Cloud Container Engine Autopilot

User Guide

Issue	01
Date	2025-02-27

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

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 Clusters

1.1 Cluster Overview

1.1.1 Basic Cluster Information

Kubernetes is an open-source container orchestration engine for automating deployment, scaling, and management of containerized applications.

For developers, Kubernetes is a cluster operating system. Kubernetes provides service discovery, scaling, load balancing, self-healing, and even leader election, freeing developers from infrastructure-related configurations.

Cluster Network

A cluster network can be divided into two parts:

- Container network: assigns IP addresses to containers in each pod of a cluster for communication. Currently, only **Cloud Native 2.0 network model** is supported.
- Service network: A Service is a Kubernetes object used to access containers. Each Service has a static IP address.

When you create a cluster, select a proper CIDR block for each network. Ensure that the CIDR blocks do not conflict with each other and have sufficient available IP addresses.

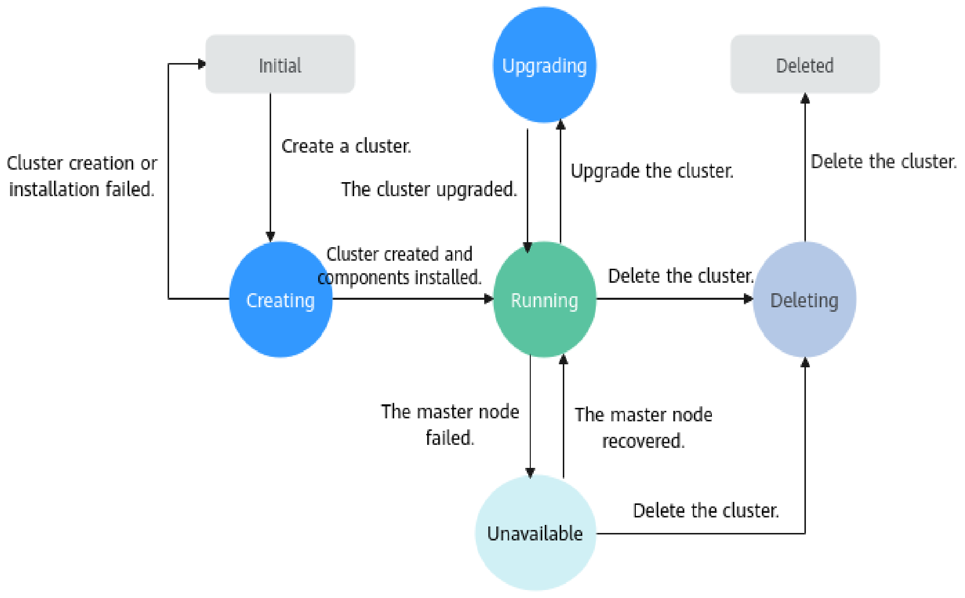
Cluster Lifecycle

Table 1-1 Cluster status

Status	Description
Creating	A cluster is being created and is requesting for cloud resources.
Running	A cluster is running properly.

Status	Description
Upgrading	A cluster is being upgraded.
Unavailable	A cluster is unavailable.
Deleting	A cluster is being deleted.

Figure 1-1 Cluster status transition



CCE Autopilot Cluster Version Release Notes

The Kubernetes community continuously introduces new features and fixes vulnerabilities. CCE Autopilot will periodically release supported Kubernetes versions and provides feature updates and maintenance based on the Kubernetes community versions. For details, see the following sections:

- **Kubernetes Version Release Notes**
- **CCE Autopilot Cluster Patch Release History**

To enable interoperability from one Kubernetes installation to the next, you must upgrade your Kubernetes clusters before the maintenance period ends.

1.1.2 Kubernetes Version Release Notes

1.1.2.1 Kubernetes 1.31 Release Notes (OBT)

CCE Autopilot has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. CCE Autopilot allows you to create Kubernetes clusters 1.31. This topic describes the changes made in Kubernetes 1.31.

Indexes

New and Enhanced Features

Kubernetes 1.31

- Start ordinal of a StatefulSet

  StatefulSet start ordinal moved to the General Availability (GA) state in Kubernetes 1.31. By default, each pod in a StatefulSet is assigned an integer ordinal from 0. With this feature, you can configure a start ordinal for each pod. For details, see **Start ordinal**.
- Elastic indexed jobs

  Elastic indexed jobs moved to GA in Kubernetes 1.31. You can scale indexed Jobs up or down by modifying fields **.spec.completions** and **.spec.parallelism**. For details, see **Elastic Indexed Jobs**.
- Pod failure policy

  Pod failure policies moved to GA in Kubernetes 1.31. This feature helps you handle pod failures based on the container exit codes and pod conditions. For details, see **Pod failure policy**.
- Pod disruption conditions

  Pod disruption conditions moved to GA in Kubernetes 1.31. The new **DisruptionTarget** condition indicates that the pod is about to be deleted due to a disruption. The **reason** field indicates one of the following reasons for the pod termination: preempted by a pod with a higher priority, the pod has been cleared due to node deletion, or the pod is terminated by kubelet. When a pod is created using a job or CronJob, you can use these pod disruption conditions as part of your job's **pod failure policy** to define the action when a pod is abnormal. For details, see **Pod disruption conditions**.
- Selectable fields for custom resources

  Selectable fields for custom resources moved to Beta in Kubernetes 1.31. You can specify the **selectableFields** field of a **CustomResourceDefinition** to define which other fields in a custom resource may be used in field selectors. Field selectors can then be used to get only resources by filtering List, Watch, and DeleteCollection requests. For details, see **Selectable fields for custom resources**.
- Job success policy

  Job success policies moved to Beta in Kubernetes 1.31. When creating an indexed Job, you can define when a job can be declared as succeeded, based on the number of pods that succeeded. For details, see **Success policy**.
- ServiceAccountTokenNodeBinding

  ServiceAccountTokenNodeBinding moved to Beta in Kubernetes 1.31. You can create a service account token that is directly bound to a node. The token defines the node information and verifies whether the node is available. The token will be valid until it expires or either the associated node is deleted. For details, see **Manually create an API token for a ServiceAccount**.

Kubernetes 1.30

- Webhook matching expression

  The Webhook matching expression feature moved to GA. This feature enables admission webhooks to be matched based on specific conditions, providing control over the triggering conditions of the webhooks in a more precise granularity. For details, see **Dynamic Admission Control**.

- Validating admission policies

  Validating admission policies moved to GA. This feature allows you to declare the validating admission policies of resources using Common Expression Language (CEL). For details, see **Validating Admission Policy**.

- Horizontal pod auto scaling based on container resource metrics

  The horizontal pod auto scaling feature based on container resource metrics advanced to GA. This feature allows HPA to configure auto scaling based on the resource usage of each container within a pod, rather than just the overall resource usage of the pod. This makes it easier to set scaling thresholds for the most critical containers in a pod. For details, see **Container resource metrics**.

- Legacy ServiceAccount token cleaner

  The legacy ServiceAccount token cleaner moved to GA. 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**, with the current date as the value. 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**.

Kubernetes 1.29

- Load balancer IP mode for Services

  The load balancer IP mode is a new alpha feature. 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 to the load balancer 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 that the load balancer is not used to distribute traffic. For details, see **Load Balancer IP Mode for Services**.

- nftables proxy mode

  The nftables proxy mode is a new alpha feature. 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

  The garbage collection for unused container images is a new alpha feature. 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**

  PodLifecycleSleepAction is a new alpha feature. 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**

  KubeletSeparateDiskGC is a new alpha feature. With this feature enabled, container images and containers can be garbage collected even if they are on separate file systems. For details, see **Feature Gates**.

- **ClusterTrustBundle** projected volumes

  clusterTrustBundle projected volumes are new alpha features. 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**.

- Image pull per runtime class

  Image pull per runtime class is a new alpha feature. 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**.

- **PodReadyToStartContainers** condition

  The **PodReadyToStartContainers** moved 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 **PodReadyToStartContainersCondition Moved to Beta**.

- Job-related features
  - Pod replacement policy

    The pod replacement policy feature moved 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

    The backoff limit per index moved 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.

For details, see **Jobs**.

- Native sidecar containers

  Native sidecar containers moved 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

  Legacy ServiceAccount token cleaner moved 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**, with the current date as the value. 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**

  DevicePluginCDIDevices moved 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. For details, see **Device Plugins**.

- **PodHostIPs**

  The **PodHostIPs** feature moved 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. For details, see **Feature Gates (removed)**.

- API priority and fairness (APF)

  APF moved 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**

  APIListChunking moved 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. For details, see **Feature Gates**.

- lastPhaseTransitionTime of PersistentVolume (PV)

  lastPhaseTransitionTime moved 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**

  ReadWriteOncePod moved 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**

  CSINodeExpandSecret moved to GA. This feature allows secret authentication data to be passed to a CSI driver for use when a node is added. For details, see **Volumes**.
- CEL-based CustomResourceDefinition (CRD) verification

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

API Changes and Removals

Kubernetes 1.31

- In Kubernetes 1.31, the **kubectl exec [POD] [COMMAND]** command cannot be executed without a **--** separator. In this case, you need to run **kubectl exec [POD] -- [COMMAND]**.
- In Kubernetes 1.31, if **caBundle** is not empty but the value is invalid or it does not define any CA certificate, the CRD does not provide services. If **caBundle** is set to a valid value, it remains unchanged if updated. Attempting direct updates results in an "invalid field value" error, ensuring uninterrupted CRD services.

Kubernetes 1.30

- kubectl replaces **prune-whitelist** with **prune-allowlist** in the **apply** command.
- SecurityContextDeny, which has been deprecated in Kubernetes 1.27, is replaced by **Pod Security Admission**.

Kubernetes 1.29

- The time zone of a newly created CronJob cannot be configured using **TZ** or **CRON_TZ** in **.spec.schedule**. Use **.spec.timeZone** instead. CronJobs 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**.
- **ServiceCIDR** is added. It allows you to specify a CIDR block for a ClusterIP Service.

- 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**.
- CEL expressions are supported by **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 upgraded 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** is not allowed in the **v1** version in 1.29. In 1.30, explicit **0** 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**.
- 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.

References

For more details about the performance comparison and function evolution between Kubernetes 1.31 and other versions, see the following documents:

- **Kubernetes v1.31 Release Notes**
- **Kubernetes v1.30 Release Notes**
- **Kubernetes v1.29 Release Notes**

1.1.2.2 Kubernetes 1.28 Release Notes

CCE has passed the Certified Kubernetes Conformance Program and is a certified Kubernetes offering. CCE allows you to create clusters of Kubernetes 1.28. This topic 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 the instructions in **GitHub**.
- The Ceph FS in-tree volume 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 volume 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.

References

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

1.1.2.3 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 topic describes the changes made in Kubernetes 1.27.

Indexes

New Features

- 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**.

Deprecations and Removals

- 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.

References

For more details about the performance comparison and function evolution between Kubernetes 1.27 and other versions, see **Kubernetes v1.27 Release Notes**.

1.1.3 CCE Autopilot Cluster Patch Release History

Indexes

v1.31 (OBT)

Table 1-2 Release notes for the v1.31 patch

CCE Autopilot Cluster Patch Version	Kubernetes Version	Feature Update	Enhancement	Vulnerability Fixing
v1.31.1-r0	v1.31.1	Supported cluster v1.31. For more information, see **Kubernetes 1.31 Release Notes**.	-	Fixed some security issues.

v1.28

Table 1-3 Release notes for the v1.28 patch

CCE Autopilot Cluster Patch Version	Kubernetes Version	Feature Update	Enhancement	Vulnerability Fixing
v1.28.7-r0	v1.28.3	• Allowed namespaces or workloads to have security groups and subnets bound. • Allowed creation of subdirectories for PVs that are dynamically provisioned from SFS Turbo file systems.	-	Fixed some security issues.
v1.28.6-r0	v1.28.3	• Supported EVS volumes. • Allowed namespaces or workloads to have security groups and subnets bound. • Supported custom disk storage capacity.	• Enhanced cluster monitoring for better O&M. • Reduced pod startup time. • Optimized the performance of large-scale clusters.	Fixed some security issues.

CCE Autopilot Cluster Patch Version	Kubernetes Version	Feature Update	Enhancement	Vulnerability Fixing
v1.28.5-r0	v1.28.3	• Supported APM probes during workload creation. • Supported access to kube-apiserver using a private IP address.	-	Fixed some security issues.
v1.28.4-r0	v1.28.3	• Custom metrics, such as network and disk metrics, can be used create HPA policies. • kube-apiserver can be accessed from a public network.	When YAML is used to create an application, the parameters that are not supported by CCE Autopilot and do not affect functionality were automatically ignored.	Fixed some security issues.
v1.28.3-r0	v1.28.3	• Hosted the Everest storage add-on at the backend. • Supported OBS volumes.	-	Fixed some security issues.
v1.28.2-r0	v1.28.3	• Supported CronHPA policies. • Supported the security context configuration for pods.	-	Fixed some security issues.
v1.28.1-r10	v1.28.3	Supported cluster v1.28. For more information, see **Kubernetes 1.28 Release Notes**.	-	-

v1.27

Table 1-4 Release notes for the v1.27 patch

CCE Autopilot Cluster Patch Version	Kubernetes Version	Feature Update	Enhancement	Vulnerability Fixing
v1.27.9-r0	v1.27.4	• Allowed namespaces or workloads to have security groups and subnets bound. • Allowed creation of subdirectories for PVs that are dynamically provisioned from SFS Turbo file systems.	-	Fixed some security issues.
v1.27.8-r0	v1.27.4	• Supported EVS volumes. • Allowed namespaces or workloads to have security groups and subnets bound. • Supported custom disk storage capacity.	• Enhanced cluster monitoring for better O&M. • Reduced pod startup time. • Optimized the performance of large-scale clusters.	Fixed some security issues.
v1.27.7-r0	v1.27.4	• Supported APM probes during workload creation. • Supported access to kube-apiserver using a private IP address.	-	Fixed some security issues.
v1.27.6-r0	v1.27.4	• Custom metrics, such as network and disk metrics, can be used create HPA policies. • kube-apiserver can be accessed from a public network.	When YAML is used to create an application, the parameters that are not supported by CCE Autopilot and do not affect functionality were automatically ignored.	Fixed some security issues.
v1.27.5-r0	v1.27.4	• Hosted the Everest storage add-on at the backend. • Supported OBS volumes.	-	Fixed some security issues.

CCE Autopilot Cluster Patch Version	Kubernetes Version	Feature Update	Enhancement	Vulnerability Fixing
v1.27.4-r0	v1.27.4	• Supported CronHPA policies. • Supported the security context configuration for pods.	-	Fixed some security issues.
v1.27.3-r30	v1.27.4	-	Supported one-click configuration of monitoring alarms.	Fixed some security issues.
v1.27.3-r20	v1.27.4	• Supported NGINX Ingress Controller. • Supported the Cloud Native Cluster Monitoring and Cloud Native Log Collection add-ons to monitor application metrics and collect application logs. • Launched the application template market. • Supported CustomResourceDefinitions (CRDs). • Interconnected with CloudShell.	Optimized the function of creating a NAT gateway by default during cluster creation so that applications can access the public network.	Fixed some security issues.
v1.27.3-r10	v1.27.4	Supported cluster v1.27. For more information, see **Kubernetes 1.27 Release Notes**.	-	-

1.2 Buying a CCE Autopilot Cluster

A CCE Autopilot cluster runs on Cloud Container Instance (CCI) and provides Kubernetes-native extended APIs, allowing you to run containers without creating or managing servers. You pay only for the resources used by your applications.

Constraints

- After a cluster is created, the following items cannot be changed:
  - Cluster type
  - Network configuration of the cluster, such as the VPC, pod subnet, Service CIDR block, and kube-proxy (request forwarding) settings.
- When using a CCE Autopilot cluster, pay attention to the quotas of related resources. The following table lists the resources required by each cluster.

Table 1-5 Cluster resource usage details

Service	Quota Item	Minimum Usage	Minimum Usage	Region Limits	Quota Increase
CCE	Cluster	1	-	Maximum number of clusters that can be created by each account in a region: 50	Increase the quota on the **My Quotas** page.
VPC	VPC	1 per cluster	Select one VPC for each cluster to provide an isolated, private virtual network environment for the cluster.	Maximum number of VPCs that can be created by each account in a region: 5	
	Subnet	1 per cluster	At least one subnet must be selected for each cluster to allocate container IP addresses. By default, the cluster control plane occupies eight IP addresses for control plane deployment and interconnection with external services.	Maximum number of subnets that can be created by each account in a region: 50	
	Security group	2 per cluster	Two security groups are automatically created for each cluster for network access control of the cluster control plane and elastic network interfaces.	Maximum number of security groups that can be created by each account in a region: 100	

Service	Quota Item	Minimum Usage	Minimum Usage	Region Limits	Quota Increase
	Security group rules	7 per cluster	Seven security group rules are automatically added for each cluster to allow traffic over specified ports and ensure normal network communication in the cluster.	Maximum number of security groups rules that can be added by each account in a region: 1,000	
VPC Endpoint	Endpoint	3 per cluster	Reserve at least three endpoints for each cluster so that the cluster can access peripheral services such as SWR and OBS.	Maximum number of VPC endpoints that can be created by each account in a region: 50	
Domain Name Service (DNS)	Private zone	2 per cluster	Each cluster requires at least two private zones for normal communication within the cluster or across clusters.	Maximum number of private zones that can be created by each account in a region: 50	
	Record set	6 per cluster	Each cluster requires at least six DNS record sets for mapping specified domain names to IP addresses or other domain names in the cluster.	Maximum number of record sets that can be added by each account in a region: 500	

Step 1: Log In to the CCE Console

Step 1 Log in to the **CCE console**.

Step 2 On the **Clusters** page, click **Buy Cluster** in the upper right corner.

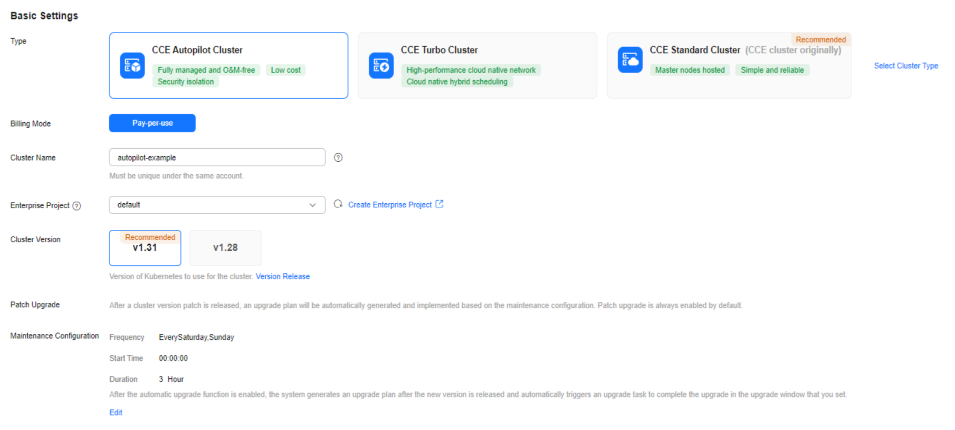
----End

Step 2: Configure the Cluster

On the **Buy Cluster** page, configure the parameters.

Basic Settings

Figure 1-2 Basic cluster information



Basic Settings

Type: CCE Autopilot Cluster (Fully managed and O&M-free, Low cost, Security isolation) | CCE Turbo Cluster (High-performance cloud native network, Cloud native hybrid scheduling) | CCE Standard Cluster (CCE cluster originally) (Recommended; Master nodes hosted, Simple and reliable) | Select Cluster Type

Billing Mode: Pay-per-use

Cluster Name: autopilot-example

Must be unique under the same account.

Enterprise Project: default | Create Enterprise Project

Cluster Version: v1.31 (Recommended) | v1.28

Version of Kubernetes to use for the cluster. Version Release

Patch Upgrade: After a cluster version patch is released, an upgrade plan will be automatically generated and implemented based on the maintenance configuration. Patch upgrade is always enabled by default.

Maintenance Configuration:

Frequency EverySaturday,Sunday

Start Time 00:00:00

Duration 3 Hour

After the automatic upgrade function is enabled, the system generates an upgrade plan after the new version is released and automatically triggers an upgrade task to complete the upgrade in the upgrade window that you set.

Edit

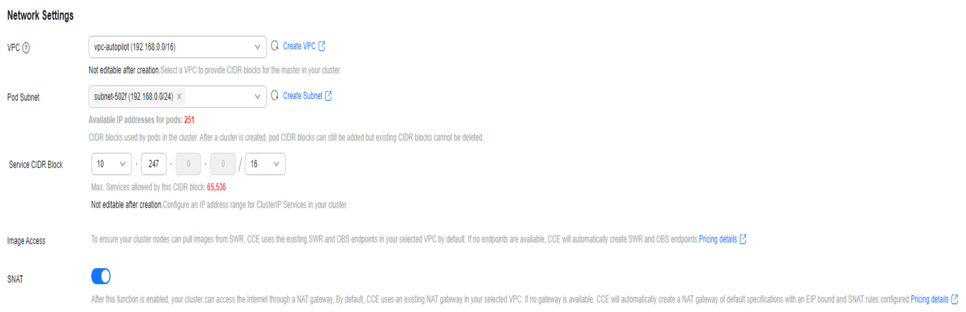
Table 1-6 Basic cluster information

Parameter	Example Value	Description
Type	CCE Autopilot Cluster	CCE allows you to create various types of clusters for diverse needs. ● CCE standard clusters provide highly reliable and secure containers for commercial use. ● CCE Turbo clusters use high-performance cloud native networks and provide cloud native hybrid scheduling. Such clusters have improved resource utilization and can be used in more scenarios. ● CCE Autopilot clusters are serverless, and you do not need to bother with server O&M. This greatly reduces O&M costs and improves application reliability and scalability. For more information about cluster types, see **Cluster Comparison**.
Cluster Name	autopilot-example	Enter a cluster name. Cluster names in the same account must be unique. Enter 4 to 128 characters. Start with a lowercase letter and do not end with a hyphen (-). Only lowercase letters, digits, and hyphens (-) are allowed.

Parameter	Example Value	Description
Enterprise Project	default	This parameter is only available for enterprise users who have enabled an enterprise project. After you select an enterprise project (for example, **default**), the cluster and resources in the cluster are created in the selected enterprise project. To manage clusters and other resources like nodes, load balancers, and node security groups, you can use the Enterprise Project Management Service (EPS). For more information, see **Enterprise Management**. If there is no special requirement, you can select **default**.
Cluster Version	v1.31	Select the Kubernetes version used by the cluster. You are advised to select the latest version.

Network Settings

Figure 1-3 Cluster network information


Network Settings

VPC: vpc-autopilot (192.168.0.0/16)
Not editable after creation.Select a VPC to provide CIDR blocks for the master in your cluster.

Pod Subnet: subnet-502f (192.168.0.0/24)
Available IP addresses for pods: 251
CIDR blocks used by pods in the cluster. After a cluster is created, pod CIDR blocks can still be added but existing CIDR blocks cannot be deleted.

Service CIDR Block: 10 . 247 . 0 . 0 / 16
Max. Services allowed by this CIDR block: 65,536
Not editable after creation.Configure an IP address range for ClusterIP Services in your cluster.

Image Access: To ensure your cluster nodes can pull images from SWR, CCE uses the existing SWR and OBS endpoints in your selected VPC by default. If no endpoints are available, CCE will automatically create SWR and OBS endpoints.Pricing details

SNAT: enabled
After this function is enabled, your cluster can access the Internet through a NAT gateway. By default, CCE uses an existing NAT gateway in your selected VPC. If no gateway is available, CCE will automatically create a NAT gateway of default specifications with an EIP bound and SNAT rules configured.Pricing details

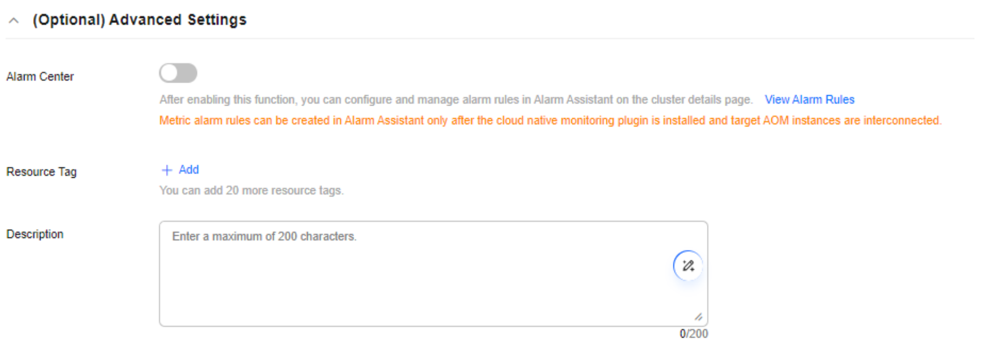
Table 1-7 Cluster network information

Parameter	Example Value	Description
VPC	**vpc-autopilot**	Select a VPC where the cluster will be running. If no VPC is available, click **Create VPC** on the right to create one. For details, see **Creating a VPC and Subnet**. The VPC cannot be changed after the cluster is created.

Parameter	Example Value	Description
Pod Subnet	**subnet-502f**	Select the subnet where the pods will be running. Each pod requires a unique IP address. The number of IP addresses in a subnet determines the maximum number of pods in a cluster and the maximum number of containers. After the cluster is created, you can add subnets. If no subnet is available, click **Create Subnet** on the right to create one. For details, see **Creating a VPC and Subnet**.
Service CIDR Block	**10.247.0.0/16**	Select a Service CIDR block, which will be used by containers in the cluster to access each other. This CIDR block determines the maximum number of Services. After the cluster is created, the Service CIDR block cannot be changed.
Image Access	**-**	To ensure that the nodes in a cluster can pull images from SoftWare Repository for Container (SWR), existing endpoints in the selected VPC are used by default. If there are no endpoints in the VPC, new endpoints will be created for you to access SWR and OBS. VPC endpoints are billed. For details, see **VPC Endpoint Price Calculator**.
SNAT	Enabled	This option is enabled by default, and the cluster can access the Internet through a NAT gateway. By default, an existing NAT gateway in the selected VPC is used. If there are no NAT gateways, CCE Autopilot automatically creates a NAT gateway with default specifications, binds an EIP to the NAT gateway, and configures SNAT rules. The NAT gateway will be billed. For details, see **NAT Gateway Billing**.

(Optional) Advanced Settings

Figure 1-4 (Optional) Advanced Settings



(Optional) Advanced Settings

Alarm Center

After enabling this function, you can configure and manage alarm rules in Alarm Assistant on the cluster details page. View Alarm Rules

Metric alarm rules can be created in Alarm Assistant only after the cloud native monitoring plugin is installed and target AOM instances are interconnected.

Resource Tag

+ Add

You can add 20 more resource tags.

Description

Enter a maximum of 200 characters.

0/200

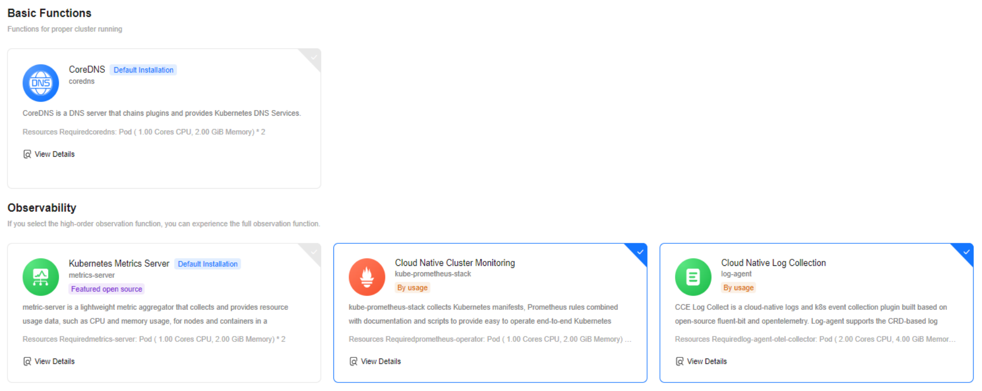
Table 1-8 (Optional) Advanced cluster settings

Parameter	Example Value	Description
Alarm Center	Disabled	Alarm Center provides comprehensive cluster alarm capabilities so that alarms can be generated in a timely manner when faults occur during cluster running, ensuring service stability. If this option is enabled, the default alarm rules will be created, and notifications will be sent to the selected contact group. For details, see **Configuring Alarms in Alarm Center**.
Resource Tag	-	You can add resource tags to classify resources. You can create **predefined tags** on the Tag Management Service (TMS) console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. ● A tag key can have no more than 128 characters and must not begin with _sys_. It can only contain letters, digits, spaces, and the following special characters: -_.:=+@. The key cannot be empty. ● A tag value can have a maximum of 255 characters. It can only contain letters, digits, spaces, and the following special characters: -_.:/=+@. The value can be empty.
Description	-	Enter a maximum of 200 characters except the following: ^~#$%&*<>()[]{}'"\.

Step 3: Select Add-ons

Click **Next: Select Add-on**. On the displayed page, select the add-ons to be installed. For details about the parameters, see **Figure 1-5**, **Table 1-9**, and **Table 1-10**.

Figure 1-5 Selecting add-ons



Basic Functions
Functions for proper cluster running

CoreDNS Default Installation
coredns
CoreDNS is a DNS server that chains plugins and provides Kubernetes DNS Services.
Resources Requiredcoredns: Pod (1.00 Cores CPU, 2.00 GiB Memory) * 2
View Details

Observability
If you select the high-order observation function, you can experience the full observation function.

Kubernetes Metrics Server Default Installation
metrics-server
Featured open source
metric-server is a lightweight metric aggregator that collects and provides resource usage data, such as CPU and memory usage, for nodes and containers in a
Resources Requiredmetrics-server: Pod (1.00 Cores CPU, 2.00 GiB Memory) * 2
View Details

✓ Cloud Native Cluster Monitoring
kube-prometheus-stack
By usage
kube-prometheus-stack collects Kubernetes manifests, Prometheus rules combined with documentation and scripts to provide easy to operate end-to-end Kubernetes
Resources Requiredprometheus-operator: Pod (1.00 Cores CPU, 2.00 GiB Memory) ...
View Details

✓ Cloud Native Log Collection
log-agent
By usage
CCE Log Collect is a cloud-native logs and k8s event collection plugin built based on open-source fluent-bit and opentelemetry. Log-agent supports the CRD-based log
Resources Requiredlog-agent-otel-collector: Pod (2.00 Cores CPU, 4.00 GiB Memor...
View Details

Table 1-9 Add-ons (basic functions)

Add-on	Example Value	Description
CoreDNS	-	This add-on (**CoreDNS**) is installed by default. It provides DNS resolution for your cluster and can be used to access the cloud DNS servers.

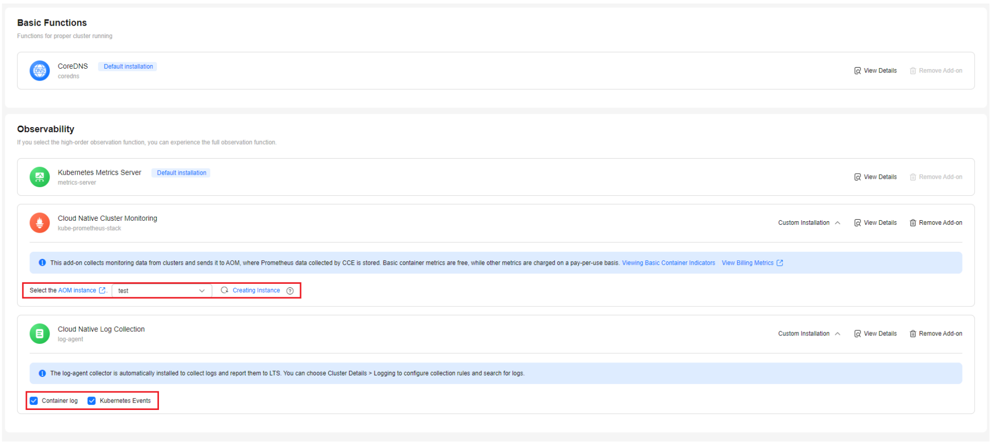
Table 1-10 Add-ons (observability)

Add-on	Example Value	Description
Kubernetes Metrics Server	-	This add-on is installed by default. It collects resource usage metrics, such as the container CPU and memory usages, for the cluster.
Cloud Native Cluster Monitoring	Select this add-on.	This is an optional add-on. If selected, this add-on (**Cloud Native Cluster Monitoring**) will be automatically installed. It collects monitoring metrics for your cluster and reports the metrics to Application Operations Management (AOM). The agent mode does not support HPA based on custom Prometheus statements. If related functions are required, install this add-on manually after the cluster is created. If this add-on is selected, pod billing is involved. You can view the prices on the console.
Cloud Native Log Collection	Select this add-on.	This is an optional add-on. If selected, this add-on (**Cloud Native Log Collection**) will be automatically installed. Cloud Native Log Collection helps report logs to LTS. After the cluster is created, you are allowed to obtain and manage collection rules on the **Logging** page of the CCE cluster console. LTS does not charge you for creating log groups and offers a free quota for log collection every month. You pay only for log volume that exceeds the quota. For details, see **Price Calculator**. For details, see **Collecting Logs**. If this add-on is selected, pod billing is involved. You can view the prices on the console.

Step 4: Configure Add-ons

Click **Next: Configure Add-on** to configure the add-ons. For details about the parameters, see **Figure 1-6** and **Table 1-11**. The add-ons that are installed by default cannot be configured. After the cluster is created, you can go to the **Add-ons** page to modify their settings.

Figure 1-6 Cluster add-on settings



This add-on collects monitoring data from clusters and sends it to AOM, where Prometheus data collected by CCE is stored. Basic container metrics are free, while other metrics are charged on a pay-per-use basis.

The log-agent collector is automatically installed to collect logs and report them to LTS. You can choose Cluster Details > Logging to configure collection rules and search for logs.

Table 1-11 Cluster add-on settings

Add-on	Example Value	Description
Cloud Native Cluster Monitoring	test	Select an AOM instance for the add-on to report metrics. If no AOM instance is available, create one first. Basic metrics are free, but custom metrics are billed based on the standard pricing of AOM. For details, see **AOM Pricing Details**.
Cloud Native Log Collection	Select **Container log** and **Kubernetes events**.	Select the logs to be collected. If enabled, a log group named **k8s-log-***{clusterId}* will be automatically created, and a log stream will be created for each selected log type. ● **Container log**: Standard output logs of containers are collected. The corresponding log stream is named in the format of **stdout-***{Cluster ID}*. ● **Kubernetes Events**: Kubernetes logs are collected. The corresponding log stream is named in the format of **event-***{Cluster ID}*. If log collection is disabled, choose **Logging** in the navigation pane of the cluster console after the cluster is created and enable this option. LTS does not charge you for creating log groups and offers a free quota for log collection every month. You pay only for log volume that exceeds the quota. For details, see **Price Calculator**. For details, see **Collecting Logs**.

Step 5: Confirm the Configuration

Click **Next: Confirm configuration**. The cluster resource list is displayed. Confirm the information and click **Submit**.

It takes about 5 to 10 minutes to create a cluster. You can click **Back to Cluster List** to perform other operations or click **Go to Cluster Events** to view the cluster details.

Related Operations

- After creating a cluster, you can use the Kubernetes command line (CLI) tool kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.
- Connect to multiple clusters using kubectl. For details, see **Connecting to Multiple Clusters Using kubectl**.

1.3 Connecting to a Cluster

1.3.1 Connecting to a Cluster Using kubectl

Scenario

You can use kubectl to connect a cluster.

Permissions

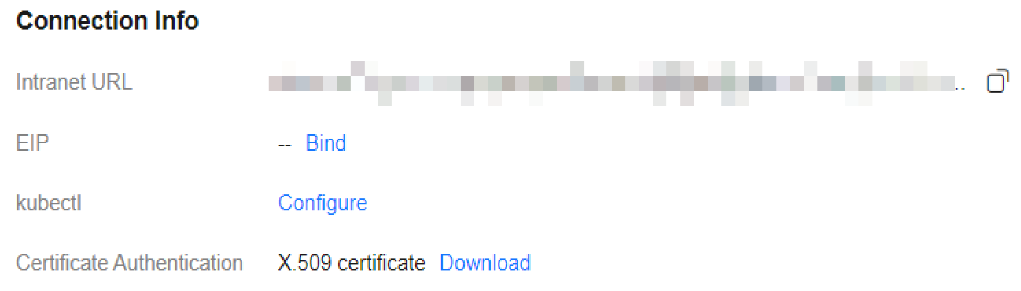
When you access a cluster using kubectl, CCE Autopilot uses **kubeconfig.json** generated on the cluster for authentication. This file contains user information, based on which CCE Autopilot determines which Kubernetes resources can be accessed by kubectl. The permissions recorded in a **kubeconfig.json** file vary from user to user.

For details about user permissions, see **Permission Overview**.

Using kubectl

kubectl is a Kubernetes command line tool for you to connect to a Kubernetes cluster from a client. You can log in to the CCE console, click the name of the cluster to be connected, and view the access address and connection steps on the cluster overview page.

Figure 1-7 Cluster connection information



Connection Info

Intranet URL	[redacted]
EIP	-- Bind
kubectl	Configure
Certificate Authentication	X.509 certificate Download

Download kubectl and the configuration file. Copy the file to your client, and configure kubectl. After the configuration is complete, you can access your Kubernetes clusters. The steps are as follows:

Step 1 **Prepare the environment.**

You need to prepare a VM that is in the same VPC as the cluster and bind an EIP to the VM for downloading kubectl.

Step 2 **Download kubectl.**

You can run the **kubectl version** command to check whether kubectl has been installed. If kubectl has been installed, skip this step.

The Linux environment is used as an example to describe how to install and configure kubectl. For details, see **Installing kubectl**.

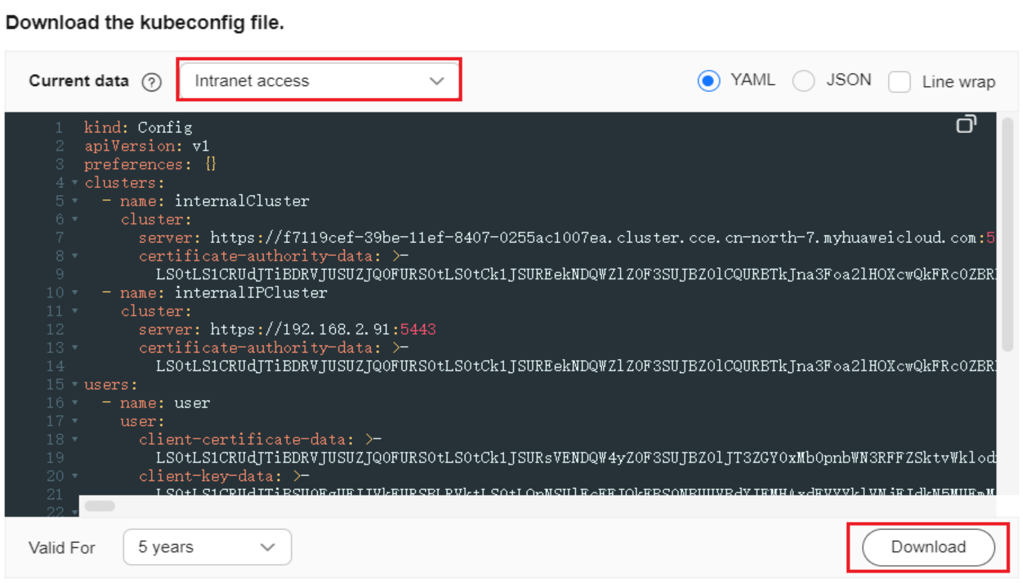
1. Log in to your client and download kubectl.
```
cd /home
curl -LO https://dl.k8s.io/release/{v1.25.0}/bin/linux/amd64/kubectl
```
{v1.25.0} specifies the version number. Replace it as required.
2. Install kubectl.
```
chmod +x kubectl
mv -f kubectl /usr/local/bin
```

Step 3 **Obtain the kubectl configuration file.**

In the **Connection Info** pane on the **Overview** page, click **Configure** next to **kubectl** to check the kubectl connection. On the displayed page, select the access method and download the configuration file.

