30, 2022**.

2.6 Optimized Key Authentication of the everest Add-on

Released: Feb 2, 2021

In everest 1.2.0, key authentication is optimized when OBS buckets are used. After the 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 buckets in the cluster. Otherwise, workloads may not be able to use OBS buckets.

For details about the Everest add-on versions, see **CCE Container Storage (Everest) Release History**.

3 Cluster Version Release Notes

3.1 End of Maintenance for Clusters 1.23

Released: Apr 25, 2024

CCE clusters 1.23 will be end of maintenance (EOM) on September 30, 2024, 00:00 GMT+08:00. After the version EOM, Huawei Cloud does not support the creation of new clusters for CCE clusters 1.23 and earlier. Upgrade your CCE clusters 1.23 and earlier versions to the latest commercial version.

For details about how to upgrade a cluster, see **Upgrade Overview**.

For details about the CCE cluster versions, see **Kubernetes Version Policy**.

3.2 End of Maintenance for Clusters 1.21

Released: Jan 22, 2024

CCE clusters 1.21 will be EOM on April 30, 2024, 00:00 GMT+08:00. After the version EOM, Huawei Cloud does not support the creation of new clusters for CCE clusters 1.21 and earlier. Upgrade your CCE clusters 1.21 and earlier versions to the latest commercial version.

For details about how to upgrade a cluster, see **Upgrade Overview**.

For details about the CCE cluster versions, see **Kubernetes Version Policy**.

3.3 End of Maintenance for Clusters 1.19

Released: Aug 2, 2023

Huawei Cloud CCE clusters 1.19 will be EOM on September 30, 2023, 00:00 GMT +08:00. After the version EOM, Huawei Cloud does not support the creation of new clusters for the CCE clusters 1.19 and earlier versions. Upgrade your CCE clusters to the latest commercial version.

For details about how to upgrade a cluster, see **Upgrade Overview**.

For details about the CCE cluster versions, see **Kubernetes Version Policy**.

4 Vulnerability Notices

4.1 Vulnerability Fixing Policies

Cluster Vulnerability Fixing SLA

- 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 operating system fixing policies and procedure. Generally, a fixing solution is provided within one month, and you need to fix the vulnerabilities by yourself.
- Other vulnerabilities:

  Other vulnerabilities can be fixed through a normal upgrade.

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.

4.2 Notice of Container Escape Vulnerability in NVIDIA Container Toolkit (CVE-2024-0132)

NVIDIA Container Toolkit is an open-source tool package from NVIDIA. It allows you to use NVIDIA GPUs to accelerate computing in a containerized environment.

The toolkit includes a container runtime library and utilities for automatically configuring containers to leverage NVIDIA GPUs.

Description

Table 4-1 Vulnerability details

Type	CVE-ID	Severity	Discovered
Container escape	CVE-2024-0132	Critical	2024-09-26

Impact

In NVIDIA Container Toolkit v1.16.1 and earlier versions, an attacker can run a malicious image, which may result in container escape and enables the attacker to obtain host permissions. Successful exploitation of this vulnerability may enable code execution, DoS, privilege escalation, information leakage, and data tampering.

Identification Method

1. If the cluster does not have the CCE AI Suite (NVIDIA GPU) add-on installed or if the add-on version is earlier than 2.0.0, this vulnerability is not relevant.

   NOTE

   In earlier versions, CCE AI Suite (NVIDIA GPU) add-on are named gpu-beta or gpu-device-plugin.

2. If CCE AI Suite (NVIDIA GPU) version is 2.0.0 or later, you can log in to the target GPU node and run the following command:

   ```
   nvidia-container-runtime --version
   ```

   - If this command is not found, this vulnerability is not relevant.
   - If the version of nvidia-container-runtime is earlier than 1.16.2, this vulnerability is present.

   ```
   root@localhost:~# nvidia-container-runtime --version
   NVIDIA Container Runtime version 1.16.2
   commit: a5a5833c14a15fd9c86bcece85d5ec6621b65652
   spec: 1.2.0

   runc version 1.1.12-0ubuntu2~22.04.1
   spec: 1.0.2-dev
   go: go1.21.1
   libseccomp: 2.5.3
   ```

Mitigation

Do not run an untrusted container image in the cluster before the vulnerability is fixed.

CCE will release a new version of the CCE AI Suite (NVIDIA GPU) add-on to fix this vulnerability. Pay attention to **CCE AI Suite (NVIDIA GPU) Release History**.

Helpful Links

https://docs.nvidia.com/datacenter/cloud-native/container-toolkit/latest/install-guide.html

4.3 Notice of Linux Remote Code Execution Vulnerability in CUPS (CVE-2024-47076, CVE-2024-47175, CVE-2024-47176, and CVE-2024-47177)

Description

Table 4-2 Vulnerability details

Type	CVE-ID	Severity	Discovered
REC	CVE-2024-47076 CVE-2024-47175 CVE-2024-47176 CVE-2024-47177	Critical	2024-09-26

Impact

The vulnerability primarily impacts Unix devices using Common Unix Printing System (CUPS) printing systems. Enabling cups-browsed simultaneously can leave Unix devices vulnerable to attack. It can compromise user device security.

Identification Method

Check whether CUPS-related services are installed:

```
systemctl status cups-browsed
```

The following is an example command output:

```
root@[redacted]:~# systemctl status cups-browsed
Unit cups-browsed.service could not be found.
```

- If the output displays "Unit cups-browsed.service could not be found.", it indicates that CUPS-related services are not present, and the system is not affected by the vulnerability.
- If the value of **Active** in the command output is **inactive (dead)**, it means that the related services have been installed but not enabled. While the vulnerability is present in the system, it does not affect the system. In this scenario, upgrading CUPS is recommended.

- If the value of **Active** in the command output is **active (running)**, it means that the related services are enabled, and the system is vulnerable to this vulnerability. Immediate implementation of workarounds is necessary.

Mitigation

The OS images of Huawei Cloud CCE cluster nodes do not include the CUPS service by default, so the vulnerability does not impact the system.

Helpful Links

https://www.evilsocket.net/2024/09/26/Attacking-UNIX-systems-via-CUPS-Part-I/

4.4 Notice of the NGINX Ingress Controller Vulnerability That Allows Attackers to Bypass Annotation Validation (CVE-2024-7646)

Description

Table 4-3 Vulnerability details

Type	CVE-ID	Severity	Discovered
Validation bypass and command injection	CVE-2024-7646	Critical	2024-08-19

Impact

Attackers with permissions to create ingresses in Kubernetes clusters (in networking.k8s.io or extensions API group) can exploit a vulnerability in ingress-nginx earlier than v1.11.2. This allows them to bypass annotation validation and inject arbitrary commands, potentially gaining access to the credentials of the ingress-nginx controller and sensitive information in a cluster.

Identification Method

This vulnerability affects CCE clusters that have NGINX Ingress Controller add-on versions earlier than 3.0.7. If the version is 3.0.7 or later, the CCE clusters are not at risk. You can check whether a cluster is affected by this vulnerability by doing as follows:

1. Use kubectl to search for pods related to **cceaddon-nginx-ingress**.
   kubectl get po -A | grep cceaddon-nginx-ingress

```
[root@192-168-53-14 paas]# kubectl get po -A|grep cceaddon-nginx-ingress
kube-system   cceaddon-nginx-ingress-controller-67bff65f66-h8xlt        1/1     Running
kube-system   cceaddon-nginx-ingress-default-backend-699d6f4578-nqqqr   1/1     Running
```


If similar information is displayed, the NGINX Ingress Controller add-on has been installed in the cluster.

2. Check the nginx-ingress image version used by the NGINX Ingress Controller add-on.

```
kubectl get deploy cceaddon-nginx-ingress-controller -nkube-system -oyaml|grep -w image
```

```
[root@192-168-53-14 paas]# kubectl get deploy cceaddon-nginx-ingress-controller -nkube-system -oyaml|grep -w image
        image: [redacted]hwofficial/nginx-ingres:v1.11.2
        image: [redacted]hwofficial/nginx-ingres:v1.11.2
```

If the installed NGINX Ingress Controller add-on has an nginx-ingress version earlier than v1.11.2, this vulnerability is present.

Mitigation

CCE will release a new version of the NGINX Ingress Controller add-on that addresses this vulnerability. Keep an eye out for **NGINX Ingress Controller Release History**. Until the issue is resolved, it is best to limit the creation and management of ingresses to trusted users who have been granted the necessary permissions based on the principle of least privilege.

NOTE

To address the vulnerability, the community has released nginx-ingress v1.11.2. However, it is important to note that this version is only compatible with Kubernetes 1.26 or later. If your CCE cluster version is earlier than v1.27, you will need to upgrade the cluster version first.

Helpful Links

Fixed version released by the community: **https://github.com/kubernetes/ingress-nginx/releases/tag/controller-v1.11.2**

4.5 Notice of Docker Engine Vulnerability That Allows Attackers to Bypass AuthZ (CVE-2024-41110)

Docker is an open-source container engine. Docker Engine serves as a portable runtime for containers. Docker's authorization plugins (AuthZ) can be used to manage and limit API requests to the Docker daemon.

Description

Table 4-4 Vulnerability details

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2024-41110	Critical	2024-07-25

Impact

An attacker can exploit this vulnerability using an API request with **Content-Length** set to **0** to bypass the permissions check. This causes the Docker daemon

to forward the request without the body to the AuthZ plugin, potentially allowing unauthorized actions and privilege escalation. Users who do not use the AuthZ plugins or who run Docker Engine of an earlier version are not affected.

CCE uses Huawei-optimized Docker containers and does not enable the AuthZ plugins, so this vulnerability will not be activated.

Identification Method

You can run commands on a node to view the plugins used by Docker.

For a node whose container engine is Docker, run the following command:

```
ps –elf | grep docker
```

The following is an example command output:

```
[root@cce-125-docker-node1 ~]# ps -elf|grep dockerd
4 S root        3978       1  0  80   0 - 326919 -     Jul25 ?        00:40:52 /usr/bin/dockerd --live-re
store --log-driver=json-file --userland-proxy=false --registry-mirror=https://100.79.8.196:20202 --bip=16
9.254.30.1/28 --exec-opt native.umask=normal
```

If **--authorization-plugin** is not configured, the AuthZ plugins are not enabled. In this case, the vulnerability will not affect this node.

Solution

Docker AuthZ plugins are not enabled in CCE clusters, so this vulnerability (CVE-2024-41110) will not affect nodes in CCE clusters. Do not enable the **--authorization-plugin** parameter. CCE is going to fix this vulnerability in the optimized Docker containers.

Helpful Links

Docker AuthZ plugins: **https://www.docker.com/blog/docker-security-advisory-docker-engine-authz-plugin**

4.6 Notice of Linux Kernel Privilege Escalation Vulnerability (CVE-2024-1086)

Description

Table 4-5 Vulnerability details

Type	CVE-ID	Severity	Discovered
Local privilege escalation	CVE-2024-1086	Critical	2024-01-31

Impact

A vulnerability was found in the **netfilter: nf_tables** component in Linux kernels 3.15 to 6.8. This vulnerability can be exploited by a local attacker to gain root

access. The **nft_verdict_init()** function allows positive values to be used as a drop error within the hook verdict. When **NF_DROP** is issued with a drop error similar to **NF_ACCEPT**, the **nf_hook_slow()** function can cause a double free vulnerability.

Although this vulnerability can be used for local privilege escalation, attackers may find it challenging to exploit as it requires initial access to a node.

Identification Method

- Nodes with a kernel version earlier than 3.15 that run CentOS 7.6 are not affected by this vulnerability.
- If EulerOS 2.9, or EulerOS 2.10 is used, you can run the following command to check the kernel version:

  ```
  uname -a
  ```

  ```
  ~]# uname -a
  5.10.0-60.18.0.50.r1083_58.hce2.aarch64 #1 SMP Tue Dec 26 06:14:18 UTC 2023 aarch64 aarch64 aarch64 GNU/Linux
  ```

  If the kernel version falls between 3.15 and 6.8, the system is affected by this vulnerability.

Mitigation

Configure **seccomp** for containerized workloads. The following shows an example:

```
  selfLink: /api/v1/namespaces/default/pods/test002-8778798bc-cjz2p
  uid: 5bc7723a-74f4-4126-b5b5-c07b45199542
spec:
  containers:
  - image: [redacted]
    imagePullPolicy: IfNotPresent
    name: network-multitool
    resources: {}
    securityContext:
      seccompProfile:
        type: RuntimeDefault
    terminationMessagePath: /dev/termination-log
    terminationMessagePolicy: File
    volumeMounts:
```

Related teams and CCE have fixed the vulnerability in EulerOS 2.9, and EulerOS 2.10. Pay attention to **OS Image Version Release Notes**.

Once an OS image with the vulnerability fixed is released, new clusters and nodes will have the vulnerability fixed by default. To fix the vulnerability on existing nodes, you can simply reset them. If the cluster version has reached EOS, you need to upgrade the version first.

Helpful Links

https://nsfocusglobal.com/linux-kernel-privilege-escalation-vulnerability-cve-2024-1086-notice

4.7 Notice of OpenSSH Remote Code Execution Vulnerability (CVE-2024-6387)

OpenSSH is a secure network communication tool based on the SSH protocol. It encrypts all traffic to eliminate eavesdropping, connection hijacking, and other

attacks. In addition, OpenSSH provides a large number of secure tunneling capabilities, multiple authentication methods, and complex configuration options. It is a necessary tool for remote server management and secure data communication.

Description

Table 4-6 Vulnerability details

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2024-6387	Critical	2024-07-01

Impact

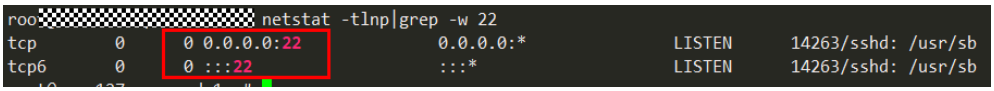
This vulnerability is caused by a signal handler race condition in OpenSSH's server (sshd). An unauthenticated attacker can exploit this vulnerability to execute arbitrary code as **root** on Linux.

Identification Method

- Check the OS and OpenSSH versions of a node:
  - If the OS of a cluster node is EulerOS or CentOS, the OpenSSH is not affected by this vulnerability.
- Run the following command to check whether the SSH port is used:

```
netstat -tlnp|grep -w 22
```

  If the query result shows that the SSH port is listening, it specifies that the SSH access is enabled on the node.



4.8 Notice of runC systemd Attribute Injection Vulnerability (CVE-2024-3154)

Security experts in the industry have revealed a vulnerability in runC related to systemd attribute injection (CVE-024-3154). This vulnerability enables attackers to insert harmful systemd attributes (such as **ExecStartPre**, **ExecStart**, and **ExecReload**) into pod annotations, granting them the ability to execute any action on the host.

Description

Table 4-7 Vulnerability details

Type	CVE-ID	Severity	Discovered
Code execution	CVE-2024-3154	Critical	2024-04-26

Impact

Attackers exploit the runC systemd cgroup functionality to insert harmful systemd attributes (such as **ExecStartPre**, **ExecStart**, and **ExecReload**) into pod annotations, allowing them to execute any action on the host.

CCE clusters are not affected by this vulnerability, because the runC systemd cgroup feature is not in use.

Identification Method

You can run commands on a node to view the cgroup used by the container engine.

- For a node whose container engine is containerd, run the following command:
  ```
  crictl info |grep -i systemdCgroup
  ```
  The following is an example command output:
  ```
  "systemdCgroup": false
  ```
- For a node whose container engine is docker, run the following command:
  ```
  docker info |grep "Cgroup"
  ```
  The following is an example command output:
  ```
  Cgroup Driver: cgroupfs
  ```

Based on the information provided, it appears that the container engine uses **cgroupfs** and not the systemd cgroup. Therefore, the container engine is not affected by this vulnerability.

Solution

The runC systemd cgroup feature is not enabled for Huawei Cloud CCE clusters. Therefore, the clusters are not affected by the vulnerability CVE-2024-3154.

Helpful Links

The runC systemd cgroup feature: **https://github.com/opencontainers/runc/blob/main/docs/systemd.md#auxiliary-properties**

4.9 Notice of the Impact of runC Vulnerability (CVE-2024-21626)

Description

runC is a lightweight tool for running containers. It implements the Open Container Initiative (OCI) specification. runC is the core and basic component of container software such as Docker, containerd, and Kubernetes. Recently, the runC community released the latest version to fix a high-risk container escape vulnerability (**CVE-2024-21626**). Due to an internal file descriptor leak, an attacker could control and set the working directory or the command path of a container process to the path under the parent directory of the file descriptor. This allows the container to read and write any files from and into the node, resulting in a container escape.

Vulnerability Exploitation Conditions

CCE clusters in normal usage are not affected by this vulnerability.

An attacker can exploit this vulnerability only when either of the following conditions is met:

1. The attacker can create or update workloads in a cluster.
2. The image source of a container that runs a workload is untrusted, which enables an attacker to modify the source image.

The following shows the common ways in which exploitation can occur:

- An attacker, with permissions to create or update workloads in a cluster, sets **WORKDIR** of a container process to **/proc/self/fd/***<num>* during workload creation to access the node file system after the container runs.
- An attacker modifies an untrusted source container image of a workload and sets the image's **WORKDIR** to **/proc/self/fd/***<num>* to access the node file system after the container built from this image runs.

Impact

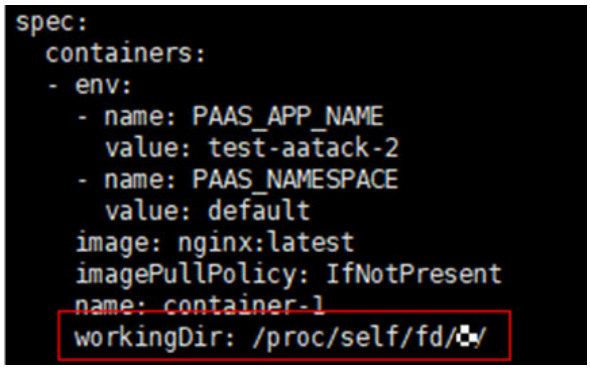
If either of the preceding exploitation conditions is met, a container process may escape to the node, resulting in node information leakage or malicious command execution.

Identification Method

The risks may be present if workload configurations or container images in clusters 1.21.1-r0 to 1.21.12-r2, 1.23.1-r0 to 1.23.11-r2, 1.25.1-r0 to 1.25.6-r2, 1.27.1-r0 to 1.27.3-r10, and 1.28.1-r0 to 1.28.1-r10 have either of the following characteristics:

- **WORKDIR** of a container process in a workload is set to **/proc/self/fd/***<num>*.

Figure 4-1 Configurations of a workload with security risks



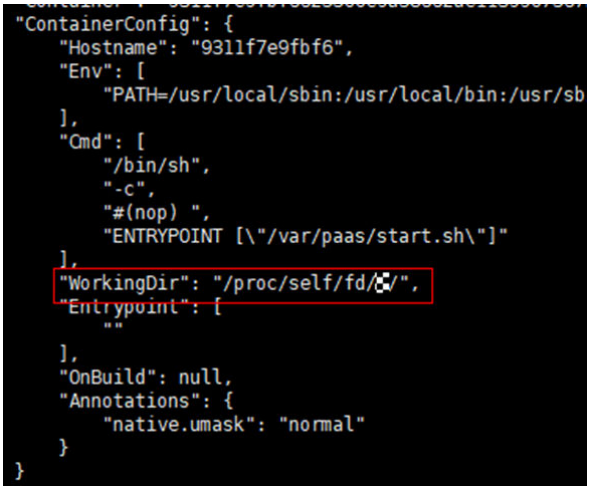
- The default **WORKDIR** or startup command of a container image in a workload contains **/proc/self/fd/***<num>*.

  Run the following command to view the container image metadata:

  – For a Docker container: **docker inspect** *<Image ID>*

  – For a containerd container: **crictl inspecti** *<Image ID>*

Figure 4-2 Configurations of an image with security risks



Solution

Workarounds:

- Set the **WORKDIR** directory of a workload to a fixed one.
- If the **WORKDIR** directory is not configured for a workload, ensure that the container images used by the workload are trusted.

NOTE

Before performing the preceding workarounds, evaluate the impact on services and perform thorough tests.

Rectification method:

We have fixed this vulnerability. Pay attention to update in **Patch Versions** and upgrade your clusters to the fixed version. For clusters that have reached EOS, upgrade them to versions under maintenance.

In clusters 1.21.12-r4, 1.23.11-r4, 1.25.6-r4, 1.27.3-r4, 1.28.1-r4, and later, this vulnerability was fixed.

NOTE

Newly started containers running in clusters which were upgraded to the versions where the vulnerability was fixed are not affected. Existing containers running in clusters of the fixed versions need to be further checked. For details, see **Identification Method**.

- In a running container, if **WORKDIR** in the container process is set to **/proc/self/fd/***<num>*, this vulnerability still exists. To minimize risks, delete the configuration and redeploy the container.
- In a running container, if **WORKDIR** in the used image is set to **/proc/self/fd/***<num>*, this vulnerability still exists. To minimize risks, use a trusted image and redeploy the container.
- In a running container, if **WORKDIR** in the container process is not set to **/proc/self/fd/***<num>*, there is no risk.

4.10 Notice on the Kubernetes Security Vulnerability (CVE-2022-3172)

Description

Kubernetes community detected a security issue in kube-apiserver. This issue allows the aggregated API server to redirect client traffic to any URL, which may cause the client to perform unexpected operations and forward the client's API server credentials to a third party.

Table 4-8 Vulnerability information

Type	CVE-ID	Severity	Discovered
SSRF	CVE-2022-3172	Medium	2022-09-09

Impact

Affected versions:

- kube-apiserver ≤ v1.23.10

CCE clusters of the preceding versions configured with the aggregated API server will be affected, especially for CCE clusters with logical multi-tenancy.

Identification Method

For CCE clusters and CCE Turbo clusters of version 1.23 or earlier, kubectl to connect to the clusters. Run the following command to check whether the aggregated API server is running:

```
kubectl get apiservices.apiregistration.k8s.io -o=jsonpath='{range .items[?(@.spec.service)]}{.metadata.name}{"\n"}{end}'
```

If the returned value is not empty, the aggregated API server exists.