- If a cluster version is v1.27.7-r0, v1.28.5-r0, or later, access via private domain names, access via private IP addresses, and public network access are supported.
- If a cluster version is earlier than v1.27.7-r0 or v1.28.5-r0, only private network access and public network access are supported. For private network access, only domain names can be used for access.

Figure 1-8 Downloading the configuration file



Current data: Intranet access
Valid For: 5 years

```
kind: Config
apiVersion: v1
preferences: {}
clusters:
  - name: internalCluster
    cluster:
      server: https://f7119cef-39be-11ef-8407-0255ac1007ea.cluster.cce.cn-north-7.myhuaweicloud.com:5
      certificate-authority-data: >-
        LS0tLS1CRUdJTiBDRVJUSUZJQ0FURS0tLS0tCk1JSUREekNDQWZlZ0F3SUJBZ0lCQURBTkJna3Foa2lHOXcwQkFRc0ZBR
  - name: internalIPCluster
    cluster:
      server: https://192.168.2.91:5443
      certificate-authority-data: >-
        LS0tLS1CRUdJTiBDRVJUSUZJQ0FURS0tLS0tCk1JSUREekNDQWZlZ0F3SUJBZ0lCQURBTkJna3Foa2lHOXcwQkFRc0ZBR
users:
  - name: user
    user:
      client-certificate-data: >-
        LS0tLS1CRUdJTiBDRVJUSUZJQ0FURS0tLS0tCk1JSURsVENDQW4yZ0F3SUJBZ0lJT3ZGY0xMb0pnbWN3RFFZSktvWklod
      client-key-data: >-
```

NOTE

- The kubectl configuration file (**kubeconfig**) is used for cluster authentication. If the file is leaked, your clusters may be attacked.
- For IAM users, the Kubernetes permissions specified in the configuration file are the same as those assigned on the CCE console.
- If the KUBECONFIG environment variable is configured in the Linux OS, kubectl preferentially loads the KUBECONFIG environment variable instead of **$home/.kube/config**.

Step 4 Configure kubectl.

A Linux OS is used as an example to describe how to configure kubectl.

1. Log in to your client and copy the configuration file (for example, **kubeconfig.yaml**) downloaded in **Step 3** to the **/home** directory on your client.
2. Configure the kubectl authentication file.
```
cd /home
mkdir -p $HOME/.kube
mv -f kubeconfig.yaml $HOME/.kube/config
```
3. Switch the kubectl access mode based on service scenarios.
   - To enable access within a VPC using domain names, run the following command:
   ```
   kubectl config use-context internal
   ```
   - Run this command to enable public access (EIP required):
   ```
   kubectl config use-context external
   ```
   - Run this command to enable public access and two-way authentication (EIP required):
   ```
   kubectl config use-context externalTLSVerify
   ```

----End

Troubleshooting

- **Error from server Forbidden**

  When you use kubectl to create or query Kubernetes resources, the following information is displayed:
  ```
  # kubectl get deploy Error from server (Forbidden): deployments.apps is forbidden: User
  "0c97ac3cb280f4d91fa7c0096739e1f8" cannot list resource "deployments" in API group "apps" in the
  namespace "default"
  ```
  This is because the user does not have permission to operate the Kubernetes resources. For details about how to grant permissions to the user, see **Namespace Permissions (Kubernetes RBAC-based)**.

- **The connection to the server localhost:8080 was refused**

  When you use kubectl to create or query Kubernetes resources, the following information is displayed:
  ```
  The connection to the server localhost:8080 was refused - did you specify the right host or port?
  ```
  This is because cluster authentication is not configured for the kubectl client. For details, see **Configure the kubectl authentication file**.

Related Operations

- **Connect to multiple clusters using kubectl**.

- Configure kubeconfig for fine-grained management on cluster resources

1.3.2 Connecting to a Cluster Using CloudShell

Scenario

You can use CloudShell to connect a cluster.

Permissions

When using kubectl in CloudShell, the kubectl permissions are determined by the login user.

Using CloudShell

CloudShell is a web shell used to manage and maintain cloud resources. CCE allows you to use CloudShell to connect to clusters and use kubectl in CloudShell to access clusters (by clicking the command line tool icon in **Figure 1-9**).

NOTE

- The kubectl certificate in CloudShell is valid for one day. You can reset the validity period by accessing CloudShell from the CCE console.
- CloudShell can be used only after CoreDNS is installed in a cluster.
- CloudShell cannot be used by delegated accounts or in sub-projects.

Figure 1-9 CloudShell

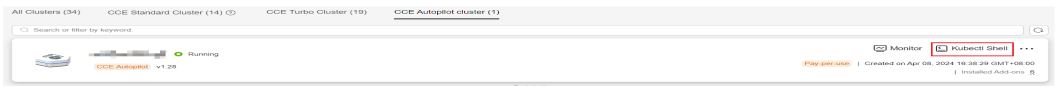
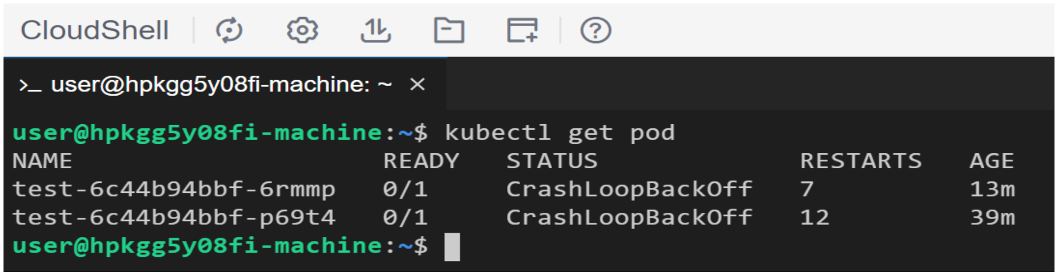

All Clusters (34)
CCE Standard Cluster (14)
CCE Turbo Cluster (19)
CCE Autopilot cluster (1)
Search or filter by keyword.
Running
CCE Autopilot
v1.28
Monitor
Kubectl Shell
Pay-per-use
Created on Apr 08, 2024 16:38:29 GMT+08:00
Installed Add-ons 6

Figure 1-10 Using kubectl in CloudShell



```
user@hpkgg5y08fi-machine:~$ kubectl get pod
NAME                     READY   STATUS             RESTARTS   AGE
test-6c44b94bbf-6rmmp    0/1     CrashLoopBackOff   7          13m
test-6c44b94bbf-p69t4    0/1     CrashLoopBackOff   12         39m
user@hpkgg5y08fi-machine:~$
```

1.3.3 Connecting to a Cluster Using an X.509 Certificate

Scenario

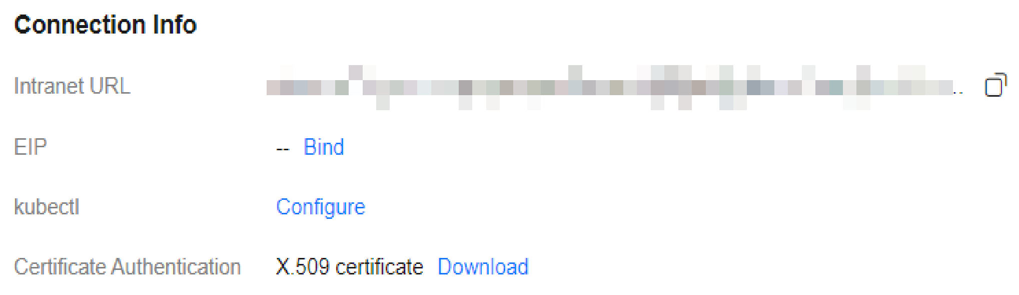
This section describes how to obtain the cluster certificate from the console and use it to access Kubernetes clusters.

Procedure

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 On the **Overview** page, locate the **Connection Info** area, and click **Download** next to **X.509 certificate**.

Figure 1-11 Downloading a cluster certificate



Connection Info

Intranet URL	[redacted]
EIP	-- Bind
kubectl	Configure
Certificate Authentication	X.509 certificate Download

Step 3 In the **Obtain Certificate** dialog box displayed, select the certificate expiration time and download the X.509 certificate of the cluster as prompted.

NOTICE

- The downloaded certificate contains three files: **client.key**, **client.crt**, and **ca.crt**. Keep these files secure.
- Certificates are not required for mutual access between containers in a cluster.

Step 4 Call native Kubernetes APIs using the cluster certificate.

For example, run the **curl** command to call an API to obtain the pod information. In the following information, ******:5443* indicates the private IP address or EIP and port number of the API server in the cluster.

```
curl --cacert ./ca.crt --cert ./client.crt --key ./client.key  https://*****:5443/api/v1/namespaces/default/pods/
```

For more cluster APIs, see **Kubernetes APIs**.

----End

1.3.4 Configuring API Server for a Cluster for Internet Access

You can bind an EIP to an API server of a Kubernetes cluster so that the API server can access the Internet.

Binding an EIP to an API Server

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 On the **Overview** page, locate the **Connection Info** area, and click **Bind** next to **EIP**.

Step 3 Select an existing EIP. If no EIP is available, click **Create EIP** to go to the EIP console and assign one.

NOTE

- Binding an EIP to an API server for Internet access can pose a risk to the cluster's security. To mitigate this risk, configure Advanced Anti-DDoS or API server access policies (**Configuring Access Policies for an API Server**) for the bound EIP.
- Binding an EIP to an API server will cause the API server to restart briefly and update the kubeconfig certificate. Do not make any changes to the cluster during this period.

Step 4 Click **OK**.

----End

Configuring Access Policies for an API Server

To ensure the security of a cluster's API server, it is important to modify the security group rules for the master nodes. This is because the EIP, which is exposed to the Internet, is at risk of being attacked.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. On the **Overview** page, copy the cluster ID in the **Basic Info** area.

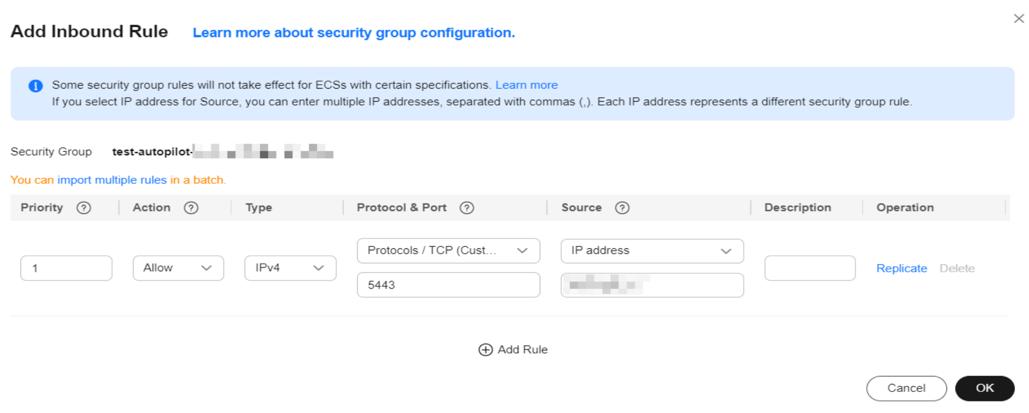
Step 2 Log in to the VPC console. In the navigation pane on the left, choose **Access Control** > **Security Groups**.

Step 3 Select **Description** as the filter criterion and paste the cluster ID to search for the target security groups.

Step 4 Locate the row that contains the security group (starting with *{CCE cluster name}*-**cce-control**) of the master node and click **Manage Rules** in the **Operation** column.

Step 5 Click **Add Rule**.

Change the source IP address that can be accessed as required. For example, if the IP address used by the client to access the API server is **100.*.*.***, you can add an inbound rule for port 5443 and set the source to **100.*.*.***.


Add Inbound Rule Learn more about security group configuration.

Some security group rules will not take effect for ECSs with certain specifications. Learn more
If you select IP address for Source, you can enter multiple IP addresses, separated with commas (,). Each IP address represents a different security group rule.

Security Group test-autopilot-[redacted]

You can import multiple rules in a batch.

Priority	Action	Type	Protocol & Port	Source	Description	Operation
1	Allow	IPv4	Protocols / TCP (Cust... 5443	IP address [redacted]		Replicate Delete

Step 6 Click **Confirm**.

----End

1.4 Managing a Cluster

1.4.1 Deleting a Cluster

Precautions

- Deleting a cluster will delete the workloads and Services in the cluster, and the deleted data cannot be recovered. Before performing this operation, ensure that data has been backed up or migrated. Deleted data cannot be restored.

  The following resources will not be deleted:

  – ELB load balancers associated with Services and Ingresses (Only the automatically created load balancers will be deleted.)

  – Resources created in the VPC, such as VPC endpoints, NAT gateways, and EIPs specified in SNAT rules

- If you delete a cluster that is not in the **Running** state (for example, **Frozen** or **Unavailable**), associated resources, such as storage and networking resources, will remain.

Deleting a Cluster

Step 1 Log in to the CCE console. In the navigation pane, choose **Clusters**.

Step 2 Locate the cluster to be deleted, click **...** to view more operations on the cluster, and choose **Delete**.

Step 3 In the displayed **Delete Cluster** dialog box, select the resources to be released.

- Delete cloud storage resources (bound to PVs) in a cluster.
- Delete network resources such as load balancers in a cluster. (Only automatically created load balancers will be deleted).
- By default, the resources (such as VPC endpoints, NAT gateways, and EIPs specified in the SNAT rules) created in the VPC of the cluster are retained. Ensure that the resources are not reused by other clusters or Services and delete them on the network console.

Step 4 Click **Yes** to start deleting the cluster.

The delete operation takes 1 to 3 minutes to complete.

----End

1.5 Upgrading a Cluster

1.5.1 Upgrade Overview

CCE strictly complies with community consistency authentication. It releases three cluster versions each year and offers a maintenance period of at least 24 months after each version is released. CCE ensures the stable running of cluster versions during the maintenance period.

To ensure your service rights and benefits, upgrade your Kubernetes clusters before a maintenance period ends. Proactive cluster upgrades help you:

- Reduce security and stability risks: During the iteration of Kubernetes versions, known security and stability vulnerabilities are continuously fixed. Long-term use of EOS clusters will result in security and stability risks to services.
- Experience the latest functions: During the iteration of Kubernetes versions, new functions and optimizations are continuously released.
- Minimize compatibility risks: During the iteration of Kubernetes versions, APIs are continuously modified and functions are deprecated. If a cluster has not been upgraded for a long time, more O&M assurance investment will be required when the cluster is upgraded. Periodic upgrades can effectively mitigate compatibility risks caused by accumulated version differences. It is a good practice to upgrade a patch version every quarter and upgrade a major version to the latest version every year.
- Obtain more effective technical support: CCE does not provide security patches or issue fixing for EOS Kubernetes cluster versions, and does not ensure technical support for the EOS versions.

You can check Kubernetes cluster versions on the cluster list page and view if there is a new version for a cluster to upgrade. If necessary, obtain the expert consultation service.

Cluster Upgrade Path

CCE Autopilot clusters evolve iteratively based on the community Kubernetes version. A CCE cluster version consists of the community Kubernetes version and the CCE patch version. Therefore, two cluster upgrade paths are provided.

- Kubernetes version

Source Kubernetes Version	Target Kubernetes Version
v1.28	/

NOTE

A Kubernetes version can be upgraded only after the patch is upgraded to the latest version. CCE will automatically generate an optimal upgrade path on the console based on the current cluster version.

- Patch version

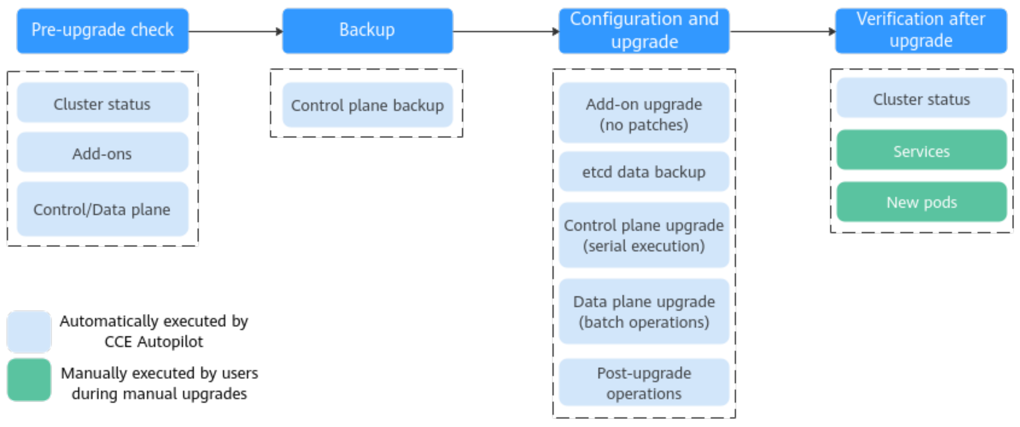
  Patch version management is available for CCE Autopilot clusters to provide new features and fix bugs and vulnerability for in-maintenance clusters without requiring a major version upgrade.

  After a new patch version is released, you can upgrade to the latest patch version at any time.

Cluster Upgrade Process

The cluster upgrade process involves pre-upgrade check, backup, upgrade, and post-upgrade verification.

Figure 1-12 Cluster upgrade process


Pre-upgrade check
Cluster status
Add-ons
Control/Data plane
Backup
Control plane backup
Configuration and upgrade
Add-on upgrade (no patches)
etcd data backup
Control plane upgrade (serial execution)
Data plane upgrade (batch operations)
Post-upgrade operations
Verification after upgrade
Cluster status
Services
New pods
Automatically executed by CCE Autopilot
Manually executed by users during manual upgrades

After determining the target version of the cluster, read the **precautions** carefully and prevent function incompatibility during the upgrade.

1. **Pre-upgrade check**

   Before a cluster upgrade, mandatory items such as the cluster status, add-ons, and workloads are automatically checked to ensure that the cluster meets the upgrade requirements. For more details, see **Troubleshooting for Pre-upgrade Check Exceptions**. If any check item is abnormal, rectify the fault on the console as prompted.

2. **Backup**

   You can use disk snapshots to back up the control plane details such as cluster component images, component configurations, and etcd data. Back up your data before an upgrade. If unexpected cases occur during an upgrade, you can use the backup to quickly restore the cluster.

Backup Method	Backup Object	Backup Type	Backup Time	Rollback Time	Description
etcd data backup	etcd data	Automatic backup during the upgrade	1–5 minutes	2 hours	Mandatory. The backup is automatically performed during the upgrade.

Backup Method	Backup Object	Backup Type	Backup Time	Rollback Time	Description
EVS snapshot backup	Control plane details, including component images, configurations, logs, and etcd data	One-click backup on web pages (manually triggered)	1–5 minutes	20 minutes	-

3. **Configuration and upgrade**

   Configure parameters before an upgrade. CCE Autopilot has provided default settings, which can be modified as needed. After the configuration, upgrade add-ons, control plane, and data plane in sequence.

   – **Add-on Upgrade Configuration**: Add-ons that have been installed in your cluster are listed. During the cluster upgrade, CCE Autopilot automatically upgrades the selected add-ons to be compatible with the target cluster version. You can click **Set** to re-define the add-on parameters.

     NOTE

     If an add-on is marked with ⚠ on its right side, the add-on cannot be compatible with both the source and target versions of the cluster upgrade. In this case, CCE will upgrade the add-on after the cluster upgrade. The add-on may be unavailable during the cluster upgrade.

4. **Post-upgrade verification**

   After an upgrade, CCE Autopilot will automatically check items such as the cluster status. You need to manually check services and new pods to ensure that the cluster functions properly after the upgrade.

1.5.2 Before You Start

Before the upgrade, you can check whether your cluster can be upgraded and which versions are available on the CCE console. For details, see **Upgrade Overview**.

Precautions

Before upgrading a cluster, note the following:

- **Perform an upgrade during off-peak hours to minimize the impact on your services.**
- Before upgrading a cluster, learn about the features and differences of each cluster version in **Kubernetes Version Release Notes** to prevent exceptions due to the use of an incompatible cluster version. For example, check whether any APIs deprecated in the target version are used in the cluster. Otherwise, calling the APIs may fail after the upgrade.

During a cluster upgrade, pay attention to the following points that may affect your services:

- Before upgrading a cluster, **ensure no high-risk operations are performed in the cluster**. If there are high-risk operations, the cluster upgrade may fail or the configuration may be lost after the upgrade. For example, modifying the configuration of a listener configured for CCE on the ELB console is a high-risk operation. It recommended that you modify configurations on the CCE console so that the modifications can be automatically inherited during the upgrade.

Constraints

If an error occurred during a cluster upgrade, the cluster can be rolled back using the backup data. If you perform other operations (for example, modifying cluster specifications) after a successful cluster upgrade, the cluster can be rolled back using the backup data.

Deprecated APIs

With the evolution of Kubernetes APIs, APIs are periodically reorganized or upgraded, and old APIs are deprecated and finally deleted. The following tables list the deprecated APIs in each Kubernetes community version. For details about more deprecated APIs, see **Deprecated API Migration Guide**.

NOTE

When an API is deprecated, the existing resources are not affected. However, when you create or edit the resources, the API version will be intercepted.

Upgrade Backup

How to back up a node:

Backup Method	Backup Object	Backup Type	Backup Time	Rollback Time	Description
etcd data backup	etcd data	Automatic backup during the upgrade	1–5 minutes	2 hours	Mandatory. The backup is automatically performed during the upgrade.

Backup Method	Backup Object	Backup Type	Backup Time	Rollback Time	Description
EVS snapshot backup	Control plane details, including component images, configurations, logs, and etcd data	One-click backup on web pages (manually triggered)	1–5 minutes	20 minutes	-

1.5.3 Manual Upgrades

You can upgrade your clusters to a newer version on the CCE console.

Before the upgrade, learn about the target version to which each CCE cluster can be upgraded in what ways, and the upgrade impacts. For details, see **Upgrade Overview** and **Before You Start**.

Precautions

- During the cluster upgrade, the system will automatically upgrade add-ons to a version compatible with the target cluster version. Do not uninstall or reinstall add-ons during the cluster upgrade.
- Before the upgrade, ensure that all add-ons are running. If an add-on fails to be upgraded, rectify the fault and try again.
- If an upgrade failure message is displayed during the cluster upgrade, rectify the fault as prompted and try again. If upgrade attempts fail again, **submit a service ticket** for assistance.

For more information, see **Before You Start**.

Procedure

The cluster upgrade goes through check, backup, configuration and upgrade, and verification.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Cluster Upgrade**.

Step 3 CCE automatically provides you with an optimal upgrade path based on the current cluster version. Select the target cluster version, check information such as version differences, and add-on versions, and click **Go to Upgrade**.

Step 4 Perform the pre-upgrade check. Click **Start Check** and confirm the check. If there are abnormal or risky items in the cluster, handle the exceptions based on the check results displayed on the page and check again.

- **Exceptions**: View the solution displayed on the page, handle the exceptions and check again.

- **Risk Items**: may affect the cluster upgrade. Check the risk description and see whether you may be impacted. If no risk exists, click **OK** next to the risk item to manually skip this risk item and check again.

After the check is passed, click **Next**.

Step 5 Back up the cluster. During the cluster upgrade, CCE automatically backs up etcd data. You can manually back up the control plane to speed up the rollback if the control plane fails to upgrade. If manual backup is not required, click **Next**.

Backup Method	Backup Object	Backup Type	Backup Time	Rollback Time	Description
etcd data backup	etcd data	Automatic backup during the upgrade	1–5 minutes	2 hours	Mandatory. The backup is automatically performed during the upgrade.
EVS snapshot backup	Control plane details, including component images, configurations, logs, and etcd data	One-click backup on web pages (manually triggered)	1–5 minutes	20 minutes	-

Step 6 Configure the upgrade parameters.

- **Add-on Upgrade Configuration**: Add-ons that have been installed in your cluster are listed. During the cluster upgrade, CCE Autopilot automatically upgrades the selected add-ons to be compatible with the target cluster version. You can click **Set** to re-define the add-on parameters.

  NOTE

  If an add-on is marked with ⚠ on its right side, the add-on cannot be compatible with both the source and target versions of the cluster upgrade. In this case, CCE will upgrade the add-on after the cluster upgrade. The add-on may be unavailable during the cluster upgrade.

Step 7 After the configuration is complete, click **Upgrade** and confirm the upgrade. The cluster starts to be upgraded. You can view the process in the lower part of the page.

NOTE

If an upgrade failure message is displayed during the cluster upgrade, rectify the fault as prompted and try again.

Step 8 After the upgrade is complete, click **Next**. Verify the upgrade based on the displayed check items. After confirming that all check items are normal, click

Complete and confirm that the post-upgrade check is complete. For details, see **Performing Post-Upgrade Verification**.

You can verify the cluster Kubernetes version on the **Clusters** page.

----End

FAQs

- **What do I do if a cluster add-on fails to be upgraded during the CCE cluster upgrade?**

1.5.4 Performing Post-Upgrade Verification

1.5.4.1 Cluster Status Check

Check Items

After a cluster is upgraded, check whether the cluster is in the **Running** state.

Procedure

CCE automatically checks your cluster status. Go to the cluster list page and confirm the cluster status based on the diagnosis result.

Solution

If your cluster malfunctions, contact technical support.

1.5.4.2 Service Check

Check Items

After a cluster is upgraded, check whether its services are running properly.

Procedure

Different services have different verification mode. Select a suitable one and verify the service before and after the upgrade.

You can verify the service from the following aspects:

- The service page is available.
- No alarm or event is generated on the normal platform.
- No error log is generated for key processes.
- The API dialing test is normal.

Solution

If your online services malfunction after the cluster upgrade, contact technical support.

1.5.4.3 New Pod Check

Check Items

Check whether pods can be created after the cluster is upgraded.

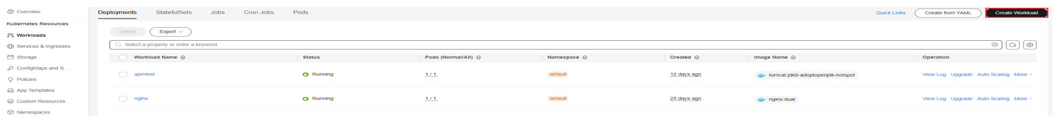
Procedure

Create pods and ensure the new and existing pods provide the same services.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. On the displayed page, click **Create Workload** or **Create from YAML** in the upper right corner. For details about how to create a workload, see **Creating a Workload**.

Figure 1-13 Creating a workload



Workload Name	Status	Pods (Normal/All)	Namespace	Created	Image Name	Operation
apmtest	Running	1 / 1	default	12 days ago	tomcat:jdk8-adoptopenjdk-hotspot	View Log Upgrade Auto Scaling More
nginx	Running	1 / 1	default	25 days ago	nginx:dual	View Log Upgrade Auto Scaling More

It is a good practice to use the image for routine tests as the base image. You can deploy minimum pods for an application by referring to the following YAML file.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: post-upgrade-check
  namespace: default
spec:
  selector:
    matchLabels:
      app: post-upgrade-check
      version: v1
  template:
    metadata:
      labels:
        app: post-upgrade-check
        version: v1
    spec:
      containers:
        - name: container-1
          image: nginx:perl
          imagePullPolicy: IfNotPresent
          resources:
            requests:
              cpu: 10m
              memory: 10Mi
            limits:
              cpu: 100m
              memory: 50Mi
```

Step 3 After the workload is created, check whether the pods of the workload are running properly.

Step 4 After the check is complete, choose **Workloads** in the navigation pane. On the displayed page, locate the **post-upgrade-check** workload and choose **More** > **Delete** in the **Operation** column to delete the test workload.

----End

Solution

If pods cannot be created or are abnormal, submit a service ticket.

1.5.5 Troubleshooting for Pre-upgrade Check Exceptions

1.5.5.1 Pre-upgrade Check

The system automatically checks a cluster before its upgrade. If the cluster does not meet the pre-upgrade check conditions, the upgrade cannot continue. To avoid risks, you can perform pre-upgrade check according to the check items and solutions described in this section.

Table 1-12 Check items

No.	Check Item	Description
1	**Upgrade Management**	Check whether the target cluster is under upgrade management.
2	**Add-ons**	● Check whether the add-on status is normal. ● Check whether the add-on supports the target version.
3	**Helm Charts**	Check whether the current HelmRelease record contains discarded Kubernetes APIs that are not supported by the target cluster version. If yes, the Helm chart may be unavailable after the upgrade.
4	**SSH Connectivity of Master Nodes**	Check whether your master nodes can be accessed using SSH.
5	**Discarded Kubernetes Resources**	Check whether there are discarded resources in the clusters.
6	**cce-hpa-controller Limitations**	Check whether there are compatibility limitations between the current and target cce-controller-hpa add-on versions.
7	**Discarded Kubernetes APIs**	The system scans the audit logs of the past day to check whether the user calls the deprecated APIs of the target Kubernetes version. **NOTE** Due to the limited time range of audit logs, this check item is only an auxiliary method. APIs to be deprecated may have been used in the cluster, but their usage is not included in the audit logs of the past day. Check the API usage carefully.
8	**HTTPS Load Balancer Certificate Consistency**	Check whether the certificate used by an HTTPS load balancer has been modified on ELB.

No.	Check Item	Description
9	**Pod Configuration**	Check whether the pod configuration has compatibility restrictions.

1.5.5.2 Upgrade Management

Check Items

Check whether the target cluster is under upgrade management.

Solution

CCE may temporarily restrict the cluster upgrade due to the following reasons:

- The cluster is identified as the core production cluster.
- Other O&M tasks are being or will be performed, for example, 3-AZ reconstruction on master nodes.
- If the cluster contains Docker nodes but their OSs are different from that of the node pool, reset these nodes and run the pre-upgrade check again.

To resolve this issue, contact technical support based on logs displayed on the console.

1.5.5.3 Add-ons

Check Items

Check the following items:

- Check whether the add-on status is normal.
- Check whether the add-on support the target version.

Solution

- **Scenario 1: The add-on malfunctions.**

  Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons** and then obtain add-ons. Then, handle malfunctional add-ons.
- **Scenario 2: The target cluster version does not support the current add-on version.**

  The add-on cannot be automatically upgraded with the cluster due to compatibility issues. In this case, log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons** and then manually upgrade the add-on.
- **Scenario 3: After the add-on is upgraded to the latest version, it is still not supported by the target cluster version.**

  Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons** and then manually uninstall the add-on. For details about the supported add-on versions and substitutions, see **Add-ons**.

- **Scenario 4: The add-on configuration does not meet the upgrade requirements. Upgrade the add-on and try again.**

  The following error information is displayed during the pre-upgrade check:

  ```
  please upgrade addon [ ] in the page of addon managecheck and try again
  ```

  In this case, log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons** and then manually upgrade the add-on.

1.5.5.4 Helm Charts

Check Items

Check whether the current HelmRelease record contains discarded Kubernetes APIs that are not supported by the target cluster version. If yes, the Helm chart may be unavailable after the upgrade.

Solution

Convert the discarded Kubernetes APIs to APIs that are compatible with both the source and target versions.

NOTE

This item has been automatically processed in the upgrade process. You can ignore this item.

1.5.5.5 SSH Connectivity of Master Nodes

Check Items

Check whether your master nodes can be accessed using SSH.

Solution

There is a low probability that the SSH connectivity check fails due to network fluctuations. Perform the pre-upgrade check again.

If the check still fails, submit a service ticket to contact technical support.

1.5.5.6 Discarded Kubernetes Resources

Check Items

Check whether there are discarded resources in the clusters.

Solution

- **Scenario 1: The Service in the clusters of v1.25 or later has discarded annotation tolerate-unready-endpoints.**

  Error log:

```
some check failed in cluster upgrade: this cluster has deprecated service list: map[***] with deprecated annotation list [tolerate-unready-endpoints]
```

Check whether the Service provided in the log information contains the annotation **tolerate-unready-endpoint**. If so, replace the annotation with the following fields in Service **spec**:

```
publishNotReadyAddresses: true
```

- **Scenario 2: The Service in the clusters of v1.27 or later has discarded annotation service.kubernetes.io/topology-aware-hints.**

  Error log:

  ```
  some check failed in cluster upgrade: this cluster has deprecated service list: map[***] with deprecated annotation list [service.kubernetes.io/topology-aware-hints]
  ```

  Check whether the Service provided in the log information contains the annotation **service.kubernetes.io/topology-aware-hints**. If so, replace it with the annotation **service.kubernetes.io/topology-mode** in the affected Service.

1.5.5.7 cce-hpa-controller Limitations

Check Items

Check whether there are compatibility limitations between the current and target cce-controller-hpa add-on versions.

Solution

There are compatibility limitations between the current and target versions of the cce-controller-hpa add-on. To address this, install an add-on that provides metrics APIs, like metrics-server, in the cluster.

Install the metrics add-on in the cluster and try again.

1.5.5.8 Discarded Kubernetes APIs

Check Items

The system scans the audit logs of the past day to check whether the user calls the deprecated APIs of the target Kubernetes version.

NOTE

Due to the limited time range of audit logs, this check item is only an auxiliary method. APIs to be deprecated may have been used in the cluster, but their usage is not included in the audit logs of the past day. Check the API usage carefully.

Solution

Check Description

Based on the check result, it is detected that your cluster calls a deprecated API of the target cluster version using kubectl or other applications. You can rectify the fault before the upgrade. Otherwise, the API will be intercepted by kube-apiserver after the upgrade.

There are no version compatibility differences when v1.27 is upgraded to v1.28.

1.5.5.9 HTTPS Load Balancer Certificate Consistency

Check Items

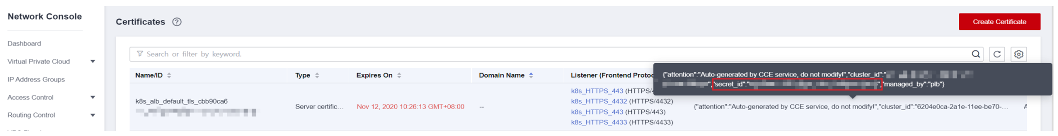
Check whether the certificate used by an HTTPS load balancer has been modified on ELB.

Solution

The certificate referenced by an HTTPS Ingress created on CCE is modified on the ELB console. This leads to inconsistent certificate content in the CCE cluster and that required by the load balancer. After the CCE cluster is upgraded, the load balancer's certificate is overwritten.

Step 1 Log in to the ELB console, choose **Elastic Load Balance** > **Certificates**, locate the certificate, and find the **secret_id** in the certificate description.

Figure 1-14 Viewing a certificate



Name/ID	Type	Expires On	Domain Name	Listener (Frontend Protocol)	
k8s_alb_default_tls_cbb90ca6 [redacted]	Server certific...	Nov 12, 2020 10:26:13 GMT+08:00	--	k8s_HTTPS_443 (HTTPS/443) k8s_HTTPS_4432 (HTTPS/4432) k8s_HTTPS_443 (HTTPS/443) k8s_HTTPS_4433 (HTTPS/4433)	{"attention":"Auto-generated by CCE service, do not modify!","cluster_id":"6204e0ca-2a1e-11ee-be70-...

{"attention":"Auto-generated by CCE service, do not modify!","cluster_id":"[redacted]","secret_id":"[redacted]","managed_by":"plb"}

The **secret_id** is the **metadata.uid** of the secret in the cluster. Use this UID to obtain the secret name in the cluster.

Run the following kubectl command to obtain the Secret name (replace *<secret_id>* with the actual value):

```
kubectl get secret --all-namespaces -o jsonpath='{range .items[*]}{"uid:"}{.metadata.uid}{" namespace:"}
{.metadata.namespace}{" name:"}{.metadata.name}{"\n"}{end}' | grep <secret_id>
```

Step 2 Replace the certificate used by an Ingress with the one used by the load balancer. Then, you can create or edit the certificate on the ELB console.

1. Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Services & Ingresses**. Click the **Ingresses** tab, locate the row containing the Ingress that uses the certificate, and choose **More** > **Update** in the **Operation** column. If multiple Ingresses are using this certificate, update the certificate for all of these Ingresses. To check which Ingresses are using a certificate, use the **secretName** parameter in **spec.tls** of the Ingress YAML files.

   Run the following kubectl command to obtain the Ingresses using a certificate (replace *<secret_name>* with the actual value):

   ```
   kubectl get ingress --all-namespaces -o jsonpath='{range .items[*]}{"namespace:"}
   {.metadata.namespace}{" name:"}{.metadata.name}{" tls:"}{.spec.tls[*]}{"\n"}{end}' | grep
   <secret_name>
   ```

2. When configuring a listener, select **ELB server certificate** for **Certificate Source** and click **OK**. In this way, the certificate can be created or edited on the ELB console.
3. On the **ConfigMaps and Secrets** page, delete the target secret. Before the deletion, back up data.

----End

1.5.5.10 Pod Configuration

Check Items

Check whether the pod configuration has compatibility restrictions.

Solution

Scenario: In clusters earlier than v1.28.6-r0 or v1.27.8-r0, if the preset **resource.cce.io/extra-ephemeral-storage-in-GiB** annotation exists, pods may fail to be created or there may be unexpected expenses after the cluster upgrade.

The error log information is as follows:

```
some check failed in cluster upgrade: this cluster has podlist: map[pod-xxx pod-xxx] with system preset
annotation list [resource.cce.io/extra-ephemeral-storage-in-GiB]
```

Check whether the pods listed in the log contain the **resource.cce.io/extra-ephemeral-storage-in-GiB** annotation. This annotation indicates the size of extra ephemeral storage capacity. For details, see **Pod Annotations**.

You can use either of the following methods to solve the problem:

- Upgrade the workload, delete the annotation, and check the pod configuration again on the cluster upgrade page. After the cluster is upgraded, new pods will not use extra ephemeral storage. The ephemeral storage for each pod is 30 GiB by default.
- Upgrade the workload, change **resource.cce.io/extra-ephemeral-storage-in-GiB** to a correct value, and skip the pod configuration check on the cluster upgrade page. After the cluster is upgraded, new pods will use extra ephemeral storage, and you will be billed for it. For details about pricing, see **Billing Description**.

2 Workloads

2.1 Workload Overview

A workload is an application running on Kubernetes. No matter how many components are there in your workload, you can run it in a group of Kubernetes pods. A workload is an abstract model of a group of pods in Kubernetes. Workloads in Kubernetes are classified as Deployments, StatefulSets, Jobs, and CronJobs.

CCE Autopilot runs on native Kubernetes capabilities for you to deploy and manage containers throughout their lifecycle, such as creation, configuration, monitoring, auto scaling, upgrade, uninstall, service discovery, and load balancing.

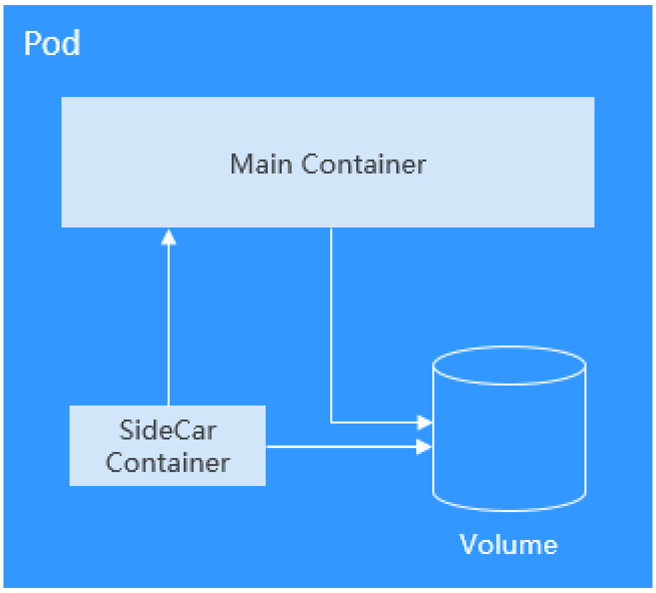
Overview of Pod

A pod is the smallest, simplest unit in the Kubernetes object model that you create or deploy. A pod is a group of one or more containers, with shared storage and network resources, and a specification for how to run the containers. Each pod has a separate IP address.

Pods can be used in either of the following ways:

- A pod runs only one container. This is the most common usage of pods in Kubernetes. You can consider a pod as a container, but Kubernetes directly manages pods instead of containers.
- A pod runs multiple containers that need to be tightly coupled. In this scenario, a pod contains a main container and several sidecar containers, as shown in **Figure 2-1**. For example, the main container is a web server that provides file services from a fixed directory, and sidecar containers periodically download files to this fixed directory.

Figure 2-1 Pod running multiple containers

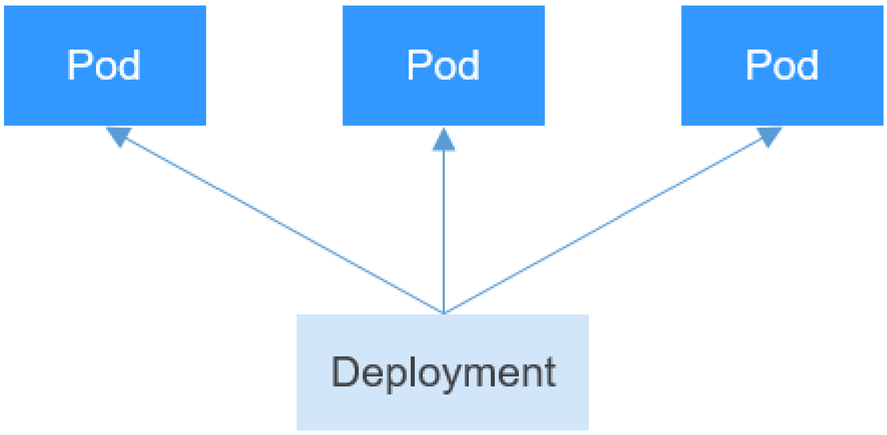


In Kubernetes, pods are rarely created directly. Instead, Kubernetes controller manages pods through pod instances such as Deployments and jobs. A controller typically uses a pod template to create pods. The controller can also manage multiple pods and provide functions such as replica management, rolling upgrade, and self-healing.

Overview of Deployment

A pod is the smallest and simplest unit that you create or deploy in Kubernetes. It is designed to be an ephemeral, one-off entity. A pod can be evicted when node resources are insufficient and disappears along with a cluster node failure. Kubernetes provides controllers to manage pods. Controllers can create and manage pods, and provide replica management, rolling upgrade, and self-healing capabilities. The most commonly used controller is Deployment.

Figure 2-2 Relationship between a Deployment and pods



A Deployment can contain one or more pods. These pods have the same role. Therefore, the system automatically distributes requests to multiple pods of a Deployment.

A Deployment integrates a lot of functions, including online deployment, rolling upgrade, replica creation, and restoration of online jobs. To some extent, Deployments can be used to realize unattended rollout, which greatly reduces difficulties and operation risks in the rollout process.

Overview of StatefulSet

All pods under a Deployment have the same characteristics except for the name and IP address. If required, a Deployment can use a pod template to create new pods. If not required, the Deployment can delete any one of the pods.

However, Deployments cannot meet the requirements in some distributed scenarios when each pod requires its own status or in a distributed database where each pod requires independent storage.

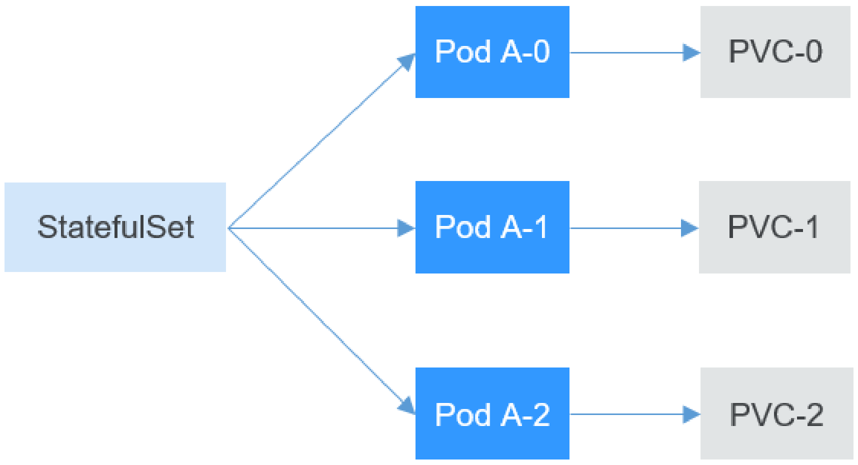
Distributed stateful applications involve different roles for different responsibilities. For example, databases work in active/standby mode, and pods depend on each other. To deploy stateful applications in Kubernetes, ensure pods meet the following requirements:

- Each pod must have a fixed identifier so that it can be recognized by other pods.
- Separate storage resources must be configured for each pod. In this way, the original data can be retrieved after a pod is deleted and restored. Otherwise, the pod status will be changed after the pod is rebuilt.

To address the preceding requirements, Kubernetes provides StatefulSets.

1. StatefulSets provide a fixed name for each pod following a fixed number ranging from 0 to N. After a pod is rescheduled, the pod name and the hostname remain unchanged.
2. StatefulSets use a headless Service to allocate a fixed domain name for each pod.
3. StatefulSets create PVCs with fixed identifiers to ensure that pods can access the same persistent data after being rescheduled.

Figure 2-3 StatefulSet



Overview of Job and CronJob

Jobs and CronJobs allow you to run short lived, one-off tasks in batch. They ensure the task pods run to completion.

- A job is a resource object used by Kubernetes to control batch tasks. Jobs are different from long-term servo tasks (such as Deployments and StatefulSets). The former is started and terminated at specific times, while the latter runs unceasingly unless being terminated. The pods managed by a job will be automatically removed after successfully completing tasks based on user configurations.
- A CronJob runs a job periodically on a specified schedule. A CronJob object is similar to a line of a crontab file in Linux.

This run-to-completion feature of jobs is especially suitable for one-off tasks, such as continuous integration (CI).

Workload Lifecycle

Table 2-1 Workload statuses

Status	Description
Running	This status is displayed when all pods are running or the number of pods is 0.
Not ready	This status is displayed when the container is abnormal or the pods are not running normally.
Processing	The workload is not running but no error is reported.
Available	For a multi-pod Deployment, some pods are abnormal but at least one pod is available.
Completed	The job is executed. This status is available only for jobs.
Deleting	After the delete operation is triggered, the workload is being deleted.

2.2 Creating a Workload

2.2.1 Creating a Deployment

Scenario

Deployments are workloads (for example, Nginx) that do not store any data or status. You can create Deployments on the CCE console or by running kubectl commands.

Prerequisites

- A cluster is available. For details about how to create a cluster, see **Buying a CCE Autopilot Cluster**.
- VPC endpoints for accessing SWR and OBS have been configured. For details, see **Configuring VPC Endpoints for Accessing SWR and OBS**.
- A Service of the LoadBalancer type has been created if the workload needs to be accessed by external networks.

  NOTE

  If a pod has multiple containers, the ports used by the containers cannot conflict with each other. If there is a conflict, the Deployment will fail to be created.

Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 Set basic information about the workload.

Basic Info

- **Workload Type**: Select **Deployment**.
- **Workload Name**: Enter a name for the workload. Enter 1 to 63 characters starting with a lowercase letter and ending with a lowercase letter or digit. Only lowercase letters, digits, and hyphens (-) are allowed.
- **Namespace**: Select a namespace. The default value is **default**. You can also click **Create Namespace** to create one. For details, see **Creating a Namespace**.
- **Pods**: Enter the number of pods of the workload.

Container Settings

- **Container Information**

  A pod can have more than one container. You can click **Add Container** on the right to configure multiple containers.

  – **Basic Info**: Configure basic information about each container.

Table 2-2 Basic container information

Parameter	Description
Container Name	Enter a name for the container.
Pull Policy	Image update or pull policy. If you select **Always**, the image is pulled from the image repository each time. If you do not select **Always**, the existing image of the node is preferentially used. If the image does not exist, the image is pulled from the image repository.

Parameter	Description
Image Name	Click **Select Image** and select the image used by the container. To use a third-party image, see **Using Third-Party Images**.
Image Tag	Select the image tag to be deployed.
CPU Quota	CPU limit, which is the maximum CPU available for the container to prevent excessive resource usage.
Memory Quota	Memory limit, which is the maximum memory available for the container. When the container's memory usage exceeds the memory limit, the container will be terminated.
(Optional) Init Container	Whether the container will be used as an init container. An init container does not support health check. An init container is a special container that runs before other app containers in a pod are started. Each pod can contain multiple containers. In addition, a pod can contain one or more init containers. Application containers in a pod are started and run only after the running of all init containers completes. For details, see **Init Containers**.

- (Optional) **Lifecycle**: Configure operations to be performed in a specific phase of the container lifecycle, such as Startup Command, Post-Start, and Pre-Stop. For details, see **Configuring the Container Lifecycle**.
- (Optional) **Health Check**: Set the liveness probe, ready probe, and startup probe as required. For details, see **Setting Health Check for a Container**.
- (Optional) **Environment Variables**: Configure variables for the container running environment using key-value pairs. These variables transfer external information to containers running in pods and can be flexibly modified after application deployment. For details, see **Configuring Environment Variables**.
- (Optional) **Data Storage**: Mount local storage or cloud storage to the container. The application scenarios and the ways for mounting the volumes vary with the storage type. For details, see **Storage**.
- (Optional) **Security Context**: Assign container permissions to protect the system and other containers from being affected. Enter the user ID to assign container permissions and prevent systems and other containers from being affected.

- **Image Access Credential**: Select the credential used for accessing the image repository. The default value is **default-secret**. You can use default-secret to access images in SWR. For details about **default-secret**, see **default-secret**.

(Optional) Service Settings

A Service provides external access for pods. With a static IP address, a Service forwards the traffic to pods and automatically balances load for these pods.

You can also create a Service after creating a workload. For details about Services of different types, see **Service**.

(Optional) Advanced Settings

- **Upgrade**: Specify the upgrade mode and upgrade parameters of the workload. **Rolling upgrade** and **Replace upgrade** are supported. For details, see **Configuring the Workload Upgrade Policy**.
- **Labels and Annotations**: Add labels or annotations for pods using key-value pairs. After entering the key and value, click **Confirm**. For details about how to use and configure labels and annotations, see **Configuring Labels and Annotations**.
- **DNS**: Configure a DNS policy for the workload. For details, see **DNS Configuration**.
- **APM Settings**: Use Application Performance Management (APM) to provide more accurate problem analysis and location for Java programs. For details, see **Configuring APM**.
- Network configuration:
  - Whether to enable the network configuration of a specified container: You can enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. For details, see **Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration**.
  - Specify the container network configuration name: Only the custom container network configuration whose associated resource type is workload can be selected.

Step 4 Click **Create Workload** in the lower right corner.

----End

Using kubectl

Nginx is used as an example here to describe how to create a workload using kubectl.

NOTICE

Node affinity and anti-affinity are not available for CCE Autopilot clusters. When you use kubectl to create a workload, do not configure the **affinity** field to prevent pod creation failures.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create and edit the **nginx-deployment.yaml** file. **nginx-deployment.yaml** is an example file name. You can rename it as required.

vi nginx-deployment.yaml

The following is an example YAML file. For more information about Deployments, see **Kubernetes documentation**.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  strategy:
    type: RollingUpdate
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
      - image: nginx    # If you use an image from an open-source image registry, enter the image name. If 
you use an image in My Images, obtain the image path from SWR.
        imagePullPolicy: Always
        name: nginx
      imagePullSecrets:
      - name: default-secret
```

For details about the parameters, see **Table 2-3**.

Table 2-3 Deployment YAML parameters

Parameter	Description	Mandatory
apiVersion	API version. **NOTE** Set this parameter based on the cluster version. ● For clusters of v1.17 or later, the apiVersion format of Deployments is **apps/v1**. ● For clusters of v1.15 or earlier, the apiVersion format of Deployments is **extensions/v1beta1**.	Yes
kind	Type of a created object.	Yes
metadata	Metadata of a resource object.	Yes
name	Name of the Deployment.	Yes
spec	Detailed description of the Deployment.	Yes
replicas	Number of pods.	Yes
selector	Determines container pods that can be managed by the Deployment.	Yes
strategy	Upgrade mode. Possible values: ● **RollingUpdate** ● **ReplaceUpdate** By default, **RollingUpdate** is used.	No
template	Detailed description of a created container pod.	Yes
metadata	Metadata.	Yes
labels	**metadata.labels**: Container labels.	No

Parameter	Description	Mandatory
spec: containers	● **image** (mandatory): Name of a container image. ● **imagePullPolicy** (optional): Policy for obtaining an image. The options include **Always** (attempting to download images each time), **Never** (only using local images), and **IfNotPresent** (using local images if they are available; downloading images if local images are unavailable). The default value is **Always**. ● **name** (mandatory): Container name.	Yes
imagePull Secrets	Name of the secret used during image pulling. If a private image is used, this parameter is mandatory. ● To pull an image from the Software Repository for Container (SWR), set this parameter to **default-secret**. ● To pull an image from a third-party image repository, set this parameter to the name of the created secret.	No

Step 3 Create a Deployment.

kubectl create -f nginx-deployment.yaml

If the following information is displayed, the Deployment is being created.

```
deployment "nginx" created
```

Step 4 Query the Deployment status.

kubectl get deployment

If the following information is displayed, the Deployment is running.

```
NAME           READY     UP-TO-DATE   AVAILABLE   AGE 
nginx         1/1       1            1           4m5s
```

Parameter description

- **NAME**: Name of the application running in the pod.
- **READY**: indicates the number of available workloads. The value is displayed as "the number of available pods/the number of expected pods".
- **UP-TO-DATE**: indicates the number of replicas that have been updated.
- **AVAILABLE**: indicates the number of available pods.
- **AGE**: period the Deployment keeps running

Step 5 If the Deployment will be accessed through a ClusterIP or NodePort Service, add the corresponding Service. For details, see **Service**.

----End

Documentation

- **Upgrading Pods Without Interrupting Services**
- **Configuring Domain Name Resolution for CCE Containers**

2.2.2 Creating a StatefulSet

Scenario

StatefulSets are a type of workloads whose data or status is stored while they are running. For example, MySQL is a StatefulSet because it needs to store new data.

A container can be migrated between different hosts, but data is not stored on the hosts. To store StatefulSet data persistently, mount HA storage volumes provided by CCE to the container.

Constraints

- When you delete or scale a StatefulSet, the system does not delete the storage volumes associated with the StatefulSet to ensure data security.
- When you delete a StatefulSet, reduce the number of replicas to **0** before deleting the StatefulSet so that pods in the StatefulSet can be stopped in order.
- When you create a StatefulSet, a headless Service is required for pod access. For details, see **Headless Service**.

Prerequisites

- A cluster is available. For details about how to create a cluster, see **Buying a CCE Autopilot Cluster**.
- VPC endpoints for accessing SWR and OBS have been configured. For details, see **Configuring VPC Endpoints for Accessing SWR and OBS**.
- A Service of the LoadBalancer type has been created if the workload needs to be accessed by external networks.

  NOTE

  If a pod has multiple containers, the ports used by the containers cannot conflict with each other. If there is a conflict, the Deployment will fail to be created.

Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 Set basic information about the workload.

Basic Info

- **Workload Type**: Select **StatefulSet**.

- **Workload Name**: Enter a name for the workload. Enter 1 to 63 characters starting with a lowercase letter and ending with a lowercase letter or digit. Only lowercase letters, digits, and hyphens (-) are allowed.
- **Namespace**: Select a namespace. The default value is **default**. You can also click **Create Namespace** to create one. For details, see **Creating a Namespace**.
- **Pods**: Enter the number of pods of the workload.

Container Settings

- Container Information

  A pod can have more than one container. You can click **Add Container** on the right to configure multiple containers.

  – **Basic Info**: Configure basic information about each container.

Table 2-4 Basic container information

Parameter	Description
Container Name	Enter a name for the container.
Pull Policy	Image update or pull policy. If you select **Always**, the image is pulled from the image repository each time. If you do not select **Always**, the existing image of the node is preferentially used. If the image does not exist, the image is pulled from the image repository.
Image Name	Click **Select Image** and select the image used by the container. To use a third-party image, see **Using Third-Party Images**.
Image Tag	Select the image tag to be deployed.
CPU Quota	CPU limit, which is the maximum CPU available for the container to prevent excessive resource usage.
Memory Quota	Memory limit, which is the maximum memory available for the container. When the container's memory usage exceeds the memory limit, the container will be terminated.
(Optional) Init Container	Whether the container will be used as an init container. An init container does not support health check. An init container is a special container that runs before other app containers in a pod are started. Each pod can contain multiple containers. In addition, a pod can contain one or more init containers. Application containers in a pod are started and run only after the running of all init containers completes. For details, see **Init Containers**.

- (Optional) **Lifecycle**: Configure operations to be performed in a specific phase of the container lifecycle, such as Startup Command, Post-Start, and Pre-Stop. For details, see **Configuring the Container Lifecycle**.
- (Optional) **Health Check**: Set the liveness probe, ready probe, and startup probe as required. For details, see **Setting Health Check for a Container**.
- (Optional) **Environment Variables**: Configure variables for the container running environment using key-value pairs. These variables transfer external information to containers running in pods and can be flexibly modified after application deployment. For details, see **Configuring Environment Variables**.
- (Optional) **Data Storage**: Mount local storage or cloud storage to the container. The application scenarios and the ways for mounting the volumes vary with the storage type.

  NOTE

  - StatefulSets allow you to mount storage volumes dynamically.

    Dynamic mounting is achieved by using the **volumeClaimTemplates** field and depends on the dynamic creation capability of StorageClass. A StatefulSet associates each pod with a PVC using the **volumeClaimTemplates** field, and the PVC is bound to the corresponding PV. Therefore, after the pod is rescheduled, the original data can still be mounted based on the PVC name.
  - After a workload is created, the storage that is dynamically mounted cannot be updated.
- (Optional) **Security Context**: Assign container permissions to protect the system and other containers from being affected. Enter the user ID to assign container permissions and prevent systems and other containers from being affected.

- **Image Access Credential**: Select the credential used for accessing the image repository. The default value is **default-secret**. You can use default-secret to access images in SWR. For details about **default-secret**, see **default-secret**.

Headless Service Parameters

A headless Service is used to solve the problem of mutual access between pods in a StatefulSet. The headless Service provides a fixed access domain name for each pod. For details, see **Headless Service**.

(Optional) Service Settings

A Service provides external access for pods. With a static IP address, a Service forwards the traffic to pods and automatically balances load for these pods.

You can also create a Service after creating a workload. For details about Services of different types, see **Service**.

(Optional) Advanced Settings

- **Upgrade**: Specify the upgrade mode and upgrade parameters of the workload. **Rolling upgrade** and **Replace upgrade** are supported. For details, see **Configuring the Workload Upgrade Policy**.
- **Pod Management Policies**

  For some distributed systems, the StatefulSet sequence is unnecessary and/or should not occur. These systems require only uniqueness and identifiers.

  - **OrderedReady**: The StatefulSet will deploy, delete, or scale pods in order and one by one. (The StatefulSet continues only after the previous pod is ready or deleted.) This is the default policy.
  - **Parallel**: The StatefulSet will create pods in parallel to match the desired scale without waiting, and will delete all pods at once.
- **Labels and Annotations**: Add labels or annotations for pods using key-value pairs. After entering the key and value, click **Confirm**. For details about how to use and configure labels and annotations, see **Configuring Labels and Annotations**.
- **DNS**: Configure a DNS policy for the workload. For details, see **DNS Configuration**.
- **APM Settings**: Use Application Performance Management (APM) to provide more accurate problem analysis and location for Java programs. For details, see **Configuring APM**.
- Network configuration:
  - Whether to enable the network configuration of a specified container: You can enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. For details, see **Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration**.
  - Specify the container network configuration name: Only the custom container network configuration whose associated resource type is workload can be selected.

Step 4 Click **Create Workload** in the lower right corner.

----End

Using kubectl

NOTICE

Node affinity and anti-affinity are not available for CCE Autopilot clusters. When you use kubectl to create a workload, do not configure the **affinity** field to prevent pod creation failures.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create and edit the **nginx-statefulset.yaml** file.

nginx-statefulset.yaml is an example file name. You can change it as required.

vi nginx-statefulset.yaml

The following provides an example of the file contents. For more information on StatefulSet, see the **Kubernetes documentation**.

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: nginx
spec:
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
        - name: container-1
          image: nginx:latest
          imagePullPolicy: IfNotPresent
          resources:
            requests:
              cpu: 250m
              memory: 512Mi
            limits:
              cpu: 250m
              memory: 512Mi
      imagePullSecrets:
        - name: default-secret
      dnsPolicy: ClusterFirst
  serviceName: nginx-svc
  replicas: 2
  updateStrategy:
    type: RollingUpdate
```

vi nginx-headless.yaml

```
apiVersion: v1
kind: Service
metadata:
  name: nginx-svc
  namespace: default
  labels:
    app: nginx
spec:
  selector:
    app: nginx
    version: v1
  clusterIP: None
  ports:
    - name: nginx
      targetPort: 80
      nodePort: 0
      port: 80
      protocol: TCP
  type: ClusterIP
```

Step 3 Create a workload and the corresponding headless service.

kubectl create -f nginx-statefulset.yaml

If the following information is displayed, the StatefulSet has been successfully created.

```
statefulset.apps/nginx created
```

kubectl create -f nginx-headless.yaml

If the following information is displayed, the headless service has been successfully created.

```
service/nginx-svc created
```

Step 4 If the workload will be accessed through a ClusterIP or NodePort Service, set the corresponding workload access type. For details, see **Service**.

----End

2.2.3 Creating a Job

Scenario

Jobs are short-lived and run for a certain time to completion. They can be executed immediately after being deployed. It is completed after it exits normally (exit 0).

A job is a resource object that is used to control batch tasks. It is different from a long-term servo workload (such as Deployment and StatefulSet).

A job is started and terminated at specific times, while a long-term servo workload runs unceasingly unless being terminated. The pods managed by a Job automatically exit after the job is completed based on user configurations. The success flag varies depending on the **spec.completions** policy.

- One-off jobs: A single pod runs once until successful termination.
- Jobs with a fixed success count: N pods run until successful termination.
- Parallel jobs in a work queue: Jobs are considered completed based on the global success confirmed by the application.

Prerequisites

- A cluster is available. For details about how to create a cluster, see **Buying a CCE Autopilot Cluster**.
- VPC endpoints for accessing SWR and OBS have been configured. For details, see **Configuring VPC Endpoints for Accessing SWR and OBS**.
- A Service of the LoadBalancer type has been created if the workload needs to be accessed by external networks.

  NOTE

  If a pod has multiple containers, the ports used by the containers cannot conflict with each other. If there is a conflict, the Deployment will fail to be created.

Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 Set basic information about the workload.

Basic Info

- **Workload Type**: Select **Job**.
- **Workload Name**: Enter a name for the workload. Enter 1 to 63 characters starting with a lowercase letter and ending with a lowercase letter or digit. Only lowercase letters, digits, and hyphens (-) are allowed.

- **Namespace**: Select a namespace. The default value is **default**. You can also click **Create Namespace** to create one. For details, see **Creating a Namespace**.
- **Pods**: Enter the number of pods of the workload.

Container Settings

- Container Information

  A pod can have more than one container. You can click **Add Container** on the right to configure multiple containers.

  – **Basic Info**: Configure basic information about each container.

Table 2-5 Basic container information

Parameter	Description
Container Name	Enter a name for the container.
Pull Policy	Image update or pull policy. If you select **Always**, the image is pulled from the image repository each time. If you do not select **Always**, the existing image of the node is preferentially used. If the image does not exist, the image is pulled from the image repository.
Image Name	Click **Select Image** and select the image used by the container. To use a third-party image, see **Using Third-Party Images**.
Image Tag	Select the image tag to be deployed.
CPU Quota	CPU limit, which is the maximum CPU available for the container to prevent excessive resource usage.
Memory Quota	Memory limit, which is the maximum memory available for the container. When the container's memory usage exceeds the memory limit, the container will be terminated.
(Optional) Init Container	Whether the container will be used as an init container. An init container does not support health check. An init container is a special container that runs before other app containers in a pod are started. Each pod can contain multiple containers. In addition, a pod can contain one or more init containers. Application containers in a pod are started and run only after the running of all init containers completes. For details, see **Init Containers**.

  – (Optional) **Lifecycle**: Configure operations to be performed in a specific phase of the container lifecycle, such as Startup Command, Post-Start, and Pre-Stop. For details, see **Configuring the Container Lifecycle**.

  – (Optional) **Environment Variables**: Configure variables for the container running environment using key-value pairs. These variables transfer external information to containers running in pods and can be flexibly modified after application deployment. For details, see **Configuring Environment Variables**.

  - (Optional) **Data Storage**: Mount local storage or cloud storage to the container. The application scenarios and the ways for mounting the volumes vary with the storage type. For details, see **Storage**.

- **Image Access Credential**: Select the credential used for accessing the image repository. The default value is **default-secret**. You can use default-secret to access images in SWR. For details about **default-secret**, see **default-secret**.

(Optional) Advanced Settings

- **Labels and Annotations**: Add labels or annotations for pods using key-value pairs. After entering the key and value, click **Confirm**. For details about how to use and configure labels and annotations, see **Configuring Labels and Annotations**.
- **Job Settings**
  - **Parallel Pods**: Maximum number of pods that can run in parallel during job execution. The value cannot be greater than the total number of pods in the job.
  - **Timeout (s)**: Once a job reaches this time, the job status becomes failed and all pods in this job will be deleted. If you leave this parameter blank, the job will never time out.
  - Completion Mode
    - **Non-indexed**: A job is considered complete when all the pods are successfully executed. Each pod completion is homologous to each other.
    - **Indexed**: Each pod gets an associated completion index from 0 to the number of pods minus 1. The job is considered complete when every pod allocated with an index is successfully executed. For an indexed job, pods are named in the format of $(job-name)-$(index).
  - **Suspend Job**: By default, a job is executed immediately after being created. The job's execution will be suspended if you enable this option, and resumed after you disable it.
- Network configuration:
  - Whether to enable the network configuration of a specified container: You can enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. For details, see **Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration**.
  - Specify the container network configuration name: Only the custom container network configuration whose associated resource type is workload can be selected.

Step 4 Click **Create Workload** in the lower right corner.

----End

Using kubectl

NOTICE

Node affinity and anti-affinity are not available for CCE Autopilot clusters. When you use kubectl to create a workload, do not configure the **affinity** field to prevent pod creation failures.

A job has the following configuration parameters:

- **.spec.completions**: indicates the number of pods that need to run successfully to end a job. The default value is **1**.
- **.spec.parallelism**: indicates the number of pods that run concurrently. The default value is **1**.
- **.spec.backoffLimit**: indicates the maximum number of retries performed if a pod fails. When the limit is exceeded, the pod will not try again.
- **.spec.activeDeadlineSeconds**: indicates the running time of pods. Once the time is reached, the job and all pods are terminated. **.spec.activeDeadlineSeconds** has a higher priority than **.spec.backoffLimit**. If a job reaches **.spec.activeDeadlineSeconds**, **spec.backoffLimit** is ignored.

Based on the **.spec.completions** and **.spec.parallelism** settings, jobs are classified into the following types.

Table 2-6 Job types

Job Type	Description	.spec.completions	.spec.parallelism
One-off jobs	A job creates one pod until it successfully completes.	1	1
Jobs with a fixed completion count	A job creates one pod in sequence and is completed when the number of successful pods reaches the value of **.spec.completions**.	>1	1
Parallel jobs with a fixed completion count	A job creates multiple pods in sequence and is completed when the number of successful pods reaches the value of **.spec.completions**.	>1	>1

Job Type	Description	.spec.completions	.spec.parallelism
Parallel jobs in a work queue	A job creates one or more pods. Each pod takes one task from the message queue, processes it, and repeats until the end of the queue is reached. Then the pod deletes the task and exists. For details, see **Fine Parallel Processing Using a Work Queue**.	Left blank	> 1 or = 1

The following is an example job, which calculates π till the 2,000th digit and prints the output.

```
apiVersion: batch/v1
kind: Job
metadata:
  name: myjob
spec:
  completions: 50        # 50 pods need to be run to finish a job. In this example, π is printed for 50 times.
  parallelism: 5        # 5 pods are run in parallel.
  backoffLimit: 5        # The maximum number of retry times is 5.
  template:
    spec:
      containers:
      - name: pi
        image: perl
        command: ["perl",  "-Mbignum=bpi", "-wle", "print bpi(2000)"]
      restartPolicy: Never
      imagePullSecrets:
        - name: default-secret
```

Description

- **apiVersion: batch/v1** indicates the version of the current job.
- **kind: Job** indicates that the current resource is a job.
- **restartPolicy: Never** indicates the current restart policy. For jobs, this parameter can only be set to **Never** or **OnFailure**. For other controllers (for example, Deployments), you can set this parameter to **Always**.

Run the job.

Step 1 Start the job.

```
[root@k8s-master k8s]# kubectl apply -f myjob.yaml
job.batch/myjob created
```

Step 2 View the job details.

kubectl get job

```
[root@k8s-master k8s]# kubectl get job
NAME    COMPLETIONS   DURATION   AGE
myjob   50/50         23s        3m45s
```

If the value of **COMPLETIONS** is **50/50**, the job is successfully executed.

Step 3 Query the pod status.

kubectl get pod

```
[root@k8s-master k8s]# kubectl get pod
NAME          READY   STATUS      RESTARTS   AGE
myjob-29qlw   0/1     Completed   0          4m5s
...
```

If the status is **Completed**, the job is complete.

Step 4 View the pod logs.

kubectl logs

```
# kubectl logs myjob-29qlw
3.141592653589793238462643383279502884197169399375105820974944592307816406286208998628034
8253421170679821480865132823066470938446095505822317253594081284811174502841027019385211
0555964462294895493038196442881097566593344612847564823378678316527120190914564856692346
0348610454326648213393607260249141273724587006606315588174881520920962829254091715364367
8925903600113305305488204665213841469519415116094330572703657595919530921861173819326117
9310511854807446237996274956735188575272489122793818301194912983367336244065664308602139
4946395224737190702179860943702770539217176293176752384674818467669405132000568127145263
5608277857713427577896091736371787214684409012249534301465495853710507922796892589235420
1995611212902196086403441815981362977477130996051870721134999999837297804995105973173281
6096318595024459455346908302642522308253344685035261931188171010003137838752886587533208
3814206171776691473035982534904287554687311595628638823537875937519577818577805321712268
0661300192787661119590921642019893809525720106548586327886593615338182796823030195203530
1852968995773622599413891249721775283479131515574857242454150695950829533116861727855889
0750983817546374649393192550604009277016711390098488240128583616035637076601047101819429
5559619894676783744944825537977472684710404753464620804668425906949129331367702898915210
4752162056966024058038150193511253382430035587640247496473263914199272604269922796782354
7816360093417216412199245863150302861829745557067498385054945885869269956909272107975093
0295532116534498720275596023648066549911988183479775356636980742654252786255181841757467
2890977772793800081647060016145249192173217214772350141441973568548161361157352552133475
7418494684385233239073941433345477624168625189835694855620992192221842725502542568876717
9049460165346680498862723279178608578438382796797668145410095388378636095068006422512520
5117392984896084128488626945604241965285022210661186306744278622039194945047123713786960
956364371917287467764657573962413890865832645995813390478027590
```

----End

Related Operations

After a one-off job is created, you can perform operations described in **Table 2-7**.

Table 2-7 Other operations

Operation	Description
Deleting a job	1. Select the job to be deleted and click **More** > **Delete** in the **Operation** column. 2. Click **Yes**. Deleted jobs cannot be recovered.

2.2.4 Creating a CronJob

Scenario

A CronJob runs on a repeating schedule. You can perform time synchronization for all active nodes at a fixed time point.

Similar to Linux crontab, a CronJob runs periodically at the specified time and has the following characteristics:

- A CronJob runs only once at the specified time.
- A CronJob runs periodically at the specified time.

A CronJob is typically used to:

- Schedule jobs at the specified time.
- Create jobs to run periodically, for example, database backup and email sending.

Prerequisites

- A cluster is available. For details about how to create a cluster, see **Buying a CCE Autopilot Cluster**.
- VPC endpoints for accessing SWR and OBS have been configured. For details, see **Configuring VPC Endpoints for Accessing SWR and OBS**.
- A Service of the LoadBalancer type has been created if the workload needs to be accessed by external networks.

  NOTE

  If a pod has multiple containers, the ports used by the containers cannot conflict with each other. If there is a conflict, the Deployment will fail to be created.

Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 Set basic information about the workload.

Basic Info

- **Workload Type**: Select **CronJob**.
- **Workload Name**: Enter a name for the workload. Enter 1 to 63 characters starting with a lowercase letter and ending with a lowercase letter or digit. Only lowercase letters, digits, and hyphens (-) are allowed.
- **Namespace**: Select a namespace. The default value is **default**. You can also click **Create Namespace** to create one. For details, see **Creating a Namespace**.

Container Settings

- Container Information

A pod can have more than one container. You can click **Add Container** on the right to configure multiple containers.

– **Basic Info**: Configure basic information about each container.

Table 2-8 Basic container information

Parameter	Description
Container Name	Enter a name for the container.
Pull Policy	Image update or pull policy. If you select **Always**, the image is pulled from the image repository each time. If you do not select **Always**, the existing image of the node is preferentially used. If the image does not exist, the image is pulled from the image repository.
Image Name	Click **Select Image** and select the image used by the container. To use a third-party image, see **Using Third-Party Images**.
Image Tag	Select the image tag to be deployed.
CPU Quota	CPU limit, which is the maximum CPU available for the container to prevent excessive resource usage.
Memory Quota	Memory limit, which is the maximum memory available for the container. When the container's memory usage exceeds the memory limit, the container will be terminated.
(Optional) Init Container	Whether the container will be used as an init container. An init container does not support health check. An init container is a special container that runs before other app containers in a pod are started. Each pod can contain multiple containers. In addition, a pod can contain one or more init containers. Application containers in a pod are started and run only after the running of all init containers completes. For details, see **Init Containers**.

– (Optional) **Lifecycle**: Configure operations to be performed in a specific phase of the container lifecycle, such as Startup Command, Post-Start, and Pre-Stop. For details, see **Configuring the Container Lifecycle**.

– (Optional) **Environment Variables**: Configure variables for the container running environment using key-value pairs. These variables transfer external information to containers running in pods and can be flexibly modified after application deployment. For details, see **Configuring Environment Variables**.

- **Image Access Credential**: Select the credential used for accessing the image repository. The default value is **default-secret**. You can use default-secret to access images in SWR. For details about **default-secret**, see **default-secret**.

Schedule

- **Concurrency Policy**: There are three options:
  - **Forbid**: A new job cannot be created before the previous job is completed.
  - **Allow**: The CronJob allows concurrently running jobs, which preempt cluster resources.
  - **Replace**: A new job replaces the previous job when it is time to create a job but the previous job is not completed.
- **Policy Settings**: Specify when a new job is executed. Policy settings in YAML are implemented using cron expressions.
  - A job is executed at a fixed interval. The unit can be minute, hour, day, or month. For example, if a job is executed every 30 minutes and the corresponding cron expression is ***/30 * * * ***, the execution time starts from 0 in the unit range, for example, **00:00:00**, **00:30:00**, **01:00:00**, and **...**.
  - A job is executed by month. For example, if a job is executed at 00:00 on the first day of each month, the cron expression is **0 0 1 */1 ***, and the execution time is ******-01-01 00:00:00**, ******-02-01 00:00:00**, and **...**.
  - A job is executed by week. For example, if a job is executed at 00:00 every Monday, the cron expression is **0 0 * * 1**, and the execution time is ******-**-01 00:00:00 on Monday**, ******-**-08 00:00:00 on Monday**, and **...**.
  - **Custom Cron Expression**: For details about how to use cron expressions, see **CronJob**.

NOTE

  - If a job is executed at a fixed time (by month) and the date in a month does not exist, the job will not be executed in this month. For example, the execution will skip February if the date is set to 30.
  - Due to the definition of cron, the fixed period is not a strict period. The time unit range is divided from 0 by period. For example, if the unit is minute, the value ranges from 0 to 59. If the value cannot be exactly divided, the last period is reset. Therefore, an accurate period can be represented only when the period can be evenly divided.

    Take a job that is executed by hour as an example. As **/2, /3, /4, /6, /8, and /12** can exactly divide 24 hours, an accurate period can be represented. If another period is used, the last period will be reset at the beginning of a new day. For example, if the cron expression is *** */12 * * ***, the execution time is **00:00:00** and **12:00:00** every day. If the cron expression is *** */13 * * ***, the execution time is **00:00:00** and **13:00:00** every day. At 00:00 on the next day, the execution time is updated even if the period does not reach 13 hours.

- **Job Records**: You can set the number of jobs that are successfully executed or fail to be executed. If the values are set to **0**, none of the jobs will be kept after they finish.

(Optional) Advanced Settings

- **Labels and Annotations**: Add labels or annotations for pods using key-value pairs. After entering the key and value, click **Confirm**. For details about how to use and configure labels and annotations, see **Configuring Labels and Annotations**.
- Network configuration:

- Whether to enable the network configuration of a specified container: You can enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. For details, see **Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration**.
- Specify the container network configuration name: Only the custom container network configuration whose associated resource type is workload can be selected.

Step 4 Click **Create Workload** in the lower right corner.

----End

Using kubectl

NOTICE

Node affinity and anti-affinity are not available for CCE Autopilot clusters. When you use kubectl to create a workload, do not configure the **affinity** field to prevent pod creation failures.

A CronJob has the following configuration parameters:

- **.spec.schedule**: takes a **Cron** format string, for example, **0 * * * *** or **@hourly**, as schedule time of jobs to be created and executed.
- **.spec.jobTemplate**: specifies jobs to be run and has the same schema as when you are **Creating a Job Using kubectl**.
- **.spec.startingDeadlineSeconds**: specifies the deadline for starting a job.
- **.spec.concurrencyPolicy**: specifies how to treat concurrent executions of a job created by the CronJob. The following options are supported:
  - **Allow** (default value): allows concurrently running jobs.
  - **Forbid**: forbids concurrent runs, skipping next run if previous has not finished yet.
  - **Replace**: cancels the currently running job and replaces it with a new one.

The following is an example CronJob, which is saved in the **cronjob.yaml** file.

NOTE

In clusters of v1.21 or later, the **apiVersion** value of CronJobs is **batch/v1**.

In clusters earlier than v1.21, the **apiVersion** value of CronJobs is **batch/v1beta1**.