Solution

Upgrades are the currently available solution. The cluster administrator must control permissions to prevent untrusted personnel from deploying and controlling the aggregated API server through the API service interface.

This vulnerability has been fixed in CCE clusters of v1.23.5-r0, v1.21.7-r0, and v1.19.16-r4.

Helpful Links

https://github.com/kubernetes/kubernetes/issues/112513

4.11 Privilege Escalation Vulnerability in Linux Kernel openvswitch Module (CVE-2022-2639)

Description

Details about the privilege escalation vulnerability in the Linux Kernel openvswitch module (CVE-2022-2639) are disclosed. The **reserve_sfa_size()** function in this module has a defect. As a result, a local user can exploit this vulnerability to escalate their privileges on the system. The POC of this vulnerability has been disclosed, and the risk is high.

Table 4-9 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2022-2639	High	2022-09-01

Impact

1. CCE clusters that use the container tunnel network model; node OS images that use EulerOS 2.9;

2. Node OS images that use Ubuntu

Cluster nodes running EulerOS 2.5 and CentOS 7.6 **are not affected by this vulnerability**.

Solution

1. If a process in a container is started by a non-root user, you can configure seccomp, the security computing mode, for the workload. You are advised to use the RuntimeDefault mode or disable system calls such as unshare. For details about the configuration, see **Restrict a Container's Syscalls with seccomp**.
2. Ubuntu images are embedded with the openvswitch kernel module. You can disable the loading of this module to avoid this problem. The procedure is as follows:


```
echo "blacklist openvswitch" >>/etc/modprobe.d/blacklist.conf
```

Then, restart the node for the settings to take effect.

Helpful Links

https://github.com/torvalds/linux/commit/cefa91b2332d7009bc0be5d951d6cbbf349f90f8

4.12 Notice on nginx-ingress Add-On Security Vulnerability (CVE-2021-25748)

Description

The Kubernetes community disclosed an ingress-nginx vulnerability. Users can obtain the credentials used by ingress-controller through the **spec.rules[].http.paths[].path** field of the ingress object. The credentials can be used to obtain the secrets of all namespaces in the cluster. This vulnerability has been assigned CVE-2021-25748.

Table 4-10 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2021-25748	Medium	2022-06-10

Impact

Users who have the permissions to create or update the **spec.rules[].http.paths[].path** field in the ingress can use a newline character to bypass the sanitization of the field to obtain the credentials of the ingress controller, with which the users can access the secrets of all namespaces in the cluster.

Identification Method

For CCE clusters of version 1.23 or earlier:

1. If you install your own nginx-ingress, check whether its image tag is earlier than 1.2.1.

2. If you use the nginx-ingress add-on provided by CCE, check whether the version is earlier than or equal to 2.1.0.

Solution

1. Upgrade ingress-nginx to version 1.2.1.

2. If you are running the "chrooted" ingress-nginx controller introduced in version 1.2.0 (gcr.io/Kubernetes-staging-ingress-nginx/controller-chroot), no action is required.

Helpful Links

1. CVE-2021-25748: **https://github.com/kubernetes/ingress-nginx/issues/8686**
2. Fixed version released by the community: **https://github.com/kubernetes/ingress-nginx/releases/tag/controller-v1.2.1**

4.13 Notice on nginx-ingress Security Vulnerabilities (CVE-2021-25745 and CVE-2021-25746)

Description

The Kubernetes open source community has disclosed two nginx-ingress vulnerabilities:

1. CVE-2021-25745: When creating or updating an ingress, a user who has permissions can use the **spec.rules[].http.paths[].path** field to obtain the credentials of the ingress controller. The credentials can be used to obtain the secrets of all namespaces in the cluster.

2. CVE-2021-25746: When creating or updating an ingress, a user who has permissions can use the **.metadata.annotations** field to obtain the credentials used by the ingress controller. The credentials can be used to obtain the secrets of all namespaces in the cluster.

Table 4-11 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2021-25745	Medium	2022-04-16
Privilege escalation	CVE-2021-25746	Medium	2022-04-16

Impact

These vulnerabilities affect multi-tenant CCE clusters where common users have permissions to create ingresses.

Identification Method

For CCE clusters of version 1.23 or earlier:

1. If you install your own nginx-ingress, check whether its image tag is earlier than 1.2.0.

2. If you use the nginx-ingress add-on provided by CCE, check whether the version is earlier than 2.1.0.

Solution

1. For CVE-2021-25745: Implement an admission policy to restrict the **spec.rules[].http.paths[].path** field in **networking.k8s.io/Ingress** to known safe

characters (see the latest **rules** in the Kubernetes community or use the suggested value in **annotation-value-word-blocklist**).

2. For CVE-2021-25746: Implement an admission policy to restrict the **metadata.annotations** values to known safe characters (see the latest **rules** in the Kubernetes community or use the suggested value in **annotation-value-word-blocklist**).

Helpful Links

1. CVE-2021-25745: **https://github.com/kubernetes/ingress-nginx/issues/8502**
2. CVE-2021-25746: **https://github.com/kubernetes/ingress-nginx/issues/8503**
3. Fixed version released by the community: **https://github.com/kubernetes/ingress-nginx/releases/tag/controller-v1.2.0**

4.14 Notice on the containerd Process Privilege Escalation Vulnerability (CVE-2022-24769)

Description

A security vulnerability has been disclosed in the containerd open source community. When non-root containers were started incorrectly with non-empty inheritable capabilities, attacker may gain access to programs with inheritable file capabilities to elevate those capabilities to the permitted set during execve. This vulnerability has been assigned CVE-2022-24769.

Table 4-12 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2022-24769	Low	2022-03-24

Impact

When a container is created using containerd, Linux process capabilities are included in the inheritable set by default. As a result, when execve() runs in a process in the container by a non-root user, the intersection of the process inheritable capabilities and the file inheritable capabilities is added to the permitted set of the process after execution, causing unexpected privilege escalation. It should be noted that the privilege escalation does not break through the process permission before execve, but only inherits the previous capabilities.

Clusters that use the following containerd versions are affected:

1. CCE Turbo clusters that use the containerd of a version earlier than 1.4.1-98 as the Kubernetes CRI runtime

2. CCE clusters that use the containerd of a version earlier than 1.5.11

Identification Method

View the containerd version by running the **containerd --version** command on a worker node as the root user.

Solution

The entry point of a container can be modified to use the capsh utility to remove inheritable capabilities.

Helpful Links

Community announcement: **https://github.com/containerd/containerd/security/advisories/GHSA-c9cp-9c75-9v8c**

4.15 Notice on CRI-O Container Runtime Engine Arbitrary Code Execution Vulnerability (CVE-2022-0811)

Description

A security vulnerability in CRI-O 1.19 was found by the crowdstrike security team. Attackers can exploit this vulnerability to bypass protection and set arbitrary kernel parameters on the host. As a result, any user with permissions to deploy a pod on a Kubernetes cluster that uses CRI-O runtime can abuse the **kernel.core_pattern** kernel parameter to achieve container escape and arbitrary code execution as root on any node in the cluster.

This vulnerability has been assigned CVE-2022-0811.

Table 4-13 Vulnerability information

Type	CVE-ID	Severity	Discovered
Container escape	CVE-2022-0811	High	2021-03-16

Impact

This vulnerability affects Kubernetes clusters that use CRI-O of versions later than 1.19. The involved patch versions include 1.19.6, 1.20.7, 1.21.6, 1.22.3, 1.23.2, and 1.24.0.

CCE clusters are not affected by this vulnerability because they do not use CRI-O.

Solution

1. For CRI-O v1.19 and v1.20, set **manage_ns_lifecycle** to **false**, and use Open Container Initiative (OCI) runtimes to configure sysctls.
2. Create a PodSecurityPolicy and set all sysctls to **false**.

3. Upgrade the CRI-O version in a timely manner.

Helpful Links

1. Red Hat community vulnerability notice: **https://access.redhat.com/security/cve/cve-2022-0811**
2. cr8escape: New Vulnerability in CRI-O Container Engine Discovered by CrowdStrike: **https://www.crowdstrike.com/blog/cr8escape-new-vulnerability-discovered-in-cri-o-container-engine-cve-2022-0811/**

4.16 Notice on the Container Escape Vulnerability Caused by the Linux Kernel (CVE-2022-0492)

Description

In some scenarios, the release_agent feature of the Linux kernel's cgroup v1 can be used to escape from the container to OS. This vulnerability has been assigned CVE-2022-0492.

Table 4-14 Vulnerability information

Type	CVE-ID	Severity	Discovered
Container escape	CVE-2022-0492	High	2021-02-07

Impact

The Linux kernel does not check whether the process is authorized to configure the **release_agent** file. On an affected node, workload processes are executed as user **root** (or the user with the **CAP_SYS_ADMIN** permission), and **seccomp** is not configured.

CCE clusters are affected by this vulnerability in the following aspects:

1. For x86 nodes, EulerOS 2.5 and CentOS images are not affected by this vulnerability.
2. EulerOS (Arm) whose kernel version is earlier than 4.19.36-vhulk1907.1.0.h962.eulerosv2r8.aarch64 is affected by this vulnerability.
3. EulerOS (x86) whose kernel version is earlier than 4.18.0-147.5.1.6.h541.eulerosv2r9.x86_64 is affected by this vulnerability.
4. Ubuntu nodes whose kernel version is 4.15.0-136-generic or earlier is affected by this vulnerability.

Solution

1. A fix version has been provided for EulerOS 2.9 images. Migrate to the 4.18.0-147.5.1.6.h541.eulerosv2r9.x86_64 nodes as soon as possible.
2. Configure **seccomp** for workloads to restrict unshare system calls. For details, see **Kubernetes documentation**.

3. Restrict the process permissions in a container and minimize the process permissions in the container. For example, use a non-root user to start processes and use the **capability** mechanism to refine the process permissions.

Helpful Links

1. Kernel repair commit: **https://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git/commit/?id=24f6008564183aa120d07c03d9289519c2fe02af**
2. Red Hat community vulnerability notice: **https://access.redhat.com/security/cve/cve-2022-0492**

4.17 Notice on the Non-Security Handling Vulnerability of containerd Image Volumes (CVE-2022-23648)

Description

A vulnerability has been disclosed in the containerd open source community. If an image has malicious attributes, processes in the container may access read-only copies of arbitrary files and directories on the host, causing sensitive information leakage on the host.

Table 4-15 Vulnerability information

Type	CVE-ID	Severity	Discovered
Container escape	CVE-2022-23648	Medium	2022-02-28

Impact

Containers launched with a specially-crafted image configuration could gain access to read-only copies of arbitrary files and directories on the host. This may expose potentially sensitive information.

The impact of this vulnerability is as follows:

1. containerd is used as the Kubernetes CRI runtime, and malicious images from unknown sources are used. This vulnerability is not involved when Docker is used as CRI.
2. The containerd version is earlier than 1.4.1-96.

Identification Method

On the new CCE console, check the value of **Runtime Version** on the **Nodes** page of the CCE Turbo cluster. If the containerd runtime is used and its version is earlier than 1.4.1-96, the vulnerability is involved.

Solution

1. Use trusted images, not third-party images from unknown sources. SoftWare Repository for Container (SWR) is recommended.
2. Migrate pods to nodes running a containerd version later than 1.4.1-96 (already available on the CCE console)

Helpful Links

A patch has been released in the community. For details, see **https://github.com/containerd/containerd/security/advisories/GHSA-crp2-qrr5-8pq7**

4.18 Linux Kernel Integer Overflow Vulnerability (CVE-2022-0185)

Description

William Liu and Jamie Hill-Daniel discovered an integer underflow vulnerability in the Linux kernel, which may lead to out-of-bounds writes. A local attacker can use this vulnerability to cause a denial of service (system crash) or execute arbitrary code. In a container scenario, a user with the CAP_SYS_ADMIN permission can escape from the container to the host machine. The vulnerability POC already exists, but no disclosed exploit code is found.

Table 4-16 Vulnerability information

Type	CVE-ID	Severity	Discovered
Resource management flaw	CVE-2022-0185	High	2022-01-27

Impact

In a container scenario, users have the CAP_SYS_ADMIN permission, and the kernel version is 5.1 or later. In a standard Docker environment, the Docker seccomp filter is used. Therefore, the system is not affected by this vulnerability by default. In the Kubernetes scenario, the seccomp filter is disabled by default. The system is affected by this vulnerability if the kernel and permission conditions are met.

The CCE is not affected by this vulnerability.

Identification Method

Run the **uname -a** command to view the kernel version.

Workarounds and Mitigation Measures

CCE clusters are not affected by this vulnerability. For a Kubernetes cluster, you are advised to:

1. Run containers with the least privilege.
2. Configure **seccomp** based on the configuration method provided by Kubernetes.

Helpful Links

https://blog.aquasec.com/cve-2022-0185-linux-kernel-container-escape-in-kubernetes

https://ubuntu.com/security/CVE-2022-0185

https://access.redhat.com/security/cve/CVE-2022-0185

https://www.openwall.com/lists/oss-security/2022/01/18/7

4.19 Linux Polkit Privilege Escalation Vulnerability (CVE-2021-4034)

Description

A security research team disclosed a privilege escalation vulnerability (CVE-2021-4034, also dubbed PwnKit) in PolKit's pkexec. Unprivileged users can gain full root privileges on a vulnerable host by exploiting this vulnerability in its default configuration. Currently, the POC/EXP of this vulnerability has been disclosed, and the risk is high.

Polkit (formerly PolicyKit) is a component for controlling system-wide privileges in Unix-like operating systems. pkexec is a part of the Polkit framework. It executes commands with elevated permissions and is an alternative to Sudo. If you are a Polkit user, check your Polkit version and implement timely security hardening.

Reference: **https://www.qualys.com/2022/01/25/cve-2021-4034/pwnkit.txt**

Table 4-17 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2021-4034	High	2022-01-28

Impact

Affected versions: all mainstream Linux versions

Secure versions: View the security bulletins of Linux vendors.

Solution

1. Linux vendors, such as Red Hat, Ubuntu, Debian, and SUSE, have released patches to fix this vulnerability. Upgrade your Linux OS to a secure version. If you are unable to update it in a timely manner, you can mitigate the risk by referring to the official suggestions provided by these vendors.

RedHat, Ubuntu: USN-5252-1, USN-5252-2; Debian, SUSE

2. EulerOS has released a patch. You only need to upgrade the polkit package (.rpm).

   The upgrade method is as follows:

   a. yum clean all
   b. yum makecache
   c. yum update polkit
   d. rpm -qa | grep polkit

      Check whether the OS has been upgraded to the corresponding version.

      - EulerOS 2.10: polkit-0.116-6.h4
      - EulerOS 2.9: polkit-0.116-5.h7
      - EulerOS 2.8: polkit-0.115-2.h14
      - EulerOS 2.5: polkit-0.112-14.h15

3. If no patch is available in your system, run the **# chmod 0755 /usr/bin/pkexec** command to delete SUID-bit from pkexec.

Before fixing vulnerabilities, back up your files and conduct a thorough test.

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

Description

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

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.

Table 4-18 Vulnerability information

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

Impact

This vulnerability affects the scenario where VolumeSubpath is enabled (enabled by default). It may have the following impacts:

- 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.

Identification Method

All clusters are affected by this vulnerability.

Log in to the node and run the following command to check BuildDate. If BuildDate 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"}
```

Solution

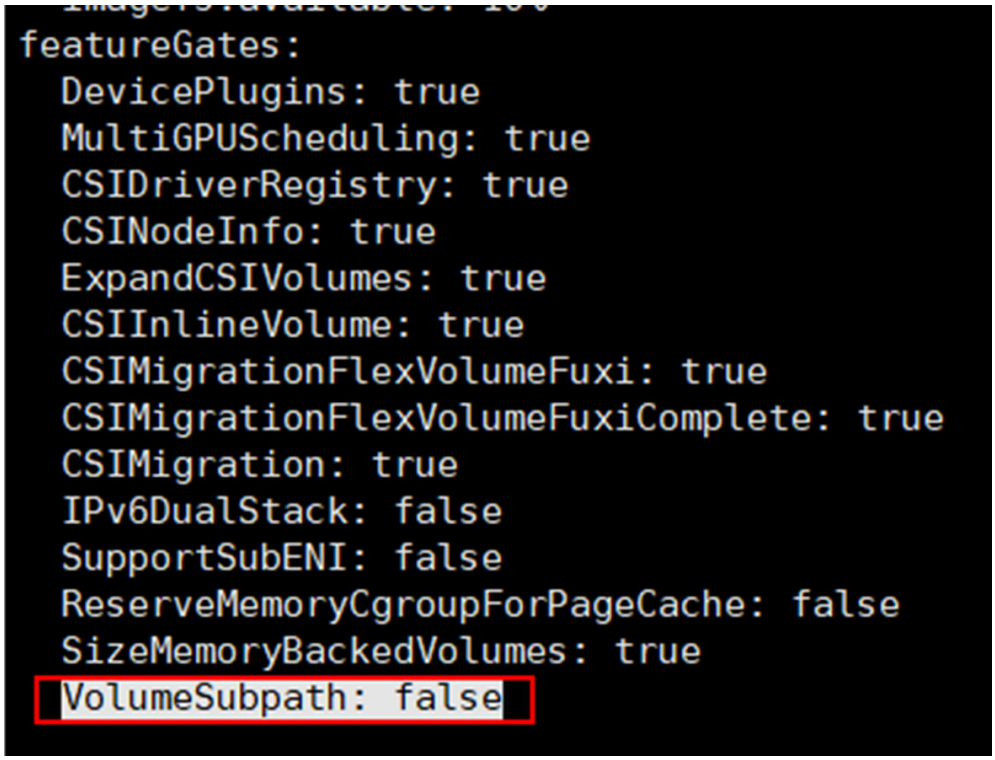
You can disable VolumeSubpath feature gate on kubelet and delete any existing pods that use the subPath function.

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.



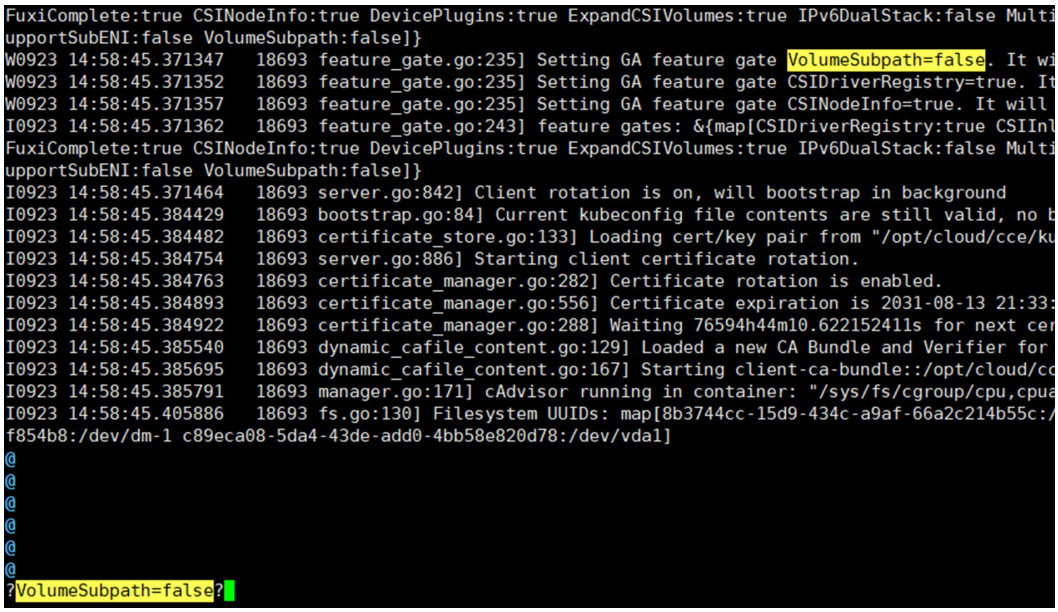
Step 3 Restart kubelet.

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.



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

Helpful Links

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

4.21 Notice of runC Vulnerability That Allows a Container Filesystem Breakout via Directory Traversal (CVE-2021-30465)

Description

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. The details and POC of this vulnerability have been disclosed and the risk is high.

Table 4-19 Vulnerability details

Type	CVE-ID	Severity	Discovered
Container escape	CVE-2021-30465	High	2021-05-31

Impact

This vulnerability is present when the runC version is 1.0.0-rc94 or earlier. 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.

You need to check whether the runC version of a node is 1.0.0-rc94 or earlier to determine whether the node is affected by the vulnerability.

Solution

- 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**.

This vulnerability has been fixed for new nodes in CCE.

You can create a node and set the old node to be unschedulable. After all pods on the old node are scheduled to the new node, delete or reset the old node.

Helpful Links

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

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

Description

Docker is an open source application container engine. It allows you to create containers (lightweight VMs) on Linux and use configuration files for automatic installation, deployment, running, and upgrade of applications. Docker versions earlier than 19.03.15 and 20.10.3 have a resource management error that may be exploited by attackers to crash the Docker daemon (dockerd).

Table 4-20 Vulnerability information

Type	CVE-ID	Severity	Discovered
Resource management flaw	CVE-2021-21285	Medium	2021-02-02

Impact

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

This vulnerability may be triggered in the following scenarios:

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

The impact of this vulnerability is as follows:

- If an image is maliciously damaged, pulling it may crash the docker 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.

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.

Solution

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

Helpful Links

The vendors have released an upgrade patch to fix the vulnerability. To obtain the patch, visit **https://github.com/moby/moby/commit/8d3179546e79065adefa67cc697c09d0ab137d30**

4.23 Notice of NVIDIA GPU Driver Vulnerability (CVE-2021-1056)

Description

NVIDIA detected a vulnerability (assigned CVE-2021-1056), which exists in the NVIDIA GPU drivers and is related to device 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**.

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 or updated the NVIDIA GPU driver on your node, this vulnerability may be involved.

Table 4-21 Vulnerability details

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2021-1056	Medium	2021-01-07

Impact

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, see the **official NVIDIA website**.

Note:

- 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.

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.

```
[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     |
|-------------------------------+----------------------+----------------------+
| 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.

Solution

Upgrade the node to the target driver version based on the **Impact**.

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.
- 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.