```
apiVersion: batch/v1
kind: CronJob
metadata:
  name: hello
spec:
  schedule: "*/1 * * * *"
  jobTemplate:
    spec:
      template:
        spec:
```

```
        containers:
        - name: hello
          image: busybox
          command:
          - /bin/sh
          - -c
          - date; echo Hello from the Kubernetes cluster
        restartPolicy: OnFailure
        imagePullSecrets:
          - name: default-secret
```

Run the job.

Step 1 Create a CronJob.

kubectl create -f cronjob.yaml

Information similar to the following is displayed:

```
cronjob.batch/hello created
```

Step 2 Query the running status of a CronJob.

kubectl get cronjob

```
NAME      SCHEDULE      SUSPEND   ACTIVE    LAST SCHEDULE   AGE
hello     */1 * * * *   False     0         <none>          9s
```

kubectl get jobs

```
NAME               COMPLETIONS   DURATION   AGE
hello-1597387980   1/1           27s        45s
```

kubectl get pod

```
NAME                           READY     STATUS      RESTARTS   AGE
hello-1597387980-tjv8f         0/1       Completed   0          114s
hello-1597388040-lckg9         0/1       Completed   0          39s
```

kubectl logs hello-1597387980-tjv8f

```
Fri Aug 14 06:56:31 UTC 2020
Hello from the Kubernetes cluster
```

kubectl delete cronjob hello

```
cronjob.batch "hello" deleted
```

NOTICE

If a CronJob is deleted, the related jobs and pods are deleted accordingly.

----End

Related Operations

After a CronJob is created, you can perform operations described in **Table 2-9**.

Table 2-9 Other operations

Operation	Description
Editing a YAML file	Click **More** > **Edit YAML** next to the CronJob name to edit the YAML file of the current job.
Stopping a CronJob	1. Select the CronJob to be stopped and click **Stop** in the **Operation** column. 2. Click **Yes**.
Deleting a CronJob	1. Select the CronJob to be deleted and click **More** > **Delete** in the **Operation** column. 2. Click **Yes**. Deleted CronJobs cannot be recovered.

2.3 Configuring a Workload

2.3.1 Configuring an Image Pull Policy

When a workload is created, the container image is pulled from the image repository to the node. The image is also pulled when the workload is restarted or upgraded.

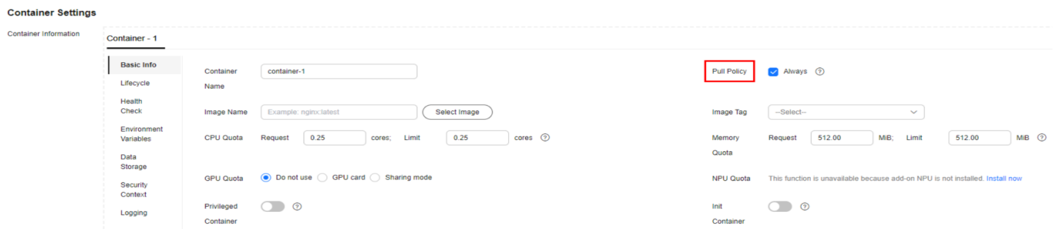
By default, **imagePullPolicy** is set to **IfNotPresent**, indicating that if the image exists on the node, the existing image is used. If the image does not exist on the node, the image is pulled from the image repository.

The image pull policy can also be set to **Always**, indicating that the image is pulled from the image repository and overwrites the image on the node regardless of whether the image exists on the node.

```
apiVersion: v1
kind: Pod
metadata:
  name: nginx
spec:
  containers:
  - image: nginx:alpine
    name: container-0
    resources:
      limits:
        cpu: 100m
        memory: 200Mi
      requests:
        cpu: 100m
        memory: 200Mi
    imagePullPolicy: Always
  imagePullSecrets:
  - name: default-secret
```

An image pull policy can also be configured on the CCE console. When creating a workload, configure **Pull Policy**. If **Always** is selected, images are always pulled. If **Always** is not selected, images are pulled as needed.

Figure 2-4 Configuring an update policy



Container Name: container-1
Image Name: Example: nginx:latest
CPU Quota: Request 0.25 cores; Limit 0.25 cores
GPU Quota: (●) Do not use ( ) GPU card ( ) Sharing mode
Privileged Container: off
Pull Policy: ✓ Always
Image Tag: --Select--
Memory Quota: Request 512.00 MiB; Limit 512.00 MiB
NPU Quota: This function is unavailable because add-on NPU is not installed. Install now
Init Container: off

NOTICE

Use a new tag each time you create an image. If you do not update the tag but only update the image, when **Pull Policy** is set to **IfNotPresent**, CCE considers that an image with the tag already exists on the current node and will not pull the image again.

2.3.2 Using Third-Party Images

Scenario

CCE allows you to create workloads using images pulled from third-party image repositories.

Generally, a third-party image repository can only be accessed after authentication (using your username and password). CCE uses the secret-based authentication to pull images. You need to create a secret for an image repository before pulling images from the repository.

Prerequisites

CCE Autopilot can access the network where the private repository is located. There are two options for this:

- Access over a private network: Ensure that the VPC of the private repository is the same as the VPC where the cluster resides.
- Access over Direct Connect or VPN: Connect the private repository network to the VPC where the cluster resides through Direct Connect or VPN.

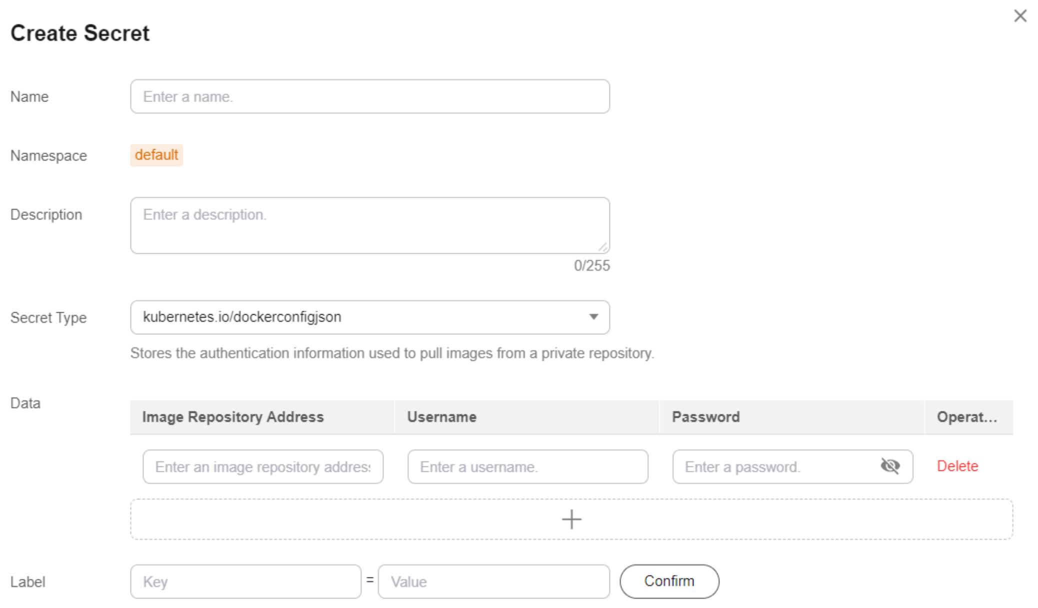
Using the Console

Step 1 Create a secret for accessing a third-party image repository.

Click the cluster name to access the cluster console. In the navigation pane on the left, choose **ConfigMaps and Secrets**. On the **Secrets** tab, click **Create Secret** in the upper right corner. Set **Secret Type** to **kubernetes.io/dockerconfigjson**. For details, see **Creating a Secret**.

Enter the username and password used to access the third-party image repository.

Figure 2-5 Creating a secret



Create Secret

Name: Enter a name.

Namespace: default

Description: Enter a description. 0/255

Secret Type: kubernetes.io/dockerconfigjson

Stores the authentication information used to pull images from a private repository.

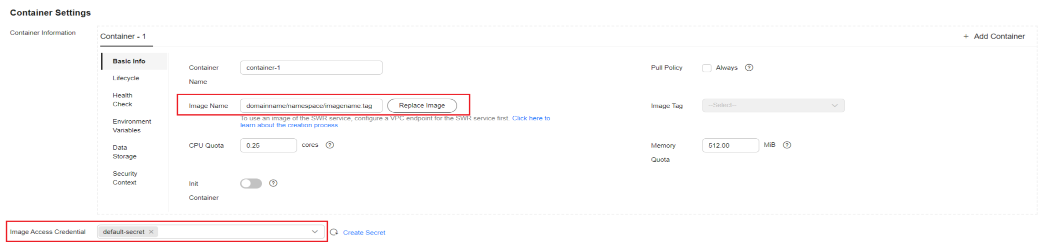
Data:

Image Repository Address	Username	Password	Operat...
Enter an image repository addres:	Enter a username.	Enter a password.	Delete

Label: Key = Value Confirm

Step 2 When creating a workload, you can enter a private image path (such as **domainname/namespace/imagename:tag**) in **Image Name** and select the key created in **Step 1**.

Figure 2-6 Private image path



Container Settings

Container Information

Container - 1 + Add Container

Basic Info / Lifecycle / Health Check / Environment Variables / Data Storage / Security Context

Container Name: container-1

Pull Policy: Always

Image Name: domainname/namespace/imagename:tag Replace Image

To use an image of the SWR service, configure a VPC endpoint for the SWR service first. Click here to learn about the creation process

Image Tag: --Select--

CPU Quota: 0.25 cores

Memory Quota: 512.00 MiB

Init Container

Image Access Credential: default-secret

Create Secret

Step 3 Set other parameters and click **Create Workload**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Use kubectl to create a secret of the kubernetes.io/dockerconfigjson.

```
kubectl create secret docker-registry myregistrykey  -n default --docker-
server=DOCKER_REGISTRY_SERVER --docker-username=DOCKER_USER --docker-
password=DOCKER_PASSWORD --docker-email=DOCKER_EMAIL
```

In the preceding command, *myregistrykey* indicates the key name, *default* indicates the namespace where the key is located, and other parameters are as follows:

- **DOCKER_REGISTRY_SERVER**: address of a third-party image repository, for example, **www.3rdregistry.com** or **10.10.10.10:443**

- **DOCKER_USER**: account used for logging in to a third-party image repository
- **DOCKER_PASSWORD**: password used for logging in to a third-party image repository
- **DOCKER_EMAIL**: email of a third-party image repository

Step 3 Use a third-party image to create a workload.

A kubernetes.io/dockerconfigjson secret is used for authentication when you obtain a private image. The following is an example of using the myregistrykey for authentication.

```
apiVersion: v1
kind: Pod
metadata:
  name: foo
  namespace: default
spec:
  containers:
    - name: foo
      image: www.3rdregistry.com/janedoe/awesomeapp:v1
  imagePullSecrets:
    - name: myregistrykey              #Use the created secret.
```

----End

2.3.3 Configuring the Container Lifecycle

Scenario

CCE provides hooks for container lifecycle management. For example, you can use a hook to make a container perform a specific operation before stopping.

CCE provides the following hooks:

- **Startup Command**: executed to start a container. For details, see **Startup Command**.
- **Post-Start**: executed immediately after a container is started. For details, see **Post-Start**.
- **Pre-Stop**: executed before a container is stopped. This hook helps you ensure that the services running on the pods can be completed in advance in case of pod upgrade or deletion. For details, see **Pre-Stop**.

Startup Command

By default, the default command is executed during image start. To run a specific command or rewrite the default image setting, you must perform specific operations.

A Docker image has metadata that stores image information. If lifecycle commands and arguments are not set, CCE runs the default commands and arguments (Docker instructions **ENTRYPOINT** and **CMD**) provided during image creation.

If the commands and arguments used to run a container are set during workload creation, the default commands **ENTRYPOINT** and **CMD** are overwritten during image build.

Table 2-10 Commands and arguments used to run a container

Image ENTRYPOINT	Image CMD	Command to Run a Container	Arguments to Run a Container	Command Executed
[touch]	[/root/test]	Not set	Not set	[touch /root/test]
[touch]	[/root/test]	[mkdir]	Not set	[mkdir]
[touch]	[/root/test]	Not set	[/opt/test]	[touch /opt/test]
[touch]	[/root/test]	[mkdir]	[/opt/test]	[mkdir /opt/test]

Step 1 Log in to the CCE console. When creating a workload, configure container information and select **Lifecycle**.

Step 2 On the **Startup Command** tab, enter a command and arguments.

Table 2-11 Container startup command

Parameter	Description
Command	Enter an executable command, for example, **/run/server**. If there are multiple executable commands, write them on different lines. **NOTE** If there are multiple commands, it is recommended that you run **/bin/sh** or other **shell** commands and use other commands as parameters.
Args	Enter an argument for the command, for example, **--port=8080**. If there are multiple arguments, write them on different lines.

----End

Post-Start

Step 1 Log in to the CCE console. When creating a workload, configure container information and select **Lifecycle**.

Step 2 On the **Post-Start** tab, configure the parameters.

Table 2-12 Post-Start parameters

Parameter	Description

CLI	The tool for running the commands for Post-Start processing. The command format is **Command Args[1] Args[2]...**. **Command** is a system command or a user-defined executable program. If no path is specified, an executable program in the default path will be selected. If multiple commands need to be executed, write them into a script for execution. **Commands that are executed in the backend or asynchronously are not supported.**
	Example command:
	`exec:`
	`  command:`
	`  - /install.sh`
	`  - install_agent`
	Enter **/install install_agent** in the script. This command indicates that **install.sh** will be executed after the container is created.

----End

Pre-Stop

Step 1 Log in to the CCE console. When creating a workload, configure container information and select **Lifecycle**.

Step 2 On the **Pre-Stop** tab, configure the parameters.

Table 2-13 Pre-Stop parameters

Parameter	Description

CLI	The tool for running the commands for Pre-Stop processing. The command format is **Command Args[1] Args[2]...**. **Command** is a system command or a user-defined executable program. If no path is specified, an executable program in the default path will be selected. If multiple commands need to be executed, write them into a script for execution.
	Example command:
	`exec:`
	`  command:`
	`  - /uninstall.sh`
	`  - uninstall_agent`
	Enter **/uninstall uninstall_agent** in the script. This command indicates that **uninstall.sh** will be executed before the container completes its execution and stops running.

----End

YAML Example

Nginx is used as an example to describe how to set the container lifecycle.

In the following configuration file, there is a Post-Start command (**install.sh**) in the **/bin/bash** directory and a PreStop command (**uninstall.sh**).

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
      - image: nginx
        command:
        - sleep 3600                      #Startup command
        imagePullPolicy: Always
        lifecycle:
          postStart:
            exec:
              command:
              - /bin/bash
              - install.sh                  #Post-Start command
          preStop:
            exec:
              command:
              - /bin/bash
              - uninstall.sh                 #Pre-Stop command
        name: nginx
      imagePullSecrets:
      - name: default-secret
```

2.3.4 Setting Health Check for a Container

Scenario

Health check regularly checks the health status of containers during container running. If the health check function is not configured, a pod cannot detect application exceptions or automatically restart the application to restore it. This will result in a situation where the pod status is normal but the application in the pod is abnormal.

Kubernetes provides the following health check probes:

- **Liveness probe** (livenessProbe): checks whether a container is still alive. It is similar to the **ps** command that checks whether a process exists. If the liveness check of a container fails, the cluster restarts the container. If the liveness check is successful, no operation is executed.
- **Readiness probe** (readinessProbe): checks whether a container is ready to process user requests. Upon that the container is detected unready, service traffic will not be directed to the container. It may take a long time for some applications to start up before they can provide services. This is because that they need to load disk data or rely on startup of an external module. In this

case, the application process is running, but the application cannot provide services. To address this issue, this health check probe is used. If the container readiness check fails, the cluster masks all requests sent to the container. If the container readiness check is successful, the container can be accessed.

- **Startup probe** (startupProbe): checks when a containerized application has started. If such a probe is configured, it disables liveness and readiness checks until it succeeds, ensuring that those probes do not interfere with the application startup. This can be used to adopt liveness checks on slow starting containers, avoiding them getting terminated by the kubelet before they are started.

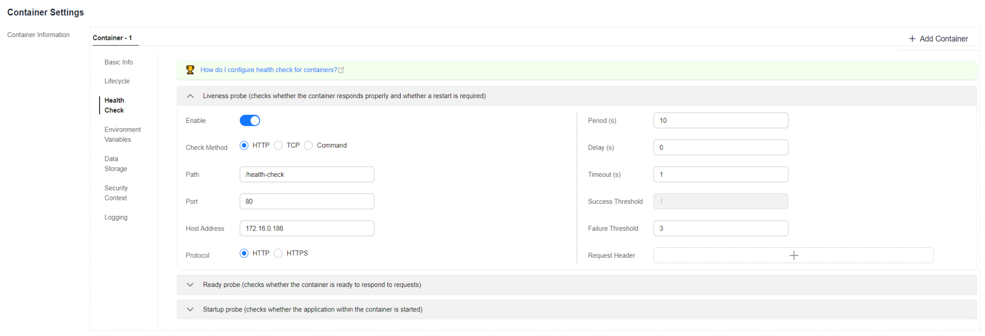
Check Method

- **HTTP request**

  This health check mode applies to containers that provide HTTP/HTTPS services. The cluster periodically initiates an HTTP/HTTPS GET request to such containers. If the return code of the HTTP/HTTPS response is within 200–399, the probe is successful. Otherwise, the probe fails. In this health check mode, you must specify a container listening port and an HTTP/HTTPS request path.

  For example, for a container that provides HTTP services, the HTTP check path is **/health-check**, the port is 80, and the host address is optional (which defaults to the container IP address). Here, 172.16.0.186 is used as an example, and we can get such a request: GET http://172.16.0.186:80/health-check. The cluster periodically initiates this request to the container. You can also add one or more headers to an HTTP request. For example, set the request header name to **Custom-Header** and the corresponding value to **example**.

  Figure 2-7 HTTP request-based check



Container Settings

Container Information

Container - 1

Add Container

Basic Info
Lifecycle
Health Check
Environment Variables
Data Storage
Security Context
Logging

How do I configure health check for containers?

Liveness probe (checks whether the container responds properly and whether a restart is required)

Enable	On
Check Method	[✓ HTTP] TCP Command
Path	/health-check
Port	80
Host Address	172.16.0.186
Protocol	[✓ HTTP] HTTPS
Period (s)	10
Delay (s)	0
Timeout (s)	1
Success Threshold	1
Failure Threshold	3
Request Header	

Ready probe (checks whether the container is ready to respond to requests)

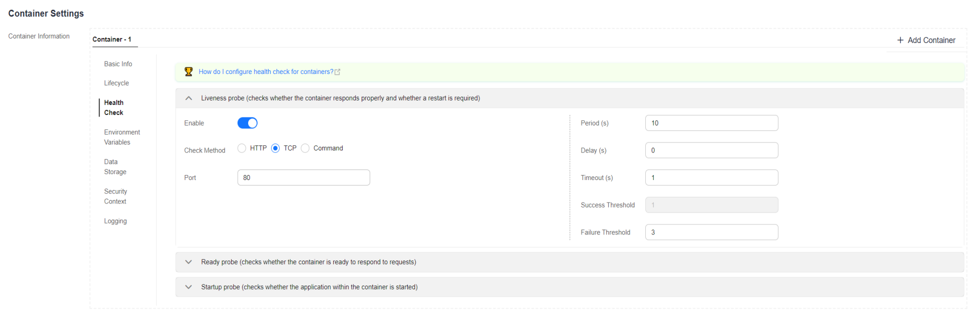
Startup probe (checks whether the application within the container is started)

- **TCP port**

  For a container that provides TCP communication services, the cluster periodically establishes a TCP connection to the container. If the connection is successful, the probe is successful. Otherwise, the probe fails. In this health check mode, you must specify a container listening port.

  For example, if you have an Nginx container with service port 80, after you specify TCP port 80 for container listening, the cluster will periodically initiate a TCP connection to port 80 of the container. If the connection is successful, the probe is successful. Otherwise, the probe fails.

Figure 2-8 TCP port-based check



Container Settings

Container Information | Container - 1 | + Add Container

Basic Info / Lifecycle / Health Check / Environment Variables / Data Storage / Security Context / Logging

How do I configure health check for containers?

Liveness probe (checks whether the container responds properly and whether a restart is required)

Enable: on
Check Method: HTTP [✓ TCP] Command
Port: 80
Period (s): 10
Delay (s): 0
Timeout (s): 1
Success Threshold: 1
Failure Threshold: 3

Ready probe (checks whether the container is ready to respond to requests)

Startup probe (checks whether the application within the container is started)

- **CLI**

  CLI is an efficient tool for health check. When using the CLI, you must specify an executable command in a container. The cluster periodically runs the command in the container. If the command output is 0, the health check is successful. Otherwise, the health check fails.

  The CLI mode can be used to replace the HTTP request-based and TCP port-based health check.

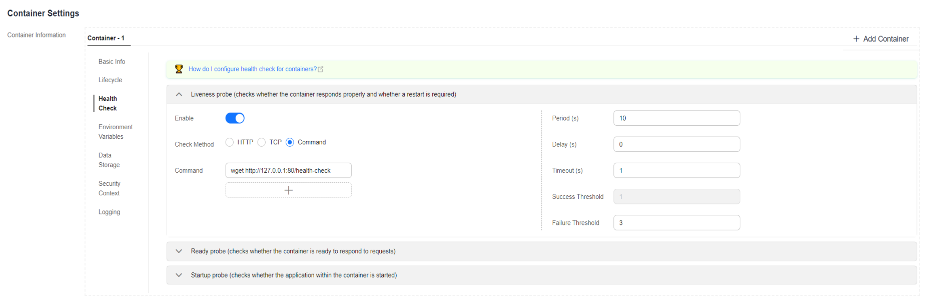
  – For a TCP port, you can use a program script to connect to a container port. If the connection is successful, the script returns **0**. Otherwise, the script returns **–1**.

  – For an HTTP request, you can use the script command to run the **wget** command to detect the container.

    wget http://127.0.0.1:80/health-check

    Check the return code of the response. If the return code is within 200–399, the script returns **0**. Otherwise, the script returns **–1**.

Figure 2-9 Check using CLI



Container Settings

Container Information | Container - 1 | + Add Container

Basic Info / Lifecycle / Health Check / Environment Variables / Data Storage / Security Context / Logging

How do I configure health check for containers?

Liveness probe (checks whether the container responds properly and whether a restart is required)

Enable: on
Check Method: HTTP TCP [✓ Command]
Command: wget http://127.0.0.1:80/health-check
Period (s): 10
Delay (s): 0
Timeout (s): 1
Success Threshold: 1
Failure Threshold: 3

Ready probe (checks whether the container is ready to respond to requests)

Startup probe (checks whether the application within the container is started)

NOTICE

- Put the program to be executed in the container image so that the program can be executed.
- If the command to be executed is a shell script, do not directly specify the script as the command, but add a script parser. For example, if the script is **/data/scripts/health_check.sh**, you must specify **sh/data/scripts/health_check.sh** for command execution. The reason is that the cluster is not in the terminal environment when executing programs in a container.

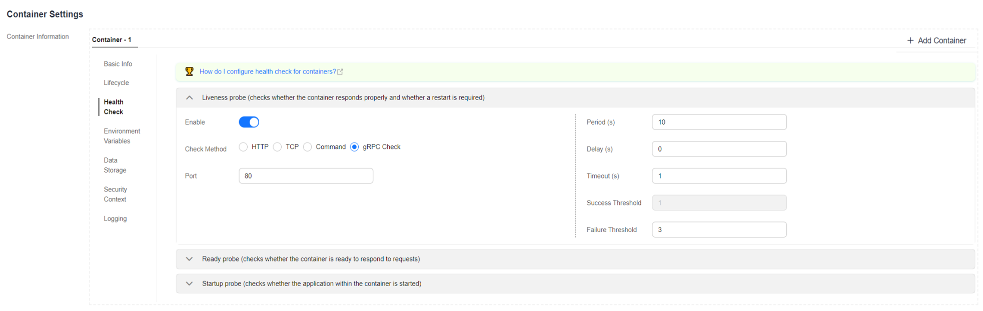
- **gRPC Check**

  gRPC checks can configure startup, liveness, and readiness probes for your gRPC application without exposing any HTTP endpoint, nor do you need an executable. Kubernetes can connect to your workload via gRPC and obtain its status.

NOTICE

- To use gRPC for check, your application must support the **gRPC health checking protocol**.
- Similar to HTTP and TCP probes, if the port is incorrect or the application does not support the health checking protocol, the check fails.

Figure 2-10 gRPC check



Liveness probe (checks whether the container responds properly and whether a restart is required)

Enable: on

Check Method: HTTP TCP Command [✓ gRPC Check]

Port: 80

Period (s): 10

Delay (s): 0

Timeout (s): 1

Success Threshold: 1

Failure Threshold: 3

Ready probe (checks whether the container is ready to respond to requests)

Startup probe (checks whether the application within the container is started)

Common Parameters

Table 2-14 Common parameter description

Parameter	Description
Period (periodSeconds)	Indicates the probe detection period, in seconds. For example, if this parameter is set to **30**, the detection is performed every 30 seconds.
Delay (initialDelaySeconds)	Check delay time in seconds. Set this parameter according to the normal startup time of services. For example, if this parameter is set to **30**, the health check will be started 30 seconds after the container is started. The time is reserved for containerized services to start.
Timeout (timeoutSeconds)	Number of seconds after which the probe times out. Unit: second. For example, if this parameter is set to **10**, the timeout wait time for performing a health check is 10s. If the wait time elapses, the health check is regarded as a failure. If the parameter is left blank or set to **0**, the default timeout time is 1s.

Parameter	Description
Success Threshold (successThreshold)	Minimum consecutive successes for the probe to be considered successful after having failed. For example, if this parameter is set to **1**, the workload status is normal only when the health check is successful for one consecutive time after the health check fails. The default value is **1**, which is also the minimum value. The value of this parameter is fixed to **1** in **Liveness Probe** and **Startup Probe**.
Failure Threshold (failureThreshold)	Number of retry times when the detection fails. Giving up in case of liveness probe means to restart the container. In case of readiness probe the pod will be marked **Unready**. The default value is **3**. The minimum value is **1**.

YAML Example

```
apiVersion: v1
kind: Pod
metadata:
  labels:
    test: liveness
  name: liveness-http
spec:
  containers:
  - name: liveness
    image: nginx:alpine
    args:
    - /server
    livenessProbe:
      httpGet:
        path: /healthz
        port: 80
        httpHeaders:
        - name: Custom-Header
          value: Awesome
      initialDelaySeconds: 3
      periodSeconds: 3
    readinessProbe:
      exec:
        command:
          - cat
          - /tmp/healthy
      initialDelaySeconds: 5
      periodSeconds: 5
    startupProbe:
      httpGet:
        path: /healthz
        port: 80
      failureThreshold: 30
      periodSeconds: 10
```

2.3.5 Configuring Environment Variables

Scenario

An environment variable is a variable whose value can affect the way a running container will behave. You can modify environment variables even after workloads are deployed, increasing flexibility in workload configuration.

The function of setting environment variables on CCE is the same as that of specifying **ENV** in a Dockerfile.

NOTICE

After a container is started, do not modify configurations in the container. If configurations in the container are modified (for example, passwords, certificates, and environment variables of a containerized application are added to the container), the configurations will be lost after the container restarts and container services will become abnormal. An example scenario of container restart is pod rescheduling due to node anomalies.

Configurations must be imported to a container as arguments. Otherwise, configurations will be lost after the container restarts.

Environment variables can be set in the following modes:

- **Custom**: Enter the environment variable name and parameter value.
- **Added from ConfigMap**: Import all key values in a ConfigMap as environment variables.
- **Added from ConfigMap key**: Import the value of a key in a ConfigMap as the value of an environment variable. As shown in **Figure 2-11**, if you import **configmap_value** of **configmap_key** in **configmap-example** as the value of environment variable **key1**, an environment variable named **key1** whose value is **configmap_value** is available in the container.
- **Added from secret**: Import all key values in a secret as environment variables.
- **Added from secret key**: Import the value of a key in a secret as the value of an environment variable. As shown in **Figure 2-11**, if you import **secret_value** of **secret_key** in **secret-example** as the value of environment variable **key2**, an environment variable named **key2** whose value is **secret_value** is available in the container.
- **Variable Value/Reference**: Use the field defined by a pod as the value of the environment variable. As shown in **Figure 2-11**, if the pod name is imported as the value of environment variable **key3**, an environment variable named **key3** whose value is the pod name is available in the container.
- **Resource Reference**: The value of **Request** or **Limit** defined by the container is used as the value of the environment variable. As shown in **Figure 2-11**, if you import the CPU limit of container-1 as the value of environment variable **key4**, an environment variable named **key4** whose value is the CPU limit of container-1 is available in the container.

Adding Environment Variables

Step 1 Log in to the CCE console.

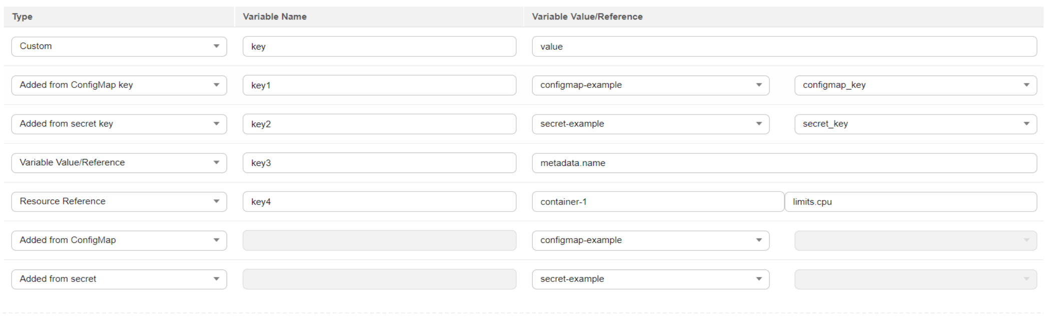
Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane, and click the **Create Workload** in the upper right corner.

Step 3 When creating a workload, modify the container information in **Container Settings** and click the **Environment Variables** tab.

Step 4 Configure environment variables.

- To add environment variables one by one, click **Adding a Variable** and configure its parameters.
- To add environment variables in batches, click **Editing Custom Variables in Batches**. Then, in the displayed dialog box, enter environment variables in the format of "Variable name=Variable or variable reference".

Figure 2-11 Configuring environment variables



Type	Variable Name	Variable Value/Reference	
Custom	key	value	
Added from ConfigMap key	key1	configmap-example	configmap_key
Added from secret key	key2	secret-example	secret_key
Variable Value/Reference	key3	metadata.name	
Resource Reference	key4	container-1	limits.cpu
Added from ConfigMap		configmap-example	
Added from secret		secret-example	

----End

YAML Example

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: env-example
  namespace: default
spec:
  replicas: 1
  selector:
    matchLabels:
      app: env-example
  template:
    metadata:
      labels:
        app: env-example
    spec:
      containers:
        - name: container-1
          image: nginx:alpine
          imagePullPolicy: Always
          resources:
            requests:
              cpu: 250m
              memory: 512Mi
            limits:
              cpu: 250m
              memory: 512Mi
          env:
            - name: key                     # Custom
              value: value
            - name: key1                   # Added from ConfigMap key
              valueFrom:
                configMapKeyRef:
                  name: configmap-example
                  key: configmap_key
            - name: key2                   # Added from secret key
              valueFrom:
                secretKeyRef:
                  name: secret-example
                  key: secret_key
            - name: key3                   # Variable reference, which uses the field defined by a pod as the value 
of the environment variable.
              valueFrom:
                fieldRef:
                  apiVersion: v1
                  fieldPath: metadata.name
            - name: key4                   # Resource reference, which uses the field defined by a container as the 
value of the environment variable.
              valueFrom:
                resourceFieldRef:
                  containerName: container1
                  resource: limits.cpu
                  divisor: 1
          envFrom:
            - configMapRef:                 # Added from ConfigMap
                name: configmap-example
            - secretRef:                   # Added from secret
                name: secret-example
      imagePullSecrets:
        - name: default-secret
```

Viewing Environment Variables

If the contents of **configmap-example** and **secret-example** are as follows:

```
$ kubectl get configmap configmap-example -oyaml
apiVersion: v1
data:
  configmap_key: configmap_value
kind: ConfigMap
...

$ kubectl get secret secret-example -oyaml
apiVersion: v1
data:
  secret_key: c2VjcmV0X3ZhbHVl              # c2VjcmV0X3ZhbHVl is the value of secret_value in Base64 
mode.
kind: Secret
...
```

The environment variables in the pod are as follows:

```
$ kubectl get pod
NAME                              READY   STATUS    RESTARTS   AGE
env-example-695b759569-lx9jp   1/1     Running   0          17m

$ kubectl exec env-example-695b759569-lx9jp  -- printenv
/ # env
key=value                             # Custom environment variable
key1=configmap_value                  # Added from ConfigMap key
key2=secret_value                     # Added from secret key
key3=env-example-695b759569-lx9jp     # metadata.name defined by the pod
key4=1                                # limits.cpu defined by container1. The value is rounded up, in unit of cores.
configmap_key=configmap_value         # Added from ConfigMap. The key value in the original ConfigMap 
key is directly imported.
secret_key=secret_value               # Added from key. The key value in the original secret is directly imported.
```

2.3.6 Configuring APM

Scenario

Application Performance Management (APM) allows you to monitor Java workloads through tracing and topology. You can install APM probes to locate and analyze problems for Java workloads.

You can configure Java workload monitoring when and after a workload is created.

Prerequisites

- You have enabled the APM service. If you have not enabled the APM service, go to the APM console and enable it as prompted.
- You have created a VPC endpoint for APM. If you have not created an endpoint, create one by following the instructions on the APM console.

Precautions

- If a VPC endpoint for APM is deleted, data will fail to be uploaded to APM. If you create this endpoint again, the metrics will not be uploaded to APM automatically. In this case, you can restart the application.
- Deleting a workload that uses APM will not delete the endpoint. To delete it, go to the network console.

Procedure

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 When creating a workload, click **APM Settings** in the **Advanced Settings** area. By default, the probe is **Disabled**. You can select **APM 2.0**. After the probe is enabled, APM can locate and analyze problems for Java programs.

NOTE

1. The APM 2.0 probe will be initialized in an automatically-created init container named **init-javaagent**. This init container will be allocated 0.25 CPU cores and 250 MiB of memory.
2. Adding an APM probe will add the following environment variables to all service containers: **PAAS_MONITORING_GROUP**, **JAVA_TOOL_OPTIONS**, **PAAS_CLUSTER_ID**, and **APM_ACCESS_ADDRESS**.
3. Adding an APM 2.0 probe will mount a local storage volume named **paas-apm2** to all service containers.

Step 4 Set probe-related parameters.

- To access APM in a CCE Autopilot cluster, you need to create a VPC endpoint for APM to connect to the VPC. For details about the VPC endpoint price, see **VPC Endpoint Pricing**.
- **Probe Version**: Select the probe version.

- **Probe Upgrade Policy**: By default, **Auto upgrade upon restart** is selected.
  - **Auto upgrade upon restart**: The system downloads the probe image each time the pod is restarted.
  - **Manual upgrade upon restart**: This policy means that if a local image is available, the local image will be used. The system downloads the probe image only when a local image is unavailable.
- **APM Environment**: (Optional) Enter an environment name.
- **APM Service**: Select an existing APM application.
- **Sub-service**: (Optional) Enter a sub-service of the APM application.
- **Access key**: The system automatically obtains the key information of APM. You can go to the APM console to view the key details.

For details about the APM 2.0 parameters, see **APM parameters**.

Step 5 Three minutes after the application is started, its data will be displayed on the APM console. You can log in to the APM console and optimize application performance through topology and tracing. For details, see **Topology**.

----End

Configuring APM Settings

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane, and click the desired workload name.

Step 3 On the displayed page, click the **APM Settings** tab and click **Edit** in the lower right corner.

For details about the parameters, see **Step 4**.

----End

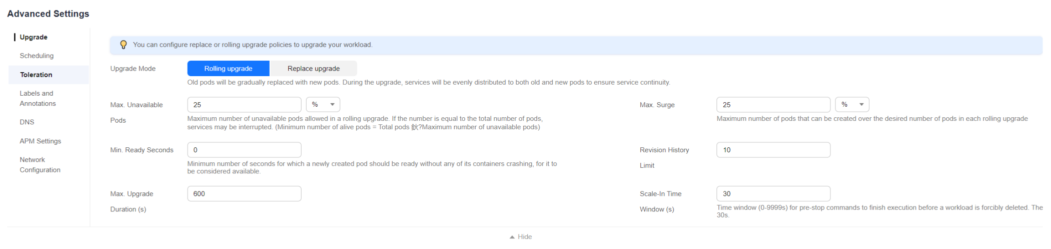
2.3.7 Configuring the Workload Upgrade Policy

Upgrades are a common operation. A Deployment, StatefulSet, or DaemonSet can easily support application upgrades.

You can set different upgrade policies:

- **Rolling upgrade**: New pods are created gradually, and then old pods are deleted. This is the default policy.
- **Replace upgrade**: The pods are deleted, and then new pods are created.

Figure 2-12 Configuring a workload upgrade policy



You can configure replace or rolling upgrade policies to upgrade your workload.

Upgrade Mode: Rolling upgrade
Old pods will be gradually replaced with new pods. During the upgrade, services will be evenly distributed to both old and new pods to ensure service continuity.

Max. Unavailable Pods: 25 %
Maximum number of unavailable pods allowed in a rolling upgrade. If the number is equal to the total number of pods, services may be interrupted. (Minimum number of alive pods = Total pods 鈥?Maximum number of unavailable pods)

Max. Surge: 25 %
Maximum number of pods that can be created over the desired number of pods in each rolling upgrade

Min. Ready Seconds: 0
Minimum number of seconds for which a newly created pod should be ready without any of its containers crashing, for it to be considered available.

Revision History Limit: 10

Max. Upgrade Duration (s): 600

Scale-In Time Window (s): 30
Time window (0-9999s) for pre-stop commands to finish execution before a workload is forcibly deleted. The 30s.

Upgrade Parameters

Parameter	Description	Constraint
Max. Surge (maxSurge)	Specifies the percentage of pods that are allowed based on the value of **spec.replicas**. The default value is **25%**. For example, if **spec.replicas** is set to **4**, there should be a maximum of five pods during the upgrade, and one pod can be added each time (at a step of 1). During the upgrade, the value is converted into a number and rounded up. The value can also be set to an absolute number.	This parameter is only supported by Deployments and DaemonSets.
Max. Unavailable Pods (maxUnavailable)	Specifies the percentage of pods that can be deleted based on the value of **spec.replicas**. The default value is **25%** For example, if **spec.replicas** is set to **4**, there should be at least three pods during the upgrade, and one pod can be deleted (at a step of 1). The value can also be set to an absolute number.	This parameter is only supported by Deployments and DaemonSets.
Min. Ready Seconds (minReadySeconds)	A pod is considered available only when the minimum readiness time is exceeded without any of its containers crashing. The default value is **0** (the pod is considered available immediately after it is ready).	-
Revision History Limit (revisionHistory-Limit)	Specifies the number of old ReplicaSets to retain to allow rollback. These old ReplicaSets consume resources in etcd and crowd the output of **kubectl get rs**. The configuration of each Deployment revision is stored in its ReplicaSets. Therefore, once the old ReplicaSet is deleted, you lose the ability to roll back to that revision of Deployment. By default, 10 old ReplicaSets will be kept, but the ideal value depends on the frequency and stability of the new Deployments.	-

Parameter	Description	Constraint
Max. Upgrade Duration (progressDeadlineSeconds)	Specifies the number of seconds that the system waits for a Deployment to make progress before reporting a Deployment progress failure. It is surfaced as a condition with Type=Progressing, Status=False, and Reason=ProgressDeadlineExceeded in the status of the resource. The Deployment controller will keep retrying the Deployment. In the future, once automatic rollback will be implemented, the Deployment controller will roll back a Deployment as soon as it observes such a condition. If this parameter is specified, the value of this parameter must be greater than that of **.spec.minReadySeconds**.	-
Scale-In Time Window (terminationGracePeriodSeconds)	Graceful deletion time. The default value is 30 seconds. When a pod is deleted, a SIGTERM signal is sent and the system waits for the applications in the container to terminate. If the application is not terminated within the time specified by **terminationGracePeriodSeconds**, a SIGKILL signal is sent to forcibly terminate the pod.	-

Upgrade Example

The Deployment can be upgraded in a declarative mode. That is, you only need to modify the YAML definition of the Deployment. For example, you can run the **kubectl edit** command to change the Deployment image to **nginx:alpine**. After the modification, query the ReplicaSet and pod. The query result shows that a new ReplicaSet is created and the pod is re-created.