Helpful Links

- NVIDIA security bulletin: **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**

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

Description

A security team 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.

Table 4-22 Vulnerability information

Type	CVE-ID	Severity	Discovered
Privilege escalation	CVE-2021-3156	High	2021-01-26

Impact

- 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)

Identification Method

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: upgrade to sudo 1.9.5p2 or later

  For more versions of sudo, see **https://www.sudo.ws/download.html**.
- For EulerOS: obtain the sudo patch package
  - 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

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

Description

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.

Table 4-23 Vulnerability information

Type	CVE-ID	Severity	Discovered
Traffic interception	CVE-2020-8554	Medium	2020-12-07

Impact

Multi-tenant clusters;

Clusters of all Kubernetes versions

Solution

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

This bug is caused by a design defect in Kubernetes. You can take precautionary measures as follows:

- **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 permissions 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

4.26 Notice of Apache containerd Security Vulnerability (CVE-2020-15257)

Description

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.

Table 4-24 Vulnerability details

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

Impact

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.

Solution

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

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

Description

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.

Table 4-25 Vulnerability information

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

Impact

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

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 **https://github.com/containernetworking/plugins/pull/484** for details) used by Kubernetes is affected. The following kubelet versions involve the affected CNI plug-in:

- kubelet v1.18.0 to v1.18.3
- kubelet v1.17.0 to v1.17.6
- kubelet < v1.16.11

Solution

- 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"]
```

4.28 Notice of Kubernetes kube-apiserver Input Verification Vulnerability (CVE-2020-8559)

Description

Kubernetes disclosed a security vulnerability in kube-apiserver. 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.

Table 4-26 Vulnerability details

Type	CVE-ID	Severity	Discovered
Others	CVE-2020-8559	Medium	2020-07-15

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.

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.

Solution

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.

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

Description

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.

Table 4-27 Vulnerability information

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

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**),or 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).

The following kubelet versions are affected by this vulnerability:

- kubelet v1.18.0 to v1.18.5
- kubelet v1.17.0 to v1.17.8
- kubelet < v1.16.13

Solution

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

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

Description

Kubernetes officially released a security notice that the core component kube-proxy has a host 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. 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.

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

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

Table 4-28 Vulnerability information

Type	CVE-ID	Severity	Discovered
Code injection	CVE-2020-8558	High	2020-07-08

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

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

- kube-proxy v1.18.0 to v1.18.3
- kube-proxy v1.17.0 to v1.17.6
- kube-proxy < 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.

Solution

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:

- kubelet/kube-proxy v1.18.4+
- kubelet/kube-proxy v1.17.7+
- kubelet/kube-proxy v1.16.11+

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"]
  ```

CAUTION

Before fixing vulnerabilities, back up your files and conduct a thorough test.

4.31 Notice on Fixing Kubernetes HTTP/2 Vulnerability

Description

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.

CCE has released the latest Kubernetes clusters of v1.13.10 to fix the vulnerability. For Kubernetes clusters of v1.13, a patch will be provided at the end of September 2019 to fix the bug. For Kubernetes clusters earlier than v1.13, upgrade them to v1.13.10.

Table 4-29 Vulnerability information

Type	CVE-ID	Severity	Discovered
DoS attack	CVE-2019-9512	High	2019-08-13
Resource management flaw	CVE-2019-9514	High	2019-08-13

Impact

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 v1.13.10 has been released to fix the vulnerability.
- If the Kubernetes cluster is earlier than v1.13, upgrade the cluster version.

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

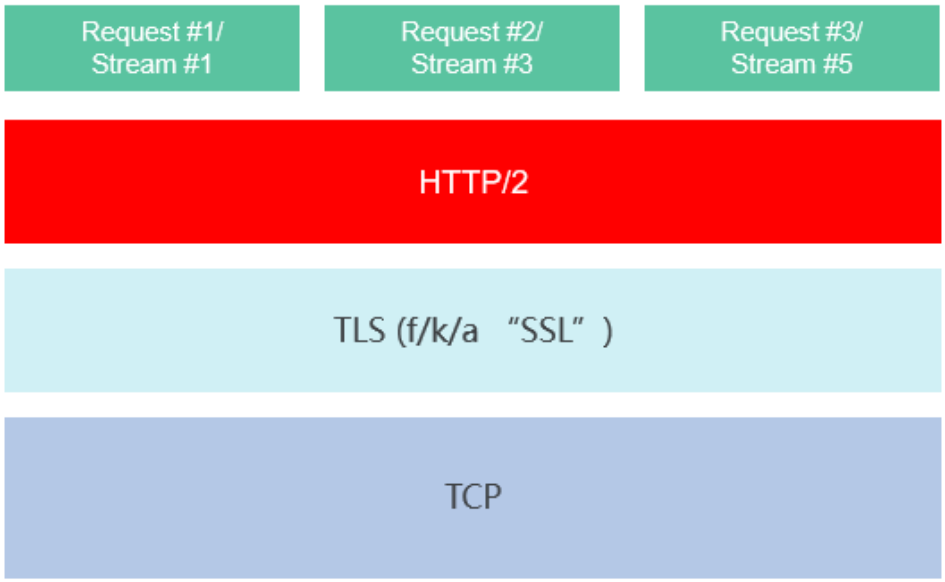
PRs in Kubernetes community:

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

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

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 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.

4.32 Notice on Fixing Linux Kernel SACK Vulnerabilities

Description

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.

The Linux Kernel SACK vulnerabilities have been fixed for Huawei Cloud CCE using the following solution.

References:

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 4-30 Vulnerability information

Type	CVE-ID	Severity	Discovered	Fixed by Huawei Cloud
Input validation flaw	CVE-2019-11477	High	2019-06-17	2019-07-06
Resource management flaw	CVE-2019-11478	High	2019-06-17	2019-07-06
Resource management flaw	CVE-2019-11479	High	2019-06-17	2019-07-06

Impact

Linux kernel version 2.6.29 and later

Solution

These issues have been resolved in stable kernel versions of 4.4.182, 4.9.182, 4.14.127, 4.19.52, and 5.1.11. You can upgrade the nodes in rolling mode.

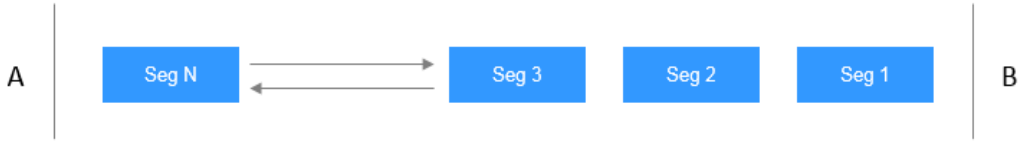
Introduction to TCP SACKs

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 permissions 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.

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

Description

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.

The runC vulnerability CVE-2019-5736 has been fixed in Huawei Cloud CCE.

Table 4-31 Vulnerability information

Type	CVE-ID	Severity	Discovered	Fixed by Huawei Cloud
Code execution	CVE-2019-5736	High	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**.

Impact

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.

The impact of this vulnerability is as follows:

- 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.
- **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 Docker containers, which are free from this vulnerability because the runC uses static compilation.
- **Cloud Container Instance (CCI)**

  CCI uses Huawei 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

  The runC vulnerability CVE-2019-5736 has been fixed in Huawei Cloud CCE.
- 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.

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

Description

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 4-32 Vulnerability information

Type	CVE-ID	Severity	Discovered	Fixed by Huawei Cloud
Privilege escalation	CVE-2018-1002105	Critical	2018-12-05	2018-12-05

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

Impact

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. The access permission of anonymous users is not prohibited in Kubernetes clusters, where the kube-apiserver startup parameter **anonymous-auth** is set to **true**. Users are granted the exec/attach/portforward permission of pods, and can also exploit this vulnerability to upgrade themselves to the cluster administrator to damage pods.

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

The impact of this vulnerability is as follows:

- 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

The affected cluster versions are as follows:

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

Solution

You do not need to worry about this vulnerability when using Huawei Cloud CCE. The reasons are as follows:

- 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**.

4.35 Notice of Fixing the Kubernetes Dashboard Security Vulnerability (CVE-2018-18264)

Description

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 obtain 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 4-33 Vulnerability details

Type	CVE-ID	Severity	Discovered	Fixed by Huawei Cloud
Access validation error	CVE-2018-18264	High	2019-01-03	2019-01-05

For details about 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

Impact

Kubernetes Dashboard v1.10 or an earlier version (v1.7.0 to 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.

5 Product Release Notes

5.1 Cluster Versions

5.1.1 Kubernetes Version Policy

CCE provides highly scalable, high-performance, enterprise-class Kubernetes clusters. As the Kubernetes community periodically releases Kubernetes versions, CCE will release cluster Open Beta Test (OBT) and commercially used versions accordingly. This section describes the Kubernetes version policy of CCE clusters.

Lifecycle of CCE Cluster Versions

Kubernetes Version	Status	Community Release In	OBT of CCE Clusters	Commercial Use of CCE Clusters	EOS of CCE Clusters
v1.29	OBT	November 2023	April 2024	June 2024	June 2026
v1.28	In commercial use[a]	August 2023	December 2023	February 2024	February 2026
v1.27	In commercial use[a]	April 2023	August 2023	October 2023	October 2025
v1.25	In commercial use[b]	August 2022	November 2022	March 2023	March 2025
v1.23	In commercial use[b]	December 2021	April 2022	September 2022	September 2024

Kubernetes Version	Status	Community Release In	OBT of CCE Clusters	Commercial Use of CCE Clusters	EOS of CCE Clusters
v1.21	End of service (EOS)	April 2021	December 2021	April 2022	April 2024
v1.19	EOS	August 2020	December 2020	March 2021	September 2023
v1.17	EOS	December 2019	/	July 2020	January 2023
v1.15	EOS	June 2019	/	December 2019	September 2022
v1.13	EOS	December 2018	/	June 2019	March 2022
v1.11	EOS	August 2018	/	October 2018	March 2021
v1.9	EOS	December 2017	/	March 2018	December 2020

NOTE

The CCE console supports clusters of the latest two commercially used versions:

- a: Clusters created using the console or APIs
- b: Clusters created only using APIs

Phases of CCE Cluster Versions

- OBT: You can experience the latest features of this cluster version. However, the stability of clusters of this version has not been completely verified, and the Service Level Agreement (SLA) of CCE is not valid for such clusters.
- In commercial use: The cluster version has been fully verified and is stable and reliable. You can use clusters of this version in the production environment, and the CCE SLA is valid for such clusters.
- EOS: After the cluster version EOS, CCE does not support the creation of new clusters or provide technical support including new feature updates, vulnerability or issue fixes, new patches, work order guidance, and online checks for the EOS cluster version. The CCE SLA is not valid for such clusters.

CCE Cluster Versions

CCE clusters are updated according to the versions available in the Kubernetes community. This means that a CCE cluster version is made up of both the Kubernetes community version number and the CCE patch version number. The CCE cluster version is in the format of **v***X.Y.Z*-**r***N*, such as **v1.28.2-r0**.

Cluster Upgrade

Periodically upgrade CCE clusters for better user experience. Using an EOS version, you cannot obtain technical support and CCE SLA assurance. Upgrade CCE clusters in a timely manner.

On the CCE console, you can easily upgrade clusters in a visualized manner, improving the stability and reliability of clusters.

5.1.2 Kubernetes Version Release Notes

5.1.2.1 Kubernetes 1.29 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. CCE allows you to create Kubernetes clusters 1.29. This section describes the changes made in Kubernetes 1.29.

Indexes

New and Enhanced Features

- The load balancer IP mode for Services is in the alpha state.

  The load balancer IP mode for Services is promoted to alpha. Kubernetes 1.29 adds the **ipMode** field to the Services' **status** field for configuring traffic forwarding from Services within a cluster to pods. If **ipMode** is set to **VIP**, traffic delivered to a node with the destination set to the load balancer's IP address and port will be redirected to the target node by kube-proxy. If it is set to **Proxy**, traffic delivered to a node will be sent to the load balancer and then redirected to the target node by the load balancer. This feature addresses the issue of load balancer functions being missed due to traffic bypassing it. For details, see **Load Balancer IP Mode for Services**.

- The nftables proxy mode is in the alpha state.

  The nftables proxy mode is promoted to alpha. This feature allows kube-proxy to run in nftables mode. In this mode, kube-proxy configures packet forwarding rules using the nftables API of the kernel netfilter subsystem. For details, see **nftables proxy mode**.

- Garbage collection for unused container images is in the alpha state.

  The garbage collection for unused container images is promoted to alpha. This feature allows you to specify the maximum time a local image can be unused for each node. If the time expires, the image will be garbage collected. To configure the setting, specify the **ImageMaximumGCAge** field for kubelet. For details, see **Garbage collection for unused container images**.

- **PodLifecycleSleepAction** is in the alpha state.

  PodLifecycleSleepAction is promoted to alpha. This feature introduces the sleep hook to the container lifecycle hooks. You can pause a container for a

specified duration after it starts or before it is stopped by enabling this feature. For details, see **Hook handler implementations**.

- **KubeletSeparateDiskGC** is in the alpha state.

  KubeletSeparateDiskGC is promoted to alpha. With this feature enabled, container images and containers can be garbage collected even if they are on separate file systems.

- **matchLabelKeys** and **mismatchLabelKeys** are in the alpha state.

  matchLabelKeys and **mismatchLabelKeys** are promoted to alpha. With these features enabled, the **matchLabelKeys** and **mismatchLabelKeys** fields are added to the pod affinity and anti-affinity configurations. This allows for configurations of more affinity and anti-affinity policies between pods. For details, see **matchLabelKeys and mismatchLabelKeys**.

- The **clusterTrustBundle** projected volumes are in the alpha state.

  The **clusterTrustBundle** projected volumes are promoted to alpha. With this feature enabled, the **clusterTrustBundle** projected volume source injects the contents of one or more ClusterTrustBundle objects as an automatically-updating file. For details, see **clusterTrustBundle projected volumes**.

- Pulling images based on runtime classes of pods is in the alpha state.

  Pulling images based on runtime classes is promoted to alpha. With this feature enabled, the kubelet references container images by a tuple (of image name or runtime handler) rather than just the image name or digest. Your container runtime may adapt its behavior based on the selected runtime handler. Pulling images based on runtime classes will be helpful for VM based containers. For details, see **Image pull per runtime class**.

- The **PodReadyToStartContainers** condition is in the beta state.

  The **PodReadyToStartContainers** condition is promoted to beta. Kubernetes 1.29 introduces the **PodReadyToStartContainers** condition to the pods' **status** field. If it is set to **true**, the sandbox of a pod is ready and service containers can be created. This feature enables cluster administrators to gain a clearer and more comprehensive view of pod sandbox creation completion and container readiness. This enhanced visibility allows them to make better-informed decisions and troubleshoot issues more effectively. For details, see **PodReadyToStartContainers Condition Moves to Beta**.

- Two job-related features are in the beta state.
  - Pod replacement policy (beta)

    The pod replacement policy feature moves to beta. This feature ensures that a pod is replaced only when it reaches the **Failed** state, which means that **status.phase** becomes **Failed**. It does not recreate a pod when the deletion timestamp is not empty and the pod is still being deleted. This prevents two pods from occupying index and node resources concurrently.
  - Backoff limit per index (beta)

    The backoff limit per index moves to beta. By default, pod failures for indexed jobs are counted and restricted by the global limit of retries, specified by **.spec.backoffLimit**. This means that if there is a consistently failing index in a job, pods specified by the job will be restarted repeatedly until pod failures exhaust the limit. Once the limit is reached, the job is marked failed and pods for other indexes in the job may never be even started. The feature allows you to complete execution of all

indexes, despite some indexes failing, and to better use the compute resources by avoiding unnecessary retries of consistently failing indexes.

- Native sidecar containers are in the beta state.

  Native sidecar containers are promoted to beta. The **restartPolicy** field is added to **initContainers**. When this field is set to **Always**, the sidecar container is enabled. The sidecar container and service container are deployed in the same pod. This cannot prolong the pod lifecycle. Sidecar containers are commonly used in scenarios such as network proxy and log collection. For details, see **Sidecar Containers**.

- The legacy ServiceAccount token cleaner is in the beta state.

  Legacy ServiceAccount token cleaner is promoted to beta. It runs as part of **kube-controller-manager** and checks every 24 hours to see if any auto-generated legacy ServiceAccount token has not been used in a specific amount of time (one year by default, specified by **--legacy-service-account-token-clean-up-period**). If so, the cleaner marks those tokens as invalid and adds the **kubernetes.io/legacy-token-invalid-since** label whose value is the current date. If an invalid token is not used for a specific period of time (one year by default, specified by **--legacy-service-account-token-clean-up-period**), the cleaner deletes it. For details, see **Legacy ServiceAccount Token Cleaner**.

- **DevicePluginCDIDevices** is in the beta state.

  DevicePluginCDIDevices moves to beta. With this feature enabled, plugin developers can use the **CDIDevices** field added to **DeviceRunContainerOptions** to pass CDI device names directly to CDI enabled runtimes.

- **PodHostIPs** is in the beta state.

  The **PodHostIPs** feature moves to beta. With this feature enabled, Kubernetes adds the **hostIPs** field to **Status** of pods and downward API to expose node IP addresses to workloads. This field specifies the dual-stack protocol version of the host IP address. The first IP address is always the same as the host IP address.

- The API Priority and Fairness feature (APF) is in the GA state.

  APF moves to GA. APF classifies and isolates requests in a more fine-grained way. It improves max-inflight limitations. It also introduces a limited amount of queuing, so that the API server does not reject any request in cases of very brief bursts. Requests are dispatched from queues using a fair queuing technique so that, for example, a poorly-behaved controller does not cause others (even at the same priority level) to become abnormal. For details, see **API Priority and Fairness**.

- **APIListChunking** is in the GA state.

  The **APIListChunking** feature moves to GA. This feature allows clients to perform pagination in List requests to avoid performance problems caused by returning too much data at a time.

- **ServiceNodePortStaticSubrange** is in the GA state.

  The **ServiceNodePortStaticSubrange** feature moves to GA. With this feature enabled, kubelet calculates the size of reserved IP addresses based on the ranges of the NodePort Services and divides node ports into static band and dynamic band. During automatic node port assignment, dynamic band is

preferentially assigned, which helps avoid port conflicts during static band assignment. For details, see **ServiceNodePortStaticSubrange**.

- The phase transition timestamp of PersistentVolume (PV) is in the beta state.

  The PV phase transition timestamp moves to beta. With this feature enabled, Kubernetes adds the **lastPhaseTransitionTime** field to the **status** field of a PV to indicate the time when the PV phase changes last time. Cluster administrators are now able to track the last time a PV transitioned to a different phase, allowing for more efficient and informed resource management. For details, see **PersistentVolume Last Phase Transition Time in Kubernetes**.

- **ReadWriteOncePod** is in the GA state.

  The **ReadWriteOncePod** feature moves to GA. With this feature enabled, you can set the access mode to **ReadWriteOncePod** in a PersistentVolumeClaim (PVC) to ensure that only one pod can modify data in the volume at a time. This can prevent data conflicts or damage. For details, see **ReadWriteOncePod**.

- **CSINodeExpandSecret** is in the GA state.

  The **CSINodeExpandSecret** feature moves to GA. This feature allows secret authentication data to be passed to a CSI driver for use when a node is added.

- The CEL-based CustomResourceDefinition (CRD) verification capability is in the GA state.

  The CEL-based CRD verification capability moves to GA. With this feature enabled, you are allowed to use the CEL to define validation rules in CRDs, which are more efficient than webhook. For details, see **CRD verification rules**.

API Changes and Removals

- The time zone of a newly created cron job cannot be configured using **TZ** or **CRON_TZ** in **.spec.schedule**. Use **.spec.timeZone** instead. Cron jobs that have been created are not affected by this change.
- The alpha API **ClusterCIDR** is removed.
- The startup parameter **--authentication-config** is added to kube-apiserver to specify the address of the **AuthenticationConfiguration** file. This startup parameter is mutually exclusive with the **--oidc-*** startup parameter.
- The API version **kubescheduler.config.k8s.io/v1beta3** of **KubeSchedulerConfiguration** is removed. Migrate **kube-scheduler** configuration files to **kubescheduler.config.k8s.io/v1**.
- The CEL expressions are added to **v1alpha1 AuthenticationConfiguration**.
- The **ServiceCIDR** type is added. It allows you to dynamically configure the IP address range used by a cluster to allocate the Service ClusterIPs.
- The startup parameters **--conntrack-udp-timeout** and **--conntrack-udp-timeout-stream** are added to **kube-proxy**. They are options for configuring the kernel parameters **nf_conntrack_udp_timeout** and **nf_conntrack_udp_timeout_stream**.
- Support for CEL expressions is added to **WebhookMatchCondition** of **v1alpha1 AuthenticationConfiguration**.

- The type of **PVC.spec.Resource** is changed from **ResourceRequirements** to **VolumeResourceRequirements**.
- **onPodConditions** in **PodFailurePolicyRule** is marked as optional.
- The API version **flowcontrol.apiserver.k8s.io/v1beta3** of **FlowSchema** and **PriorityLevelConfiguration** has been promoted to **flowcontrol.apiserver.k8s.io/v1**, and the following changes have been made:
  - **PriorityLevelConfiguration**:
    The **.spec.limited.nominalConcurrencyShares** field defaults to **30** if the field is omitted. To ensure compatibility with 1.28 API servers, specifying an explicit **0** value is not allowed in the **v1** version in 1.29. In 1.30, explicit **0** values will be allowed in this field in the **v1** API. The **flowcontrol.apiserver.k8s.io/v1beta3** APIs are deprecated and will no longer be served in 1.32.
- The kube-proxy command line document is updated. kube-proxy does not bind any socket to the IP address specified by **--bind-address**.
- The **selectorSpread** scheduler plugin is replaced by **podTopologySpread**.
- If CSI-Node-Driver is not running, NodeStageVolume calls will be retried.
- **ValidatingAdmissionPolicy** type checking now supports CRDs. To use this feature, the **ValidatingAdmissionPolicy** feature gate must be enabled.
- The startup parameter **--nf-conntrack-tcp-be-liberal** is added to **kube-proxy**. You can configure it by setting the kernel parameter **nf_conntrack_tcp_be_liberal**.
- The startup parameter **--init-only** is added to **kube-proxy**. Setting the flag makes **kube-proxy** init container run in the privileged mode, perform its initial configuration, and then exit.
- The **fileSystem** field of container is added to the response body of CRI. It specifies the file system usage of a container. Originally, the **fileSystem** field contains only the file system of the container images.
- All built-in cloud providers are disabled by default. If you still need to use them, you can configure the **DisableCloudProviders** and **DisableKubeletCloudCredentialProvider** feature gates to disable or enable cloud providers.
- **--node-ips** can be used in kubelet to configure IPv4/IPv6 dual-stack. If **--cloud-provider** is set to **external**, you are allowed to use **--node-ips** to configure IPv4/IPv6 dual-stack for node IP addresses. To use **--node-ips**, you need to enable the **CloudDualStackNodeIPs** feature gate.