```
$ kubectl edit deploy nginx

$ kubectl get rs
NAME               DESIRED   CURRENT   READY     AGE
nginx-6f9f58dffd   2         2         2         1m
nginx-7f98958cdf   0         0         0         48m

$ kubectl get pods
NAME                     READY     STATUS    RESTARTS   AGE
nginx-6f9f58dffd-tdmqk   1/1       Running   0          1m
nginx-6f9f58dffd-tesqr   1/1       Running   0          1m
```

The Deployment can use the **maxSurge** and **maxUnavailable** parameters to control the proportion of pods to be re-created during the upgrade, which is useful in many scenarios. The configuration is as follows:

```
spec:
  strategy:
    rollingUpdate:
      maxSurge: 1
      maxUnavailable: 0
    type: RollingUpdate
```

In the preceding example, the value of **spec.replicas** is **2**. If both **maxSurge** and **maxUnavailable** are the default value 25%, **maxSurge** allows a maximum of three pods to exist (2 x 1.25 = 2.5, rounded up to 3), and **maxUnavailable** does not allow a maximum of two pods to be unavailable (2 x 0.75 = 1.5, rounded up to 2). That is, during the upgrade process, there will always be two pods running. Each time a new pod is created, an old pod is deleted, until all pods are new.

Rollback

Rollback is to roll an application back to the earlier version when a fault occurs during the upgrade. A Deployment can be easily rolled back to the earlier version.

For example, if the upgraded image is faulty, you can run the **kubectl rollout undo** command to roll back the Deployment.

```
$ kubectl rollout undo deployment nginx
deployment.apps/nginx rolled back
```

A Deployment can be easily rolled back because it uses a ReplicaSet to control a pod. After the upgrade, the previous ReplicaSet still exists. The Deployment is rolled back by using the previous ReplicaSet to re-create the pod. The number of ReplicaSets stored in a Deployment can be restricted by the **revisionHistoryLimit** parameter. The default value is **10**.

2.3.8 Configuring Labels and Annotations

Pod Annotations

CCE Autopilot clusters provide multiple advanced functions during pod creation. To enable these functions, you can add annotations to pods on the console or using YAML. **Table 2-15** describes the available annotations.

Table 2-15 Pod annotations

Function and Reference	Annotation	Example Value	Description
Setting AZ Affinity	node.cce.io/node-az-list	cn-east-3a,cn-east-3b	Lists the AZs for pod affinity. For a CCE Autopilot cluster, you can set workload annotations to implement AZ affinity and schedule pods to specified AZs.

Function and Reference	Annotation	Example Value	Description
Increasing the Ephemeral Storage Space of a Pod	resource.cce.io/extra-ephemeral-storage-in-GiB	14	Indicates the extra ephemeral storage capacity for a pod, in GiB. The value ranges from **0** to **994**. Each pod can have 30-GiB free ephemeral storage and specified extra storage. **NOTE** Only clusters v1.28.6-r0, v1.27.8-r0, and later support this function. Clusters of other versions need to be upgraded to support this function.

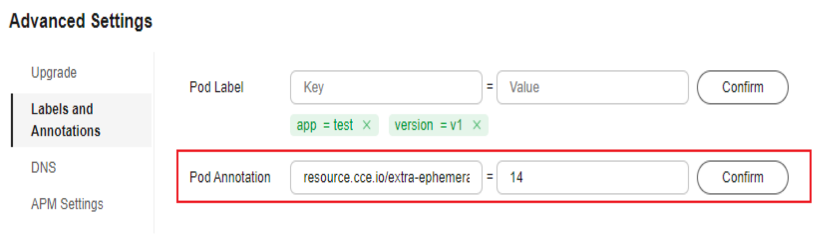
You can add an annotation on the console or using YAML.

Adding an Annotation on the Console

When creating a workload on the console, you can choose **Advanced Settings** > **Labels and Annotations** and add pod annotations to enable advanced functions.

For example, to add 14 GiB of ephemeral storage capacity for a pod, set **Pod Annotation** to **resource.cce.io/extra-ephemeral-storage-in-GiB=14** and click **Confirm**.

Figure 2-13 Adding ephemeral storage capacity for a pod



Adding an Annotation Using YAML

When creating a workload using YAML, you can use the **annotations** parameter to enable advanced pod functions.

For example, when creating an Nginx workload, you can use **annotations** to add 14 GiB of ephemeral storage capacity for a pod.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx # Workload name
spec:
  replicas: 1 # Number of pods
  selector:
    matchLabels: # Selector for selecting resources with specific labels.
      app: nginx
  template:
```

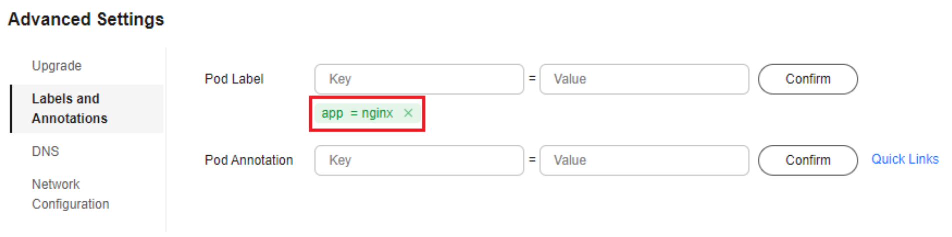
```
  metadata:
    labels:     # Labels
      app: nginx
    annotations:
      resource.cce.io/extra-ephemeral-storage-in-GiB: '14'
  spec:
    containers:
    - image: nginx:latest   # {Image name}:{Image tag}
      name: nginx
    imagePullSecrets:
    - name: default-secret
```

Pod Labels

You can use pod labels to organize, select, and manage resources for pods to make it easy to use and maintain resources.

When you create a workload on the console, the following labels are added to the pod by default. The value of **app** is the workload name.

Figure 2-14 Pod Label



Example YAML:

```
...
spec:
  selector:
    matchLabels:
      app: nginx
      version: v1
  template:
    metadata:
      labels:
        app: nginx
        version: v1
    spec:
      ...
```

2.3.9 Setting AZ Affinity

Scenario

An availability zone (AZ) is a physical region in a data center where resources use independent power supplies and networks. Nodes in the same AZ can quickly communicate with each other through a high-speed network, while nodes in different AZs need to communicate with each other across physical distances, which may cause latency and risks.

Scheduling pods to different AZs improves the availability and fault tolerance of applications.

Procedure

For a CCE Autopilot cluster, you can set workload annotations to implement AZ affinity and schedule pods to specified AZs.

Step 1 Log in to the CCE console.

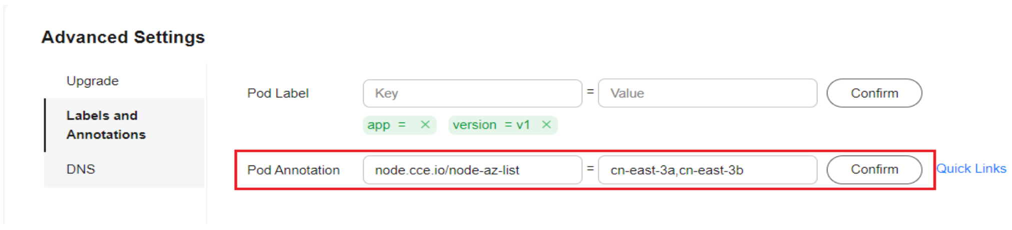
Step 2 Click the cluster name to go to the cluster console, choose **Workloads** in the navigation pane on the left, and click the **Create Workload** in the upper right corner.

Step 3 In the **Advanced Settings** area, select **Labels and Annotations** and set the following annotation:

- **Key**: **node.cce.io/node-az-list**
- **Value**: AZ name. Use commas (,) to separate multiple AZs.

  For details about AZ names in different regions, see **Regions and Endpoints**.

Figure 2-15 Setting AZ affinity



Pod Annotation: node.cce.io/node-az-list = cn-east-3a,cn-east-3b

Step 4 Configure other workload parameters and click **Create Workload**.

----End

2.4 Logging In to a Container

Scenario

If you encounter unexpected problems when using a container, you can log in to the container to debug it.

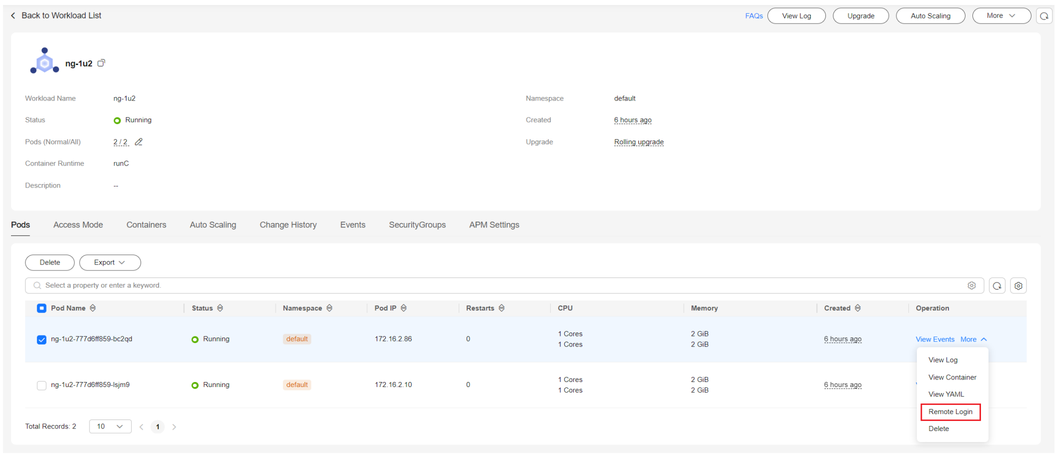
Logging In to a Container Using CloudShell

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. Click the name of the target workload to view its pods.

Step 3 Locate the target pod and choose **More** > **Remote Login** in the **Operation** column.

Figure 2-16 Accessing a container

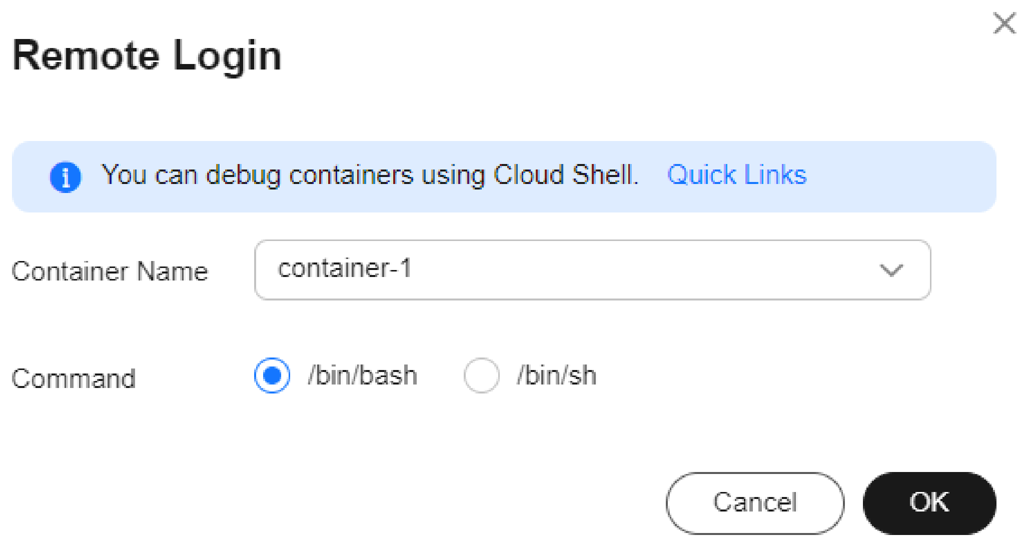


Workload Name	ng-1u2
Status	Running
Pods (Normal/All)	2 / 2
Container Runtime	runC
Description	--
Namespace	default
Created	6 hours ago
Upgrade	Rolling upgrade

Pod Name	Status	Namespace	Pod IP	Restarts	CPU	Memory	Created	Operation
ng-1u2-777d6ff859-bc2qd	Running	default	172.16.2.86	0	1 Cores 1 Cores	2 GiB 2 GiB	6 hours ago	View Events More
ng-1u2-777d6ff859-lsjm9	Running	default	172.16.2.10	0	1 Cores 1 Cores	2 GiB 2 GiB	6 hours ago	

Step 4 In the displayed dialog box, select the container you want to access and the command, and click **OK**.

Figure 2-17 Selecting a container and login command



You can debug containers using Cloud Shell.

Container Name: container-1

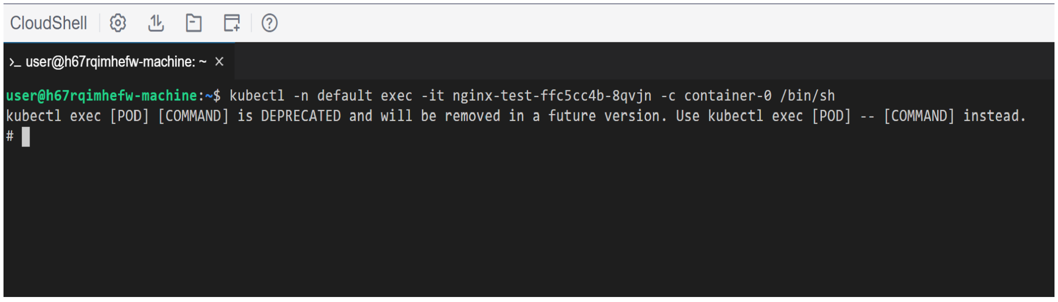
Command: [✓ /bin/bash] /bin/sh

Step 5 You will be automatically redirected to CloudShell. Then, kubectl is initialized and runs the **kubectl exec** command to log in to the container.

NOTE

Wait for 5 to 10 seconds until the **kubectl exec** command is automatically executed.

Figure 2-18 CloudShell page



```
user@h67rqimhefw-machine:~$ kubectl -n default exec -it nginx-test-ffc5cc4b-8qvjn -c container-0 /bin/sh
kubectl exec [POD] [COMMAND] is DEPRECATED and will be removed in a future version. Use kubectl exec [POD] -- [COMMAND] instead.
#
```

----End

Logging In to a Container Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Run the following command to view the created pod:

```
kubectl get pod
```

The example output is as follows:

```
NAME                     READY   STATUS    RESTARTS     AGE
nginx-59d89cb66f-mhljr          1/1     Running   0          11m
```

Step 3 Query the container name in the pod.

```
kubectl get po nginx-59d89cb66f-mhljr -o jsonpath='{range .spec.containers[*]}{.name}{end}{"\n"}'
```

The example output is as follows:

```
container-1
```

Step 4 Run the following command to log in to the **container-1** container in the **nginx-59d89cb66f-mhljr** pod:

```
kubectl exec -it nginx-59d89cb66f-mhljr -c container-1 -- /bin/sh
```

Step 5 To exit the container, run the **exit** command.

----End

2.5 Managing Workloads and Jobs

Scenario

After a workload is created, you can upgrade, monitor, roll back, or delete the workload, as well as edit its YAML file.

Table 2-16 Workload/Job management

Operation	Description
View Log	You can view the logs of workloads.
Upgrade	You can replace images or image tags to quickly upgrade Deployments and StatefulSets without interrupting services.

Operation	Description
Edit YAML	You can modify and download the YAML files of Deployments, StatefulSets, CronJobs, and pods on the CCE console. YAML files of Jobs can only be viewed, copied, and downloaded. **NOTE** If an existing CronJob is modified, the new configuration will only be applied for the new pods, and existing pods continue to run without any change.
Roll Back	Only Deployments can be rolled back.
Redeploy	You can redeploy a workload. After the workload is redeployed, all pods in the workload will be restarted.
Enable/Disable Upgrade	Only Deployments support this operation.
Manage Label	Labels are attached to workloads as key-value pairs to manage and select workloads. Jobs and CronJobs do not support this operation.
Delete	You can delete a workload or job that is no longer needed. Deleted workloads or jobs cannot be recovered.
View Events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time.
Stop/Start	You can only start or stop a CronJob.

Viewing Logs

You can view logs of Deployments, StatefulSets, and jobs. This section uses a Deployment as an example to describe how to view logs.

NOTICE

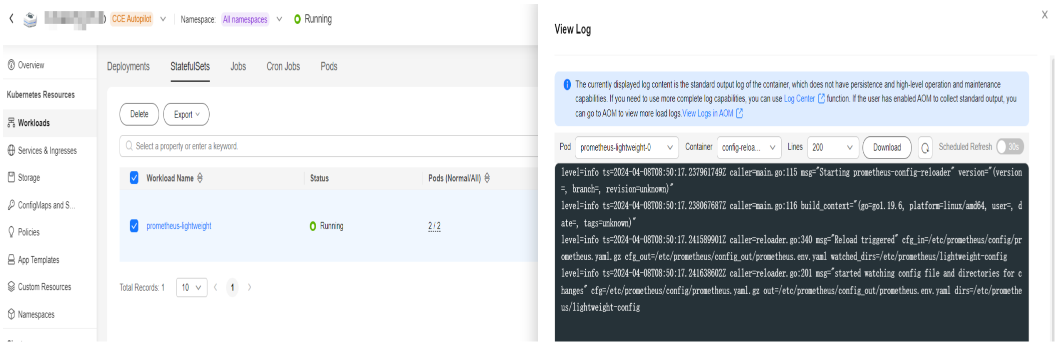
Before viewing logs, ensure that the time of the browser is the same as that on the backend server.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab and click the **View Log** of the target workload.

On the displayed **View Log** window, you can view logs.

Figure 2-19 Viewing logs of a workload



The currently displayed log content is the standard output log of the container, which does not have persistence and high-level operation and maintenance capabilities. If you need to use more complete log capabilities, you can use Log Center function. If the user has enabled AOM to collect standard output, you can go to AOM to view more load logs. View Logs in AOM

```
level=info ts=2024-04-08T08:50:17.237961749Z caller=main.go:115 msg="Starting prometheus-config-reloader" version="(version=, branch=, revision=unknown)"
level=info ts=2024-04-08T08:50:17.238067687Z caller=main.go:116 build_context="(go=go1.19.6, platform=linux/amd64, user=, date=, tags=unknown)"
level=info ts=2024-04-08T08:50:17.241589901Z caller=reloader.go:340 msg="Reload triggered" cfg_in=/etc/prometheus/config/prometheus.yaml.gz cfg_out=/etc/prometheus/config_out/prometheus.env.yaml watched_dirs=/etc/prometheus/lightweight-config
level=info ts=2024-04-08T08:50:17.241638602Z caller=reloader.go:201 msg="started watching config file and directories for changes" cfg=/etc/prometheus/config/prometheus.yaml.gz out=/etc/prometheus/config_out/prometheus.env.yaml dirs=/etc/prometheus/lightweight-config
```

----End

Upgrading a Workload

You quickly upgrade Deployments and StatefulSets on the console.

This section uses a Deployment as an example to describe how to upgrade a workload.

Before replacing an image or image version, upload the new image to the SWR service. For details, see **Uploading an Image Through a Container Engine Client**.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab. Locate the target workload and click **Upgrade** in the **Operation** column.

NOTE

- Workloads cannot be upgraded in batches.
- Before performing an in-place StatefulSet upgrade, you must manually delete old pods. Otherwise, the upgrade status is always displayed as **Processing**.

Step 3 Upgrade the workload based on service requirements. The method for setting parameter is the same as that for creating a workload.

Step 4 After the update is complete, click **Upgrade Workload**, manually confirm the YAML file, and submit the upgrade.

----End

Editing a YAML file

You can modify and download the YAML files of Deployments, StatefulSets, CronJobs, and pods on the CCE console. YAML files of Jobs can only be viewed, copied, and downloaded. This section uses a Deployment as an example to describe how to edit the YAML file.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab. Locate the target workload and choose **More** > **Edit YAML** in the **Operation** column. In the dialog box that is displayed, modify the YAML file.

Step 3 Click **OK**.

Step 4 (Optional) In the **Edit YAML** window, click **Download** to download the YAML file.

----End

Rolling Back a Workload (Available Only for Deployments)

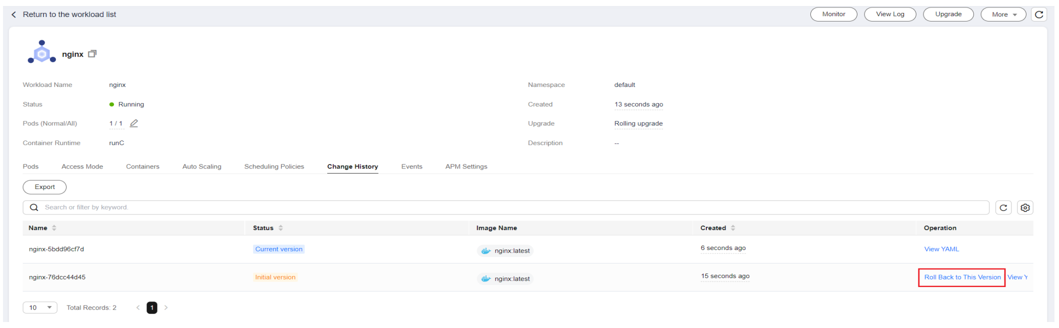
CCE records the release history of all Deployments. You can roll back a Deployment to a specified version.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab. Locate the target workload and click **More > Roll Back** in the **Operation** column.

Step 3 Switch to the **Change History** tab. Locate the target workload and click **Roll Back to This Version**. Manually confirm the YAML file and click **OK**.

Figure 2-20 Rolling back a workload version



Workload Name	nginx	Namespace	default
Status	Running	Created	13 seconds ago
Pods (Normal/All)	1 / 1	Upgrade	Rolling upgrade
Container Runtime	runC	Description	--

Name	Status	Image Name	Created	Operation
nginx-5bdd96cf7d	Current version	nginx:latest	6 seconds ago	View YAML
nginx-76dcc44d45	Initial version	nginx:latest	15 seconds ago	Roll Back to This Version View Y

----End

Redeploying a Workload

After you redeploy a workload, all pods in the workload will be restarted. This section uses Deployments as an example to illustrate how to redeploy a workload.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab. Locate the target workload and click **More** > **Redeploy** in the **Operation** column.

Step 3 In the dialog box that is displayed, click **Yes**.

----End

Disabling/Enabling Upgrade (Available Only for Deployments)

Only Deployments support this operation.

- After the upgrade is disabled, the upgrade command can be delivered but will not be applied to the pods.

If you are performing a rolling upgrade, the rolling upgrade stops after the disabling upgrade command is delivered. In this case, the new and old pods co-exist.

- If a Deployment is being upgraded, it can be upgraded or rolled back. Its pods will inherit the latest updates of the Deployment. If they are inconsistent, the pods are upgraded automatically according to the latest information of the Deployment.

NOTICE

The workload cannot be rolled back when the upgrade is disabled.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab. Locate the target workload and choose **More** > **Disable/Enable Upgrade** in the **Operation** column.

Step 3 In the dialog box that is displayed, click **Yes**.

----End

Managing Labels

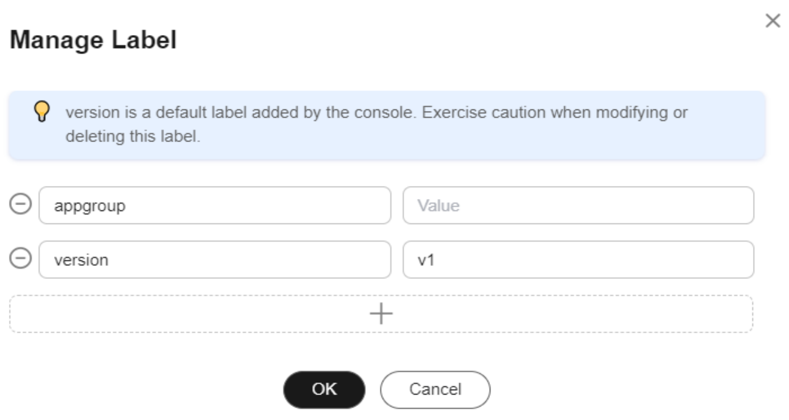
Labels are key-value pairs and can be attached to workloads. You can manage and select workloads by labels. You can add labels to multiple workloads or a specified workload.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Click the **Deployments** tab and choose **More** > **Manage Label** in the **Operation** column of the workload.

Step 3 Click **Add**, enter a key and a value, and click **OK**.

Figure 2-21 Managing labels



Manage Label

version is a default label added by the console. Exercise caution when modifying or deleting this label.

appgroup	Value
version	v1

NOTE

A key-value pair must contain 1 to 63 characters starting and ending with a letter or digit. Only letters, digits, hyphens (-), underscores (_), and periods (.) are allowed.

----End

Deleting a Workload/Job

You can delete a workload or job that is no longer needed. Deleted workloads or jobs cannot be recovered. This section uses a Deployment as an example to describe how to delete a workload.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 Locate the workload you will delete and choose **More** > **Delete** in the **Operation** column.

Read the prompts carefully. A workload cannot be recovered after it is deleted.

Step 3 Click **Yes**.

NOTE

- If the node where the pod is located is unavailable or shut down and the workload cannot be deleted, you can forcibly delete the pod from the pod list on the workload details page.
- Ensure that the storage volumes to be deleted are not used by other workloads. If these volumes are imported or have snapshots, you can only unbind them.

----End

Events

This section uses Deployments as an example to illustrate how to view events of a workload. To view the event of a job or CronJob, click **View Event** in the **Operation** column of the target workload.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Workloads**.

Step 2 On the **Deployments** tab, click the target workload. In the **Pods** tab, click the **View Events** to view the event name, event type, number of occurrences, Kubernetes event, first occurrence time, and last occurrence time.

NOTE

Event data will be retained for one hour and then automatically deleted.

----End

2.6 Managing kernel Options

CCE Autopilot is a serverless cluster and isolated from the kernel of physical machines. kernel tuning is a common practice in advanced service deployment scenarios. In a safe situation, CCE Autopilot allows you to configure kernel parameters through a security context of a pod based on the solution recommended by the Kubernetes community, greatly improving the flexibility of service deployment. For details of security contexts, see **Configure a Security Context for a Pod or Container**.

In Linux, sysctl is the most common method of modifying kernel parameters. In Kubernetes, kernel parameters are configured through the sysctl security context of the pod. If you are not familiar with the sysctl concept, see **Using sysctls in a Kubernetes Cluster**. A security context applies to all containers in the same pod.

CCE Autopilot allows you to modify the following non-secure sysctl parameters:

```
kernel.shm*,
kernel.msg*,
kernel.sem,
fs.mqueue.*,
net.*
```

NOTICE

To avoid affecting the stability of the OS, modify the sysctl parameters after understanding the consequences of the modification.

The sysctl parameters with a namespace may change in future Linux kernel versions.

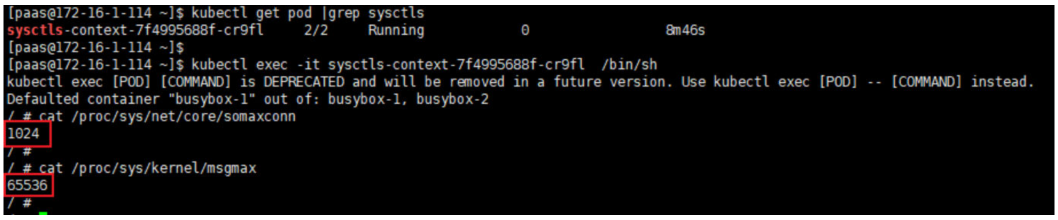
Non-secure sysctl parameters are unstable. Using non-secure sysctl parameters may cause some serious problems, such as container errors. You need to take care of the risks.

In the following example, the pod's security context is used to set two sysctl parameters: **kernel.msgmax** and **net.core.somaxconn**.

```
apiVersion: v1
kind: Pod
metadata:
  name: sysctls-context-example
spec:
  securityContext:
    sysctls:
    - name: kernel.msgmax
      value: "65536"
    - name: net.core.somaxconn
      value: "1024"
  ...
```

Go to the container to check whether the configuration takes effect.

```
kubectl exec -it podname -c container-1 -- /bin/sh
```



```
[paas@172-16-1-114 ~]$ kubectl get pod |grep sysctls
sysctls-context-7f4995688f-cr9fl     2/2     Running            0                 8m46s
[paas@172-16-1-114 ~]$
[paas@172-16-1-114 ~]$ kubectl exec -it sysctls-context-7f4995688f-cr9fl  /bin/sh
kubectl exec [POD] [COMMAND] is DEPRECATED and will be removed in a future version. Use kubectl exec [POD] -- [COMMAND] instead.
Defaulted container "busybox-1" out of: busybox-1, busybox-2
/ # cat /proc/sys/net/core/somaxconn
1024
/ #
/ # cat /proc/sys/kernel/msgmax
65536
/ #
```

2.7 Managing Custom Resources

Custom Resource Definition (CRD) is an extension of Kubernetes APIs. When default Kubernetes resources cannot meet service requirements, you can use CRDs to define new resource types. According to CRD, you can create custom resources in a cluster to meet service requirements. CRD allows you to create new resource types without adding new Kubernetes API servers. This makes cluster management more flexible.

Creating a CRD

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console, choose **Custom Resources** in the navigation pane, and click the **Create from YAML** in the upper right corner.

Step 3 Customize the YAML file to create a CRD based on service requirements. For details, see **Extend the Kubernetes API with CustomResourceDefinitions**.

Step 4 Click **OK**.

----End

Viewing CRDs and Their Resources

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to access the cluster console. Choose **Custom Resources** in the navigation pane.

Step 3 On the **Custom Resources** page, view CRDs and their resources.

- View a CRD and its YAML.

  All CRDs in the cluster as well as their API groups, API versions, and resource application scopes are listed. Click **View YAML** in the **Operation** column of a CRD to view its YAML.

  You can enter a keyword in the search box to search for target resource types.

- View the resources of a CRD.

  Locate a CDR in the list and click **View Details** in the **Operation** column to view the resources.

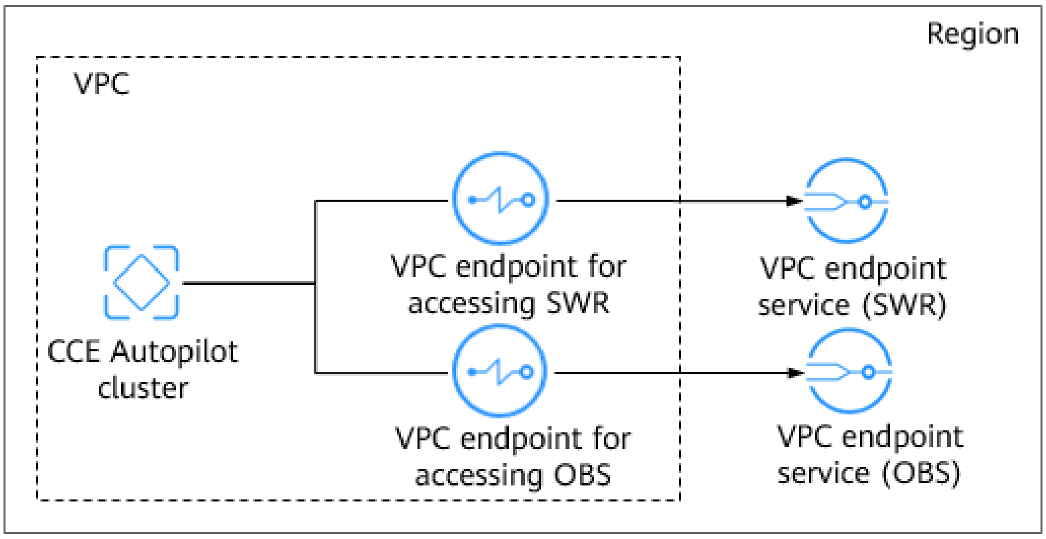
----End

2.8 Configuring VPC Endpoints for Accessing SWR and OBS

When deploying a workload in a CCE Autopilot cluster, you need to use the VPC endpoints of SWR and OBS to pull images.

For details about VPC Endpoint, see **What Is VPC Endpoint?**

Figure 2-22 A CCE Autopilot cluster accessing SWR and OBS


Region
VPC
CCE Autopilot cluster
VPC endpoint for accessing SWR
VPC endpoint for accessing OBS
VPC endpoint service (SWR)
VPC endpoint service (OBS)

Configuring the VPC Endpoint for Accessing SWR

Step 1 Log in to the **VPC Endpoint** console.

Step 2 On the displayed page, click **Buy VPC Endpoint**.

Step 3 Configure the parameters.

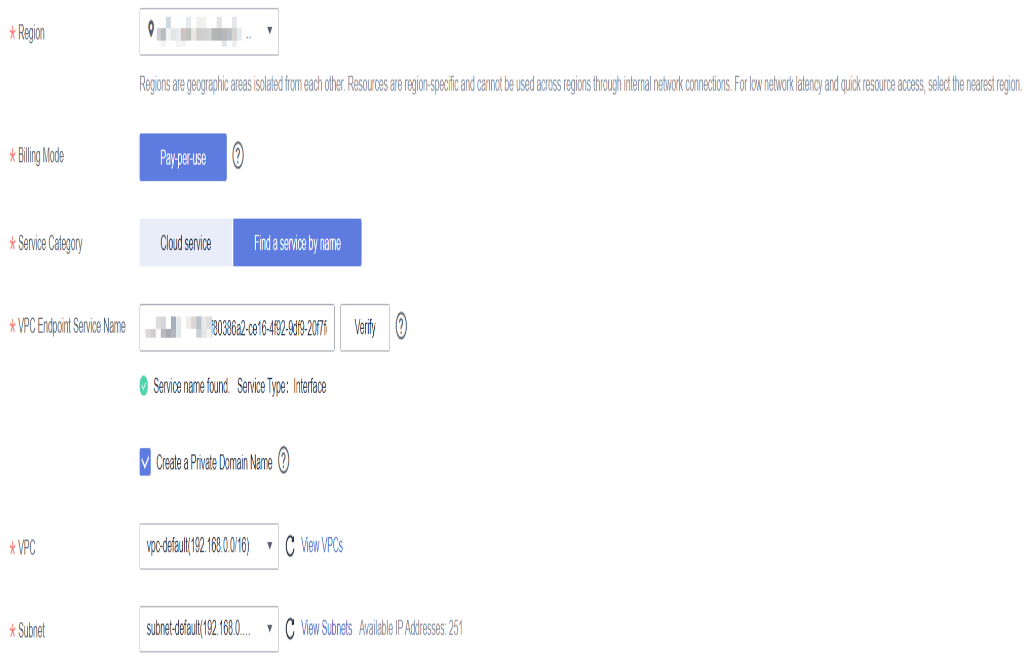
Table 2-17 Parameters for creating a VPC endpoint

Parameter	Description
Region	Select the region where the VPC endpoint is located. The VPC endpoint must be in the same region as the cluster.
Billing Mode	Select **Pay-per-use**.
Service Category	Select **Find a service by name**.
VPC Endpoint Service Name	Enter a service name based on **Table 2-18** and click **Verify**.
VPC	Select the VPC where the cluster is located.
Subnet	Select a subnet.
IPv4 Address	By default, **Automatically assign IPv4 address** is selected. You can also select **Manually specify an IP address** as required.

Table 2-18 VPC endpoint service names for SWR

Region	Name
CN South-Guangzhou-InvitationOnly	cn-south-4.SWR.f80386a2-ce16-4f92-9df9-20f7fc01e7a2
CN Southwest-Guiyang1	com.myhuaweicloud.cn-southwest-2.swr
CN South-Guangzhou	swr.cn-south-1.myhuaweicloud.com
CN East-Shanghai1	com.myhuaweicloud.cn-east-3.swr
CN North-Beijing4	com.myhuaweicloud.cn-north-4.swr
AP-Bangkok	ap-southeast-2.SWR.ac7067e1-f8d1-4f5c-abe1-0f78960e5d4c
AP-Singapore	com.myhuaweicloud.ap-southeast-3.swr

Figure 2-23 Creating a VPC endpoint for SWR



* Region: [redacted]

Regions are geographic areas isolated from each other. Resources are region-specific and cannot be used across regions through internal network connections. For low network latency and quick resource access, select the nearest region.