Enhanced Kubernetes 1.29 on CCE

During a version maintenance period, CCE periodically updates Kubernetes 1.29 and provides enhanced functions.

For details about cluster version updates, see **Patch Versions**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.29 and other versions, see **Kubernetes v1.29 Release Notes**.

5.1.2.2 Kubernetes 1.28 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. CCE allows you to create Kubernetes clusters 1.28. This section describes the changes made in Kubernetes 1.28.

Indexes

Important Notes

- In Kubernetes 1.28, the scheduling framework is improved to reduce useless retries. The overall scheduling performance is enhanced. If a custom scheduler plugin is used in a cluster, you can perform the adaptation upgrade following instructions in **GitHub**.
- The Ceph FS in-tree plugin has been deprecated in Kubernetes 1.28 and will be removed in Kubernetes 1.31. (The community does not plan to support CSI migration.) Use **Ceph CSI driver** instead.
- The Ceph RBD in-tree plugin has been deprecated in Kubernetes 1.28 and will be removed in Kubernetes 1.31. (The community does not plan to support CSI migration.) Use RBD **Ceph CSI driver** instead.

New and Enhanced Features

Features in alpha stage are disabled by default, those in beta stage are enabled by default, and those in GA stage are always enabled and they cannot be disabled. The function of turning on or off the features in GA stage will be removed in later Kubernetes versions. CCE policies for new features are the same as those in the community.

- The version skew policy is expanded to three versions.

  Starting with control planes 1.28 and worker nodes 1.25, the Kubernetes skew policy expands the supported control plane and worker node skew to three versions. This enables annual minor version upgrades of nodes while staying on supported minor versions. For details, see **Version Skew Policy**.
- Retroactive Default StorageClass moves to GA.

  The retroactive default StorageClass assignment graduates to GA. This enhancement brings a significant improvement to how default StorageClasses are assigned to PersistentVolumeClaims (PVCs).

  The PV controller has been modified to automatically assign a default StorageClass to any unbound PVC with **storageClassName** not configured. Additionally, the PVC admission validation mechanism within the API server has been adjusted to allow changing values from an unset state to an actual StorageClass name. For details, see **Retroactive default StorageClass assignment**.

- Native sidecar containers are introduced.

  The native sidecar containers are available in alpha. Kubernetes 1.28 adds **restartPolicy** to Init containers. This field is available when the SidecarContainers feature gate is enabled. However, there are still some problems to be solved in the native sidecar containers. Therefore, the Kubernetes community recommends only using this feature gate in **short lived testing clusters** at the alpha phase. For details, see **Introducing native sidecar containers**.

- Mixed version proxy is introduced.

  A new mechanism (mixed version proxy) is released to improve cluster upgrade. It is an alpha feature in Kubernetes 1.28. When a cluster undergoes an upgrade, API servers of different versions in the cluster can serve different sets (groups, versions, or resources) of built-in resources. A resource request made in this scenario may be served by any of the available API servers, potentially resulting in the request ending up at an API server that may not be aware of the requested resource. As a result, the request fails. This feature can solve this problem. (Note that CCE provides hitless upgrade. Therefore, this feature is not used in CCE clusters.) For details, see **A New (alpha) Mechanism For Safer Cluster Upgrades**.

- Non-graceful node shutdown moves to GA.

  The non-graceful node shutdown is now GA in Kubernetes 1.28. When a node was shut down and that shutdown was not detected by the Kubelet's Node Shutdown Manager, the StatefulSet pods that run on this node will stay in the terminated state and cannot be moved to a running node. If you have confirmed that the shutdown node is unrecoverable, you can add an **out-of-service** taint to the node. This ensures that the StatefulSet pods and VolumeAttachments on this node can be forcibly deleted and the corresponding pods will be created on a healthy node. For details, see **Non-Graceful Node Shutdown Moves to GA**.

- NodeSwap moves to beta.

  Support for NodeSwap goes to beta in Kubernetes 1.28. NodeSwap is disabled by default and can be enabled using the NodeSwap feature gate. NodeSwap allows you to configure swap memory usage for Kubernetes workloads running on Linux on a per-node basis. Note that although NodeSwap has reached beta, there are still some problems to be solved and security risks to be enhanced. For details, see **Beta Support for Using Swap on Linux**.

- Two job-related features are added.

  Two alpha features are introduced: **delayed creation of replacement pods** and **backoff limit per index**.

  - Delayed creation of replacement pods

    By default, when a pod enters the terminating state (for example, due to the preemption or eviction), Kubernetes immediately creates a replacement pod. Therefore, both pods are running concurrently.

    In Kubernetes 1.28, this feature can be enabled by turning on the JobPodReplacementPolicy feature gate. With this feature gate enabled, you can set the **podReplacementPolicy** field under **spec** of a job to **Failed**. In this way, pods would only be replaced when they reached the failed phase, and not when they are terminating. Additionally, you can check the **.status.termination** field of a job. The value of this field is the number of pods owned by the job that are currently terminating.

  - Backoff limit per index

    By default, pod failures for indexed jobs are recorded and restricted by the global limit of retries, specified by **.spec.backoffLimit**. This means that if there is a consistently failing index in a job, pods specified by the job will be restarted repeatedly until pod failures exhaust the limit. Once the limit is reached, the job is marked failed and pods for other indexes in the job may never be even started.

    In Kubernetes 1.28, this feature can be enabled by turning on the JobBackoffLimitPerIndex feature gate of a cluster. With this feature gate enabled, **.spec.backoffLimitPerIndex** can be specified when an indexed job is created. Only if the failures of pods with all indexes specified in this job exceed the upper limit, pods specified by the job will not be restarted.

- Some CEL related features are improved.

  CEL related capabilities are enhanced.

  - CEL used to validate CRDs moves to beta.

    This feature has been upgraded to beta since Kubernetes 1.25. By embedding CEL expressions into CRDs, developers can solve most of the CR validation use cases without using webhooks. More CEL functions, such as support for default value and CRD conversion, will be developed in later Kubernetes versions.

  - CEL admission control graduates to beta.

    CEL admission control is customizable. With CEL expressions, you can decide whether to accept or reject requests received by kube-apiserver. CEL expressions can also serve as a substitute for admission webhooks. Kubernetes 1.28 has upgraded CEL admission control to beta and introduced new functions, such as:

    - ValidatingAdmissionPolicy can correctly handle the **authorizer** variable.
    - ValidatingAdmissionPolicy can have the **messageExpression** field checked.
    - The ValidatingAdmissionPolicy controller is added to kube-controller-manager to check the type of the CEL expression in ValidatingAdmissionPolicy and save the reason in the **status** field.
    - CEL expressions can contain a combination of one or more variables, which can be defined in ValidatingAdmissionPolicy. These variables can be used to define other variables.
    - CEL library functions can be used to parse resources specified by **resource.Quantity** in Kubernetes.

- Other features
  - The ServiceNodePortStaticSubrange feature gate moves to beta. With this feature enabled, static port range can be reserved to avoid conflicts with dynamically allocated ports. For details, see **Avoiding Collisions Assigning Ports to NodePort Services**.
  - The alpha feature ConsistentListFromCache is added to allow the API server to serve consistent lists from cache. Get and list requests can read data from the cache instead of etcd.

  - In Kubernetes 1.28, kubelet can configure the drop-in directory (alpha). This feature allows you to add support for the **--config-dir** flag to kubelet so that you can specify an insert directory that overwrites the kubelet configuration in **/etc/kubernetes/kubelet.conf**.
  - ExpandedDNSConfig moves to GA and is enabled by default. With this feature enabled, DNS configurations can be expanded.
  - The alpha feature CRDValidationRatcheting is added. This feature allows CRs with failing validations to pass if a Patch or Update request does not alter any of the invalid fields.
  - **--concurrent-cron-job-syncs** is added to kube-controller-manager to configure the number of workers for the cron job controller.

API Changes and Removals

- **NetworkPolicyStatus** is removed. There is no status attribute in a network policy.
- **annotationbatch.kubernetes.io/cronJob-scheduled-timestamp** is added to job objects to indicate the creation time of a job.
- The **podReplacementPolicy** and **terminating** fields are added to job APIs. With these fields specified, once a previously created pod is terminated in a job, the job immediately starts a new pod to replace the pod. The new fields allow you to specify whether to replace the pod immediately after the previous pod is terminated (original behavior) or replace the pod after the existing pod is completely terminated (new behavior). This is an alpha feature, and you can enable it by turning on the **JobPodReplacementPolicy** feature gate in your cluster.
- The **BackoffLimitPerIndex** field is available in a job. Pods specified by a job share a backoff mechanism. When backoff times of the job reach the limit, this job is marked as failed and resources, including indexes that are not running, are cleared up. This field allows you to configure backoff limit for a single index. For details, see **Backoff limit per index**.
- The **ServedVersions** field is added to the **StorageVersion** API. This change is introduced by mixed version proxy. The new field is used to indicate a version that can be provided by the API server.
- **SelfSubjectReview** is added to **authentication.k8s.io/v1**, and **kubectl auth whoami** goes to GA.
- **LastPhaseTransitionTime** is added to **PersistentVolume**. The new field is used to store the last time when a volume changes to a different phase.
- **resizeStatus** in **PVC.Status** is replaced by **AllocatedResourceStatus**. The new field indicates the statuses of the storage resize operation. The default value is an empty string.
- If **hostNetwork** is set to **true** and ports are specified for a pod, the **hostport** field will be automatically configured.
- StatefulSet pods have the pod index set as a pod label **statefulset.kubernetes.io/pod-index**.
- **PodHasNetwork** in the **Condition** field of pods has been renamed to **PodReadyToStartContainers**. The new field specifies that containers are ready to start after the network, volumes, and sandbox pod have been created.

- A new configuration option **delayCacheUntilActive** is added to **KubeSchedulerConfiguration**. If **delayCacheUntilActive** is set to **true**, kube-scheduler on the leader will not cache scheduling information. This reduces the memory pressure of other master nodes, but slows down the failover speed after the leader failed.
- The **namespaceParamRef** field is added to **admissionregistration.k8s.io/v1alpha1.ValidatingAdmissionPolicy**.
- The **reason** and **fieldPath** fields are added to CRD validation rules to allow you to specify reason and field path after verification failed.
- The CEL expression of ValidatingAdmissionPolicy supports namespace access via namespaceObject.
- API groups ValidatingAdmissionPolicy and ValidatingAdmissionPolicyBinding are promoted to betav1.
- A ValidatingAdmissionPolicy now has its **messageExpression** field checked against resolved types.

Feature Gate and Command Line Parameter Changes and Removals

- **–short** is removed from kubelet. Therefore, the default output of **kubectl version** is the same as that of **kubectl version –short**.
- **--volume-host-cidr-denylist** and **--volume-host-allow-local-loopback** are removed from kube-controller-manager. **--volume-host-cidr-denylist** is a comma-separated list of CIDR ranges. Volume plugins at these IP addresses are not allowed. If **--volume-host-allow-local-loopback** is set to **false**, the local loopback IP address and the CIDR ranges specified in **--volume-host-cidr-denylist** are disabled.
- **--azure-container-registry-config** is deprecated in kubelet and will be deleted in later Kubernetes versions. Use **--image-credential-provider-config** and **--image-credential-provider-bin-dir** instead.
- **--lock-object-namespace** and **--lock-object-name** are removed from kube-scheduler. Use **--leader-elect-resource-namespace** and **--leader-elect-resource-name** or **ComponentConfig** instead. (**--lock-object-namespace** is used to define the namespace of a lock object, and **--lock-object-name** is used to define the name of a lock object.)
- KMS v1 is deprecated and will only receive security updates. Use KMS v2 instead. In later Kubernetes versions, use **--feature-gates=KMSv1=true** to configure a KMS v1 provider.
- The DelegateFSGroupToCSIDriver, DevicePlugins, KubeletCredentialProviders, MixedProtocolLBService, ServiceInternalTrafficPolicy, ServiceIPStaticSubrange, and EndpointSliceTerminatingCondition feature gates are removed.

Enhanced Kubernetes 1.28 on CCE

During a version maintenance period, CCE periodically updates Kubernetes 1.28 and provides enhanced functions.

For details about cluster version updates, see **Release Notes for CCE Cluster Versions**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.28 and other versions, see **Kubernetes v1.28 Release Notes**.

5.1.2.3 Kubernetes 1.27 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. CCE allows you to create clusters of Kubernetes 1.27. This section describes the changes made in Kubernetes 1.27 compared with Kubernetes 1.25.

Indexes

New Features

Kubernetes 1.27

- SeccompDefault is stable.

  To use SeccompDefault, add the **--seccomp-default command line flag** using kubelet on each node. If this feature is enabled, the **RuntimeDefault** profile will be used for all workloads by default, instead of the **Unconfined** (seccomp disabled) profile.

- Jobs' scheduling directives are configurable.

  This feature was introduced in Kubernetes 1.22 and is stable in Kubernetes 1.27. In most cases, you use a job to influence where the pods will run, like all in the same AZ. This feature allows scheduling directives to be modified before a job starts. You can use the **suspend** field to suspend a job. In the suspension phase, the scheduling directives (such as the node selector, node affinity, anti-affinity, and tolerations) in the job's pod template can be modified. For details, see **Mutable Scheduling Directives**.

- Downward API hugepages are stable.

  In Kubernetes 1.20, **requests.hugepages-**<*pagesize*> and **limits.hugepages-**<*pagesize*> were introduced to the **downward API**. Requests and limits can be configured for hugepages like other resources.

- Pod scheduling readiness moves to beta.

  After a pod is created, the Kubernetes scheduler selects an appropriate node to run the pod in the pending state. In practice, some pods may stay in the pending state for a long period due to insufficient resources. These pods may affect the running of other components like Cluster Autoscaler in the cluster. By specifying or deleting **.spec. schedulingGates** for a pod, you can control when the pod is ready for scheduling. For details, see **Pod Scheduling Readiness**.

- Accessing node logs using Kubernetes APIs is supported.

This function is in the alpha phase. The cluster administrator can directly query node logs to help debug malfunctioning services running on the node. To use this function, ensure that the NodeLogQuery **feature gate** is enabled for that node and the kubelet configuration options **enableSystemLogHandler** and **enableSystemLogQuery** are set to **true**.

- ReadWriteOncePod access mode moves to beta.

  Kubernetes 1.22 introduced a ReadWriteOncePod access mode for PVs and PVCs. This feature has evolved into the beta phase. A volume can be mounted to a single pod in read/write mode. Use this access mode if you want to ensure that only one pod in the cluster can read that PVC or write to it. For details, see **Access Modes**.

- The **matchLabelKeys** field in the pod topology spread constraint moves to beta.

  matchLabelKeys is a list of pod label keys. It is used to select a group of pods over which spreading will be calculated. With **matchLabelKeys**, you do not need to update **pod.spec** between different revisions. The controller or operator just needs to set different values to the same label key for different revisions. The scheduler will automatically determine the values based on **matchLabelKeys**. For details, see **Pod Topology Distribution Constraints**.

- The function of efficiently labeling SELinux volumes moves to beta.

  By default, the container runtime recursively assigns the SELinux label to all files on all pod volumes. To speed up this process, Kubernetes uses the mount option **-o context=***<label>* to immediately change the SELinux label of the volume. For details, see **Efficient SELinux volume relabeling**.

- VolumeManager reconstruction goes to beta.

  After the VolumeManager is reconstructed, if the NewVolumeManagerReconstruction **feature gate** is enabled, mounted volumes will be obtained in a more effective way during kubelet startup.

- Server side field validation and OpenAPI V3 are stable.

  OpenAPI V3 was added in Kubernetes 1.23. In Kubernetes 1.24, it moved to beta. In Kubernetes 1.27, it is stable.

- StatefulSet start ordinal moves to beta.

  Kubernetes 1.26 introduced a new, alpha-level feature for StatefulSets to control the ordinal numbering of pod replicas. Since Kubernetes 1.27, this feature moves to beta. The ordinals can start from arbitrary non-negative numbers. For details, see **Kubernetes 1.27: StatefulSet Start Ordinal Simplifies Migration**.

- **ContainerResource** metric in HorizontalPodAutoscaler moves to beta.

  Kubernetes 1.20 introduced the **ContainerResource** metric in HorizontalPodAutoscaler (HPA). In Kubernetes 1.27, this feature moves to beta, and the HPAContainerMetrics feature gate is enabled by default.

- StatefulSet PVC auto deletion moves to beta.

  Kubernetes 1.27 provides a new policy to control the lifecycle of PVCs of StatefulSets. This policy allows users to specify if the PVCs generated from the StatefulSet spec template should be automatically deleted or retained when the StatefulSet is deleted or replicas in the StatefulSet are scaled down. For details, see **PersistentVolumeClaim retention**.

- Volume group snapshots are introduced.

  Volume group snapshots are introduced as an alpha feature in Kubernetes 1.27. This feature allows users to create snapshots for multiple volumes to ensure data consistency when a fault occurs. It uses a label selector to group multiple PVCs for snapshot. This feature only supports CSI volume drivers. For details, see **Kubernetes 1.27: Introducing an API for Volume Group Snapshots**.

- **kubectl apply** pruning is more secure and efficient.

  In Kubernetes 1.5, the **--prune** flag was introduced in **kubectl apply** to delete resources that are no longer needed. This allowed **kubectl apply** to automatically clear resources removed from the current configuration. However, the existing implementation of **--prune** has design defects that degrade its performance and lead to unexpected behaviors. In Kubernetes 1.27, **kubectl apply** provides ApplySet-based pruning, which is in the alpha phase. For details, see **Declarative Management of Kubernetes Objects Using Configuration Files**.

- Conflicts during port allocation to NodePort Service can be avoided.

  In Kubernetes 1.27, you can enable a new **feature gate** ServiceNodePortStaticSubrange to use different port allocation policies for NodePort Services. This mitigates the risk of port conflicts. This feature is in the alpha phase.

- Resizing resources assigned to pods without restarting the containers is supported.

  Kubernetes 1.27 allows users to resize CPU and memory resources assigned to pods without restarting the container. This feature is in the alpha phase. For details, see **Kubernetes 1.27: In-place Resource Resize for Kubernetes Pods (alpha)**.

- Pod startup is accelerated.

  A series of parameter adjustments like parallel image pulls and increased default API query limit for kubelet per second are made in Kubernetes 1.27 to accelerate pod startup. For details, see **Kubernetes 1.27: updates on speeding up Pod startup**.

- KMS V2 moves to beta.

  The key management KMS V2 API goes to beta. This has greatly improved the performance of the KMS encryption provider. For details, see **Using a KMS provider for data encryption**.

Kubernetes 1.26

- CRI v1alpha2 is removed.

  Kubernetes 1.26 does not support CRI v1alpha2 any longer. Use CRI v1 (containerd version must be later than or equal to 1.5.0). containerd 1.5.x or earlier is not supported by Kubernetes 1.26. Update the containerd version to 1.6.x or later before upgrading kubelet to 1.26.

  NOTE

  The containerd version used by CCE is 1.6.14, which meets the requirements. If the existing nodes do not meet the containerd version requirements, reset them to the latest version.

- Alpha API for dynamic resource allocation is added.

  In Kubernetes 1.26, **Dynamic Resource Allocation** is added to request and share resources between pods and between containers in a pod. Resources can be initialized based on parameters provided by the user. This function is still in the alpha phase. You need to enable the DynamicResourceAllocation feature gate and the **resource.k8s.io/v1alpha1** API group. You need to install drivers for specific resources to be managed. For details, see **Kubernetes 1.26: Alpha API for Dynamic Resource Allocation**.

- The non-graceful node shutdown feature goes to beta.

  In Kubernetes 1.26, the non-graceful node shutdown feature goes to beta and is enabled by default. A node shutdown can be graceful only if the kubelet's node shutdown manager can detect the upcoming node shutdown action. For details, see **Non-graceful node shutdown handling**.

- Passing pod **fsGroup** to CSI drivers during mounting is supported.

  In Kubernetes 1.22, delegation of **fsGroup** to CSI drivers was first introduced as an alpha feature. In Kubernetes 1.25, it moved to beta. In Kubernetes 1.26, this feature enters the official release phase. For details, see **Delegating volume permission and ownership change to CSI driver**.

- Pod scheduling readiness is introduced.

  Kubernetes 1.26 introduces a new feature schedulingGates, which enables the scheduler to detect when pod scheduling can be performed. For details, see **Pod Scheduling Readiness**.

- CPU manager is officially released.

  The CPU manager is a part of kubelet. Since Kubernetes 1.10, it has moved to **beta**. The CPU manager can allocate exclusive CPUs to containers. This feature is stable in Kubernetes 1.26. For details, see **Control CPU Management Policies on the Node**.

- Kubernetes traffic engineering is advanced.

  Internal node-local traffic optimization and **EndpointSlice conditions** are upgraded to the official release version. **ProxyTerminatingEndpoints** moves to beta.

- Cross-namespace volume data sources are supported.

  This feature allows you to specify a data source that belongs to different namespaces for a PVC. This feature is in the alpha phase. For details, see **Cross namespace data sources**.

- Retroactive default StorageClass assignment moves to beta.

  In Kubernetes 1.25, an alpha feature was introduced to change the way how a default StorageClass is allocated to a PVC. After this feature is enabled, you no longer need to create a default StorageClass and then create a PVC to assign the class. Additionally, any PVCs without a StorageClass assigned can be updated later. This feature moves to beta in Kubernetes 1.26. For details, see **Retroactive default StorageClass assignment**.

- PodDisruptionBudget allows users to specify the eviction policies for unhealthy pods.

  You are allowed to specify unhealthy pod eviction policies for **PodDisruptionBudget** (PDB). This feature helps ensure node availability during node management. This feature is in the beta phase. For details, see **Unhealthy Pod Eviction Policy**.

- The number of Horizontal Pod Autoscaler (HPA) can be configured.

  kube-controller-manager allows **--concurrent-horizontal-pod-autoscaler-syncs** to configure the number of worker nodes of the pod autoscaler for horizontal scaling.

Deprecations and Removals

Kubernetes 1.27

- In Kubernetes 1.27, the feature gates that are used for volume extension and in the GA status, including ExpandCSIVolumes, ExpandInUsePersistentVolumes, and ExpandPersistentVolumes are removed and can no longer be referenced in the **--feature-gates** flag.
- The **--master-service-namespace** parameter is removed. This parameter specifies where to create a Service named **kubernetes** to represent the API server. This parameter was deprecated in Kubernetes 1.26 and is removed from Kubernetes 1.27.
- The ControllerManagerLeaderMigration feature gate is removed. **Leader Migration** provides a mechanism for HA clusters to safely migrate "cloud specific" controllers using a resource lock shared between kube-controller-manager and cloud-controller-manager when upgrading the replicated control plane. This feature has been enabled unconditionally since its release in Kubernetes 1.24. In Kubernetes 1.27, this feature is removed.
- The **--enable-taint-manager** parameter is removed. The feature that it supports, taint-based eviction, is enabled by default and will continue to be implicitly enabled when the flag is removed.
- The **--pod-eviction-timeout** parameter is removed from kube-controller-manager.
- The CSIMigration feature gate is removed. The **CSI migration** program allows smooth migration from the in-tree volume plug-ins to the out-of-tree CSI drivers. This feature was officially released in Kubernetes 1.16.
- The CSIInlineVolume feature gate is removed. The feature (**CSI Ephemeral Volume**) allows CSI volumes to be specified directly in the pod specification for ephemeral use cases. They can be used to inject arbitrary states, such as configuration, secrets, identity, variables, or similar information, directly inside the pod using a mounted volume. This feature graduated to GA in Kubernetes 1.25 and is removed in Kubernetes 1.27.
- The EphemeralContainers feature gate is removed. For Kubernetes 1.27, API support for ephemeral containers is unconditionally enabled.
- The LocalStorageCapacityIsolation feature gate is removed. This feature gate (**Local Ephemeral Storage Capacity Isolation**) moved to GA in Kubernetes 1.25. The feature provides support for capacity isolation of local ephemeral storage between pods, such as emptyDir volumes, so that a pod can be limited in its consumption of shared resources. kubelet will evict a pod if its consumption of local ephemeral storage exceeds the configured limit.
- The NetworkPolicyEndPort feature gate is removed. In Kubernetes 1.25, **endPort** in NetworkPolicy moved to GA. NetworkPolicy providers that support the **endPort** field can be used to specify a range of ports to apply NetworkPolicy.
- The StatefulSetMinReadySeconds feature gate is removed. For a pod that is part of a StatefulSet, Kubernetes marks the pod as read-only when the pod is

available (and passes the check) at least within the period specified in **minReadySeconds**. This feature was officially released in Kubernetes 1.25. It is locked to **true** and removed from Kubernetes 1.27.

- The IdentifyPodOS feature gate is removed. If this feature is enabled, you can specify an OS for a pod. It has been stable since Kubernetes 1.25. This feature is removed from Kubernetes 1.27.
- The DaemonSetUpdateSurge feature gate is removed. In Kubernetes 1.25, this feature was stable. It was implemented to minimize DaemonSet downtime during deployment, but it is removed from Kubernetes 1.27.
- The **--container-runtime** parameter is removed. kubelet accepts a deprecated parameter **--container-runtime**, and the only valid value will be **remote** after the dockershim code is removed. This parameter was deprecated in 1.24 and later versions and is removed from Kubernetes 1.27.

Kubernetes 1.26

- HorizontalPodAutoscaler API for v2beta2 is removed.

  The autoscaling/v2beta2 API of HorizontalPodAutoscaler is no longer available in Kubernetes 1.26. For details, see **Removed APIs by release**. Use autoscaling/v2 API instead.
- The **flowcontrol.apiserver.k8s.io/v1beta1** API is removed.

  In Kubernetes 1.26 and later versions, the API of the **flowcontrol.apiserver.k8s.io/v1beta1** version for FlowSchema and PriorityLevelConfiguration is no longer served. For details, see **Removed APIs by release**. The **flowcontrol.apiserver.k8s.io/v1beta2** version is available in Kubernetes 1.23 and later versions, and the **flowcontrol.apiserver.k8s.io/v1beta3** version is available in Kubernetes 1.26 and later versions.
- The cloud service vendors' in-tree storage drivers are removed.
- The kube-proxy userspace mode is removed.

  The deprecated userspace mode is no longer supported by Linux or Windows. Linux users can use Iptables or IPVS, and Windows users can use the Kernelspace mode. Errors are returned if you use **--mode userspace**.

  – Windows winkernel kube-proxy no longer supports Windows HNS v1 APIs.
- **--prune-whitelist** flag is deprecated.

  The **--prune-whitelist** flag is **deprecated** and replaced by **--prune-allowlist** to support **Inclusive Naming Initiative**. This deprecated flag will be completely removed in later versions.
- The DynamicKubeletConfig feature gate is removed.

  The kubelet configuration of nodes can be dynamically updated through the API. The feature gate is removed from the kubelet in Kubernetes 1.24 and removed from the API server in Kubernetes 1.26. This simplifies the code and improves stability. It is recommended that you modify the kubelet configuration file instead and then restart the kubelet. For details, see **Remove DynamicKubeletConfig feature gate from the code**.
- A kube-apiserver command line parameter is removed.

  The **--master-service-namespace** parameter is deprecated. It is unused in the API Server.

- Several **kubectl run** parameters are deprecated.

  Several unused kubectl subcommands are marked as **deprecated** and will be removed in later versions. These subcommands include **--cascade**, **--filename**, **--force**, **--grace-period**, **--kustomize**, **--recursive**, **--timeout**, and **--wait**.

- Some command line parameters related to logging are removed.

  Some logging-related command line parameters are **removed**. These parameters were **deprecated** in earlier versions.

Enhanced Kubernetes 1.27 on CCE

During a version maintenance period, CCE periodically updates Kubernetes 1.27 and provides enhanced functions.

For details about cluster version updates, see **Release Notes for CCE Cluster Versions**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.27 and other versions, see the following documents:

- **Kubernetes v1.27 Release Notes**
- **Kubernetes v1.26 Release Notes**

5.1.2.4 Kubernetes 1.25 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the changes made in Kubernetes 1.25 compared with Kubernetes 1.23.

Indexes

New Features

Kubernetes 1.25

- Pod Security Admission is stable. PodSecurityPolicy is deprecated.

  PodSecurityPolicy is replaced by Pod Security Admission. For details about the migration, see **Migrate from PodSecurityPolicy to the Built-In PodSecurity Admission Controller**.

- The ephemeral container is stable.

  An **ephemeral container** is a container that runs temporarily in an existing pod. It is useful for troubleshooting, especially when kubectl exec cannot be used to check a container that breaks down or its image lacks a debugging tool.

- Support for cgroups v2 enters the stable phase.

Kubernetes supports cgroups v2. cgroups v2 provides some improvements over cgroup v1. For details, see **About cgroup v2**.

- SeccompDefault moves to beta.

  To enable this feature, add the startup parameter **--seccomp-default=true** to kubelet. In this way, **seccomp** is set to **RuntimeDefault** by default, improving system security. Clusters of v1.25 no longer support **seccomp.security.alpha.kubernetes.io/pod** and **container.seccomp.security.alpha.kubernetes.io/annotation**. Replace them with the **securityContext.seccompProfile** field in pods or containers. For details, see **Configure a Security Context for a Pod or Container**.

  NOTE

  After this feature is enabled, the system calls required by the application may be restricted by the runtime. Ensure that the debugging is performed in the test environment, so that application is not affected.

- The EndPort in the network policy moves to stable.

  EndPort in Network Policy is stable. This feature is incorporated in version 1.21. EndPort is added to NetworkPolicy. You can specify a port range.

- Local ephemeral storage capacity isolation is stable.

  This feature provides support for capacity isolation of local ephemeral storage between pods, such as EmptyDir. If a pod's consumption of shared resources exceeds the limit, it will be evicted.

- The CRD verification expression language moves to beta.

  This makes it possible to declare how to validate custom resources using **CEL**. For details, see **Extend the Kubernetes API with CustomResourceDefinitions**.

- KMS v2 APIs are introduced.

  The KMS v2 alpha1 API is introduced to add performance, rotation, and observability improvements. This API uses AES-GCM to replace AES-CBC and uses DEK to encrypt data at rest (Kubernetes Secrets). No additional operation is required during this process. Additionally, data can be read through AES-GCM and AES-CBC. For details, see **Using a KMS provider for data encryption**.

- Pod network readiness is introduced.

  Kubernetes 1.25 introduces Alpha support for PodHasNetwork. This status is in the **status** field of the pod. For details, see **Pod network readiness**.

- The two features used for application rollout are stable.
  - In Kubernetes 1.25, **minReadySeconds** for StatefulSets is stable. It allows each pod to wait for an expected period of time to slow down the rollout of a StatefulSet. For details, see **Minimum ready seconds**.
  - In Kubernetes 1.25, **maxSurge** for DaemonSets is stable. It allows a DaemonSet workload to run multiple instances of the same pod on one node during a rollout. This minimizes DaemonSet downtime for users. DaemonSet does not allow **maxSurge** and **hostPort** to be used at the same time because two active pods cannot share the same port on the same node. For details, see **Perform a Rolling Update on a DaemonSet**.

- Alpha support for running pods with user namespaces is provided.

  This feature maps the **root** user in a pod to a non-zero ID outside the container. In this way, the container runs as the **root** user and the node runs

as a regular unprivileged user. This feature is still in the internal test phase. The UserNamespacesStatelessPodsSupport gate needs to be enabled, and the container runtime must support this function. For details, see **Kubernetes 1.25: alpha support for running Pods with user namespaces**.

Kubernetes 1.24

- Dockershim is removed from kubelet.

  Dockershim was marked deprecated in Kubernetes 1.20 and officially removed from kubelet in Kubernetes 1.24. If you want to use Docker container, switch to cri-dockerd or other runtimes that support CRI, such as containerd and CRI-O.

  For details about how to switch from Docker to containerd, see **Migrating Nodes from Docker to containerd**.

  NOTE

  Check whether there are agents or applications that depend on Docker Engine. For example, if **docker ps**, **docker run**, and **docker inspect** are used, ensure that multiple runtimes are compatible and switch to the standard CRI.

- Beta APIs are disabled by default.

  The Kubernetes community found 90% cluster administrators did not care about the beta APIs and left them enabled. However, the beta features are not recommended because these APIs enabled in the production environment by default incur risks. Therefore, in 1.24 and later versions, beta APIs are disabled by default, but the existing beta APIs will retain the original settings.

- OpenAPI v3 is supported.

  In Kubernetes 1.24 and later versions, OpenAPI V3 is enabled by default.

- Storage capacity tracking is stable.

  In Kubernetes 1.24 and later versions, the CSIStorageCapacity API supports exposing the available storage capacity. This ensures that pods are scheduled to nodes with sufficient storage capacity, which reduces pod scheduling delay caused by volume creation and mounting failures. For details, see **Storage Capacity**.

- gRPC container probe moves to beta.

  In Kubernetes 1.24 and later versions, the gRPC probe goes to beta. The feature gate GRPCContainerProbe is available by default. For details about how to use this probe, see **Configure Probes**.

- LegacyServiceAccountTokenNoAutoGeneration is enabled by default.

  LegacyServiceAccountTokenNoAutoGeneration moves to beta. By default, this feature is enabled, where no secret token is automatically generated for a service account. To use a token that never expires, create a secret to hold the token. For details, see **Service account token Secrets**.

- IP address conflict is prevented.

  In Kubernetes 1.24, **an IP address pool is soft reserved for the static IP addresses of Services**. After you manually enable this function, Service IP addresses will be automatically from the IP address pool to minimize IP address conflict.

- Clusters are compiled based on Go 1.18.

  Kubernetes clusters of versions later than 1.24 are compiled based on Go 1.18. By default, the SHA-1 hash algorithm, such as SHA1WithRSA and

ECDSAWithSHA1, is no longer supported for certificate signature verification. Use the certificate generated by the SHA256 algorithm instead.

- The maximum number of unavailable StatefulSet replicas is configurable.

  In Kubernetes 1.24 and later versions, the **maxUnavailable** parameter can be configured for StatefulSets so that pods can be stopped more quickly during a rolling update.

- Alpha support for non-graceful node shutdown is introduced.

  The non-graceful node shutdown is introduced as alpha in Kubernetes v1.24. A node shutdown is considered graceful only if kubelet's node shutdown manager can detect the upcoming node shutdown action. For details, see **Non-graceful node shutdown handling**.

Deprecations and Removals

Kubernetes 1.25

- The iptables chain ownership is cleared up.

  Kubernetes typically creates iptables chains to ensure data packets can be sent to the destination. These iptables chains and their names are for internal use only. These chains were never intended to be part of any Kubernetes API/ABI guarantees. For details, see **Kubernetes's IPTables Chains Are Not API**.

  In versions later than Kubernetes 1.25, Kubelet uses IPTablesCleanup to migrate the Kubernetes-generated iptables chains used by the components outside of Kubernetes in phases so that iptables chains such as KUBE-MARK-DROP, KUBE-MARK-MASQ, and KUBE-POSTROUTING will not be created in the NAT table. For more details, see **Cleaning Up IPTables Chain Ownership**.

- In-tree volume drivers from cloud service vendors are removed.

Kubernetes 1.24

- In Kubernetes 1.24 and later versions, Service.Spec.LoadBalancerIP is deprecated because it cannot be used for dual-stack protocols. Instead, use custom annotations.
- In Kubernetes 1.24 and later versions, the **--address**, **--insecure-bind-address**, **--port**, and **--insecure-port=0** parameters are removed from **kube-apiserver**.
- In Kubernetes 1.24 and later versions, startup parameters **--port=0** and **--address** are removed from **kube-controller-manager** and **kube-scheduler**.
- In Kubernetes 1.24 and later versions, **kube-apiserver --audit-log-version** and **--audit-webhook-version** support only **audit.k8s.io/v1**. In Kubernetes 1.24, **audit.k8s.io/v1[alpha|beta]1** is removed, and only **audit.k8s.io/v1** can be used.
- In Kubernetes 1.24 and later versions, the startup parameter **--network-plugin** is removed from kubelet. This Docker-specific parameter is available only when the container runtime environment is **Docker** and it is deleted with Dockershim.
- In Kubernetes 1.24 and later versions, dynamic log clearance has been discarded and removed accordingly. A log filter is introduced to the logs of all Kubernetes system components to prevent sensitive information from being leaked through logs. However, this function may block logs and therefore is discarded. For more details, see **Dynamic log sanitization** and **KEP-1753**.

- VolumeSnapshot v1beta1 CRD is discarded in Kubernetes 1.20 and removed in Kubernetes 1.24. Use VolumeSnapshot v1 instead.
- In Kubernetes 1.24 and later versions, **service annotation tolerate-unready-endpoints** discarded in Kubernetes 1.11 is replaced by **Service.spec.publishNotReadyAddresses**.
- In Kubernetes 1.24 and later versions, the **metadata.clusterName** field is discarded and will be deleted in the next version.
- In Kubernetes 1.24 and later versions, the logic for kube-proxy to listen to NodePorts is removed. If NodePorts conflict with **kernel net.ipv4.ip_local_port_range**, TCP connections may fail occasionally, which leads to a health check failure or service exception. Before the upgrade, ensure that cluster NodePorts do not conflict with **net.ipv4.ip_local_port_range** of all nodes in the cluster. For more details, see **Kubernetes PR**.

Enhanced Kubernetes 1.25 on CCE

During a version maintenance period, CCE periodically updates Kubernetes 1.25 and provides enhanced functions.

For details about cluster version updates, see **Release Notes for CCE Cluster Versions**.

References

For more details about the performance comparison and function evolution between Kubernetes 1.25 and other versions, see the following documents:

- **Kubernetes v1.25 Release Notes**
- **Kubernetes v1.24 Release Notes**

5.1.2.5 Kubernetes 1.23 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.23.

Resource Changes and Deprecations

Kubernetes 1.23 Release Notes

- FlexVolume is deprecated. Use CSI.
- HorizontalPodAutoscaler v2 is promoted to GA, and HorizontalPodAutoscaler API v2 is gradually stable in version 1.23. The HorizontalPodAutoscaler v2beta2 API is not recommended. Use the v2 API.
- **PodSecurity** moves to beta, replacing the deprecated PodSecurityPolicy. PodSecurity is an admission controller that enforces pod security standards on pods in the namespace based on specific namespace labels that set the enforcement level. PodSecurity is enabled by default in version 1.23.

Kubernetes 1.22 Release Notes

- Ingresses no longer support networking.k8s.io/v1beta1 and extensions/v1beta1 APIs. If you use the API of an earlier version to manage ingresses, an application cannot be exposed to external services. Use networking.k8s.io/v1.
- CustomResourceDefinitions no longer support the apiextensions.k8s.io/v1beta1 API. If you use the API of an earlier version to create a CRD, the creation will fail, which affects the controller that reconciles this CRD. Use apiextensions.k8s.io/v1.
- ClusterRoles, ClusterRoleBindings, Roles, and RoleBindings no longer support the rbac.authorization.k8s.io/v1beta1 API. If you use the API of an earlier version to manage RBAC resources, application permission control is affected and even cannot work in the cluster. Use rbac.authorization.k8s.io/v1.
- The Kubernetes release cycle is changed from four releases a year to three releases a year.
- StatefulSets support **minReadySeconds**.
- During scale-in, pods are randomly selected and deleted based on the pod UID by default (LogarithmicScaleDown). This feature enhances the randomness of the pods to be deleted and alleviates the problems caused by pod topology spread constraints. For more information, see **KEP-2185** and **issues 96748**.
- The **BoundServiceAccountTokenVolume** feature is stable, which has changed the method of mounting tokens into pods for enhanced token security of the service account. This feature is enabled by default in Kubernetes clusters of v1.21 and later versions.

References

For more details about the performance comparison and function evolution between Kubernetes 1.23 and other versions, see the following documents:

- **Kubernetes v1.23 Release Notes**
- **Kubernetes v1.22 Release Notes**

5.1.2.6 Kubernetes 1.21 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.21.

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/**.
- The **BoundServiceAccountTokenVolume** feature is in beta testing, which has changed the method of mounting tokens into pods for enhanced token security of the service account. This feature will be enabled by default in Kubernetes clusters of v1.21 and later versions.

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.
- 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**

5.1.2.7 Kubernetes 1.19 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.19.

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. Use **apiextensions.k8s.io/v1** instead.
- **apiregistration.k8s.io/v1beta1** has been deprecated. Use **apiregistration.k8s.io/v1** instead.
- **authentication.k8s.io/v1beta1** and **authorization.k8s.io/v1beta1** have been deprecated and will be removed from Kubernetes 1.22. Use **authentication.k8s.io/v1** and **authorization.k8s.io/v1** instead.
- **autoscaling/v2beta1** has been deprecated. Use **autoscaling/v2beta2** instead.
- **coordination.k8s.io/v1beta1** has been deprecated in Kubernetes 1.19 and will be removed from version 1.22. Use **coordination.k8s.io/v1** instead.
- 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.
- **storage.k8s.io/v1beta1** has been deprecated. Use **storage.k8s.io/v1** instead.

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.
    - 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 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**

5.1.2.8 Kubernetes 1.17 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. 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.
- 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**.

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**

5.1.2.9 Kubernetes 1.15 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.15.

Upgrade your Kubernetes clusters before the version EOM for more stable, reliable cluster running.

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.
- 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.

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:

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

5.1.2.10 Kubernetes 1.13 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.13.

Table 5-1 Version 1.13 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

5.1.2.11 Kubernetes 1.11 (EOM) Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.11.

Table 5-2 Version 1.11 description

Kubernetes (CCE Enhanced Version)	Description
v1.11.7-r2	**Highlights**: • Support for GPU V100 is provided. • Support for permission management is provided.
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 is provided. • 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.

Kubernetes (CCE Enhanced Version)	Description
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.

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

5.1.2.12 Kubernetes 1.9 (EOM) and Earlier Versions Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. This section describes the updates in CCE Kubernetes 1.9 and earlier versions.

Table 5-3 Description of v1.9 and earlier versions

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 configured.

Kubernetes (CCE Enhanced Version)	Description
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
v1.9.7-r1	**Highlights**: ● The mechanism for reporting PVC and PersistentVolume (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.

Kubernetes (CCE Enhanced Version)	Description
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
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.

Kubernetes (CCE Enhanced Version)	Description
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.
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 ELB. ● CCE nodes can use SFS for storage. ● BMS clusters support SFS.

Kubernetes (CCE Enhanced Version)	Description
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.
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

Kubernetes (CCE Enhanced Version)	Description
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.