* Billing Mode: Pay-per-use

* Service Category: Cloud service | Find a service by name

* VPC Endpoint Service Name: [redacted]f80386a2-ce16-4f92-9df9-20f7f Verify

Service name found. Service Type: Interface

✓ Create a Private Domain Name

* VPC: vpc-default(192.168.0.0/16) View VPCs

* Subnet: subnet-default(192.168.0.... View Subnets Available IP Addresses: 251

Step 4 Click **Next** to confirm the configuration.

- If the configuration is correct, click **Submit**.
- If any parameter is incorrect, click **Previous** to modify it as needed and then click **Submit**.

Step 5 Go back to the VPC endpoint list.

If the status of the VPC endpoint changes to **Accepted**, the VPC endpoint is connected to the VPC endpoint service.

----End

Configuring the VPC Endpoint for Accessing OBS

Step 1 Log in to the **VPC Endpoint** console.

Step 2 On the displayed page, click **Buy VPC Endpoint**.

Step 3 Configure the parameters.

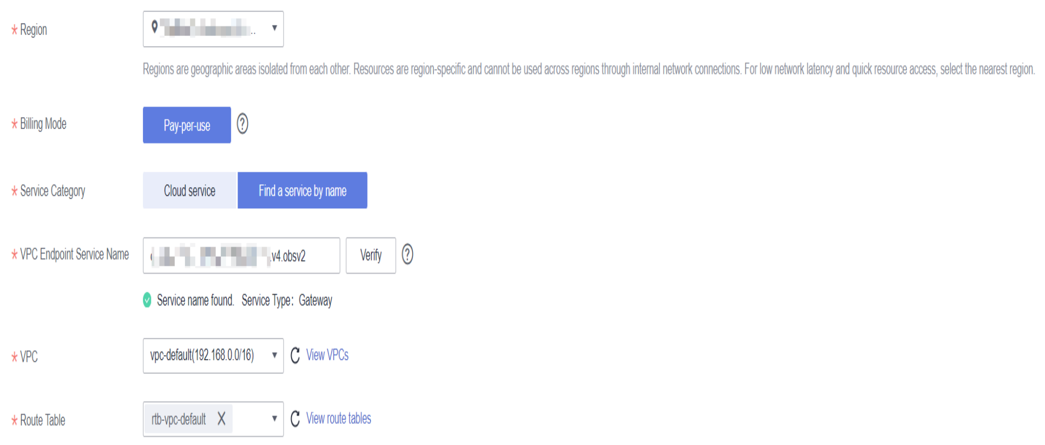
Table 2-19 Parameters for creating a VPC endpoint

Parameter	Description
Region	Select the region where the VPC endpoint is located. The VPC endpoint must be in the same region as the cluster.
Billing Mode	Select **Pay-per-use**.
Service Category	Select **Find a service by name**.
VPC Endpoint Service Name	Enter a service name based on **Table 2-20** and click **Verify**.
VPC	Select the VPC where the cluster is located.
Route Table	Select a route table.

Table 2-20 VPC endpoint service names for OBS

Region	Name
CN South-Guangzhou-InvitationOnly	cn-south-4.com.myhuaweicloud.v4.obsv2
CN Southwest-Guiyang1	cn-southwest-2.com.myhuaweicloud.v4.obsv2
CN South-Guangzhou	cn-south-1.com.myhuaweicloud.v4.obsv2
CN East-Shanghai1	cn-east-3.com.myhuaweicloud.v4.global.obsv2
CN North-Beijing4	cn-north-4.com.myhuaweicloud.v4.obsv2
AP-Bangkok	ap-southeast-2.myhuaweicloud.v4.obsv2
AP-Singapore	ap-southeast-3.com.myhuaweicloud.v4.obsv2

Figure 2-24 Creating a VPC endpoint for OBS



* Region: [redacted]

Regions are geographic areas isolated from each other. Resources are region-specific and cannot be used across regions through internal network connections. For low network latency and quick resource access, select the nearest region.

* Billing Mode: Pay-per-use

* Service Category: Cloud service | Find a service by name

* VPC Endpoint Service Name: [redacted]v4.obsv2 Verify

Service name found. Service Type: Gateway

* VPC: vpc-default(192.168.0.0/16) View VPCs

* Route Table: rtb-vpc-default View route tables

Step 4 Click **Next** to confirm the configuration.

- If the configuration is correct, click **Submit**.
- If any parameter is incorrect, click **Previous** to modify it as needed and then click **Submit**.

Step 5 Go back to the VPC endpoint list.

If the status of the VPC endpoint changes to **Accepted**, the VPC endpoint is connected to the VPC endpoint service.

----End

3 Network

3.1 Network Overview

You can learn about a cluster network from the following:

- Cluster network types and their functions. For details, see **Cluster Network Structure**.
- Access to a pod or container n a cluster. Kubernetes provides **Service** and **Ingress** to enable pod access. This section summarizes common network access scenarios. You can select the proper scenario based on site requirements. For details about the network access scenarios, see **Network Access Scenarios**.

Cluster Network Structure

A cluster network can be divided into two parts:

- **Container Network**

  A container network assigns IP addresses to pods in a cluster. CCE Autopilot inherits the IP-Per-Pod-Per-Network network model of Kubernetes. That means each pod has an independent IP address on a network plane and all containers in a pod share the same network namespace. All pods in a cluster are running in a directly connected flat network. They can access each other through their IP addresses without using NAT. Kubernetes only provides a network mechanism for pods, but does not directly configure pod networks. You can use specific container network add-ons to configure networks for pods and manages container IP addresses.

  CCE Autopilot supports only the Cloud Native 2.0 model. Cloud Native 2.0, built on CCE, is a next-generation container network. It integrates network interfaces and supplementary network interfaces of Virtual Private Cloud (VPC).IP addresses can be assigned to containers from a VPC CIDR block. Load balancers, security groups, and EIPs can be used in a container network to deliver high performance.

- **Service Network**

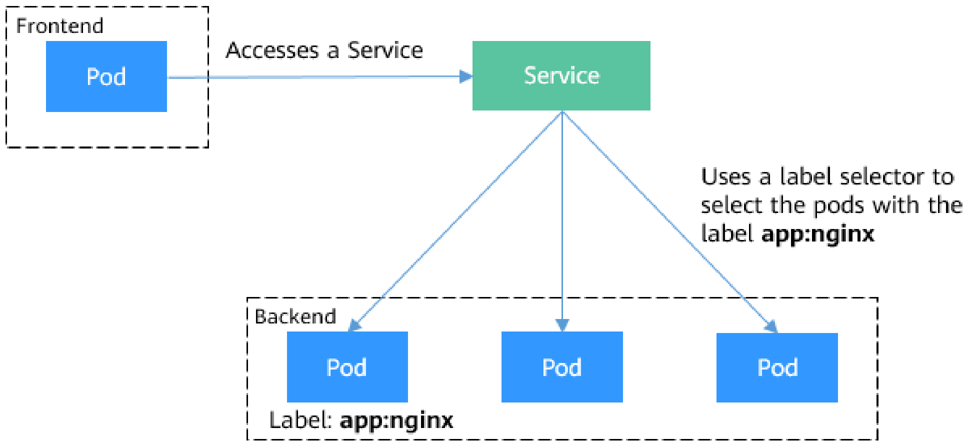
  Service is a Kubernetes concept. Each Service has an IP address. When creating a cluster on CCE, you can specify a CIDR block for ClusterIP Service.

The Service CIDR block cannot overlap with the container subnet CIDR block. IP addresses in the Service CIDR block can only be assigned to the containers in the cluster.

Service

Each Service has a fixed IP address and routes traffic across the pods.

Figure 3-1 Accessing pods through a Service



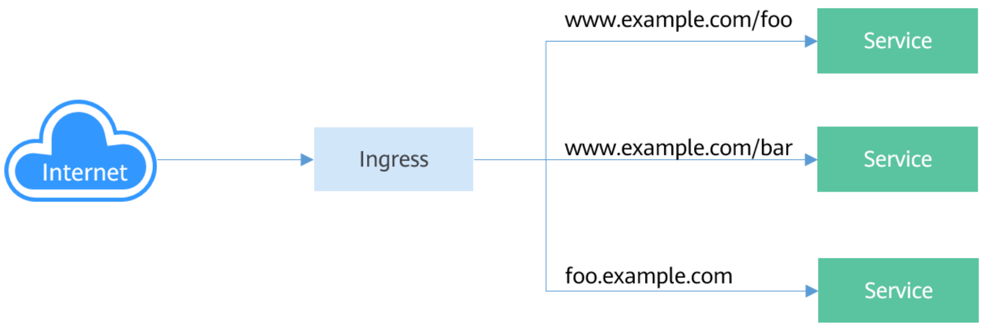
You can configure the following types of Services:

- ClusterIP: used to make the Service only reachable from within a cluster.
- LoadBalancer: used for access from outside a cluster. A load balancer distributes external requests across the supplementary network interfaces of nodes.

Ingress

Services forward requests using layer-4 TCP and UDP protocols. Ingresses forward requests using layer-7 HTTP and HTTPS protocols. Domain names and paths can be used to achieve finer granularities.

Figure 3-2 Ingress and Service

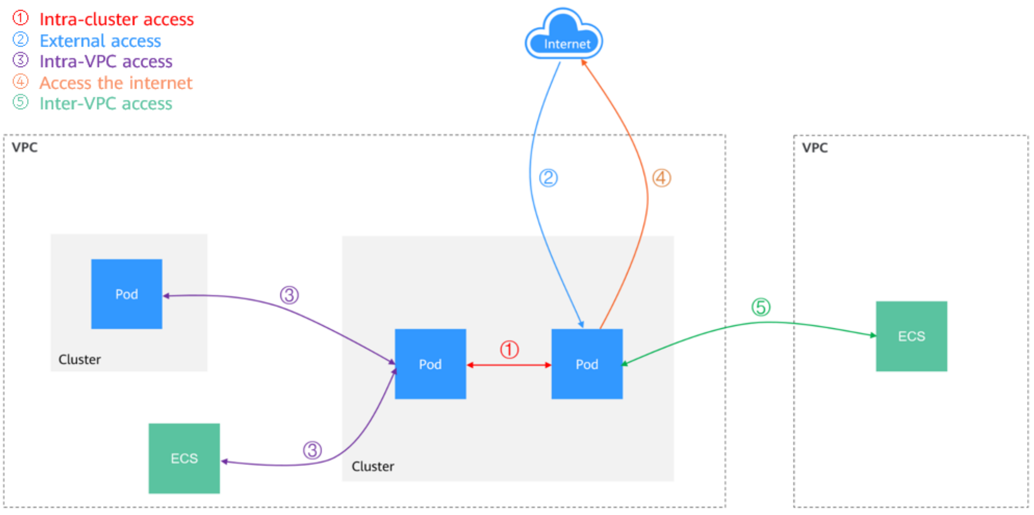


Network Access Scenarios

Workload access scenarios can be categorized as follows:

- Intra-cluster access: A ClusterIP Service is used for workloads in the same cluster to access each other.
- Access from outside a cluster: A Service (LoadBalancer) or an Ingress is recommended for a workload outside a cluster to access workloads in a cluster.
  - Access through a public network requires an EIP to be bound to the load balancer.
  - Access through an intranet requires the workload to be accessed through the internal IP address of the load balancer. If workloads are located in different VPCs, a peering connection is required to enable communication between different VPCs.
- The workload can access the external network as follows:
  - Accessing an intranet: How workloads access an intranet address varies depending on container network models. Ensure that the peer security group rules allow the access from the container CIDR block.
  - Accessing the internet: You can bind an EIP to the pod by referring to **Configuring an EIP for a Pod**, or configure SNAT rules for NAT gateway by referring to **Accessing Public Networks from a Container**.

Figure 3-3 Network access diagram



3.2 Container Network

3.2.1 Adding or Deleting a Container Subnet

Scenario

If the container subnet fails to meet your service requirements, you can add or delete container subnets as needed.

NOTE

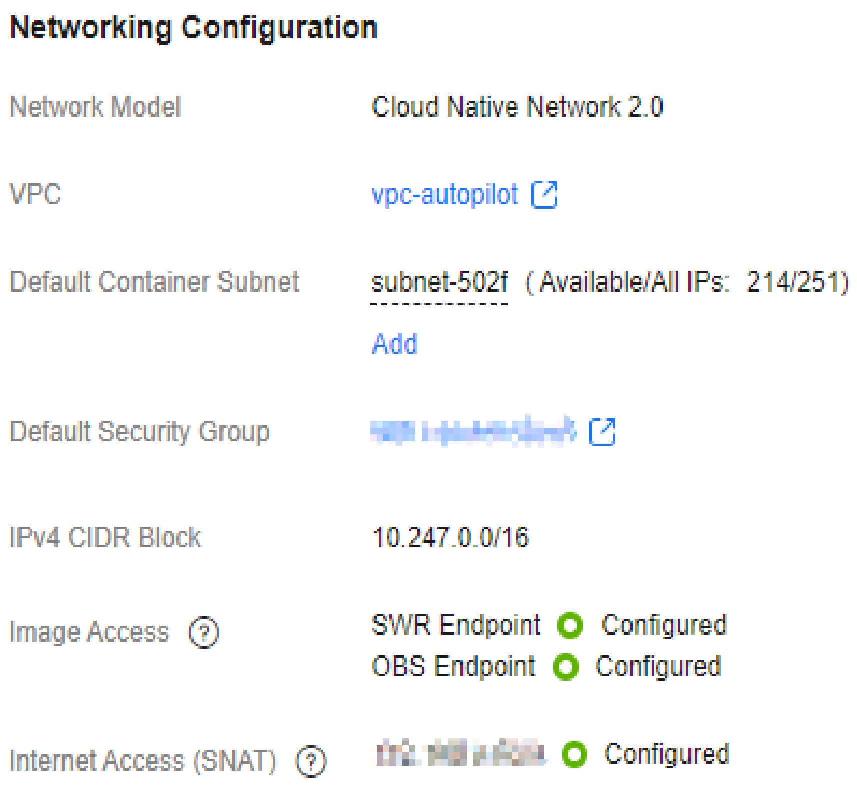
Container subnets can be deleted only when the cluster version is v1.27.8-r0, v1.28.6-r0, or later.

Adding Container Subnets

Step 1 Log in to the CCE console, locate the target CCE Autopilot cluster, and click its name.

Step 2 On the **Overview** page, locate the **Networking Configuration** area and click **Add** under **Default Container Subnet**.

Figure 3-4 Adding subnets

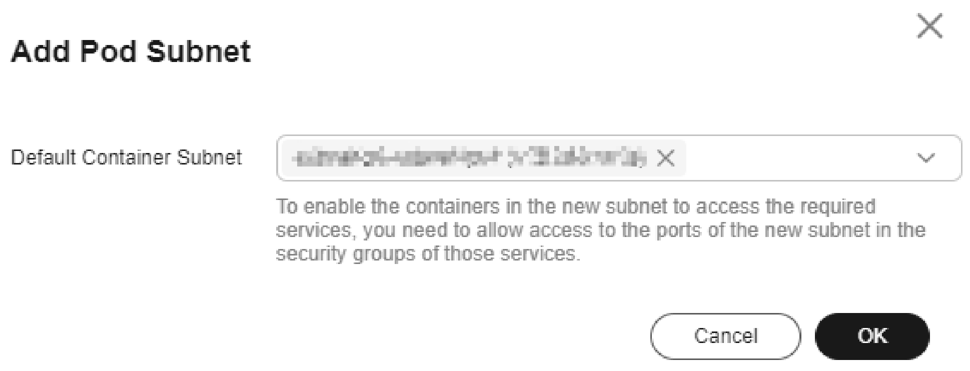


Networking Configuration

Network Model	Cloud Native Network 2.0
VPC	vpc-autopilot
Default Container Subnet	subnet-502f（Available/All IPs: 214/251） Add
Default Security Group	[illegible]
IPv4 CIDR Block	10.247.0.0/16
Image Access	SWR Endpoint ● Configured OBS Endpoint ● Configured
Internet Access (SNAT)	[illegible] ● Configured

Step 3 Select one or more subnets in the same VPC. If no subnet is available, go to the VPC console and create one first.

Figure 3-5 Selecting subnets



Add Pod Subnet

Default Container Subnet: [illegible]

To enable the containers in the new subnet to access the required services, you need to allow access to the ports of the new subnet in the security groups of those services.

Step 4 Click **OK**.

----End

Deleting Container Subnets

NOTICE

Deleting a container subnet is a high-risk operation. Ensure that no network interfaces are using the subnet to be deleted.

Procedure: On the cluster overview page, view the cluster ID in the **Basic Info** area. In the **Networking Configuration** > **Default Container Subnet** area, copy the subnet ID of the container subnet to be deleted. Then, use this ID to filter the network interfaces and network interfaces associated with this subnet on the **Network Interfaces** tab. If the name or description of any of the filtered network interfaces contains a cluster ID, those network interfaces are being used by the cluster.

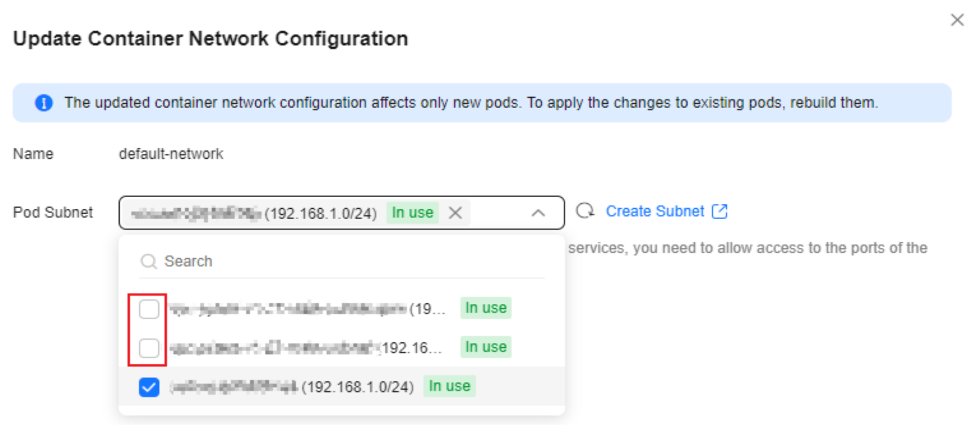
Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Cluster** > **Settings** and switch to the **Network** tab.

Step 3 In the **Container Network** area, click **Update** in the **Operation** column of **default-network (Default Container Subnet)**.

Step 4 In the **Pod Subnet** area, deselect the subnets and click **OK**.

Figure 3-6 Deselecting subnets



Update Container Network Configuration

The updated container network configuration affects only new pods. To apply the changes to existing pods, rebuild them.

Name default-network

Pod Subnet (192.168.1.0/24) In use

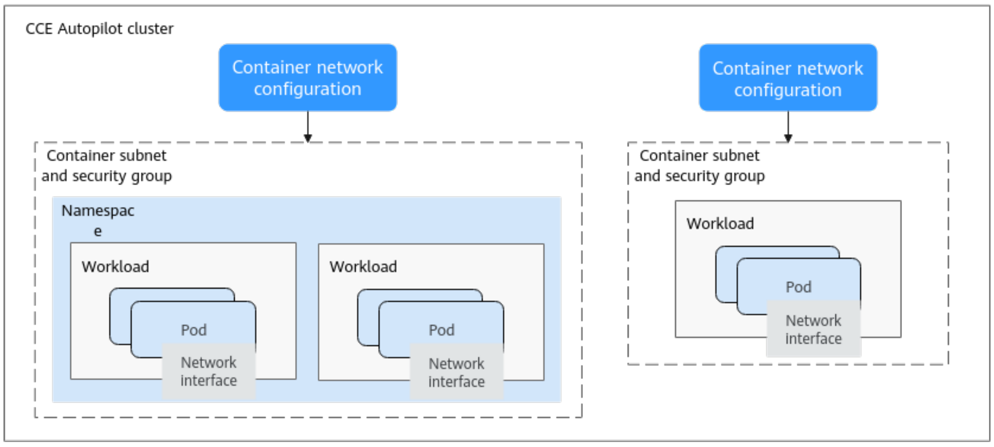
----End

3.2.2 Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration

Scenario

In a CCE Autopilot cluster, you can configure subnets and security groups for containers by namespace or workload using NetworkAttachmentDefinition, which is a **CRD** resource. To configure a particular container subnet and security group for a specified namespace or workload, create a container network configuration (NetworkAttachmentDefinition) and associate it with the target namespace or workload. In this way, service subnets can be planned or services can be securely isolated.

Figure 3-7 Custom container network settings


CCE Autopilot cluster
Container network configuration
Container network configuration
Container subnet and security group
Container subnet and security group
Namespace
Workload
Workload
Workload
Pod
Pod
Pod
Network interface
Network interface
Network interface

The following table lists the resources that a container network configuration (NetworkAttachmentDefinition) can be associated with.

Table 3-1 Associated resources

Dimension	Resources That a Container Network Configuration Can Be Associated With	
	Namespace	Workload
Subnet and security group configurations	All workloads created in the namespace associated with a container network configuration will use the same subnet and security group configurations.	The workloads using the same container network configuration will use the same subnet and security group configurations.
Supported CCE Autopilot cluster version	v1.27.8-r0, v1.28.6-r0, and later	v1.27.8-r0, v1.28.6-r0, and later

Dimension	Resources That a Container Network Configuration Can Be Associated With	
	Namespace	Workload
Constraint	A namespace can only be associated with one container network configuration.	Only custom container network configurations that are not associated with any namespace can be specified.

NOTE

- The priorities (in descending order) of container network configurations used by a pod are as follows: Container network configuration directly associated with the pod > Container network configuration associated with the pod namespace > Default container network configuration of the cluster (**default-network**)
- If **default-network** is available in a cluster, it will be applied on all pods when no custom container network configuration has been configured. The default container subnet in the network settings on the **Overview** page is the container subnet in **default-network**. **default-network** cannot be deleted.
- If there is only one container network configuration in a cluster, it is the default container network configuration. If there are multiple container network configurations in a cluster, the one with annotation **yangtse.io/default-network: true** is the default container network configuration. If there is no default container network configuration in a cluster, the pod that is not associated with any container network configuration fails to start after being created because no network interface can be allocated for the pod.

Constraints

Before deleting a custom container network configuration, delete the pods (with the **cni.yangtse.io/network-status** annotation) created using the configuration in the target namespace. For details, see **Deleting a Container Network Configuration**.

Creating a Container Network Configuration for a Namespace

After a container network configuration is created and associated with a namespace, all workloads created in the namespace will use the same subnet and security group configurations.

Creating a Container Network Configuration for a Namespace on the CCE Console

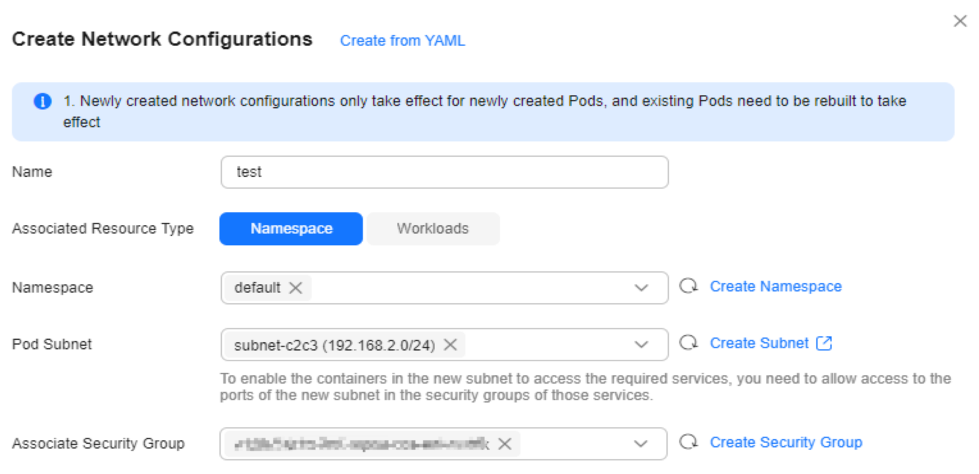
Step 1 Log in to the CCE console.

Step 2 Click the cluster name to access the cluster console. Choose **Settings** in the navigation pane on the left and click the **Network** tab.

Step 3 In the **Container Network** area, click **Add**. In the window that slides out from the right, configure parameters such as the pod subnet and security group.

- **Name**: Enter a name that contains a maximum of 63 characters. Do not use **default-network**, **default**, **mgnt0**, or **mgnt1**. They are reserved for system use.
- **Associated Resource Type**: Select a resource type that will be associated with the custom container network configuration. For details, see **Table 3-1**. To create a container network configuration for a namespace, select **Namespace**.
- **Namespace**: Select the namespace to be associated. The namespaces associated with different container network configurations must be unique. If no namespace is available, click **Create Namespace** to create one.
- **Pod Subnet**: Select one or more subnets. A maximum of 100 subnets can be selected. If no subnet is available, create one first. After the subnet is created, click the refresh button.
- **Associate Security Group**: The default value is the security group associated with the container network interface. You can also create one. After the security group is created, click the refresh button. A maximum of five security groups can be selected.

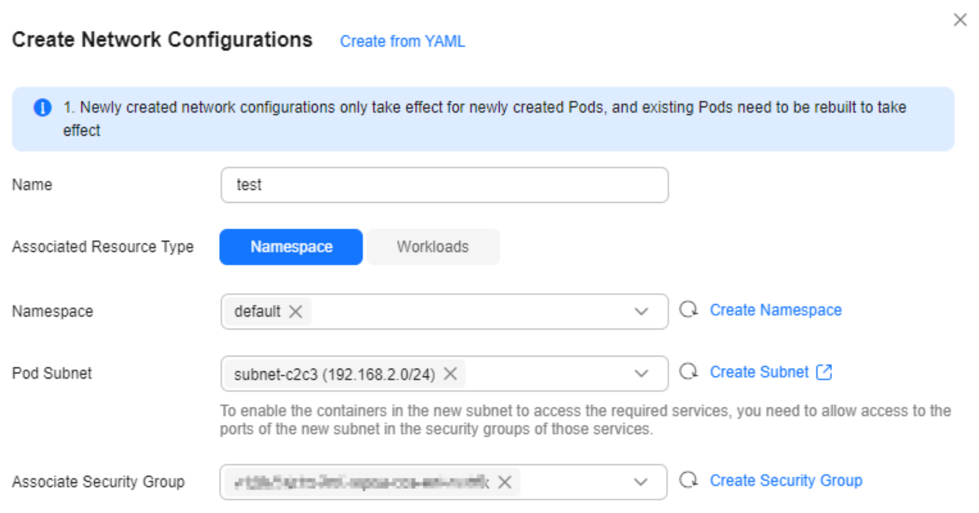
Figure 3-8 Creating a container network configuration for a namespace



1. Newly created network configurations only take effect for newly created Pods, and existing Pods need to be rebuilt to take effect

Name	test
Associated Resource Type	[✓ Namespace] Workloads
Namespace	default
Pod Subnet	subnet-c2c3 (192.168.2.0/24)
Associate Security Group	[redacted]

To enable the containers in the new subnet to access the required services, you need to allow access to the ports of the new subnet in the security groups of those services.

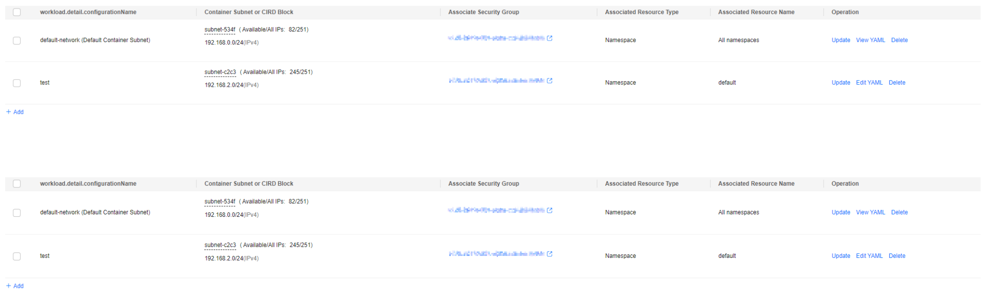

Create Network Configurations
Create from YAML

1. Newly created network configurations only take effect for newly created Pods, and existing Pods need to be rebuilt to take effect

Name: test
Associated Resource Type: [✓ Namespace] Workloads
Namespace: default
Pod Subnet: subnet-c2c3 (192.168.2.0/24)
To enable the containers in the new subnet to access the required services, you need to allow access to the ports of the new subnet in the security groups of those services.
Associate Security Group: [redacted]
Create Namespace
Create Subnet
Create Security Group

Step 4 Click **OK**. You will be redirected to the custom container network configuration list, where you can view new container network configuration.

Figure 3-9 Container network configuration list



	workload.detail.configurationName	Container Subnet or CIRD Block	Associate Security Group	Associated Resource Type	Associated Resource Name	Operation
☐	default-network (Default Container Subnet)	subnet-534f (Available/All IPs: 82/251) 192.168.0.0/24(IPv4)	[redacted]	Namespace	All namespaces	Update View YAML Delete
☐	test	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Namespace	default	Update Edit YAML Delete

+ Add

	workload.detail.configurationName	Container Subnet or CIRD Block	Associate Security Group	Associated Resource Type	Associated Resource Name	Operation
☐	default-network (Default Container Subnet)	subnet-534f (Available/All IPs: 82/251) 192.168.0.0/24(IPv4)	[redacted]	Namespace	All namespaces	Update View YAML Delete
☐	test	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Namespace	default	Update Edit YAML Delete

+ Add

----End

Using Kubectl to Create a Container Network Configuration for a Namespace

After a container network configuration is created and associated with a namespace, all workloads created in the namespace will use the same subnet and security group configurations.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a YAML file for creating a custom container network configuration and associating the configuration with a namespace. You can use any file name. In this example, the file name is **networkattachment-test.yaml**.

```
vim networkattachment-test.yaml
```

Example file content:

```
apiVersion: k8s.cni.cncf.io/v1
kind: NetworkAttachmentDefinition
metadata:
  annotations:
    yangtse.io/project-id: 05e38**
  name: example
  namespace: kube-system
spec:
  config: |
   {
    "type":"eni-neutron",
    "args":{
      "securityGroups":"41891**",
      "subnets":[
        {
          "subnetID":"27d95**"
        }
      ]
    },
    "selector":{
      "namespaceSelector":{
        "matchLabels":{
          "kubernetes.io/metadata.name":"default"
        }
      }
    }
   }
```

Table 3-2 Key parameters

Parameter	Mandatory	Type	Description
apiVersion	Yes	String	API version. The value is fixed at **k8s.cni.cncf.io/v1**.
kind	Yes	String	Type of the object to be created. The value is fixed at **NetworkAttachmentDefinition**.
yangtse.io/project-id	Yes	String	Project ID in the current region. For details, see **Obtaining a Project ID**.
name	Yes	String	Configuration name, which can be customized. Enter up to 63 characters. Start and end with a lowercase letter or digit. Only lowercase letters, digits, hyphens (-), and periods (.) are allowed. Do not use **default-network**, **default**, **mgnt0**, or **mgnt1**. They are reserved for system use.
namespace	Yes	String	Namespace of the configuration. The value is fixed at **kube-system**.
config	Yes	Object	Configuration content, which is a string in JSON format. For details about the data structure of the **config** field, see **Table 3-3**.

Table 3-3 Data structure of the **config** field

Parameter	Mandatory	Type	Description
type	Yes	String	Network type. The value is fixed at **eni-neutron**.
args	Yes	Object	Security group and subnet configurations. For details about the data structure of the **args** field, see **Table 3-4**.
selector	No	Object	Namespace where the configuration will take effect. For details about the data structure of the **selector** field, see **Table 3-5**.

Table 3-4 Data structure of the **args** field

Parameter	Mandatory	Type	Description
securityGroups	No	String	ID of the security group to be configured. A maximum of five security groups can be associated. If no security group is planned, ensure that the security group is the same as that in **default-network**. How to obtain: Log in to the VPC console. In the navigation pane on the left, choose **Access Control** > **Security Groups**. Click the target security group name and copy the ID on the **Summary** tab.
subnets	Yes	Array of subnetID objects	Container subnet ID list. Enter 1 to 100 subnet IDs. The format is as follows: `[{"subnetID":"27d95**"},{"subnetID":"827bb**"},{"subnetID":"bdd6b**"}]` The container subnet must be in the same VPC as the cluster. How to obtain: Log in to the VPC console. In the navigation pane on the left, choose **Virtual Private Cloud** > **Subnets**. Click the target subnet name and copy the **Subnet ID** on the **Summary** tab.

Table 3-5 Data structure of the **selector** field

Parameter	Mandatory	Type	Description
namespaceSelector	No	matchLabels object	A Kubernetes standard selector. Enter the namespace label in the following format: `"matchLabels":{ "kubernetes.io/metadata.name":"default" }` The namespaces for different configurations must be unique.

Step 3 Create the container network configuration.

```
kubectl create -f networkattachment-test.yaml
```

If information similar to the following is displayed, the configuration has been created:

```
networkattachmentdefinition.k8s.cni.cncf.io/example created
```

----End

Creating a Container Network Configuration for a Workload

After such a container network configuration is created, all workloads associated with it will use the same subnet and security group configurations.

Creating a Container Network Configuration for a Workload on the CCE Console

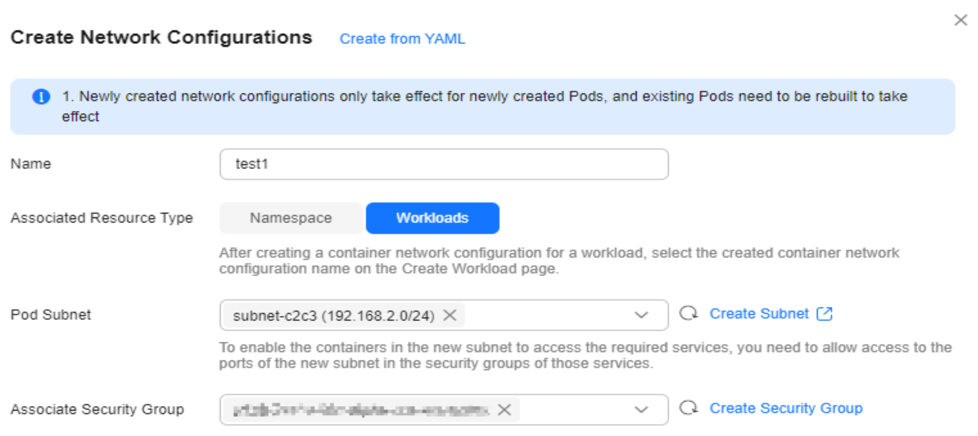
Step 1 Log in to the CCE console.

Step 2 Click the cluster name to access the cluster console. Choose **Settings** in the navigation pane on the left and click the **Network** tab.

Step 3 In the **Container Network** area, click **Add**. In the window that slides out from the right, configure parameters such as the pod subnet and security group.

- **Name**: Enter a name that contains a maximum of 63 characters. Do not use **default-network**, **default**, **mgnt0**, or **mgnt1**. They are reserved for system use.
- **Associated Resource Type**: Select a resource type that will be associated with the custom container network configuration. For details, see **Table 3-1**. To create a container network configuration for a workload, select **Workload**.
- **Pod Subnet**: Select one or more subnets. A maximum of 100 subnets can be selected. If no subnet is available, create one first. After the subnet is created, click the refresh button.
- **Associate Security Group**: The default value is the security group associated with the container network interface. You can also create one. After the security group is created, click the refresh button. A maximum of five security groups can be selected.

Figure 3-10 Creating a container network configuration for a workload



1. Newly created network configurations only take effect for newly created Pods, and existing Pods need to be rebuilt to take effect

Name: test1

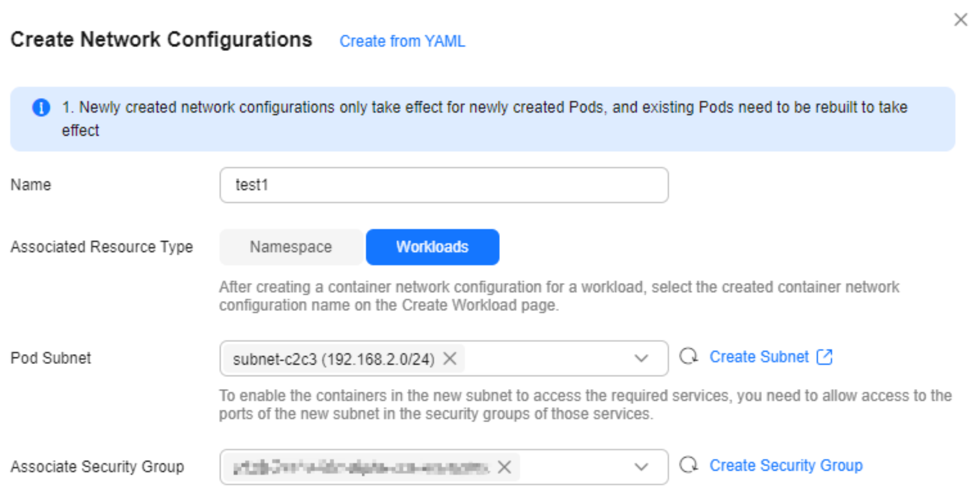
Associated Resource Type: Workloads

After creating a container network configuration for a workload, select the created container network configuration name on the Create Workload page.

Pod Subnet: subnet-c2c3 (192.168.2.0/24)

To enable the containers in the new subnet to access the required services, you need to allow access to the ports of the new subnet in the security groups of those services.

Associate Security Group:



Create Network Configurations   Create from YAML

1. Newly created network configurations only take effect for newly created Pods, and existing Pods need to be rebuilt to take effect

Name: test1

Associated Resource Type: Namespace [✓ Workloads]

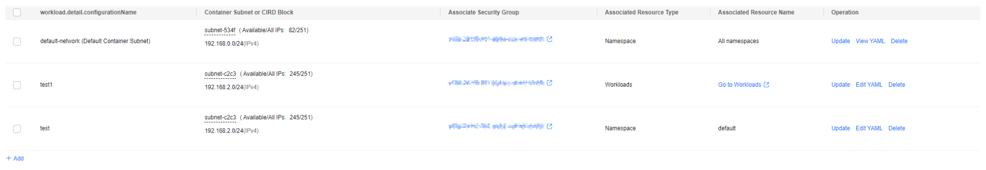
After creating a container network configuration for a workload, select the created container network configuration name on the Create Workload page.

Pod Subnet: subnet-c2c3 (192.168.2.0/24)   Create Subnet

To enable the containers in the new subnet to access the required services, you need to allow access to the ports of the new subnet in the security groups of those services.

Associate Security Group: [redacted]   Create Security Group

Step 4 Click **OK**. You will be redirected to the custom container network configuration list, where you can view new container network configuration.



workload.detail.configurationName	Container Subnet or CIRD Block	Associate Security Group	Associated Resource Type	Associated Resource Name	Operation
default-network (Default Container Subnet)	subnet-534f (Available/All IPs: 82/251) 192.168.0.0/24(IPv4)	[redacted]	Namespace	All namespaces	Update View YAML Delete
test1	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Workloads	Go to Workloads	Update Edit YAML Delete
test	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Namespace	default	Update Edit YAML Delete

+ Add

workload.detail.configurationName	Container Subnet or CIRD Block	Associate Security Group	Associated Resource Type	Associated Resource Name	Operation
default-network (Default Container Subnet)	subnet-534f (Available/All IPs: 82/251) 192.168.0.0/24(IPv4)	[redacted]	Namespace	All namespaces	Update View YAML Delete
test1	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Workloads	Go to Workloads	Update Edit YAML Delete
test	subnet-c2c3 (Available/All IPs: 245/251) 192.168.2.0/24(IPv4)	[redacted]	Namespace	default	Update Edit YAML Delete

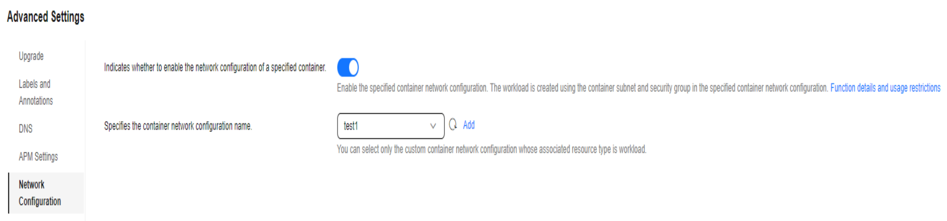
+ Add

Figure 3-11 Container network configuration list

Step 5 When creating a workload, select a custom container network configuration.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. Click **Create Workload** in the upper right corner of the page. In the **Advanced Settings** area, choose **Network Configuration** and determine whether to enable a specified container network configuration.
3. Select an existing container network configuration. If no configuration is available, click **Add** to create one.

Figure 3-12 Selecting a container network configuration



Advanced Settings

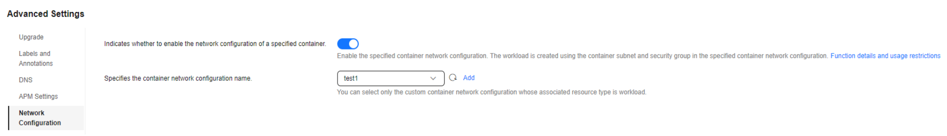
Upgrade | Labels and Annotations | DNS | APM Settings | Network Configuration

Indicates whether to enable the network configuration of a specified container. [✓]

Enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. Function details and usage restrictions

Specifies the container network configuration name.: test1   Add

You can select only the custom container network configuration whose associated resource type is workload.


Advanced Settings

- Upgrade
- Labels and Annotations
- DNS
- APM Settings
- Network Configuration

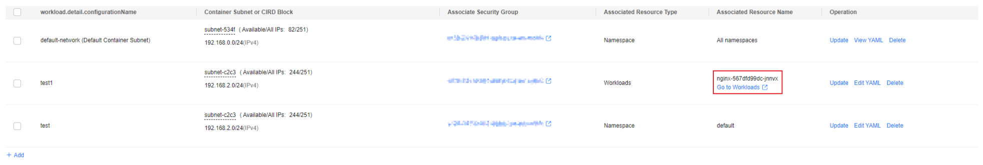
Indicates whether to enable the network configuration of a specified container.: enabled
Enable the specified container network configuration. The workload is created using the container subnet and security group in the specified container network configuration. Function details and usage restrictions

Specifies the container network configuration name.: test1
You can select only the custom container network configuration whose associated resource type is workload.

4. Configure other parameters and click **Create Workload**.

   Return to the **Settings** page. In the container network configuration list, the name of the resource associated with the created container network configuration is displayed.

   Figure 3-13 Resources associated with a container network configuration



workload.detail.configurationName	Container Subnet or CIRD Block	Associate Security Group	Associated Resource Type	Associated Resource Name	Operation
default-network (Default Container Subnet)	subnet-534f (Available/All IPs: 82/251) 192.168.0.0/24(IPv4)	[redacted]	Namespace	All namespaces	Update View YAML Delete
test1	subnet-c2c3 (Available/All IPs: 244/251) 192.168.2.0/24(IPv4)	[redacted]	Workloads	nginx-567dfd99dc-jnnvx Go to Workloads	Update Edit YAML Delete
test	subnet-c2c3 (Available/All IPs: 244/251) 192.168.2.0/24(IPv4)	[redacted]	Namespace	default	Update Edit YAML Delete

+ Add

----End

Using Kubectl to Create a Container Network Configuration for a Workload

After a container network configuration is created, all workloads associated with it will use the same subnet and security group configurations.

Step 1 Connect to a cluster using kubectl.

Step 2 Create a YAML file for creating a custom container network configuration. You can use any file name. In this example, the file name is **networkattachment-test.yaml**.

```
vim networkattachment-test.yaml
```

Example file content:

```
apiVersion: k8s.cni.cncf.io/v1
kind: NetworkAttachmentDefinition
metadata:
  annotations:
    yangtse.io/project-id: 80d5a**
  name: example
  namespace: kube-system
spec:
  config: |
   {
     "type":"eni-neutron",
     "args":{
       "securityGroups":"f4983**",
       "subnets":[
         {
           "subnetID":"5594b**"
         }
       ]
     }
   }
```

Table 3-6 Key parameters

Parameter	Mandatory	Type	Description
apiVersion	Yes	String	API version. The value is fixed at **k8s.cni.cncf.io/v1**.
kind	Yes	String	Type of the object to be created. The value is fixed at **NetworkAttachmentDefinition**.
yangtse.io/project-id	Yes	String	Project ID in the current region. For details, see **Obtaining a Project ID**.
name	Yes	String	Configuration name, which can be customized. Enter up to 63 characters. Start and end with a lowercase letter or digit. Only lowercase letters, digits, hyphens (-), and periods (.) are allowed. Do not use **default-network**, **default**, **mgnt0**, or **mgnt1**. They are reserved for system use.
namespace	Yes	String	Namespace of the configuration. The value is fixed at **kube-system**.
config	Yes	Object	Configuration content, which is a string in JSON format. For details about the data structure of the **config** field, see **Table 3-7**.

Table 3-7 Data structure of the **config** field

Parameter	Mandatory	Type	Description
type	Yes	String	Network type. The value is fixed at **eni-neutron**.
args	Yes	Object	Security group and subnet configurations. For details about the data structure of the **args** field, see **Table 3-8**.

Table 3-8 Data structure of the **args** field

Parameter	Mandatory	Type	Description
securityGroups	No	String	ID of the security group to be configured. A maximum of five security groups can be associated. If no security group is planned, ensure that the security group is the same as that in **default-network**. How to obtain: Log in to the VPC console. In the navigation pane on the left, choose **Access Control** > **Security Groups**. Click the target security group name and copy the ID on the **Summary** tab.
subnets	Yes	Array of subnetID objects	Container subnet ID list. Enter 1 to 100 subnet IDs. The format is as follows: `[{"subnetID":"27d95**"},{"subnetID":"827bb**"},{"subnetID":"bdd6b**"}]` The container subnet must be in the same VPC as the cluster. How to obtain: Log in to the VPC console. In the navigation pane on the left, choose **Virtual Private Cloud** > **Subnets**. Click the target subnet name and copy the **Subnet ID** on the **Summary** tab.

Step 3 Create the container network configuration.

```
kubectl create -f networkattachment-test.yaml
```

If information similar to the following is displayed, the configuration has been created:

```
networkattachmentdefinition.k8s.cni.cncf.io/example created
```

Step 4 Create a YAML file for creating a workload (for example, a Deployment) and associating it with the new container network configuration. You can use any file name. In this example, the file name is **deploy.yaml**.

```
vim deploy.yaml
```

Example file content:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
```

```
 selector:
   matchLabels:
     app: nginx
 template:
   metadata:
     labels:
       app: nginx
       yangtse.io/network: "example"  # Name of the custom container network configuration, which can
be used to obtain all pods associated with the container network configuration by label
     annotations:
       yangtse.io/network: "example"  # Name of the custom container network configuration
   spec:
     containers:
       - name: container-0
         image: nginx:alpine
         resources:
           limits:
             cpu: 100m
             memory: 200Mi
           requests:
             cpu: 100m
             memory: 200Mi
     imagePullSecrets:
       - name: default-secret
```

- **yangtse.io/network**: name of the specified custom container network configuration. Only a container network configuration that is not associated with any namespace can be specified. Add this parameter to the label so that you can use the label to obtain all pods associated with this container network configuration.

  NOTICE

  The name of the custom container network configuration specified by **yangtse.io/network** is the same as the value of **name** in **networkattachment-test.yaml**.

Step 5 Create a Deployment.

```
kubectl create -f deploy.yaml
```

Information similar to the following is displayed:

```
deployment.apps/nginx created
```

----End

Verify That a Container Network Configuration Is Associated with a Namespace or Workload

The following are steps to verify that a subnet and security group specified in the container network configuration have been associated with a workload. The steps for verifying that a container network configuration has been associated with a namespace are the same as those provided here.

Step 1 Verify that a subnet is associated.

1. Check the IP address of a pod for a workload.

   ```
   kubectl get pod -o wide
   ```

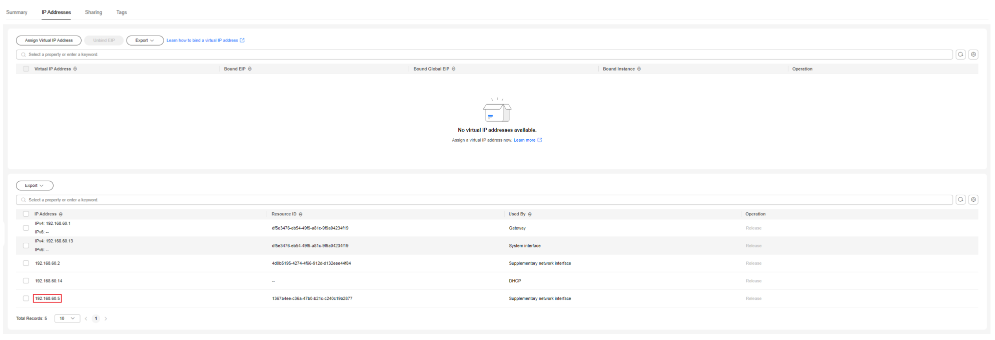
   Information similar to the following is displayed:

   ```
   NAME                READY   STATUS    RESTARTS   AGE    IP            NODE
   NOMINATED NODE   READINESS GATES
   ```

```
nginx-85dbdb8c5d-ng5bb   1/1     Running   0         5d18h   192.168.60.5   ca50a5ae-e1ef-41c6-
b3fc-6ebcd10a1e07   <none>           <none>
```

2. Log in to **Network Console**. In the navigation pane on the left, choose **Virtual Private Cloud** > **Subnets**. On the displayed page, click the subnet name.
3. Click **IP Addresses**. If the IP address of the pod is displayed, the subnet is associated.

Figure 3-14 Viewing the subnet associated with a pod


Summary
IP Addresses
Sharing
Tags
Assign Virtual IP Address
Unbind EIP
Export
Learn how to bind a virtual IP address
Select a property or enter a keyword.
No virtual IP addresses available.
Assign a virtual IP address now. Learn more
Export
Select a property or enter a keyword.
Total Records: 5
10
1

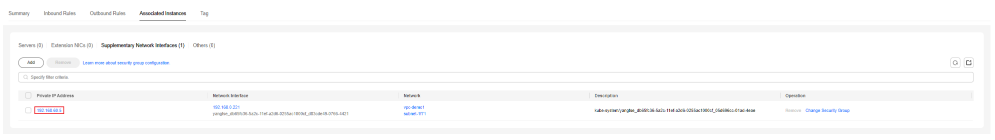
Virtual IP Address	Bound EIP	Bound Global EIP	Bound Instance	Operation

IP Address	Resource ID	Used By	Operation
IPv4: 192.168.60.1 IPv6: --	df5e3476-eb54-49f9-a81c-9f9a04234f19	Gateway	Release
IPv4: 192.168.60.13 IPv6: --	df5e3476-eb54-49f9-a81c-9f9a04234f19	System interface	Release
192.168.60.2	4d0b5195-4274-4f66-912d-d132eee44f84	Supplementary network interface	Release
192.168.60.14	--	DHCP	Release
192.168.60.5	1367a4ee-c36a-47b0-b21c-c240c19a2877	Supplementary network interface	Release

Step 2 Verify that a security group is associated.

1. Log in to **Network Console**. In the navigation pane on the left, choose **Access Control** > **Security Groups**. On the displayed page, click the security group name.
2. On the **Associate Instances** tab, click **Supplementary Network Interfaces**.
3. View the private IP address of the pod in the **Private IP Address** column. If the private IP address of the pod is displayed, the security group is associated.

Figure 3-15 Viewing the IP address associated with a security group



Private IP Address	Network Interface	Network	Description	Operation
192.168.60.5	192.168.0.221 yangtse_db65fc36-5a2c-11ef-a2d6-0255ac1000cf_d83cde49-0766-4421	vpc-demo1 subnet-1f71	kube-system/yangtse_db65fc36-5a2c-11ef-a2d6-0255ac1000cf_05d696cc-01ad-4eae	Remove Change Security Group

----End

Deleting a Container Network Configuration

You can delete a new container network configuration or view its YAML file.

Before deleting a container network configuration, delete all pods using the configuration, or the deletion will fail.

1. Run the following command to filter out the pods that use the configuration (**example** is used as an example here) in the cluster:

```
kubectl get pod -A -o=jsonpath="{.items[?(@.metadata.annotations.cni\.yangtse\.io/network-
status=='[{\"name\":\"example\"}]')]['metadata.namespace', 'metadata.name']}"
```

   The name of each pod and the namespace associated with the configuration will be displayed.

2. Delete the workload that each pod is running. The workload can be a Deployment, StatefulSet, Job, or CronJob.
   ```
   kubectl delete deployment nginx
   ```
   Information similar to the following is displayed:
   ```
   deployment.apps "nginx" deleted
   ```
3. Delete the container network configuration.
   ```
   kubectl delete -f networkattachment-test.yaml -n kube-system
   ```
   Information similar to the following is displayed:
   ```
   networkattachmentdefinition.k8s.cni.cncf.io "example" deleted
   ```

3.3 Service

3.3.1 ClusterIP

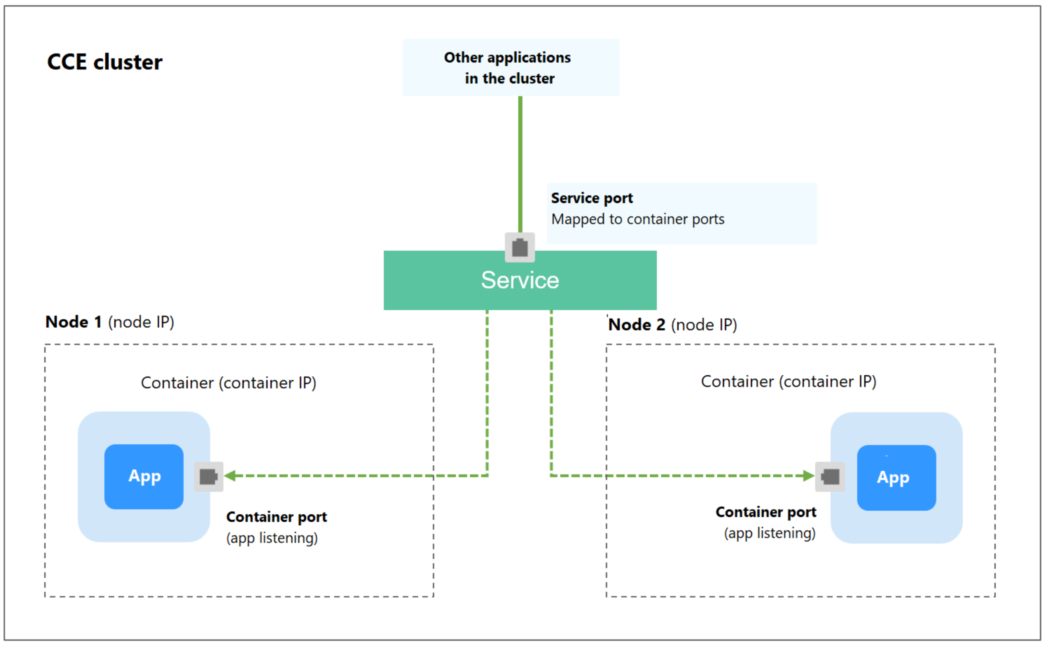
Scenario

ClusterIP Services allow workloads in the same cluster to use their cluster-internal domain names to access each other.

The cluster-internal domain name format is *<Service name>*.*<Namespace of the workload>***.svc.cluster.local:***<Port>*, for example, **nginx.default.svc.cluster.local:80**.

Figure 3-16 shows the access channel and mappings between the access port and container ports.

Figure 3-16 Intra-cluster access (ClusterIP)



Creating a ClusterIP Service

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Services & Ingresses**. In the upper right corner, click **Create Service**.

Step 3 Configure intra-cluster access parameters.

- **Service Name**: Specify a Service name, which can be the same as the workload name.
- **Service Type**: Select **ClusterIP**.
- **Namespace**: Select the namespace that the workload belongs to.
- **Selector**: Add a label and click **Confirm**. A Service selects a pod based on the added label. You can also click **Reference Workload Label** to reference the label of an existing workload. In the dialog box that is displayed, select a workload and click **OK**.
- **Ports**
  - **Protocol**: Select the protocol used by the Service.
  - **Service Port**: Specify the port used by the Service. The port number ranges from 1 to 65535.
  - **Container Port**: Specify the port on which the workload listens. For example, Nginx uses port 80 by default.

Step 4 Click **OK**.

----End

Adding a Service Using kubectl

You can run kubectl commands to add a Service. An Nginx workload is used as an example to describe how to implement intra-cluster access using kubectl.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create the **nginx-deployment.yaml** and **nginx-clusterip-svc.yaml** files and edit them.

The file names are user-defined. **nginx-deployment.yaml** and **nginx-clusterip-svc.yaml** are merely example file names.

vi nginx-deployment.yaml

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
```

```
   - image: nginx:latest
     name: nginx
   imagePullSecrets:
   - name: default-secret
```

vi nginx-clusterip-svc.yaml

```
apiVersion: v1
kind: Service
metadata:
  labels:
    app: nginx
  name: nginx-clusterip
spec:
  ports:
  - name: service0
    port: 8080              # Port for accessing a Service
    protocol: TCP           # Protocol used for accessing a Service. The value can be TCP or UDP.
    targetPort: 80          # Port used by a Service to access the target container. This port is closely related 
to the applications running in a container. In this example, the Nginx image uses port 80 by default.
  selector:                 # Label selector. A Service selects a pod based on the label and forwards the requests 
for accessing the Service to the pod. In this example, select the pod with the app:nginx label.
    app: nginx
  type: ClusterIP           # Type of a Service. ClusterIP indicates that a Service is only reachable from within 
the cluster.
```

Step 3 Create a workload.

kubectl create -f nginx-deployment.yaml

If information similar to the following is displayed, the workload has been created.

```
deployment "nginx" created
```

kubectl get po

If information similar to the following is displayed, the workload is running.

```
NAME                     READY     STATUS             RESTARTS   AGE
nginx-2601814895-znhbr   1/1       Running            0          15s
```

Step 4 Create a Service.

kubectl create -f nginx-clusterip-svc.yaml

If information similar to the following is displayed, the Service is being created.

```
service "nginx-clusterip" created
```

kubectl get svc

If information similar to the following is displayed, the Service has been created, and a cluster-internal IP address has been assigned to the Service.

```
# kubectl get svc
NAME              TYPE        CLUSTER-IP     EXTERNAL-IP   PORT(S)    AGE
kubernetes        ClusterIP   10.247.0.1     <none>        443/TCP    4d6h
nginx-clusterip   ClusterIP   10.247.74.52   <none>        8080/TCP   14m
```

Step 5 Access the Service.

A Service can be accessed from containers or nodes in a cluster.

Create a pod, access the pod, and run the **curl** command to access the Service over *IP address:Port* or the domain name.

The domain name suffix can be omitted. In the same namespace, you can use **nginx-clusterip:8080** for access. In other namespaces, you can use **nginx-clusterip.default:8080** for access.

```
# kubectl run -i --tty --image nginx:alpine test --rm /bin/sh
If you do not see a command prompt, try pressing Enter.
/ # curl 10.247.74.52:8080
<!DOCTYPE html>
<html>
<head>
<title>Welcome to nginx!</title>
<style>
    body {
        width: 35em;
        margin: 0 auto;
        font-family: Tahoma, Verdana, Arial, sans-serif;
    }
</style>
</head>
<body>
<h1>Welcome to nginx!</h1>
<p>If you see this page, the nginx web server is successfully installed and
working. Further configuration is required.</p>

<p>For online documentation and support please refer to
<a href="http://nginx.org/">nginx.org</a>.<br/>
Commercial support is available at
<a href="http://nginx.com/">nginx.com</a>.</p>

<p><em>Thank you for using nginx.</em></p>
</body>
</html>
/ # curl nginx-clusterip.default.svc.cluster.local:8080
...
<h1>Welcome to nginx!</h1>
...
/ # curl nginx-clusterip.default:8080
...
<h1>Welcome to nginx!</h1>
...
/ # curl nginx-clusterip:8080
...
<h1>Welcome to nginx!</h1>
...
```

----End

3.3.2 LoadBalancer

3.3.2.1 Creating a LoadBalancer Service

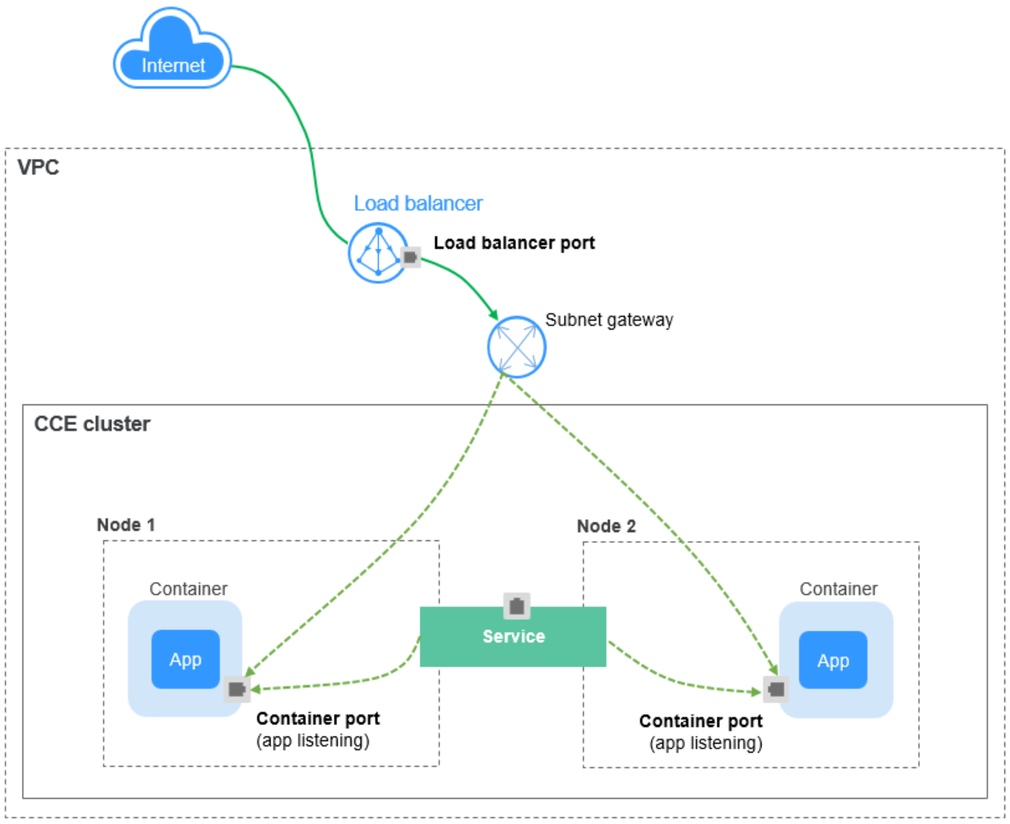
Scenario

LoadBalancer Services can access workloads from the public network through ELB, which is more reliable than EIP-based access. The LoadBalancer access address is in the format of *IP address of public network load balancer*:*Access port*, for example, **10.117.117.117:80**.

If dedicated load balancers are deployed for CCE Autopilot clusters, passthrough networking is supported to reduce the network latency and ensure zero performance loss.

External access requests are directly forwarded from a load balancer to pods. Internal access requests can be forwarded to a pod through a Service.

Figure 3-17 Passthrough networking



Constraints

- Automatically created load balancers should not be used by other resources. If they are used by other resources, they cannot be deleted completely.
- Dedicated load balancers must have private IP addresses bound. If a Service needs to support HTTP, dedicated load balancers must also support application load balancing (load balancing over HTTP or HTTPS).

Creating a LoadBalancer Service

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Services & Ingresses**. In the upper right corner, click **Create Service**.

Step 3 Configure parameters.

- **Service Name**: Specify a Service name, which can be the same as the workload name.
- **Service Type**: Select **LoadBalancer**.
- **Namespace**: Select the namespace that the workload belongs to.
- **Service Affinity**

  Cluster level: The IP addresses and access ports of all nodes in a cluster can access the workload associated with the Service. Service access will cause performance loss due to route redirection, and the source IP address of the client cannot be obtained.

- **Selector**: Add a label and click **Confirm**. A Service selects a pod based on the added label. You can also click **Reference Workload Label** to reference the label of an existing workload. In the dialog box that is displayed, select a workload and click **OK**.
- **Load Balancer**: Select the load balancer type and whether to use an existing load balancer or create a new one.

  You can select **Dedicated**. A dedicated load balancer supports **Network (TCP/UDP)**, **Application (HTTP/HTTPS)**, or **Network (TCP/UDP) & Application (HTTP/HTTPS)**.

  Select either **Use existing** or **Auto create**. For more information, see **Table 3-9**.

Table 3-9 Load balancer configurations

Option	Description
Use existing	Only the load balancers in the same VPC as the cluster can be selected. If no load balancer is available, click **Create Load Balancer** to create one on the ELB console.
Auto create	– **Instance Name**: Enter a load balancer name. – **Enterprise Project**: This parameter is only available for enterprise users who have enabled an enterprise project. Enterprise projects facilitate project-level management and grouping of cloud resources and users. – **AZ**: This parameter is only available for dedicated load balancers. You can create load balancers in multiple AZs to improve service availability. If diaster recovery is required, you are advised to select multiple AZs. – **Frontend Subnet**: This parameter is used to allocate IP addresses to load balancers to receive traffic from clients. – **Backend Subnet**: This parameter is used to allocate IP addresses for load balancers to routing traffic to pods. – **Network/Application-oriented Specifications** ▪ **Elastic**: applies to fluctuating traffic, billed based on the total traffic. ▪ **Fixed**: applies to stable traffic, billed based on specifications. – **EIP**: If you select **Auto create**, you can select a bandwidth billing option and set the bandwidth. – **Resource Tag**: You can add resource tags to classify resources. You can create predefined tags on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. This is supported by clusters of v1.27.5-r0, v1.28.3-r0, and later versions.

You can click **Edit** in the **Set ELB** area to set the load balancing algorithm and sticky session.

- **Algorithm**: Three algorithms are available: weighted round robin, weighted least connections algorithm, or source IP hash.

  NOTE

  - **Weighted round robin**: Requests are forwarded to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests. This algorithm is often used for short connections, such as HTTP services.
  - **Weighted least connections**: In addition to the weight assigned to each server, the number of connections processed by each backend server is considered. Requests are forwarded to the server with the lowest connections-to-weight ratio. This algorithm is based on the least connections algorithm and assigns a weight to each server based on their processing capability. This algorithm is often used for persistent connections, such as database connections.
  - **Source IP hash**: The source IP address of each request is calculated using the hash algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key allocates the client to a particular server. This enables requests from different clients to be distributed in load balancing mode and ensures that requests from the same client are forwarded to the same server. This algorithm applies to TCP connections without cookies.
- **Sticky Session**: This feature is disabled by default.

  After this feature is enabled, the sticky session type will be automatically selected based on the protocol supported by the listener.

  - If the listener's frontend protocol is TCP or UDP, sticky sessions can use source IP addresses. This means that access requests from the same IP address will be directed to the same pod.
  - If the listener's frontend protocol is HTTP or HTTPS, sticky sessions can use load balancer cookies.

  NOTE

  When **Source IP hash** is used for load balancing, sticky sessions are not available.

- **Health Check**: Configure health check for the load balancer.
  - **Global health check**: applies only to ports of the same protocol. You are advised to select **Custom health check**.
  - **Custom health check**: applies to **ports** used by different protocols.

Table 3-10 Health check parameters

Parameter	Description
Protocol	When the protocol is set to **TCP**, both TCP and HTTP are supported. When the protocol is set to **UDP**, only UDP is supported. **Check Path**: This parameter is only available for HTTP health check. It specifies the URL for health check. The check path must start with a slash (/) and contain 1 to 80 characters.
Port	By default, the service ports are used for health check. You can also specify another port for health check. If a port is specified, a service port named **cce-healthz** will be added for the Service. **Container Port**: When a dedicated load balancer has an elastic network interface associated, the container port is used for health check. The value ranges from 1 to 65535.
Check Period (s)	Specifies the maximum interval between health checks. The value ranges from **1** to **50**.
Timeout (s)	Specifies the maximum timeout duration for each health check. The value ranges from **1** to **50**.
Max. Retries	Specifies the maximum number of health check retries. The value ranges from **1** to **10**.

- **Port**
  - **Protocol**: Select the protocol used by the Service.
  - **Service Port**: Specify the port used by the Service. The port number ranges from 1 to 65535.
  - **Container Port**: Specify the port on which the workload listens. For example, Nginx uses port 80 by default.
  - **Frontend Protocol**: Set the protocol of the load balancer listener for establishing a connection with the clients. For a dedicated load balancer, to use HTTP/HTTPS, the type of the load balancer must be **Application (HTTP/HTTPS)**.
  - **Health Check**: If **Health Check** is set to **Custom health check**, you can configure health check for ports that come with different protocols. For details, see **Table 3-10**.

  NOTE

  When a LoadBalancer Service is created, a random node port number (NodePort) is automatically generated.
- **Annotation**: A LoadBalancer Service has some advanced features, which are implemented by annotations. For details, see **Using Annotations to Configure Load Balancing**.

Step 4 Click **OK**.

----End

Using kubectl to Create a Service (Using an Existing Load Balancer)

You can set the Service when creating a workload using kubectl. This section uses an Nginx workload as an example to describe how to add a LoadBalancer Service using kubectl.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create the **nginx-deployment.yaml** and **nginx-elb-svc.yaml** files and edit them.

The file names are user-defined. **nginx-deployment.yaml** and **nginx-elb-svc.yaml** are merely example file names.

vi nginx-deployment.yaml

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
      - image: nginx
        name: nginx
      imagePullSecrets:
      - name: default-secret
```

vi nginx-elb-svc.yaml

```
apiVersion: v1
kind: Service
metadata:
  name: nginx
  annotations:
    kubernetes.io/elb.id: <your_elb_id>                          # ELB ID. Replace it with the actual value.
    kubernetes.io/elb.class: performance                     # Load balancer type
    kubernetes.io/elb.lb-algorithm: ROUND_ROBIN                   # Load balancer algorithm
    kubernetes.io/elb.session-affinity-mode: SOURCE_IP          # The sticky session type is source IP address.
    kubernetes.io/elb.session-affinity-option: '{"persistence_timeout": "30"}'     # Stickiness duration (min)
    kubernetes.io/elb.health-check-flag: 'on'                   # Enable the ELB health check function.
    kubernetes.io/elb.health-check-option: '{
      "protocol":"TCP",
      "delay":"5",
      "timeout":"10",
      "max_retries":"3"
    }'
spec:
  selector:
     app: nginx
  ports:
  - name: service0
    port: 80     # Port for accessing the Service, which is also the listener port on the load balancer.
    protocol: TCP
    targetPort: 80  # Port used by a Service to access the target container. This port is closely related to the
applications running in a container.
    nodePort: 31128  # Port number on the node. If this parameter is not specified, a random port number
ranging from 30000 to 32767 is generated.
  type: LoadBalancer
```

This example uses annotations to implement some advanced features of load balancing, such as sticky sessions and health check. For details, see **Table 3-11**.

For more annotations and examples related to advanced features, see **Using Annotations to Configure Load Balancing**.

Table 3-11 annotations parameters

Parameter	Mandatory	Type	Description
kubernetes.io/elb.id	Yes	String	ID of an enhanced load balancer. Mandatory when an existing load balancer is to be associated. **How to obtain**: On the management console, click **Service List**, and choose **Networking** > **Elastic Load Balance**. Click the name of the load balancer. On the **Summary** tab, find and copy the ID.
kubernetes.io/elb.class	Yes	String	The value can be: ● **performance**: dedicated load balancer **NOTE** When an existing dedicated load balancer is used for a LoadBalancer Service, the private network must be supported.
kubernetes.io/elb.lb-algorithm	No	String	Specifies the load balancing algorithm of the backend server group. The default value is **ROUND_ROBIN**. Options: ● **ROUND_ROBIN**: weighted round robin algorithm ● **LEAST_CONNECTIONS**: weighted least connections algorithm ● **SOURCE_IP**: source IP hash algorithm **NOTE** If this parameter is set to **SOURCE_IP**, the weight setting (**weight** field) of backend servers bound to the backend server group is invalid, and sticky session cannot be enabled.

Parameter	Mandatory	Type	Description
kubernetes.io/elb.session-affinity-mode	No	String	In source IP address-based sticky sessions, requests from the same IP address are forwarded to the same backend server. ● Disabling sticky session: Do not configure this parameter. ● Enabling sticky session: Set this parameter to **SOURCE_IP**, indicating that the sticky session is based on the source IP address. **NOTE** When **kubernetes.io/elb.lb-algorithm** is set to **SOURCE_IP** (source IP hash), sticky session cannot be enabled.
kubernetes.io/elb.session-affinity-option	No	**Table 3-12** Object	Sticky session timeout.
kubernetes.io/elb.health-check-flag	No	String	Whether to enable the ELB health check. ● Enabling health check: Leave blank this parameter or set it to **on**. ● Disabling health check: Set this parameter to **off**. If this option is enabled, the **kubernetes.io/elb.health-check-option** field must also be specified at the same time.
kubernetes.io/elb.health-check-option	No	**Table 3-13** Object	ELB health check configuration items.

Table 3-12 elb.session-affinity-option data structure

Parameter	Mandatory	Type	Description
persistence_timeout	Yes	String	Sticky session timeout, in minutes. This parameter is valid only when **elb.session-affinity-mode** is set to **SOURCE_IP**. Value range: 1 to 60. Default value: **60**

Table 3-13 elb.health-check-option data structure

Parameter	Mandatory	Type	Description
delay	No	String	Health check interval (s) Value range: 1 to 50. Default value: **5**
timeout	No	String	Health check timeout, in seconds. Value range: 1 to 50. Default value: **10**
max_retries	No	String	Maximum number of health check retries. Value range: 1 to 10. Default value: **3**
protocol	No	String	Health check protocol. Value options: TCP or HTTP
path	No	String	Health check URL. This parameter needs to be configured when the protocol is **HTTP**. Default value: **/** Value range: 1-80 characters

Step 3 Create a workload.

kubectl create -f nginx-deployment.yaml

If information similar to the following is displayed, the workload has been created.

```
deployment/nginx created
```

kubectl get pod

If information similar to the following is displayed, the workload is running.

```
NAME                     READY     STATUS             RESTARTS   AGE
nginx-2601814895-c1xhw   1/1       Running            0          6s
```

Step 4 Create a Service.

kubectl create -f nginx-elb-svc.yaml

If information similar to the following is displayed, the Service has been created.

```
service/nginx created
```

kubectl get svc

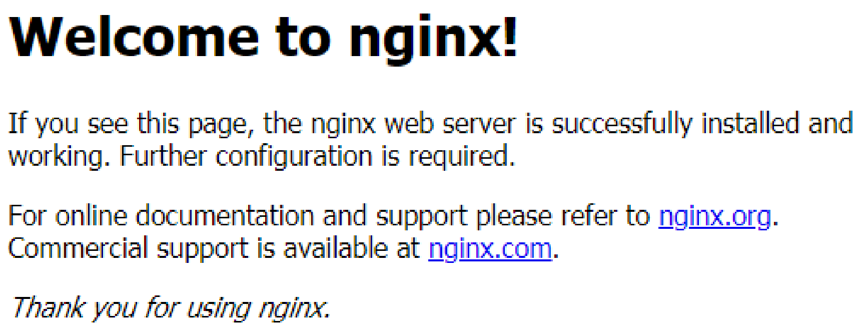
If information similar to the following is displayed, the access type has been set, and the workload is accessible.

```
NAME         TYPE           CLUSTER-IP       EXTERNAL-IP    PORT(S)        AGE
kubernetes   ClusterIP      10.247.0.1       <none>         443/TCP        3d
nginx        LoadBalancer   10.247.130.196   10.78.42.242   80:31540/TCP   51s
```

Step 5 Enter the URL in the address box of the browser, for example, **10.78.42.242:80**. **10.78.42.242** indicates the IP address of the load balancer, and **80** indicates the access port displayed on the CCE console.

Nginx is accessible.

Figure 3-18 Accessing Nginx through the LoadBalancer Service



Welcome to nginx!

If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to nginx.org.
Commercial support is available at nginx.com.

Thank you for using nginx.

----End

Using kubectl to Create a Service (Automatically Creating a Load Balancer)

You can set the Service when creating a workload using kubectl. This section uses an Nginx workload as an example to describe how to add a LoadBalancer Service using kubectl.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create the **nginx-deployment.yaml** and **nginx-elb-svc.yaml** files and edit them.

The file names are user-defined. **nginx-deployment.yaml** and **nginx-elb-svc.yaml** are merely example file names.

vi nginx-deployment.yaml

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
      - image: nginx
        name: nginx
      imagePullSecrets:
      - name: default-secret
```

vi nginx-elb-svc.yaml

Example of a Service using a dedicated load balancer on a public network:

```
apiVersion: v1
kind: Service
metadata:
  name: nginx
  labels:
    app: nginx
  namespace: default
  annotations:
    kubernetes.io/elb.class: performance
    kubernetes.io/elb.autocreate: '{
      "type": "public",
      "bandwidth_name": "cce-bandwidth-1626694478577",
      "bandwidth_chargemode": "bandwidth",
      "bandwidth_size": 5,
      "bandwidth_sharetype": "PER",
      "eip_type": "5_bgp",
      "vip_subnet_cidr_id": "*****",
      "vip_address": "**.**.**.**",
      "elb_virsubnet_ids": ["*****"],
      "available_zone": [
        ""
      ],
      "l4_flavor_name": "L4_flavor.elb.s1.small"
    }'
    kubernetes.io/elb.enterpriseID: '0'       # ID of the enterprise project to which the load balancer belongs
    kubernetes.io/elb.lb-algorithm: ROUND_ROBIN                  # Load balancer algorithm
    kubernetes.io/elb.session-affinity-mode: SOURCE_IP         # The sticky session type is source IP address.
    kubernetes.io/elb.session-affinity-option: '{"persistence_timeout": "30"}'     # Stickiness duration (min)
    kubernetes.io/elb.health-check-flag: 'on'                   # Enable the ELB health check function.
    kubernetes.io/elb.health-check-option: '{
      "protocol":"TCP",
      "delay":"5",
      "timeout":"10",
      "max_retries":"3"
    }'
    kubernetes.io/elb.tags: key1=value1,key2=value2           # ELB resource tags
spec:
  selector:
    app: nginx
  ports:
  - name: cce-service-0
    targetPort: 80
    nodePort: 0
    port: 80
    protocol: TCP
  type: LoadBalancer
```

This example uses annotations to implement some advanced features of load balancing, such as sticky sessions and health check. For details, see **Table 3-14**.