5.1.3 Patch Versions

Version 1.29

Table 5-4 Release notes for the v1.29 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.29.2-r4	v1.29.3	None	The stability of ELB has been improved during upgrades that span across multiple versions.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.29.2-r0	v1.29.3	• CCE ingresses support traffic distribution based on custom HTTP headers. • Scaling priority policies can be configured for third-party workloads. • You can configure a security group for a pod using annotations. This feature is only available for CCE Turbo clusters. • You can bind an existing EIP to a pod. This feature is only available for CCE Turbo clusters.	• An in-progress node drainage can be canceled. • When updating a node pool, you can change its agency name, prefix, and suffix. • Kubernetes labels and taints of a node are retained after the node is reset. • Both the Kubernetes service account token volume projection and the load scaling controller can be configured.	Fixed some security issues.
v1.29.1-r0	v1.29.1	CCE clusters of v1.29 are released for the first time. For more information, see **Kubernetes 1.29 Release Notes**.	None	None

Version 1.28

Table 5-5 Release notes for the v1.28 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.28.6-r4	v1.28.8	None	The stability of ELB has been improved during upgrades that span across multiple versions.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.28.6-r0	v1.28.8	• CCE ingresses support traffic distribution based on custom HTTP headers. • Scaling priority policies can be configured for third-party workloads. • You can configure a security group for a pod using annotations. This feature is only available for CCE Turbo clusters. • You can bind an existing EIP to a pod. This feature is only available for CCE Turbo clusters.	• An in-progress node drainage can be canceled. • When updating a node pool, you can change its agency name, prefix, and suffix. • Kubernetes labels and taints of a node are retained after the node is reset. • Both the Kubernetes service account token volume projection and the load scaling controller can be configured.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.28.4-r0	v1.28.5	Docker can be selected when you create a node.	The configurations of frequently used cluster parameters and node pool parameters are publicly available.	Fixed some security issues.
v1.28.3-r0	v1.28.3	LoadBalancer Services and ingresses allow you to: • Configure SNI. • Enable HTTP/2. • Configure idle timeout, request timeout, and response timeout.	None	Fixed some security issues.
v1.28.2-r0	v1.28.3	• You can configure an ELB blocklist/trustlist for access control when creating a Service or ingress.	None	Fixed some security issues.
v1.28.1-r4	v1.28.3	None	None	Fixed CVE-2024-21626 issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.28.1-r0	v1.28.3	CCE clusters of v1.28 are released for the first time. For more information, see **Kubernetes 1.28 Release Notes**. • The prefix and suffix of a node name can be customized in node pools. • In CCE Turbo clusters, you can create container networks for workloads and specify pod subnets. • LoadBalancer ingresses support gRPC. • LoadBalancer Services allow you to specify a private IP address for a load balancer during Service creation using YAML.	• Accelerated the startup speed for creating a large number of Kata containers in a CCE Turbo cluster. • Improved the stability when Kata containers are repeatedly created or deleted in a CCE Turbo cluster.	None

Version 1.27

NOTICE

dockershim has been removed since Kubernetes v1.24, and Docker is not supported in v1.24 and later versions by default. Use containerd. To migrate nodes from Docker to containerd, follow the operations described in **Migrating Nodes from Docker to containerd**.

Table 5-6 Release notes for the v1.27 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.27.8-r4	v1.27.12	None	The stability of ELB has been improved during upgrades that span across multiple versions.	Fixed some security issues.
v1.27.8-r0	v1.27.12	• CCE ingresses support traffic distribution based on custom HTTP headers. • Scaling priority policies can be configured for third-party workloads. • You can configure a security group for a pod using annotations. This feature is only available for CCE Turbo clusters. • You can bind an existing EIP to a pod. This feature is only available for CCE Turbo clusters.	• An in-progress node drainage can be canceled. • When updating a node pool, you can change its agency name, prefix, and suffix. • Kubernetes labels and taints of a node are retained after the node is reset. • Both the Kubernetes service account token volume projection and the load scaling controller can be configured.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.27.6-r0	v1.27.9	Docker can be selected when you create a node.	The configurations of frequently used cluster parameters and node pool parameters are publicly available.	Fixed some security issues.
v1.27.3-r4	v1.27.4	None	None	Fixed CVE-2024-21626 issues.
v1.27.2-r0	v1.27.2	• Volcano supports node pool affinity scheduling. • Volcano supports workload rescheduling.	None	Fixed some security issues.
v1.27.1-r10	v1.27.2	None	Optimized the events generated during node pool scaling.	Fixed some security issues.
v1.27.1-r0	v1.27.2	CCE clusters of v1.27 are released for the first time. For more information, see Kubernetes 1.27 Release Notes. • Both soft eviction and hard eviction are supported in node pool configurations.	None	None

Version 1.25

NOTICE

All nodes in the CCE clusters of version 1.25, except the ones running EulerOS 2.5, use containerd by default.

Table 5-7 Release notes for the v1.25 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.25.11-r4	v1.25.16	None	The stability of ELB has been improved during upgrades that span across multiple versions.	Fixed some security issues.
v1.25.11-r0	v1.25.16	• CCE ingresses support traffic distribution based on custom HTTP headers. • Scaling priority policies can be configured for third-party workloads. • You can configure a security group for a pod using annotations. This feature is only available for CCE Turbo clusters. • You can bind an existing EIP to a pod. This feature is only available for CCE Turbo clusters.	• An in-progress node drainage can be canceled. • When updating a node pool, you can change its agency name, prefix, and suffix. • Kubernetes labels and taints of a node are retained after the node is reset. • Both the Kubernetes service account token volume projection and the load scaling controller can be configured.	Fixed some security issues.
v1.25.9-r0	v1.25.16	Docker can be selected when you create a node.	The configurations of frequently used cluster parameters and node pool parameters are publicly available.	Fixed some security issues.
v1.25.6-r4	v1.25.10	None	None	Fixed **CVE-2024-21626** issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.25.5-r0	v1.25.5	• Volcano supports node pool affinity scheduling. • Volcano supports workload rescheduling.	None	Fixed some security issues.
v1.25.4-r10	v1.25.5	None	Optimized the events generated during node pool scaling.	Fixed some security issues.
v1.25.4-r0	v1.25.5	• Both soft eviction and hard eviction are supported in node pool configurations. • TMS tags can be added to automatically created EVS disks to facilitate cost management.	None	Fixed some security issues.
v1.25.3-r10	v1.25.5	The timeout interval can be configured for a load balancer.	High-frequency parameters of kube-apiserver are configurable.	Fixed some security issues.
v1.25.3-r0	v1.25.5	None	Enhanced network stability of CCE Turbo clusters when their specifications are modified.	Fixed some security issues.
v1.25.1-r0	v1.25.5	CCE clusters of v1.25 are released for the first time. For more information, see **Kubernetes 1.25 Release Notes**.	None	None

Version 1.23

Table 5-8 Release notes for the v1.23 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.23.16-r4	v1.23.17	None	The stability of ELB has been improved during upgrades that span across multiple versions.	Fixed some security issues.
v1.23.16-r0	v1.23.17	• CCE ingresses support traffic distribution based on custom HTTP headers. • Scaling priority policies can be configured for third-party workloads. • You can configure a security group for a pod using annotations. This feature is only available for CCE Turbo clusters. • You can bind an existing EIP to a pod. This feature is only available for CCE Turbo clusters.	• An in-progress node drainage can be canceled. • When updating a node pool, you can change its agency name, prefix, and suffix. • Kubernetes labels and taints of a node are retained after the node is reset. • Both the Kubernetes service account token volume projection and the load scaling controller can be configured.	Fixed some security issues.
v1.23.14-r0	v1.23.17	None	The configurations of frequently used cluster parameters and node pool parameters are publicly available.	Fixed some security issues.
v1.23.11-r4	v1.23.17	None	None	Fixed CVE-2024-21626 issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.23.10-r0	v1.23.11	• Volcano supports node pool affinity scheduling. • Volcano supports workload rescheduling.	None	Fixed some security issues.
v1.23.9-r10	v1.23.11	None	Optimized the events generated during node pool scaling.	Fixed some security issues.
v1.23.9-r0	v1.23.11	• Both soft eviction and hard eviction are supported in node pool configurations. • TMS tags can be added to automatically created EVS disks to facilitate cost management.	None	Fixed some security issues.
v1.23.8-r10	v1.23.11	The timeout interval can be configured for a load balancer.	High-frequency parameters of kube-apiserver are configurable.	Fixed some security issues.
v1.23.8-r0	v1.23.11	None	• Enhanced Docker reliability during upgrades. • Optimized node time synchronization.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.23.5-r0	v1.23.11	• Fault detection and isolation are supported on GPU nodes. • Security groups can be customized by cluster. • CCE Turbo clusters support ENIs pre-binding by node. • containerd is supported.	• Upgraded the etcd version of the master node to the Kubernetes version 3.5.6. • Optimized scheduling so that pods are evenly distributed across AZs after pods are scaled in. • Optimized the memory usage of kube-apiserver when CRDs are frequently updated.	Fixed some security issues and the following CVE vulnerabilities: • CVE-2022-3294 • CVE-2022-3162 • CVE-2022-3172 • CVE-2021-25749
v1.23.1-r0	v1.23.4	CCE clusters of v1.23 are released for the first time. For more information, see Kubernetes 1.23 Release Notes.	None	None

Version 1.21

Table 5-9 Release notes for the v1.21 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.21.15-r0	v1.21.14	None	The configurations of frequently used cluster parameters and node pool parameters are publicly available.	Fixed some security issues.
v1.21.14-r0	v1.21.14	A PVC can be used to dynamically create and mount an SFS Turbo subdirectory.	None	Fixed some security issues.
v1.21.12-r4	v1.21.14	None	None	Fixed CVE-2024-21626 issues.
v1.21.11-r20	v1.21.14	• Volcano supports node pool affinity scheduling. • Volcano supports workload rescheduling.	None	Fixed some security issues.
v1.21.11-r10	v1.21.14	None	Optimized the events generated during node pool scaling.	Fixed some security issues.
v1.21.11-r0	v1.21.14	• Both soft eviction and hard eviction are supported in node pool configurations. • TMS tags can be added to automatically created EVS disks to facilitate cost management.	None	Fixed some security issues.
v1.21.10-r10	v1.21.14	The timeout interval can be configured for a load balancer.	High-frequency parameters of kube-apiserver are configurable.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.21.10-r0	v1.21.14	None	• Enhanced Docker reliability during upgrades. • Optimized node time synchronization. • Enhanced the stability of the Docker runtime for pulling images after nodes are restarted.	Fixed some security issues.
v1.21.7-r0	v1.21.14	• Fault detection and isolation are supported on GPU nodes. • Security groups can be customized by cluster. • CCE Turbo clusters support ENIs pre-binding by node. • Control plane logs can be collected.	Improved the stability of LoadBalancer Services/ingresses with a large number of connections.	Fixed some security issues and the following CVE vulnerabilities: • CVE-2022-3294 • CVE-2022-3162 • CVE-2022-3172
v1.21.1-r0	v1.21.7	CCE clusters of v1.21 are released for the first time. For more information, see Kubernetes 1.21 Release Notes.	None	None

Version 1.19

Table 5-10 Release notes for the v1.19 patch

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
1.19.16-r84	v1.19.16	None	None	Fixed CVE-2024-21626 issues.
v1.19.16-r60	v1.19.16	• Volcano supports node pool affinity scheduling. • Volcano supports workload rescheduling.	None	Fixed some security issues.
v1.19.16-r50	v1.19.16	None	Optimized the events generated during node pool scaling.	Fixed some security issues.
v1.19.16-r40	v1.19.16	• Both soft eviction and hard eviction are supported in node pool configurations. • TMS tags can be added to automatically created EVS disks to facilitate cost management.	None	Fixed some security issues.
v1.19.16-r30	v1.19.16	The timeout interval can be configured for a load balancer.	High-frequency parameters of kube-apiserver are configurable.	Fixed some security issues.

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.19.16-r20	v1.19.16	None	• Cloud Native 2.0 Networks allow you to specify subnets for a namespace. • Enhanced the stability of the Docker runtime for pulling images after nodes are restarted. • Optimized the performance of CCE Turbo clusters in allocating ENIs if not all ENIs are pre-bound.	Fixed some security issues.
v1.19.16-r4	v1.19.16	• Containers support SFS 3.0 for storage. • Fault detection and isolation are supported on GPU nodes. • Security groups can be customized by cluster. • CCE Turbo clusters support ENIs pre-binding by node.	• Scheduling is optimized on taint nodes. • Enhanced the long-term running stability of containerd when cores are bound. • Improved the stability of LoadBalancer Services/ingresses with a large number of connections. • Optimized the memory usage of kube-apiserver when CRDs are frequently updated.	Fixed some security issues and the following CVE vulnerabilities: • **CVE-2022-3294** • **CVE-2022-3162** • **CVE-2022-3172**

CCE Cluster Patch Version	Kubernetes Version	Feature Updates	Optimization	Vulnerability Fixing
v1.19.16-r0	v1.19.16	None	Enhanced the stability in updating LoadBalancer Services when workloads are upgraded and nodes are scaled in or out.	Fixed some security issues and the following CVE vulnerabilities: • CVE-2021-25741 • CVE-2021-25737
v1.19.10-r0	v1.19.10	CCE clusters of v1.19 are released for the first time. For more information, see Kubernetes 1.19 Release Notes.	None	None

5.2 OS Images

5.2.1 OS Version Support Mechanism

Synchronization Mechanism

CCE cluster node components are updated with the release of CCE cluster versions.

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

Mappings Between Cluster Versions and OS Versions

The following table lists the mappings between released cluster versions and OS versions.

Table 5-11 ECS (VM) OS

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
EulerOS release 2.9	v1.29	√	√	√	4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64
	v1.28	√	√	√	4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64
	v1.27	√	√	√	4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64
	v1.25	√	√	√	4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64
	v1.23	√	√	√	4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64
	v1.21 (end of maintenance)	√	√	√	4.18.0-147.5.1.6.h1071.eulerosv2r9.x86_64
	v1.19 (end of maintenance)	√	√	√	4.18.0-147.5.1.6.h1071.eulerosv2r9.x86_64
EulerOS release 2.9 (Arm)	v1.29	√	√	√	4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64
	v1.28	√	√	√	4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64
	v1.27	√	√	√	4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
	v1.25	√	√	√	4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64
	v1.23	√	√	√	4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64
	v1.21 (end of maintenance)	√	√	√	4.19.90-vhulk2103.1.0.h1060.eulerosv2r9.aarch64
	v1.19 (end of maintenance)	√	√	√	4.19.90-vhulk2103.1.0.h1060.eulerosv2r9.aarch64
CentOS Linux release 7.6	v1.29	√	√	√	3.10.0-1160.108.1.el7.x86_64
	v1.28	√	√	√	3.10.0-1160.108.1.el7.x86_64
	v1.27	√	√	√	3.10.0-1160.108.1.el7.x86_64
	v1.25	√	√	√	3.10.0-1160.108.1.el7.x86_64
	v1.23	√	√	√	3.10.0-1160.108.1.el7.x86_64
	v1.21 (end of maintenance)	√	√	√	3.10.0-1160.92.1.el7.x86_64
	v1.19.16 (end of maintenance)	√	√	√	3.10.0-1160.92.1.el7.x86_64
Ubuntu 22.04	v1.29	√	x	√	5.15.0-92-generic

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
	v1.28	√	x	√	5.15.0-92-generic
	v1.27	√	x	√	5.15.0-92-generic
	v1.25	√	x	√	5.15.0-92-generic
	v1.23	√	x	√	5.15.0-92-generic

Table 5-12 ECS (PM) OS

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
EulerOS release 2.10	v1.29	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64
	v1.28	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64
	v1.27	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64
	v1.25	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64
	v1.23	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64
	v1.21 (end of maintenance)	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
	v1.19.16 (end of maintenance)	√	√	√	4.18.0-147.5.2.15.h1109.eulerosv2r10.x86_64

Table 5-13 BMS OS

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
EulerOS release 2.9 (restricted use. Submit a service ticket to apply for it.)	v1.29	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
	v1.28	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
	v1.27	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
	v1.25	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
	v1.23	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
	v1.21 (end of maintenance)	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64

OS	Cluster Version	CCE Standard Cluster		CCE Turbo Cluster	Latest Kernel
		VPC Network	Tunnel Network	Cloud Native 2.0 Network	
	v1.19 (end of maintenance)	√	√	x	4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64
EulerOS release 2.3 (end of maintenance)	v1.27 or later	x	x	x	None
	v1.25 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64
	v1.23 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64
	v1.21 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64
	v1.19 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64
	v1.17 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64
	v1.15.11 (end of maintenance)	√	√	x	3.10.0-514.41.4.28.h62.x86_64

5.2.2 OS Image Version Release Notes

This section describes the latest updates on CCE cluster OS versions.

For more information, see **Mappings Between Cluster Versions and OS Versions**.

Ubuntu 22.04

Kernel Version	Release Date	Release Note
5.15.0-113-generic	September 2024	The system kernel is updated to fix the CVE-2024-1086 security vulnerability.
5.15.0-92-generic	April 2024	The system kernel is updated to fix security vulnerabilities.
5.15.0-86-generic	December 2023	The system kernel is updated to fix security vulnerabilities.
5.15.0-60-generic	April 2023	The system kernel is updated to fix security vulnerabilities.
5.15.0-53-generic	January 2023	CCE supports Ubuntu 22.04.

EulerOS 2.9

Kernel Version	Release Date	Release Note
4.18.0-147.5.1.6.h1305.eulerosv2r9.x86 64	September 2024	• The system kernel is updated to fix the CVE-2024-1086 security vulnerability. • Fixed the issue that the VM may be suspended due to a kernel defect.
4.18.0-147.5.1.6.h1152.eulerosv2r9.x86_64	April 2024	The system kernel is updated to fix security vulnerabilities.
4.18.0-147.5.1.6.h1071.eulerosv2r9.x86_64	December 2023	The system kernel is updated to fix security vulnerabilities.
4.18.0-147.5.1.6.h1017.eulerosv2r9.x86_64	June 2023	• The system kernel is updated to fix security vulnerabilities. • Resolved the issue that resolution occasionally times out after CoreDNS is upgraded on EulerOS 2.9 nodes in IPVS mode.
4.18.0-147.5.1.6.h841.eulerosv2r9.x86_64	January 2023	The system kernel is updated to fix security vulnerabilities.

Kernel Version	Release Date	Release Note
4.18.0-147.5.1.6.h766.eulerosv2r9.x86_64	December 2022	The system kernel is updated to fix security vulnerabilities.

EulerOS 2.9 (Arm)

Kernel Version	Release Date	Release Note
4.19.90-vhulk2103.1.0.h1263.eulerosv2r9.aarch64	September 2024	• The system kernel is updated to fix the CVE-2024-1086 security vulnerability. • Fixed the issue that the VM may be suspended due to a kernel defect.
4.19.90-vhulk2103.1.0.h1144.eulerosv2r9.aarch64	April 2024	The system kernel is updated to fix security vulnerabilities.
4.19.90-vhulk2103.1.0.h1060.eulerosv2r9.aarch64	December 2023	The system kernel is updated to fix security vulnerabilities.
4.19.90-vhulk2103.1.0.h990.eulerosv2r9.aarch64	June 2023	• The system kernel is updated to fix security vulnerabilities. • Resolved the issue that resolution occasionally times out after CoreDNS is upgraded on EulerOS 2.9 nodes in IPVS mode.
4.19.90-vhulk2103.1.0.h848.eulerosv2r9.aarch64	January 2023	The system kernel is updated to fix security vulnerabilities.

CentOS 7.6

Kernel Version	Release Date	Release Note
3.10.0-1160.119.1.el7.x86 64	September 2024	The system kernel is updated to fix security vulnerabilities.
3.10.0-1160.108.1.el7.x86_64	April 2024	The system kernel is updated to fix security vulnerabilities.

Kernel Version	Release Date	Release Note
3.10.0-1160.92.1.el7.x86_64	December 2023	The system kernel is updated to fix security vulnerabilities.
3.10.0-1160.90.1.el7.x86_64	August 2023	The system kernel is updated to fix security vulnerabilities.
3.10.0-1160.66.1.el7.x86_64	January 2023	• The system kernel is updated to fix security vulnerabilities. • Resolved the issue that the ext4 file system is suspended occasionally when container OOM occurs on a CentOS node.

5.3 Add-on Versions

5.3.1 CoreDNS Release History

Table 5-14 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.29.4	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	CCE clusters 1.29 are supported.	1.10.1
1.28.4	v1.21 v1.23 v1.25 v1.27 v1.28	CCE clusters 1.28 are supported.	1.10.1
1.27.4	v1.19 v1.21 v1.23 v1.25 v1.27	None	1.10.1

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.25.14	v1.19 v1.21 v1.23 v1.25	• Supports association between add-on specifications and cluster specifications. • Synchronized time zones used by the add-on and the node.	1.10.1
1.25.11	v1.19 v1.21 v1.23 v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Upgrades to its community version 1.10.1.	1.10.1
1.25.1	v1.19 v1.21 v1.23 v1.25	CCE clusters 1.25 are supported.	1.8.4
1.23.3	v1.15 v1.17 v1.19 v1.21 v1.23	Regular upgrade of add-on dependencies	1.8.4
1.23.2	v1.15 v1.17 v1.19 v1.21 v1.23	Regular upgrade of add-on dependencies	1.8.4
1.23.1	v1.15 v1.17 v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.	1.8.4
1.17.15	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.	1.8.4

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.17.9	v1.15 v1.17 v1.19	Regular upgrade of add-on dependencies	1.8.4
1.17.7	v1.15 v1.17 v1.19	Updated the add-on to its community version v1.8.4.	1.8.4
1.17.4	v1.17 v1.19	CCE clusters 1.19 are supported.	1.6.5
1.17.3	v1.17	Supported clusters 1.17 and fixed stub domain configuration issues.	1.6.5
1.17.1	v1.17	Clusters 1.17 are supported.	1.6.5

5.3.2 CCE Container Storage (Everest) Release History

Table 5-15 Release history

Add-on Version	Supported Cluster Version	New Feature
2.4.28	v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.
2.3.14	v1.21 v1.23 v1.25 v1.27 v1.28	CCE clusters 1.28 are supported.
2.1.51	v1.19 v1.21 v1.23 v1.25 v1.27	Supported Huawei Cloud EulerOS 2.0.

Add-on Version	Supported Cluster Version	New Feature
2.1.38	v1.19 v1.21 v1.23 v1.25	Supports association between add-on specifications and cluster specifications.
2.1.30	v1.19 v1.21 v1.23 v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Adapts the obsfs package to Ubuntu 22.04.
2.1.13	v1.19 v1.21 v1.23 v1.25	Optimized the performance of creating subpath PVCs in batches for SFS Turbo volumes.
2.1.9	v1.19 v1.21 v1.23 v1.25	• Supported graceful exit of the controller. • CCE clusters 1.25 are supported.
2.0.9	v1.19 v1.21 v1.23	• Rebuilt certain code and architecture of everest to improve its scalability and stability. • Enabled graceful exit. • Supported OBS process monitoring.
1.3.28	v1.19 v1.21 v1.23	• Enabled graceful exit. • Supported OBS process monitoring.
1.3.22	v1.19 v1.21 v1.23	Handled occasional read and write failures after repeated disk mounting.
1.3.20	v1.19 v1.21 v1.23	Handled occasional read and write failures after repeated disk mounting.

Add-on Version	Supported Cluster Version	New Feature
1.3.17	v1.19 v1.21 v1.23	• Updated the rollingUpdates.maxUnavailable of everest-csi-driver from 10 to 10%. • Supported user-defined pod anti-affinity rules. • Counted the maximum number of SCSI volumes that can be managed by the CSI plug-in on a node. • Drivers can be deployed based on customized resource specifications.
1.3.8	v1.23	CCE clusters 1.23 are supported.
1.3.6	v1.23	CCE clusters 1.23 are supported.
1.2.78	v1.15 v1.17 v1.19 v1.21	Supported anti-affinity scheduling of add-on pods on nodes in different AZs.
1.2.70	v1.15 v1.17 v1.19 v1.21	Optimized the performance of creating subpath PVCs in batches for SFS Turbo volumes.
1.2.67	v1.15 v1.17 v1.19 v1.21	• Supported graceful exit of the controller. • Supported OBS process monitoring.
1.2.61	v1.15 v1.17 v1.19 v1.21	• Enabled graceful exit. • Supported OBS process monitoring.
1.2.55	v1.15 v1.17 v1.19 v1.21	Handled occasional read and write failures after repeated disk mounting.
1.2.53	v1.15 v1.17 v1.19 v1.21	Handled occasional read and write failures after repeated disk mounting.

Add-on Version	Supported Cluster Version	New Feature
1.2.51	v1.15 v1.17 v1.19 v1.21	• Updated the rollingUpdates.maxUnavailable of everest-csi-driver from 10 to 10%. • Supported user-defined pod anti-affinity rules. • Counted the maximum number of SCSI volumes that can be managed by the CSI plug-in on a node.
1.2.44	v1.15 v1.17 v1.19 v1.21	• Enterprise projects can be selected for EVS and OBS volumes. • By default, the **enable_noobj_cache** parameter is no longer used for mounting OBS buckets.
1.2.42	v1.15 v1.17 v1.19 v1.21	• Enterprise projects can be selected for EVS and OBS volumes. • By default, the **enable_noobj_cache** parameter is no longer used for mounting OBS buckets.
1.2.30	v1.15 v1.17 v1.19 v1.21	Supported emptyDir.
1.2.28	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.
1.2.27	v1.15 v1.17 v1.19 v1.21	Supports ultra-fast SSD (ESSD) and general-purpose SSD (GPSSD) EVS disks.
1.2.13	v1.15 v1.17 v1.19	Supported EulerOS 2.10.

Add-on Version	Supported Cluster Version	New Feature
1.2.9	v1.15 v1.17 v1.19	• Enhances the reliability of PV resource lifecycle maintenance. • Attach/Detach Controller can be used to attach or detach volumes in clusters 1.19.10. • Improves SFS mounting stability. • Changes the default EVS creation type of a new cluster to SAS.
1.2.5	v1.15 v1.17 v1.19	• Improves the reliability of mounting-related capabilities. • Optimizes the authentication function of using OBS, which requires you to upload the access key. • Improves the compatibility of the everest add-on with FlexVolume volumes. • Improves running stability of the add-on.
1.1.12	v1.15 v1.17	Enhances the reliability of the everest-csi-controller component.
1.1.11	v1.15 v1.17	• Supports security hardening. • Supports third-party OBS storage. • Switches to the EVS query API with better performance. • Uses snapshots to create disks in clone mode by default. • Optimizes and enhances disk status detection and log output for attaching and detaching operations. • Improves the reliability of determining authentication expiration.
1.1.8	v1.15 v1.17	Supported CCE 1.17. If CCE 1.13 is upgraded to 1.15, Everest can take over all functions of the FlexVolume driver.
1.1.7	v1.15 v1.17	Supported CCE 1.17. If CCE 1.13 is upgraded to 1.15, Everest can take over all functions of the FlexVolume driver.

5.3.3 CCE Node Problem Detector Release History

Table 5-16 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.19.1	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.	**0.8.10**
1.18.46	v1.21 v1.23 v1.25 v1.27 v1.28	CCE clusters 1.28 are supported.	**0.8.10**
1.18.22	v1.19 v1.21 v1.23 v1.25 v1.27	None	**0.8.10**
1.18.14	v1.19 v1.21 v1.23 v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Allows adding a taint to a node before the release of a spot ECS for the node to repel a set of pods. • Synchronized time zones used by the add-on and the node.	**0.8.10**

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.18.10	v1.19 v1.21 v1.23 v1.25	• Optimizes the configuration page. • Adds threshold configuration to the DiskSlow check item. • Added threshold configuration to the NTPProblem check item. • Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Supports interruption detection for spot ECSs and evicts pods on nodes before the interruption.	0.8.10
1.17.4	v1.17 v1.19 v1.21 v1.23 v1.25	Optimizes DiskHung check item.	0.8.10
1.17.3	v1.17 v1.19 v1.21 v1.23 v1.25	• The maximum number of taint nodes that can be added to the NPC can be configured by percentage. • Added the ProcessZ check item. • Added the time deviation detection to the NTPProblem check item. • Fixed the processes consistently in the D state (exist in the BMS node).	0.8.10

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.17.2	v1.17 v1.19 v1.21 v1.23 v1.25	• Added the DiskHung check item for disk I/O. • Added the DiskSlow check item for disk I/O. • Added the ProcessD check item. • Added MountPointProblem to check the health of mount points. • To avoid conflicts with the service port range, the default health check listening port is changed to **19900**, and the default Prometheus metric exposure port is changed to **19901**. • Supports clusters 1.25.	**0.8.10**
1.16.4	v1.17 v1.19 v1.21 v1.23	• Adds the beta check item ScheduledEvent to detect cold and live VM migration events caused by host machine exceptions using the metadata API. This check item is disabled by default.	**0.8.10**
1.16.3	v1.17 v1.19 v1.21 v1.23	Adds the function of checking the ResolvConf configuration file.	**0.8.10**
1.16.1	v1.17 v1.19 v1.21 v1.23	• Adds node-problem-controller. Supports basic fault isolation. • Adds the PID, FD, disk, memory, temporary volume pool, and PV pool check items.	**0.8.10**

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.15.0	v1.17 v1.19 v1.21 v1.23	• Hardens check items comprehensively to avoid false positives. • Supports kernel check. Supports reporting of OOMKilled and TaskHung events.	0.8.10
1.14.11	v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.	0.7.1
1.14.5	v1.17 v1.19	Fixes the issue that monitoring metrics cannot be obtained.	0.7.1
1.14.4	v1.17 v1.19	• Supported containerd nodes.	0.7.1
1.14.2	v1.17 v1.19	• CCE clusters 1.19 are supported. • Supported Ubuntu OS and Kata containers.	0.7.1
1.13.8	v1.15.11 v1.17	• Fixes the CNI health check issue on the container tunnel network. • Adjusts resource quotas.	0.7.1
1.13.6	v1.15.11 v1.17	Fixes the issue that zombie processes are not reclaimed.	0.7.1
1.13.5	v1.15.11 v1.17	Added taint tolerance configuration.	0.7.1
1.13.2	v1.15.11 v1.17	Added resource limits and enhanced the detection capability of the cni add-on.	0.7.1

5.3.4 Kubernetes Dashboard Release History

Table 5-17 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
3.0.2	v1.27 v1.28 v1.29	• Supported clusters of v1.27, v1.28, and v1.29. • Updated the add-on to its community version 7.3.2.	7.3.2
2.2.27	v1.21 v1.23 v1.25	Fixed some issues.	2.7.0
2.2.7	v1.21 v1.23 v1.25	None	2.7.0
2.2.5	v1.21 v1.23 v1.25	Synchronized time zones used by the add-on and the node.	2.7.0
2.2.3	v1.21 v1.23 v1.25	None	2.7.0
2.1.1	v1.19 v1.21 v1.23	• CCE clusters 1.23 are supported. • Updated the add-on to its community version v2.5.0.	2.5.0
2.0.10	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.	2.0.0
2.0.4	v1.15 v1.17 v1.19	Adds the default seccomp profile.	2.0.0
2.0.3	v1.15 v1.17 v1.19	CCE clusters 1.15 are supported.	2.0.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.0.2	v1.17 v1.19	CCE clusters 1.19 are supported.	2.0.0
2.0.1	v1.15 v1.17		2.0.0
2.0.0	v1.17	Enables interconnection with CCE v1.17	2.0.0

5.3.5 CCE Cluster Autoscaler Release History

Table 5-18 Release history for add-on adapted to clusters 1.29

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.29.17	v1.29	Optimized events.	1.29.1
1.29.13	v1.29	Clusters 1.29 are supported.	1.29.1

Table 5-19 Release history for add-on adapted to clusters 1.28

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.28.55	v1.28	Optimized events.	1.28.1
1.28.51	v1.28	Optimized the logic for generating alarms when resources in a bode pool are sold out.	1.28.1
1.28.22	v1.28	Fixed some issues.	1.28.1
1.28.20	v1.28	Fixed some issues.	1.28.1
1.28.17	v1.28	Fixed the issue that scale-in cannot be performed when there are custom pod controllers in a cluster.	1.28.1

Table 5-20 Release history for add-on adapted to clusters 1.27

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.27.88	v1.27	Optimized events.	1.27.1
1.27.84	v1.27	Optimized the logic for generating alarms when resources in a bode pool are sold out.	1.27.1
1.27.55	v1.27	Fixed some issues.	1.27.1
1.27.51	v1.27	Fixed some issues.	1.27.1
1.27.14	v1.27	Fixed the scale-in failure of nodes of different specifications in the same node pool and unexpected **PreferNoSchedule** taint issues.	1.27.1

Table 5-21 Release history for add-on adapted to clusters 1.25

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.25.120	v1.25	Optimized events.	1.25.0
1.25.116	v1.25	Optimized the logic for generating alarms when resources in a bode pool are sold out.	1.25.0
1.25.88	v1.25	Fixed some issues.	1.25.0
1.25.84	v1.25	Fixed some issues.	1.25.0
1.25.34	v1.25	• Optimized the method of identifying GPUs and NPUs. • Used the remaining node quota of a cluster for the extra nodes that are beyond the cluster scale.	1.25.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.25.21	v1.25	• Fixed the issue that the autoscaler's least-waste is disabled by default. • Fixed the issue that the node pool cannot be switched to another pool for scaling out after a scale-out failure and the add-on has to restart. • The default taint tolerance duration is changed to 60s. • Fixed the issue that scale-out is still triggered after the scale-out rule is disabled.	1.25.0
1.25.11	v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when AS group resources are insufficient.	1.25.0
1.25.7	v1.25	• CCE clusters 1.25 are supported. • Modified the memory request and limit of a customized flavor. • Enabled to report an event indicating that scaling cannot be performed in a node pool with auto scaling disabled. • Fixed the bug that NPU node scale-out is triggered again during scale-out.	1.25.0

Table 5-22 Release history for add-on adapted to clusters 1.23

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.23.125	v1.23	Optimized events.	1.23.0
1.23.121	1.23	Optimized the logic for generating alarms when resources in a bode pool are sold out.	1.23.0
1.23.95	v1.23	Fixed some issues.	1.23.0
1.23.91	v1.23	Fixed some issues.	1.23.0
1.23.54	v1.23	Fixed the scale-in failure of nodes of different specifications in the same node pool and unexpected **PreferNoSchedule** taint issues.	1.23.0
1.23.44	v1.23	• Optimized the method of identifying GPUs and NPUs. • Used the remaining node quota of a cluster for the extra nodes that are beyond the cluster scale.	1.23.0
1.23.31	v1.23	• Fixed the issue that the autoscaler's least-waste is disabled by default. • Fixed the issue that the node pool cannot be switched to another pool for scaling out after a scale-out failure and the add-on has to restart. • The default taint tolerance duration is changed to 60s. • Fixed the issue that scale-out is still triggered after the scale-out rule is disabled.	1.23.0
1.23.21	v1.23	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when AS group resources are insufficient.	1.23.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.23.17	v1.23	• Supported NPUs and security containers. • Supported node scaling policies without a step. • Fixed a bug so that deleted node pools are automatically removed. • Supported scheduling by priority. • Supported the emptyDir scheduling policy. • Fixed a bug so that scale-in can be triggered on the nodes whose capacity is lower than the scale-in threshold when the node scaling policy is disabled. • Modified the memory request and limit of a customized flavor. • Enabled to report an event indicating that scaling cannot be performed in a node pool with auto scaling disabled. • Fixed the bug that NPU node scale-out is triggered again during scale-out.	1.23.0
1.23.10	v1.23	• Optimized logging. • Supported scale-in waiting so that operations such as data dump can be performed before a node is deleted.	1.23.0
1.23.9	v1.23	Added the **nodenetworkconfigs.crd.yangtse.cni** resource object permission.	1.23.0
1.23.8	v1.23	Fixed the issue that scale-out fails when the number of nodes to be added at a time exceeds the upper limit in periodic scale-outs.	1.23.0
1.23.7	v1.23		1.23.0
1.23.3	v1.23	CCE clusters 1.23 are supported.	1.23.0

Table 5-23 Release history for add-on adapted to clusters 1.21

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.21.114	v1.21	Optimized the logic for generating alarms when resources in a bode pool are sold out.	1.21.0
1.21.89	v1.21	Fixed some issues.	1.21.0
1.21.86	v1.21	Fixed the issue that the node pool auto scaling cannot meet expectations after AZ topology constraints are configured for nodes.	1.21.0
1.21.51	v1.21	Fixed the scale-in failure of nodes of different specifications in the same node pool and unexpected **PreferNoSchedule** taint issues.	1.21.0
1.21.43	v1.21	• Optimized the method of identifying GPUs and NPUs. • Used the remaining node quota of a cluster for the extra nodes that are beyond the cluster scale.	1.21.0
1.21.29	v1.21	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when scaling group resources are insufficient. • Fixed the issue that the node pool cannot be switched to another pool for scaling out after a scale-out failure and the add-on has to restart. • The default taint tolerance duration is changed to 60s. • Fixed the issue that scale-out is still triggered after the scale-out rule is disabled.	1.21.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.21.20	v1.21	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when scaling group resources are insufficient.	1.21.0
1.21.16	v1.21	• Supported NPUs and security containers. • Supported node scaling policies without a step. • Fixed a bug so that deleted node pools are automatically removed. • Supported scheduling by priority. • Supported the emptyDir scheduling policy. • Fixed a bug so that scale-in can be triggered on the nodes whose capacity is lower than the scale-in threshold when the node scaling policy is disabled. • Modified the memory request and limit of a customized flavor. • Enabled to report an event indicating that scaling cannot be performed in a node pool with auto scaling disabled. • Fixed the bug that NPU node scale-out is triggered again during scale-out.	1.21.0
1.21.9	v1.21	• Optimized logging. • Supported scale-in waiting so that operations such as data dump can be performed before a node is deleted.	1.21.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.21.8	v1.21	Added the **nodenetworkconfigs.crd.yangtse.cni** resource object permission.	**1.21.0**
1.21.6	v1.21	Fixed the issue that authentication fails due to incorrect signature in the add-on request retries.	**1.21.0**
1.21.4	v1.21	Fixed the issue that authentication fails due to incorrect signature in the add-on request retries.	**1.21.0**
1.21.2	v1.21	Fixed the issue that auto scaling may be blocked due to a failure in deleting an unregistered node.	**1.21.0**
1.21.1	v1.21	Fixed the issue that the node pool modification in the existing periodic auto scaling rule does not take effect.	**1.21.0**

Table 5-24 Release history for add-on adapted to clusters 1.19

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.19.56	v1.19	Fixed the scale-in failure of nodes of different specifications in the same node pool and unexpected **PreferNoSchedule** taint issues.	**1.19.0**
1.19.48	v1.19	• Optimized the method of identifying GPUs and NPUs. • Used the remaining node quota of a cluster for the extra nodes that are beyond the cluster scale.	**1.19.0**

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.19.35	v1.19	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when scaling group resources are insufficient. • Fixed the issue that the node pool cannot be switched to another pool for scaling out after a scale-out failure and the add-on has to restart. • The default taint tolerance duration is changed to 60s. • Fixed the issue that scale-out is still triggered after the scale-out rule is disabled.	1.19.0
1.19.27	v1.19	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • Fixed the issue that the number of node pools cannot be restored when scaling group resources are insufficient.	1.19.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.19.22	v1.19	• Supported NPUs and security containers. • Supported node scaling policies without a step. • Fixed a bug so that deleted node pools are automatically removed. • Supported scheduling by priority. • Supported the emptyDir scheduling policy. • Fixed a bug so that scale-in can be triggered on the nodes whose capacity is lower than the scale-in threshold when the node scaling policy is disabled. • Modified the memory request and limit of a customized flavor. • Enabled to report an event indicating that scaling cannot be performed in a node pool with auto scaling disabled. • Fixed the bug that NPU node scale-out is triggered again during scale-out.	1.19.0
1.19.14	v1.19	• Optimized logging. • Supported scale-in waiting so that operations such as data dump can be performed before a node is deleted.	1.19.0
1.19.13	v1.19	Fixed the issue that scale-out fails when the number of nodes to be added at a time exceeds the upper limit in periodic scale-outs.	1.19.0
1.19.12	v1.19	Fixed the issue that authentication fails due to incorrect signature in the add-on request retries.	1.19.0
1.19.11	v1.19	Fixed the issue that authentication fails due to incorrect signature in the add-on request retries.	1.19.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.19.9	v1.19	Fixed the issue that auto scaling may be blocked due to a failure in deleting an unregistered node.	1.19.0
1.19.8	v1.19	Fixed the issue that the node pool modification in the existing periodic auto scaling rule does not take effect.	1.19.0
1.19.7	v1.19	Regular upgrade of add-on dependencies	1.19.0
1.19.6	v1.19	Fixed the issue that repeated scale-out is triggered when taints are asynchronously updated.	1.19.0
1.19.3	v1.19	Supports scheduled scaling policies based on the total number of nodes, CPU limit, and memory limit. Fixes other functional defects.	1.19.0

Table 5-25 Release history for add-on adapted to clusters 1.17

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.17.27	v1.17	• Optimized logging. • Fixed a bug so that deleted node pools are automatically removed. • Supported scheduling by priority. • Fixed the issue that taints on newly added nodes are overwritten. • Fixed a bug so that scale-in can be triggered on the nodes whose capacity is lower than the scale-in threshold when the node scaling policy is disabled. • Modified the memory request and limit of a customized flavor. • Enabled to report an event indicating that scaling cannot be performed in a node pool with auto scaling disabled.	1.17.0
1.17.22	v1.17	Optimized logging.	1.17.0
1.17.21	v1.17	Fixed the issue that scale-out fails when the number of nodes to be added at a time exceeds the upper limit in periodic scale-outs.	1.17.0
1.17.19	v1.17	Fixed the issue that authentication fails due to incorrect signature in the add-on request retries.	1.17.0
1.17.17	v1.17	Fixed the issue that auto scaling may be blocked due to a failure in deleting an unregistered node.	1.17.0
1.17.16	v1.17	Fixed the issue that the node pool modification in the existing periodic auto scaling rule does not take effect.	1.17.0
1.17.15	v1.17	Unified resource specification configuration unit.	1.17.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.17.14	v1.17	Fixed the issue that repeated scale-out is triggered when taints are asynchronously updated.	1.17.0
1.17.8	v1.17	Fixed bugs.	1.17.0
1.17.7	v1.17	Added log content and fixed bugs.	1.17.0
1.17.5	v1.17	Supported clusters 1.17 and allowed scaling events to be displayed on the CCE console.	1.17.0
1.17.2	v1.17	Clusters 1.17 are supported.	1.17.0

5.3.6 NGINX Ingress Controller Release History

Table 5-26 Release history for NGINX Ingress Controller 2.6.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.6.5	v1.25 v1.27 v1.28 v1.29	Metric collection can be disabled in the startup command.	1.9.6
2.6.4	v1.25 v1.27 v1.28 v1.29	CCE clusters 1.29 are supported.	1.9.6

Table 5-27 Release history for NGINX Ingress Controller 2.4.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.4.6	v1.25 v1.27 v1.28	• CCE clusters 1.28 are supported. • Supported admission verification. • Supported graceful shutdown and hitless upgrade. • Supported equivalent distribution of add-on instances in multi-AZ deployment mode. • Fixed the CVE-2023-44487 vulnerability.	1.9.3

Table 5-28 Release history for NGINX Ingress Controller 2.3.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.3.5	v1.27	None	1.8.0

Table 5-29 Release history for NGINX Ingress Controller 2.2.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.2.53	v1.23 v1.25	Fixed some issues.	1.5.1
2.2.42	v1.23 v1.25	• Supported graceful shutdown and hitless upgrade. • Supported equivalent distribution of add-on instances in multi-AZ deployment mode.	1.5.1
2.2.7	v1.25	• Synchronized time zones used by the add-on and the node. • Supports IPv4 and IPv6 dual stack.	1.5.1

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.2.3	v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • Added the tolerance time during which the pods with temporary storage volumes cannot be scheduled. • The default taint tolerance duration is changed to 60s.	1.5.1
2.2.1	v1.25	• CCE clusters 1.25 are supported. • Updated the add-on to its community version v1.5.1.	1.5.1

Table 5-30 Release history for NGINX Ingress Controller 2.1.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.1.33	v1.19 v1.21 v1.23	• Supported graceful shutdown and hitless upgrade. • Supported equivalent distribution of add-on instances in multi-AZ deployment mode.	1.2.1
2.1.9	v1.19 v1.21 v1.23	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • The default taint tolerance duration is changed to 60s. • Synchronized time zones used by the add-on and the node. • Supports IPv4 and IPv6 dual stack.	1.2.1

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.1.5	v1.19 v1.21 v1.23	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • The default taint tolerance duration is changed to 60s.	1.2.1
2.1.3	v1.19 v1.21 v1.23	Enables publishService for nginx-ingress.	1.2.1
2.1.1	v1.19 v1.21 v1.23	Updated the add-on to its community version v1.2.1.	1.2.1
2.1.0	v1.19 v1.21 v1.23	• Updated the add-on to its community version v1.2.0. • Fixed the CVE-2021-25746 vulnerability and added rules to disable some annotations values that may cause unauthorized operations. • Fixed the CVE-2021-25745 vulnerability and added rules to disable some access paths that may cause unauthorized operations.	1.2.0

Table 5-31 Release history for NGINX Ingress Controller 2.0.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.0.1	v1.19 v1.21 v1.23	• CCE clusters 1.23 are supported. • Updated the add-on to its community version v1.1.1.	1.1.1

Table 5-32 Release history for NGINX Ingress Controller 1.3.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.3.2	v1.15 v1.17 v1.19 v1.21	• CCE clusters 1.21 are supported. • Updated the add-on to its community version v0.49.3.	0.49.3

Table 5-33 Release history for NGINX Ingress Controller 1.2.x

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.2.6	v1.15 v1.17 v1.19	Adds the default seccomp profile.	0.46.0
1.2.5	v1.15 v1.17 v1.19	Updated the add-on to its community version v0.46.0.	0.46.0
1.2.3	v1.15 v1.17 v1.19	CCE clusters 1.19 are supported.	0.43.0
1.2.2	v1.15 v1.17	Updated the add-on to its community version v0.43.0.	0.43.0

5.3.7 Kubernetes Metrics Server Release History

Table 5-34 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.3.60	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	CCE clusters 1.29 are supported.	0.6.2

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.3.37	v1.21 v1.23 v1.25 v1.27 v1.28	CCE clusters 1.28 are supported.	0.6.2
1.3.12	v1.19 v1.21 v1.23 v1.25 v1.27	None	0.6.2
1.3.8	v1.19 v1.21 v1.23 v1.25	Synchronized time zones used by the add-on and the node.	0.6.2
1.3.6	v1.19 v1.21 v1.23 v1.25	• Supported anti-affinity scheduling of add-on pods on nodes in different AZs. • The default taint tolerance duration is changed to 60s.	0.6.2
1.3.3	v1.19 v1.21 v1.23 v1.25	• CCE clusters 1.25 are supported. • Allowed CronHPA to adjust the number of Deployment pods with the skip scenario supported.	0.6.2
1.3.2	v1.19 v1.21 v1.23 v1.25	CCE clusters 1.25 are supported.	0.6.2
1.2.1	v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.	0.4.4
1.1.10	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.	0.4.4

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.1.4	v1.15 v1.17 v1.19	Unified resource specification configuration unit.	0.4.4
1.1.2	v1.15 v1.17 v1.19	Updated the add-on to its community version v0.4.4.	0.4.4
1.1.1	v1.13 v1.15 v1.17 v1.19	Allows you to change the maximum number of invalid pods to 1.	0.3.7
1.1.0	v1.13 v1.15 v1.17 v1.19	CCE clusters 1.19 are supported.	0.3.7
1.0.5	v1.13 v1.15 v1.17	Updated the add-on to its community version v0.3.7.	0.3.7

5.3.8 CCE Advanced HPA Release History

Table 5-35 Release history

Add-on Version	Supported Cluster Version	New Feature
1.4.3	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.
1.4.2	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	CCE clusters 1.29 are supported.

Add-on Version	Supported Cluster Version	New Feature
1.3.42	v1.21 v1.23 v1.25 v1.27 v1.28	CCE clusters 1.28 are supported.
1.3.14	v1.19 v1.21 v1.23 v1.25 v1.27	CCE clusters 1.27 are supported.
1.3.10	v1.19 v1.21 v1.23 v1.25	Periodic scaling is not affected by the cooldown period.
1.3.7	v1.19 v1.21 v1.23 v1.25	Supported anti-affinity scheduling of add-on pods on nodes in different AZs.
1.3.3	v1.19 v1.21 v1.23 v1.25	• CCE clusters 1.25 are supported. • Allowed CronHPA to adjust the number of Deployment pods with the skip scenario supported.
1.3.1	v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.
1.2.12	v1.15 v1.17 v1.19 v1.21	Optimizes the add-on performance to reduce resource consumption.
1.2.11	v1.15 v1.17 v1.19 v1.21	• Enables the Kubernetes metrics API to obtain resource metrics. • Takes not-ready pods into consideration when calculating resource usage.

Add-on Version	Supported Cluster Version	New Feature
1.2.10	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.
1.2.4	v1.15 v1.17 v1.19	• Regular upgrade of add-on dependencies • Allows custom add-on resource specifications.
1.2.3	v1.15 v1.17 v1.19	Supports ARM64 nodes.
1.2.2	v1.15 v1.17 v1.19	Enhances the health check function.
1.2.1	v1.15 v1.17 v1.19	• CCE clusters 1.19 are supported. • Updates the add-on to a stable version.
1.1.3	v1.15 v1.17	Supports periodic scaling rules.

5.3.9 CCE Cloud Bursting Engine for CCI Release History

Table 5-36 Release history

Add-on Version	Supported Cluster Version	New Feature
1.3.25	v1.17 v1.19 v1.21 v1.23	• Supports Downward API volumes. • Supports Projected volumes. • Supports custom StorageClass.
1.3.19	v1.17 v1.19 v1.21 v1.23	Supports schedule profile.

Add-on Version	Supported Cluster Version	New Feature
1.3.7	v1.17 v1.19 v1.21 v1.23	Clusters 1.21 and 1.23 are supported.
1.2.12	v1.13 v1.15 v1.17 v1.19	• Adds some metrics. • Supports HPA and CustomedHPA. • Enables the hostPath in the pod that is scaled to CCI to be converted to other types of storage. • Fixes an issue that the Kubernetes dashboard cannot run on terminals.
1.2.5	v1.13 v1.15 v1.17 v1.19	• Automatically clears CCI resources that are no longer used by pods. • **Requests** and **Limits** can be set to different values. When CCI is scaled, the number of applied resources is subject to **Limits**. • Fixes the issue that the add-on fails to be uninstalled when the CCI namespace does not exist. • Adds the function of intercepting creation requests when the pod specifications exceed the CCI limit.

Add-on Version	Supported Cluster Version	New Feature
1.2.0	v1.13 v1.15 v1.17 v1.19	• Clusters 1.19 are supported. • Supported SFS and SFS Turbo storage. • Supported CronJobs. • Supported envFrom configuration. • Supports automatic logs dumping. • Shields TCPSocket health check. • Supports resource tags (pod-tag). • Improves performance and reliability. • Resolves some known issues.
1.0.5	v1.13 v1.15 v1.17	Clusters 1.17 are supported.

5.3.10 CCE AI Suite (NVIDIA GPU) Release History

Table 5-37 Release history

Add-on Version	Supported Cluster Version	New Feature
2.6.4	v1.28 v1.29	Updated the isolation logic of GPU cards.
2.5.4	v1.28	Clusters 1.28 are supported.
2.0.46	v1.21 v1.23 v1.25 v1.27	• Supported Nvidia driver 535. • Non-root users can use xGPUs. • Optimized startup logic.
2.0.18	v1.21 v1.23 v1.25 v1.27	Supported Huawei Cloud EulerOS 2.0.

Add-on Version	Supported Cluster Version	New Feature
1.2.28	v1.19 v1.21 v1.23 v1.25	• Optimized the automatic mounting of the GPU driver directory.
1.2.24	v1.19 v1.21 v1.23 v1.25	• Enabled a node pool to configure GPU driver versions. • Supported GPU metric collection.
1.2.20	v1.19 v1.21 v1.23 v1.25	Set the add-on alias to **gpu**.
1.2.17	v1.15 v1.17 v1.19 v1.21 v1.23	Added the nvidia-driver-install pod limits configuration.
1.2.15	v1.15 v1.17 v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.
1.2.11	v1.15 v1.17 v1.19 v1.21	Supported EulerOS 2.10.
1.2.10	v1.15 v1.17 v1.19 v1.21	CentOS supports the GPU driver of the new version.
1.2.9	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.

Add-on Version	Supported Cluster Version	New Feature
1.2.2	v1.15 v1.17 v1.19	Supported the new EulerOS kernel.
1.2.1	v1.15 v1.17 v1.19	• CCE clusters 1.19 are supported. • Added taint tolerance configuration.
1.1.13	v1.13 v1.15 v1.17	Supported kernel-3.10.0-1127.19.1.el7.x86_64 for CentOS 7.6.
1.1.11	v1.15 v1.17	• Allowed users to customize driver addresses to download drivers. • Clusters 1.15 and 1.17 are supported.

5.3.11 CCE AI Suite (Ascend NPU) Release History

Table 5-38 Release history

Add-on Version	Supported Cluster Version	New Feature
2.1.7	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.
2.1.5	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	• CCE clusters 1.29 are supported. • Added silent fault codes.

Add-on Version	Supported Cluster Version	New Feature
2.0.5	v1.21 v1.23 v1.25 v1.27 v1.28	• CCE clusters 1.28 are supported. • Supported liveness probe.
1.2.14	v1.19 v1.21 v1.23 v1.25 v1.27	Supported NPU monitoring.
1.2.6	v1.19 v1.21 v1.23 v1.25	Supports automatical installation of NPU drivers.
1.2.5	v1.19 v1.21 v1.23 v1.25	Supports automatical installation of NPU drivers.
1.2.4	v1.19 v1.21 v1.23 v1.25	CCE clusters 1.25 are supported.
1.2.2	v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.
1.2.1	v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.
1.1.8	v1.15 v1.17 v1.19 v1.21	CCE clusters 1.21 are supported.
1.1.2	v1.15 v1.17 v1.19	Adds the default seccomp profile.

Add-on Version	Supported Cluster Version	New Feature
1.1.1	v1.15 v1.17 v1.19	CCE clusters 1.15 are supported.
1.1.0	v1.17 v1.19	CCE clusters 1.19 are supported.
1.0.8	v1.13 v1.15 v1.17	Adapts to the D310 C75 driver.
1.0.6	v1.13 v1.15 v1.17	Supports the Ascend C75 driver.
1.0.5	v1.13 v1.15 v1.17	Allows containers to use Huawei NPU add-ons.
1.0.3	v1.13 v1.15 v1.17	Allows containers to use Huawei NPU add-ons.

5.3.12 Volcano Scheduler Release History

Table 5-39 Release history

Add-on Version	Supported Cluster Version	New Feature
1.13.3	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	• Supported scale-in of customized resources based on node priorities. • Optimized the association between preemption and node scale-out.

Add-on Version	Supported Cluster Version	New Feature
1.13.1	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	Optimized scheduler memory usage.
1.12.18	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	• CCE clusters 1.29 are supported. • The preemption function is enabled by default.
1.11.21	v1.19.16 v1.21 v1.23 v1.25 v1.27 v1.28	• Supported Kubernetes 1.28. • Supported load-aware scheduling. • Updated image OS to HCE 2.0. • Optimized CSI resource preemption. • Optimized load-aware rescheduling. • Optimized preemption in hybrid deployment scenarios.
1.11.6	v1.19.16 v1.21 v1.23 v1.25 v1.27	• Supported Kubernetes 1.27. • Supported rescheduling. • Supported affinity scheduling of nodes in the node pool. • Optimized the scheduling performance.
1.10.7	v1.19.16 v1.21 v1.23 v1.25	Fixes the issue that the local PV add-on fails to calculate the number of pods pre-bound to the node.

Add-on Version	Supported Cluster Version	New Feature
1.10.5	v1.19.16 v1.21 v1.23 v1.25	• The volcano agent supports resource oversubscription. • Adds the verification admission for GPUs. The value of **nvidia.com/gpu** must be less than 1 or a positive integer, and the value of **volcano.sh/gpu-core.percentage** must be less than 100 and a multiple of 5. • Fixes the issue that pod scheduling is slow after PVC binding fails. • Fixes the issue that newly added pods cannot run when there are terminating pods on a node for a long time. • Fixes the issue that volcano restarts when creating or mounting PVCs to pods.
1.9.1	v1.19.16 v1.21 v1.23 v1.25	• Fixes the issue that the counting pipeline pod of the networkresource add-on occupies supplementary network interfaces (Sub-ENI). • Fixes the issue where the binpack add-on scores nodes with insufficient resources. • Fixes the issue of processing resources in the pod with unknown end status. • Optimizes event output. • Supports HA deployment by default.
1.7.2	v1.19.16 v1.21 v1.23 v1.25	• Adapts to clusters 1.25. • Improves scheduling performance of volcano.
1.7.1	v1.19.16 v1.21 v1.23 v1.25	Adapts to clusters 1.25.
1.4.7	v1.15 v1.17 v1.19 v1.21	Deletes the pod status Undetermined to adapt to cluster Autoscaler.

Add-on Version	Supported Cluster Version	New Feature
1.4.5	v1.17 v1.19 v1.21	Changes the deployment mode of volcano-scheduler from **statefulset** to **deployment**, and fixes the issue that pods cannot be automatically migrated when the node is abnormal.
1.4.2	v1.15 v1.17 v1.19 v1.21	• Resolves the issue that cross-GPU allocation fails. • Supports the updated EAS API.
1.3.7	v1.15 v1.17 v1.19 v1.21	• Supports hybrid deployment of online and offline jobs and resource oversubscription. • Optimizes the scheduling throughput for clusters. • Fixes the issue where the scheduler panics in certain scenarios. • Fixes the issue that the volumes.secret verification of the volcano job in the CCE clusters 1.15 fails. • Fixes the issue that jobs fail to be scheduled when volumes are mounted.
1.3.3	v1.15 v1.17 v1.19 v1.21	Fixes the scheduler crash caused by GPU exceptions and the privileged init container admission failure.
1.3.1	v1.15 v1.17 v1.19	• Upgrades the volcano framework to the latest version. • Supported Kubernetes 1.19. • Adds the numa-aware add-on. • Fixes the deployment scaling issue in the multi-queue scenario. • Adjusts the algorithm add-on enabled by default.

Add-on Version	Supported Cluster Version	New Feature
1.2.5	v1.15 v1.17 v1.19	• Fixes the OutOfcpu issue in some scenarios. • Fixes the issue that pods cannot be scheduled when some capabilities are set for a queue. • Makes the log time of the volcano component consistent with the system time. • Fixes the issue of preemption between multiple queues. • Fixes the issue that the result of the ioaware add-on does not meet the expectation in some extreme scenarios. • Supports hybrid clusters.
1.2.3	v1.15 v1.17 v1.19	• Fixes the training task OOM issue caused by insufficient precision. • Fixes the GPU scheduling issue in CCE v1.15 and later versions. Rolling upgrade of CCE versions during task distribution is not supported. • Fixes the issue where the queue status is unknown in certain scenarios. • Fixes the issue where a panic occurs when a PVC is mounted to a job in a specific scenario. • Fixes the issue that decimals cannot be configured for GPU jobs. • Adds the ioaware add-on. • Adds the ring controller.

5.3.13 CCE Secrets Manager for DEW Release History

Table 5-40 Release history

Add-on Version	Supported Cluster Version	New Feature
1.1.3	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.
1.1.2	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29	• CCE clusters 1.29 are supported. • Secrets can be created during CSMS secret synchronization.
1.0.31	v1.21 v1.23 v1.25 v1.27 v1.28	• CCE clusters 1.27 are supported. • CCE clusters 1.28 are supported.
1.0.6	v1.19 v1.21 v1.23 v1.25	None
1.0.3	v1.19 v1.21 v1.23 v1.25	CCE clusters 1.25 are supported.
1.0.2	v1.19 v1.21 v1.23	CCE clusters 1.23 are supported.
1.0.1	v1.19 v1.21	Actively detects SecretProviderClass object changes.

5.3.14 CCE Network Metrics Exporter Release History

Table 5-41 Release history

Add-on Version	Supported Cluster Version	New Feature
1.4.7	v1.23 v1.25 v1.27 v1.28 v1.29	Fixed some issues.
1.4.5	v1.23 v1.25 v1.27 v1.28 v1.29	• Supported pod-based UDP, TCP drop, and TCP connect fail monitoring for common containers. • Supported flow-based UDP and TCP drop monitoring for common containers. • Supported Huawei Cloud EulerOS 2.0 on x86 or Arm. • CCE clusters 1.29 are supported.
1.3.8	v1.23 v1.25 v1.27 v1.28	• Supported pod-based IP address and TCP monitoring of containers. • Supported flow-based IP address and TCP monitoring of containers. • CCE clusters 1.27 are supported. • CCE clusters 1.28 are supported.
1.2.27	v1.19 v1.21 v1.23 v1.25	None
1.2.4	v1.19 v1.21 v1.23 v1.25	• Added the description that only EulerOS is supported.

Add-on Version	Supported Cluster Version	New Feature
1.2.2	v1.19 v1.21 v1.23 v1.25	• Supported VPC network health checks in a local pod.
1.1.8	v1.19 v1.21 v1.23 v1.25	• CCE clusters 1.25 are supported.
1.1.6	v1.19 v1.21 v1.23	None
1.1.5	v1.19 v1.21 v1.23	• Optimizes liveness health check.
1.1.2	v1.19 v1.21 v1.23	• Supports wide matching of operating system types.
1.0.1	v1.19 v1.21	• Supports traffic statistics persistence and local socket communications.

5.3.15 NodeLocal DNSCache Release History

Table 5-42 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.6.8	v1.23 v1.25 v1.27 v1.28 v1.29	Optimized custom injection configuration experience.	1.22.20

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.6.7	v1.23 v1.25 v1.27 v1.28 v1.29	Added custom injection configurations.	1.22.20
1.5.0	v1.21 v1.23 v1.25 v1.27 v1.28	• CCE clusters 1.28 are supported. • Supported equivalent distribution of add-on instances in multi-AZ deployment mode. • Changed the basic image OS of the add-on pods to Huawei Cloud EulerOS 2.0.	1.22.20
1.4.0	v1.19 v1.21 v1.23 v1.25 v1.27	Resolved the issue that CCI pods took an excessively long time to access the Internet.	1.22.20
1.3.1	v1.19 v1.21 v1.23 v1.25	• Enables automatic DNS Config injection for namespaces. • Synchronized time zones used by the add-on and the node.	1.22.20
1.2.7	v1.19 v1.21 v1.23 v1.25	Supported anti-affinity scheduling of add-on pods on nodes in different AZs.	1.21.1
1.2.4	v1.19 v1.21 v1.23 v1.25	CCE clusters 1.25 are supported.	1.21.1
1.2.2	v1.19 v1.21 v1.23	Supports customized NodeLocal DNSCache specifications.	1.21.1

5.3.16 Cloud Native Cluster Monitoring Release History

Table 5-43 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
3.11.0	v1.21 v1.23 v1.25 v1.27 v1.28 v1.29 v1.30	CCE clusters 1.30 are supported.	2.37.8
3.7.3	v1.17 v1.19 v1.21 v1.23 v1.25	None	2.35.0
3.7.2	v1.17 v1.19 v1.21 v1.23 v1.25	Supported collection of Virtual Kubelet pod metrics.	2.35.0
3.7.1	v1.17 v1.19 v1.21 v1.23 v1.25	Supports PrometheusAgent.	2.35.0
3.6.6	v1.17 v1.19 v1.21 v1.23 v1.25	• Grafana is upgraded to 7.5.17. • Supported containerd nodes.	2.35.0
3.5.1	v1.17 v1.19 v1.21 v1.23	None	2.35.0

Add-on Version	Supported Cluster Version	New Feature	Community Version
3.5.0	v1.17 v1.19 v1.21 v1.23	Updated the add-on to its community version v2.35.0.	2.35.0

5.3.17 Cloud Native Logging Release History

Table 5-44 Release history

Add-on Version	Supported Cluster Version	New Feature
1.3.2	v1.17 v1.19 v1.21 v1.23 v1.25	Supports reporting Kubernetes events to AOM.
1.3.0	v1.17 v1.19 v1.21 v1.23 v1.25	Clusters 1.25 are supported.
1.2.3	v1.17 v1.19 v1.21 v1.23	None
1.2.2	v1.17 v1.19 v1.21 v1.23	log-agent is a cloud native log collection add-on built on open source Fluent Bit and OpenTelemetry and supports CRD-based log collection policies. It collects and forwards standard container output logs, container file logs, node logs, and Kubernetes event logs in a cluster following your rules.

5.3.18 CCE Cluster Backup & Recovery (End of Maintenance) Release History

Table 5-45 Release history

Add-on Version	Supported Cluster Version	New Feature
1.2.0	v1.15 v1.17 v1.19 v1.21	• Supports EulerOS 2.0 (SP5, SP9). • Supports security hardening. • Optimizes functions.

5.3.19 Kubernetes Web Terminal (End of Maintenance) Release History

Table 5-46 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.1.12	v1.15 v1.17 v1.19 v1.21	• CCE clusters 1.21 are supported.	**0.6.6**
1.1.6	v1.15 v1.17 v1.19	• Adds the default seccomp profile.	**0.6.6**
1.1.5	v1.15 v1.17 v1.19	• CCE clusters 1.15 are supported.	**0.6.6**
1.1.3	v1.17 v1.19	• CCE clusters 1.19 are supported.	**0.6.6**
1.0.6	v1.15 v1.17	• Adds pod security policies.	**0.6.6**

Add-on Version	Supported Cluster Version	New Feature	Community Version
1.0.5	v1.9 v1.11 v1.13 v1.15 v1.17	• Clusters 1.17 are supported.	0.6.6

5.3.20 Prometheus (End of Maintenance) Release History

Table 5-47 Release history

Add-on Version	Supported Cluster Version	New Feature	Community Version
2.23.32	v1.17 v1.19 v1.21	None	2.10.0
2.23.31	v1.15	• CCE clusters 1.15 are supported.	2.10.0
2.23.30	v1.17 v1.19 v1.21	• CCE clusters 1.21 are supported.	2.10.0
2.21.14	v1.17 v1.19 v1.21	• CCE clusters 1.21 are supported.	2.10.0
2.21.12	v1.15	• CCE clusters 1.15 are supported.	2.10.0
2.21.11	v1.17 v1.19	• CCE clusters 1.19 are supported.	2.10.0
1.15.1	v1.15 v1.17	• The add-on is a monitoring system and time series library.	2.10.0