For more annotations and examples related to advanced features, see **Using Annotations to Configure Load Balancing**.

Table 3-14 annotations parameters

Parameter	Mandatory	Type	Description
kubernetes.io/elb.class	Yes	String	Select a proper load balancer type. The value can be: ● **performance**: dedicated load balancer

Parameter	Mandatory	Type	Description
kubernetes.io/elb.autocreate	Yes	**elb.autocreate** object	Whether to automatically create a load balancer for the Service. **NOTE** Automatic creation of a public network load balancer: You need to purchase an EIP for the load balancer to allow access over the public network. The load balancer can also be accessed over a private network. Automatic creation of a private network load balancer: You do not need to purchase an EIP for the load balancer. The load balancer can only be accessed over a private network. **Example** ● If a public network load balancer will be automatically created, set this parameter to the following value: '{"type":"public","bandwidth_name":"cce-bandwidth-1551163379627","bandwidth_chargemode":"bandwidth","bandwidth_size":5,"bandwidth_sharetype":"PER","eip_type":"5_bgp","name":"james","available_zone": ["cn-east-3a"],"l4_flavor_name": "L4_flavor.elb.s1.small"}' ● If a private network load balancer will be automatically created, set this parameter to the following value: '{"type":"inner","available_zone": ["cn-east-3a"],"l4_flavor_name": "L4_flavor.elb.s1.small"}'
kubernetes.io/elb.subnet-id	-	String	ID of the subnet where the cluster is located. The value can contain 1 to 100 characters. ● Mandatory when a cluster of v1.11.7-r0 or earlier is to be automatically created. ● Optional for clusters later than v1.11.7-r0. For details about how to obtain the value, see **What Is the Difference Between the VPC Subnet API and the OpenStack Neutron Subnet API?**

Parameter	Mandatory	Type	Description
kubernetes.io/elb.enterpriseID	No	String	This parameter indicates the ID of the enterprise project in which the ELB load balancer will be created. If this parameter is not specified or is set to **0**, resources will be bound to the default enterprise project. **How to obtain**: Log in to the management console and choose **Enterprise** > **Project Management** on the top menu bar. In the list displayed, click the name of the target enterprise project and copy the ID on the enterprise project details page.
kubernetes.io/elb.lb-algorithm	No	String	Specifies the load balancing algorithm of the backend server group. The default value is **ROUND_ROBIN**. Options: ● **ROUND_ROBIN**: weighted round robin algorithm ● **LEAST_CONNECTIONS**: weighted least connections algorithm ● **SOURCE_IP**: source IP hash algorithm **NOTE** If this parameter is set to **SOURCE_IP**, the weight setting (**weight** field) of backend servers bound to the backend server group is invalid, and sticky session cannot be enabled.
kubernetes.io/elb.session-affinity-mode	No	String	In source IP address-based sticky sessions, requests from the same IP address are forwarded to the same backend server. ● Disabling sticky session: Do not configure this parameter. ● Enabling sticky session: Set this parameter to **SOURCE_IP**, indicating that the sticky session is based on the source IP address. **NOTE** When **kubernetes.io/elb.lb-algorithm** is set to **SOURCE_IP** (source IP hash), sticky session cannot be enabled.
kubernetes.io/elb.session-affinity-option	No	Table 3-12 Object	Sticky session timeout.

Parameter	Mandatory	Type	Description
kubernetes.io/elb.health-check-flag	No	String	Whether to enable the ELB health check. ● Enabling health check: Leave blank this parameter or set it to **on**. ● Disabling health check: Set this parameter to **off**. If this option is enabled, the **kubernetes.io/elb.health-check-option** field must also be specified at the same time.
kubernetes.io/elb.health-check-option	No	**Table 3-13** Object	ELB health check configuration items.
kubernetes.io/elb.tags	No	String	Add resource tags to a load balancer. This parameter can be configured only when a load balancer is automatically created. A tag is in the format of "key=value". Use commas (,) to separate multiple tags.

Table 3-15 elb.autocreate data structure

Parameter	Mandatory	Type	Description
name	No	String	Name of the automatically created load balancer. The value can contain 1 to 64 characters. Only letters, digits, underscores (_), hyphens (-), and periods (.) are allowed. Default: **cce-lb+service.UID**
type	No	String	Network type of the load balancer. ● **public**: public network load balancer ● **inner**: private network load balancer Default: **inner**
bandwidth_name	Yes for public network load balancers	String	Bandwidth name. The default value is **cce-bandwidth-*******. The value can contain 1 to 64 characters. Only letters, digits, underscores (_), hyphens (-), and periods (.) are allowed.

Parameter	Mandatory	Type	Description
bandwidth_chargemode	No	String	Bandwidthbilling option. ● **bandwidth**: billed by bandwidth ● **traffic**: billed by traffic Default: **bandwidth**
bandwidth_size	Yes for public network load balancers	Integer	Bandwidth size. The default value is 1 to 2,000 Mbit/s. Configure this parameter based on the bandwidth range allowed in your region. The minimum increment for bandwidth adjustment varies depending on the bandwidth range. ● The minimum increment is 1 Mbit/s if the allowed bandwidth does not exceed 300 Mbit/s. ● The minimum increment is 50 Mbit/s if the allowed bandwidth ranges from 300 Mbit/s to 1,000 Mbit/s. ● The minimum increment is 500 Mbit/s if the allowed bandwidth exceeds 1,000 Mbit/s.
bandwidth_sharetype	Yes for public network load balancers	String	Bandwidth sharing mode. ● **PER**: dedicated bandwidth
eip_type	Yes for public network load balancers	String	EIP type. ● **5_telcom**: China Telecom ● **5_union**: China Unicom ● **5_bgp**: dynamic BGP ● **5_sbgp**: static BGP
vip_subnet_cidr_id	No	String	Specifies the subnet where a load balancer is located. The subnet must belong to the VPC where the cluster resides. If this parameter is not specified, the ELB load balancer and the cluster are in the same subnet.

Parameter	Mandatory	Type	Description
vip_address	No	String	Specifies the private IP address of the load balancer. Only IPv4 addresses are supported. The IP address must be in the CIDR block of the load balancer. If this parameter is not specified, an IP address will be automatically assigned from the CIDR block of the load balancer.
available_zone	Yes	Array of strings	AZ where the load balancer is located. You can obtain all supported AZs by **querying the AZ list**. This parameter is available only for dedicated load balancers.
l4_flavor_name	Yes	String	Flavor name of the Layer-4 load balancer. You can obtain all supported types by **querying the flavor list**. This parameter is available only for dedicated load balancers.
l7_flavor_name	No	String	Flavor name of the Layer-7 load balancer. You can obtain all supported types by **querying the flavor list**. This parameter is available only for dedicated load balancers. The value of this parameter must be the same as that of **l4_flavor_name**, that is, both are elastic specifications or fixed specifications.

elb_virsubnet_ids	No	Array of strings	Subnet where the backend server of the load balancer is located. If this parameter is left blank, the default cluster subnet is used. Load balancers occupy different number of subnet IP addresses based on their specifications. Do not use the subnet CIDR blocks of other resources (such as clusters) as the load balancer CIDR block.
			This parameter is available only for dedicated load balancers.
			Example:
			`"elb_virsubnet_ids": [`
			`  "14567f27-8ae4-42b8-ae47-9f847a4690dd"`
			`]`

Step 3 Create a workload.

kubectl create -f nginx-deployment.yaml

If information similar to the following is displayed, the workload is being created.

```
deployment/nginx created
```

kubectl get pod

If information similar to the following is displayed, the workload is running.

```
NAME                    READY     STATUS        RESTARTS   AGE
nginx-2601814895-c1xhw   1/1       Running          0         6s
```

Step 4 Create a Service.

kubectl create -f nginx-elb-svc.yaml

If information similar to the following is displayed, the Service has been created.

```
service/nginx created
```

kubectl get svc

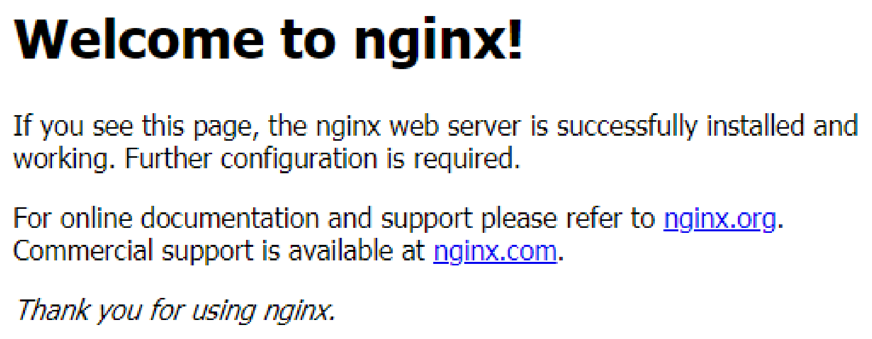
If information similar to the following is displayed, the access type has been set, and the workload is accessible.

```
NAME         TYPE           CLUSTER-IP       EXTERNAL-IP    PORT(S)         AGE
kubernetes   ClusterIP      10.247.0.1       <none>         443/TCP         3d
nginx        LoadBalancer   10.247.130.196   10.78.42.242   80:31540/TCP    51s
```

Step 5 Enter the URL in the address box of the browser, for example, **10.78.42.242:80**. **10.78.42.242** indicates the IP address of the load balancer, and **80** indicates the access port displayed on the CCE console.

Nginx is accessible.

Figure 3-19 Accessing Nginx through the LoadBalancer Service



Welcome to nginx!

If you see this page, the nginx web server is successfully installed and working. Further configuration is required.

For online documentation and support please refer to nginx.org.
Commercial support is available at nginx.com.

Thank you for using nginx.

----End

3.3.2.2 Configuring Security Group Rules for ICMP Traffic

Scenario

Dedicated load balancers are used for CCE Autopilot clusters. If both the listener protocol and the health check protocol are UDP, you need to allow ICMP traffic from the backend subnet of the load balancer in the security group associated with the elastic network interfaces.

Procedure

Step 1 Log in to the CCE console, choose **Service List** > **Networking** > **Virtual Private Cloud**. In the navigation pane on the left, choose **Access Control** > **Security Groups**.

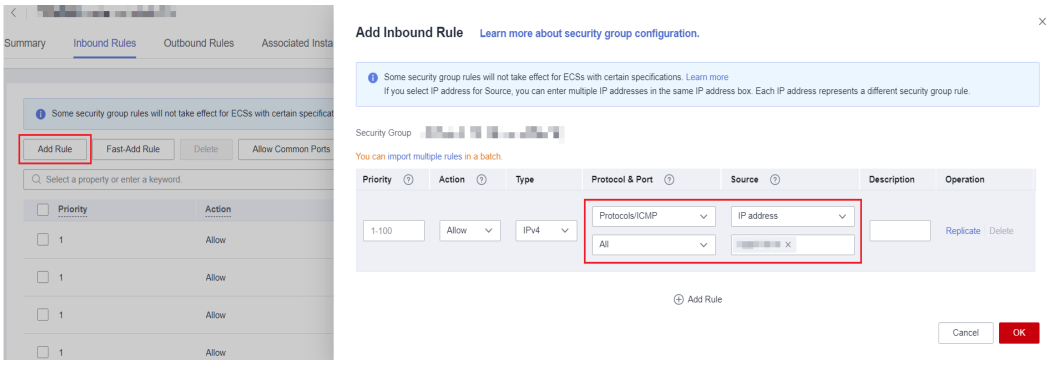
Step 2 In the security group list, locate the security group associated with the elastic network interfaces. The default security group is named *{Cluster name}-cce-eni-{Random ID}*.

If you specify a custom security group for the cluster, select this security group.

Step 3 Click the name of the security group. On the **Inbound Rules** tab, click **Add Rule** to add an inbound rule. For details, see **Figure 3-20**.

- **Protocol & Port**: Select all ICMP ports.
- **Source**: Enter the backend subnet of the load balancer.

Figure 3-20 Adding a security group rule



[redacted]

Summary | Inbound Rules | Outbound Rules | Associated Insta

Some security group rules will not take effect for ECSs with certain specificat

Add Rule | Fast-Add Rule | Delete | Allow Common Ports

Select a property or enter a keyword.

Priority	Action
1	Allow
1	Allow
1	Allow
1	Allow

Add Inbound Rule Learn more about security group configuration.

Some security group rules will not take effect for ECSs with certain specifications. Learn more
If you select IP address for Source, you can enter multiple IP addresses in the same IP address box. Each IP address represents a different security group rule.

Security Group [redacted]

You can import multiple rules in a batch.

Priority	Action	Type	Protocol & Port	Source	Description	Operation
1-100	Allow	IPv4	Protocols/ICMP All	IP address [redacted]		Replicate Delete

Add Rule

Cancel OK

Step 4 Click **OK**.

----End

3.3.3 Headless Service

The preceding types of Services allow internal and external pod access, but not the following scenarios:

- Accessing all pods at the same time
- Pods in a Service accessing each other

This is where headless Service come into service. A headless Service does not create a cluster IP address, and the DNS records of all pods are returned during query. In this way, the IP addresses of all pods can be queried. StatefulSets use headless Services to support mutual access between pods.

```
apiVersion: v1
kind: Service      # Object type (Service)
metadata:
  name: nginx-headless
  labels:
    app: nginx
spec:
  ports:
    - name: nginx     #     - name: nginx     # Name of the port for communication between pods
      port: 80        # Port number for communication between pods
```

```
 selector:
   app: nginx       # Select the pod whose label is app:nginx.
 clusterIP: None     # Set this parameter to None, indicating that a headless Service is to be created.
```

Run the following command to create a headless Service:

```
# kubectl create -f headless.yaml
service/nginx-headless created
```

After the Service is created, you can query the Service.

```
# kubectl get svc
NAME            TYPE        CLUSTER-IP   EXTERNAL-IP   PORT(S)   AGE
nginx-headless   ClusterIP   None         <none>        80/TCP    5s
```

Create a pod to query the DNS. You can view the records of all pods. In this way, all pods can be accessed.

```
$ kubectl run -i --tty --image tutum/dnsutils dnsutils --restart=Never --rm /bin/sh
If you do not see a command prompt, try pressing Enter.
/ # nslookup nginx-0.nginx
Server:         10.247.3.10
Address:        10.247.3.10#53
Name:   nginx-0.nginx.default.svc.cluster.local
Address: 172.16.0.31

/ # nslookup nginx-1.nginx
Server:         10.247.3.10
Address:        10.247.3.10#53
Name:   nginx-1.nginx.default.svc.cluster.local
Address: 172.16.0.18

/ # nslookup nginx-2.nginx
Server:         10.247.3.10
Address:        10.247.3.10#53
Name:   nginx-2.nginx.default.svc.cluster.local
Address: 172.16.0.19
```

3.4 Ingresses

3.4.1 LoadBalancer Ingresses

3.4.1.1 Creating a LoadBalancer Ingress on the Console

Prerequisites

- A workload is available in the cluster (because an Ingress enables network access for workloads). If no workload is available, deploy a workload by referring to **Creating a Workload**.
- A Service has been configured for the workload.

Precautions

- Other resources should not use the load balancer automatically created by an Ingress. If the load balancer is used by other resources, there will be residual resources when the Ingress is deleted.
- After an Ingress is created, you can only upgrade and maintain the configuration of each load balancer on the CCE console. Do not modify the

configuration on the ELB console. If you modify the configuration on the ELB console, the Ingress may be abnormal.

- The URL specified in an Ingress forwarding policy must be the same as that used to access the backend Service. If the URL is not the same, 404 will be returned.
- Dedicated load balancers must be of the application type (HTTP/HTTPS) and each have a private IP address bound.
- If multiple Ingresses are used to connect to the same load balancer port in the same cluster, the listener configuration (such as the certificate associated with the listener and HTTP/2) for the load balancer of the first Ingress is used to configure the listeners of load balancers for other Ingresses.

Adding a LoadBalancer Ingress

An Nginx workload is used as an example to describe how to add a LoadBalancer Ingress.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Services & Ingresses**. On the **Ingresses** tab, click **Create Ingress** in the upper right corner.

Step 3 Configure the parameters.

- **Name**: Enter a name for the Ingress, for example, **ingress-demo**.
- **Load Balancer**: Select the load balancer type and whether to use an existing load balancer or create one.

  Only dedicated load balancers are allowed. A dedicated load balancer must be of the application (HTTP/HTTPS) type and work on a private network.

  Select either **Use existing** or **Auto create**. For more information, see **Table 3-16**.

Table 3-16 Load balancer configurations

Option	Description
Use existing	Only the load balancers in the same VPC as the cluster can be selected. If no load balancer is available, click **Create Load Balancer** to create one on the ELB console.

Option	Description
Auto create	– **Instance Name**: Enter a load balancer name. – **Enterprise Project**: This parameter is only available for enterprise users who have enabled an enterprise project. Enterprise projects facilitate project-level management and grouping of cloud resources and users. – **AZ**: This parameter is only available for dedicated load balancers. You can create load balancers in multiple AZs to improve service availability. If diaster recovery is required, you are advised to select multiple AZs. – **Frontend Subnet**: This parameter is used to allocate IP addresses to load balancers to receive traffic from clients. – **Backend Subnet**: This parameter is used to allocate IP addresses for load balancers to routing traffic to pods. – **Network/Application-oriented Specifications** ▪ **Elastic**: applies to fluctuating traffic, billed based on the total traffic. ▪ **Fixed**: applies to stable traffic, billed based on specifications. – **EIP**: If you select **Auto create**, you can select a bandwidth billing option and set the bandwidth. – **Resource Tag**: You can add resource tags to classify resources. You can create predefined tags on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. This is supported by clusters of v1.27.5-r0, v1.28.3-r0, and later versions.

- **Listener**: An Ingress configures a listener for the load balancer, and this listener listens to and distributes requests. After the configuration is complete, a listener will be created on the load balancer. The default listener name is in the format of k8s__<Protocol>_<Port>, for example, *k8s_HTTP_80*.
  - **External Protocol**: **HTTP** and **HTTPS** are available.
  - **External Port**: Port for the load balancer to receive requests. You can specify any port.
  - **Access Control**
    - **Inherit ELB Configurations**: Use the existing access control settings configured on ELB.
    - **Allow all IP addresses**: No access control is configured.
    - **Trustlist**: Only the selected IP address group can access the load balancer.
    - **Blocklist**: The selected IP address group cannot access the load balancer.

- **Certificate Source**: TLS secrets and ELB server certificates are supported.
- **Server Certificate**: When an HTTPS listener is added to the load balancer, bind a certificate to the load balancer for encrypted transmission for HTTPS data.
  - **TLS secret**: For details about how to create a secret certificate, see **Creating a Secret**.
  - **ELB server certificate**: Use the certificate created in the ELB service.

  NOTE

  If there is already an HTTPS Ingress for the chosen port on the load balancer, the certificate of the new HTTPS Ingress must be the same as the certificate of the existing Ingress. This means that a listener has only one certificate. If two certificates, each with a different Ingress, are added to the same listener of the same load balancer, only the certificate added earliest takes effect on the load balancer.
- **SNI**: Server Name Indication (SNI) is an extended protocol of TLS. It allows multiple TLS-compliant domain names for external access using the same IP address and port, and different domain names can use different security certificates. If SNI is enabled, the client is allowed to submit the requested domain name when initiating a TLS handshake request. After receiving the TLS request, the load balancer searches for the certificate based on the domain name in the request. If the certificate corresponding to the domain name is found, the load balancer returns the certificate for authorization. Otherwise, the default certificate (server certificate) is returned for authorization.

  NOTE

  - SNI is available only when **HTTPS** is selected.
  - You need to specify the domain name for the SNI certificate. Only one domain name can be specified for each certificate. Wildcard-domain certificates are supported.
- **Advanced Options**

Description	Description
Transfer Listener Port Number	If this option is enabled, the listening port on the load balancer can be transferred to backend servers through the HTTP header of the packet.
Transfer Port Number in the Request	If this option is enabled, the source port of the client can be transferred to backend servers through the HTTP header of the packet.
Rewrite X-Forwarded-Host	If this option is enabled, **X-Forwarded-Host** will be rewritten using the **Host** field in the request and transferred to backend servers.

Description	Description
Idle Timeout	Timeout for an idle client connection. If there are no requests reaching the load balancer during the timeout duration, the load balancer will disconnect the connection from the client and establish a new connection when there is a new request.
Request Timeout	Timeout for waiting for a request from a client. There are two cases: ■ If the client fails to send a request header to the load balancer during the timeout duration, the request will be interrupted. ■ If the interval between two consecutive request bodies reaching the load balancer is greater than the timeout duration, the connection will be disconnected.
Response Timeout	Timeout for waiting for a response from a backend server. After a request is forwarded to the backend server, if the backend server does not respond during the timeout duration, the load balancer will stop waiting and return HTTP 504 Gateway Timeout.
HTTP2	Whether to use HTTP/2 for a client to communicate with a load balancer. Request forwarding using HTTP/2 improves the access performance between your application and the load balancer. However, the load balancer still uses HTTP/1.x to forward requests to the backend server.

- **Forwarding Policy**: When the access address of a request matches the forwarding policy (that consists of a domain name and URL, for example, 10.117.117.117:80/helloworld), the request is forwarded to the corresponding target Service for processing. You can click + to add multiple forwarding policies.
  - **Domain Name**: Enter an actual domain name to be accessed. If it is left blank, the Ingress can be accessed through the IP address. Ensure that the domain name has been registered and licensed. Once the domain name is specified, it must be used for access.
  - **Path Matching Rule**
    - **Prefix match**: If the URL is set to **/healthz**, the URL that meets **/healthz** can be accessed, for example, **/healthz/v1** and **/healthz/v2**.
    - **Exact match**: The URL can be accessed only when it is fully matched. For example, if the URL is set to **/healthz**, only **/healthz** can be accessed.

        - **RegEX match**: The URL is matched based on the regular expression. If the regular expression is **/[A-Za-z0-9_.-]+/test**, all URLs that comply with this rule can be accessed, for example, **/abcA9/test** and **/v1-Ab/test**. Two regular expression standards are supported: POSIX and Perl.
    - **Path**: access path, for example, **/healthz**.

      NOTE

      The access path added here must exist in the backend applications. If it does not exist, requests will fail to be forwarded.

      For example, the default access URL of the Nginx application is **/usr/share/nginx/html**. When adding **/test** to the Ingress forwarding policy, ensure the access URL of your Nginx application contains **/usr/share/nginx/html/test**. Otherwise, 404 will be returned.
    - **Destination Service**: Select an existing Service or create a Service. Services that do not meet search criteria are automatically filtered out.
    - **Destination Service Port**: Select the access port of the destination Service.
    - **Set ELB**:
        - **Algorithm**: Three algorithms are available: weighted round robin, weighted least connections algorithm, or source IP hash.

          NOTE

            - **Weighted round robin**: Requests are forwarded to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests. This algorithm is often used for short connections, such as HTTP services.
            - **Weighted least connections**: In addition to the weight assigned to each server, the number of connections processed by each backend server is considered. Requests are forwarded to the server with the lowest connections-to-weight ratio. Building on **least connections**, the **weighted least connections** algorithm assigns a weight to each server based on their processing capability. This algorithm is often used for persistent connections, such as database connections.
            - **Source IP hash**: The source IP address of each request is calculated using the hash algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key allocates the client to a particular server. This enables requests from different clients to be distributed in load balancing mode and ensures that requests from the same client are forwarded to the same server. This algorithm applies to TCP connections without cookies.
        - **Type**: This feature is disabled by default. There are two options:
            - **Load balancer cookie**: Enter the **Stickiness Duration** , which ranges from **1** to **1440** minutes.

          NOTE

          When **Source IP hash** is used for load balancing, sticky sessions are not available.

- **Health Check**: Set the health check configuration of the load balancer. If this feature is enabled, you need to configure the following parameters.

Parameter	Description
Protocol	When the protocol of the target Service is TCP, gRPC, and HTTP are supported. ○ **Check Path**: This parameter is only available for HTTP and GRPC health check. It specifies the URL for health check. The check path must start with a slash (/) and contain 1 to 80 characters.
Port	By default, the service port (container port of the Service) is used for health check. You can also specify another port for health check. If a port is specified, a service port named **cce-healthz** will be added for the Service. ○ **Node Port**: If this parameter is not specified, a random port is used. The value ranges from **30000** to **32767**. ○ **Container Port**: When a dedicated load balancer has an elastic network interface associated, the container port is used for health check. The value ranges from **1** to **65535**.
Check Period (s)	Specifies the maximum interval between health checks. The value ranges from **1** to **50**.
Timeout (s)	Specifies the maximum timeout duration for each health check. The value ranges from **1** to **50**.
Max. Retries	Specifies the maximum number of health check retries. The value ranges from **1** to **10**.

– **Operation**: Click **Delete** to delete the configuration.

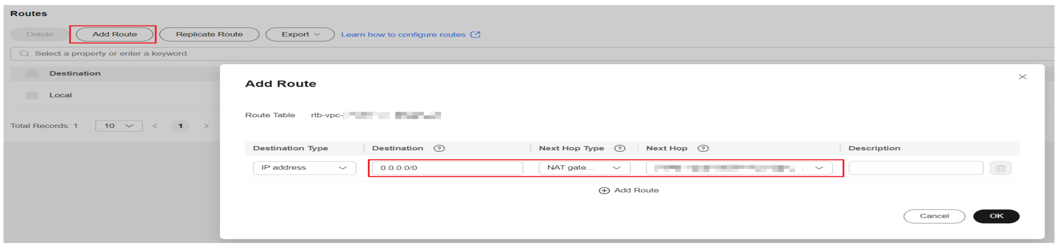
- **Annotation**: Ingresses provide some advanced CCE functions, which are implemented by annotations. When you use kubectl to create a container, annotations will be used. For details, see **Creating an Ingress - Automatically Creating a Load Balancer** and **Creating an Ingress - Interconnecting with an Existing Load Balancer**.

Step 4 Click **OK**. After the Ingress is created, it is displayed in the Ingress list.

On the ELB console, you can view the ELB automatically created through CCE. The default name is **cce-lb-ingress.UID**. Click the load balancer name to access its details page. On the **Listeners** tab, view the route settings of the Ingress, such as the URL, listener port, and backend port.

NOTICE

After an Ingress is created, upgrade and maintain the selected load balancer on the CCE console. Do not modify the configuration on the ELB console, or the Ingress may be abnormal.



Route Table rtb-vpc-

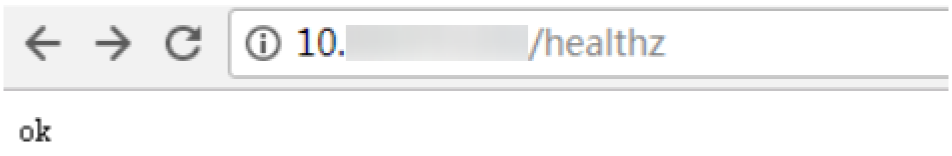
Destination Type	Destination	Next Hop Type	Next Hop	Description
IP address	0.0.0.0/0	NAT gate...		

Figure 3-21 ELB routing configuration

Step 5 Access the /healthz interface of the workload, for example, workload **defaultbackend**.

1. Obtain the access address of the **/healthz** interface of the workload. The access address consists of the load balancer IP address, external port, and mapping URL, for example, 10.**.**.**:80/healthz.
2. Enter the URL of the /healthz interface, for example, http://10.**.**.**:80/healthz, in the address box of the browser to access the workload.

Figure 3-22 Accessing the /healthz interface of defaultbackend



```
10.          /healthz
```

```
ok
```

----End

Related Operations

The Kubernetes Ingress structure does not contain the **property** field. Therefore, the Ingress created by the API called by client-go does not contain the **property** field. CCE provides a solution to ensure compatibility with the Kubernetes client-go. For details about the solution, see **How Can I Achieve Compatibility Between Ingress's property and Kubernetes client-go?**

3.4.1.2 Using kubectl to Create a LoadBalancer Ingress

Scenario

An Nginx workload is used as an example to describe how to create a LoadBalancer Ingress using kubectl.

- If no load balancer is available in the same VPC, CCE can automatically create a load balancer when creating an Ingress. For details, see **Creating an Ingress - Automatically Creating a Load Balancer**.
- If a load balancer is available in the same VPC, perform the operation by referring to **Creating an Ingress - Interconnecting with an Existing Load Balancer**.

Prerequisites

- A workload is available in the cluster (because an Ingress enables network access for workloads). If no workload is available, deploy an Nginx workload by referring to **Creating a Workload**.
- A Service has been configured for the workload.
- Dedicated load balancers must be of the application type (HTTP/HTTPS) and each have a private IP address bound.

Creating an Ingress - Automatically Creating a Load Balancer

The following describes how to run the kubectl command to automatically create a load balancer when creating an Ingress.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a YAML file named **ingress-test.yaml**. The file name can be customized.

vi ingress-test.yaml

Example of using a dedicated load balancer on a public network:

```
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: ingress-test
  namespace: default
  annotations:
    kubernetes.io/elb.class: performance
    kubernetes.io/elb.port: '80'
    kubernetes.io/elb.autocreate:
      '{
          "type": "public",
          "bandwidth_name": "cce-bandwidth-******",
          "bandwidth_chargemode": "bandwidth",
          "bandwidth_size": 5,
          "bandwidth_sharetype": "PER",
          "eip_type": "5_bgp",
          "elb_virsubnet_ids":["*****"],
          "available_zone": [
              "ap-southeast-1a"
          ],
          "l7_flavor_name": "L7_flavor.elb.s1.small"
       }'
spec:
  rules:
  - host: ''
    http:
      paths:
      - path: '/'
        backend:
          service:
            name: <your_service_name>  # Replace it with the name of your target Service.
            port:
              number: <your_service_port>  # Replace it with the port number of your target Service.
        property:
          ingress.beta.kubernetes.io/url-match-mode: STARTS_WITH
        pathType: ImplementationSpecific
  ingressClassName: cce
```

Table 3-17 Key parameters

Parameter	Mandatory	Type	Description
kubernetes.io/elb.class	Yes	String	The value can be: ● **performance**: dedicated load balancer
ingressClassName	Yes	String	**cce**: A LoadBalancer Ingress is used. This parameter is mandatory when an Ingress is created by calling the API.
kubernetes.io/elb.port	Yes	Integer	This parameter indicates the external port registered with the address of the LoadBalancer Service. Supported range: 1 to 65535 **NOTE** Some ports are high-risk ports and are blocked by default, for example, port 21.
kubernetes.io/elb.subnet-id	No	String	ID of the subnet where the cluster is located. The value can contain 1 to 100 characters. For details about how to obtain the value, see **What Is the Difference Between the VPC Subnet API and the OpenStack Neutron Subnet API?**
kubernetes.io/elb.enterpriseID	No	String	ID of the enterprise project in which the load balancer will be created. The value contains 1 to 100 characters. **How to obtain**: Log in to the management console and choose **Enterprise** > **Project Management** on the top menu bar. In the list displayed, click the name of the target enterprise project and copy the ID on the enterprise project details page.

Parameter	Mandatory	Type	Description
kubernetes.io/elb.autocreate	Yes	elb.autocreate object	Whether to automatically create a load balancer associated with an Ingress. For details about the field description, see **Table 3-18**. **Example** ● If a public network load balancer will be automatically created, set this parameter to the following value: {"type":"public","bandwidth_name":"cce-bandwidth-******","bandwidth_chargemode":"bandwidth","bandwidth_size":5,"bandwidth_sharetype":"PER","eip_type":"5_bgp","name":"james"} ● If a private network load balancer will be automatically created, set this parameter to the following value: {"type":"inner","name":"A-location-d-test"}
kubernetes.io/elb.tags	No	String	Add resource tags to a load balancer. This parameter can be configured only when a load balancer is automatically created. A tag is in the format of "key=value". Use commas (,) to separate multiple tags.
host	No	String	Domain name for accessing the Service. By default, this parameter is left blank, and the domain name needs to be fully matched. Ensure that the domain name has been registered and licensed. Once the domain name is specified, it must be used for access.
path	Yes	String	User-defined route path. All external access requests must match **host** and **path**. **NOTE** The access path added here must exist in the backend application. Otherwise, the forwarding fails. For example, the default access URL of the Nginx application is **/usr/share/nginx/html**. When adding **/test** to the ingress forwarding policy, ensure the access URL of your Nginx application contains **/usr/share/nginx/html/test**. Otherwise, error 404 will be returned.

Parameter	Mandatory	Type	Description
ingress.beta.kubernetes.io/url-match-mode	No	String	Route matching policy. Default: **STARTS_WITH** (prefix match) Options: ● **EQUAL_TO**: exact match ● **STARTS_WITH**: prefix match ● **REGEX**: regular expression match
pathType	Yes	String	Path type. The options are as follows: ● **ImplementationSpecific**: The matching method depends on Ingress Controller. The matching method defined by **ingress.beta.kubernetes.io/url-match-mode** is used in CCE. ● **Exact**: exact matching of the URL, which is case-sensitive. ● **Prefix**: prefix matching, which is case-sensitive. With this method, the URL path is separated into multiple elements by slashes (/) and the elements are matched one by one. If each element in the URL matches the path, the subpaths of the URL can be routed normally. **NOTE** – During prefix matching, each element must be exactly matched. If the last element of the URL is the substring of the last element in the request path, no matching is performed. For example, **/foo/bar** matches **/foo/bar/baz** but does not match **/foo/barbaz**. – When elements are separated by slashes (/), if the URL or request path ends with a slash (/), the slash (/) at the end is ignored. For example, **/foo/bar** matches **/foo/bar/**. See **examples** of Ingress path matching.

Table 3-18 elb.autocreate data structure

Parameter	Mandatory	Type	Description
name	No	String	Name of the automatically created load balancer. The value can contain 1 to 64 characters. Only letters, digits, underscores (_), hyphens (-), and periods (.) are allowed. Default: **cce-lb+service.UID**
type	No	String	Network type of the load balancer. ● **public**: public network load balancer ● **inner**: private network load balancer Default: **inner**
bandwidth_name	Yes for public network load balancers	String	Bandwidth name. The default value is **cce-bandwidth-*******. The value can contain 1 to 64 characters. Only letters, digits, underscores (_), hyphens (-), and periods (.) are allowed.
bandwidth_chargemode	No	String	Bandwidthbilling option. ● **bandwidth**: billed by bandwidth ● **traffic**: billed by traffic Default: **bandwidth**
bandwidth_size	Yes for public network load balancers	Integer	Bandwidth size. The default value is 1 to 2,000 Mbit/s. Configure this parameter based on the bandwidth range allowed in your region. The minimum increment for bandwidth adjustment varies depending on the bandwidth range. ● The minimum increment is 1 Mbit/s if the allowed bandwidth does not exceed 300 Mbit/s. ● The minimum increment is 50 Mbit/s if the allowed bandwidth ranges from 300 Mbit/s to 1,000 Mbit/s. ● The minimum increment is 500 Mbit/s if the allowed bandwidth exceeds 1,000 Mbit/s.

Parameter	Mandatory	Type	Description
bandwidth_sharetype	Yes for public network load balancers	String	Bandwidth sharing mode. ● **PER**: dedicated bandwidth
eip_type	Yes for public network load balancers	String	EIP type. ● **5_telcom**: China Telecom ● **5_union**: China Unicom ● **5_bgp**: dynamic BGP ● **5_sbgp**: static BGP
vip_subnet_cidr_id	No	String	Specifies the subnet where a load balancer is located. The subnet must belong to the VPC where the cluster resides. If this parameter is not specified, the ELB load balancer and the cluster are in the same subnet.
vip_address	No	String	Specifies the private IP address of the load balancer. Only IPv4 addresses are supported. The IP address must be in the CIDR block of the load balancer. If this parameter is not specified, an IP address will be automatically assigned from the CIDR block of the load balancer.
available_zone	Yes	Array of strings	AZ where the load balancer is located. You can obtain all supported AZs by **querying the AZ list**. This parameter is available only for dedicated load balancers.
l4_flavor_name	Yes	String	Flavor name of the Layer-4 load balancer. You can obtain all supported types by **querying the flavor list**. This parameter is available only for dedicated load balancers.

Parameter	Mandatory	Type	Description
l7_flavor_name	No	String	Flavor name of the Layer-7 load balancer. You can obtain all supported types by **querying the flavor list**. This parameter is available only for dedicated load balancers. The value of this parameter must be the same as that of **l4_flavor_name**, that is, both are elastic specifications or fixed specifications.

elb_virsubnet_ids	No	Array of strings	Subnet where the backend server of the load balancer is located. If this parameter is left blank, the default cluster subnet is used. Load balancers occupy different number of subnet IP addresses based on their specifications. Do not use the subnet CIDR blocks of other resources (such as clusters) as the load balancer CIDR block.
			This parameter is available only for dedicated load balancers.
			Example:
			`"elb_virsubnet_ids": [`
			`  "14567f27-8ae4-42b8-ae47-9f847a4690dd"`
			` ]`

Step 3 Create an Ingress.

kubectl create -f ingress-test.yaml

If information similar to the following is displayed, the Ingress has been created.

```
ingress/ingress-test created
```

kubectl get ingress

If information similar to the following is displayed, the Ingress has been created, and the workload is accessible.

```
NAME           HOSTS     ADDRESS          PORTS   AGE
ingress-test    *        121.**.**.**     80      10s
```

Step 4 Enter **http://*121.**.**.***:80** in the address box of the browser to access the workload.

*121.**.**.*** indicates the IP address of the unified load balancer.

----End

Creating an Ingress - Interconnecting with an Existing Load Balancer

CCE allows you to connect to an existing load balancer when creating an Ingress.

NOTE

- Dedicated load balancers must be of the application type (HTTP/HTTPS) and each have a private IP address bound.

The YAML file is configured as follows:

```
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: ingress-test
  annotations:
    kubernetes.io/elb.id: <your_elb_id>  # Replace it with the ID of your existing load balancer.
    kubernetes.io/elb.ip: <your_elb_ip>  # Replace it with the IP of your existing load balancer.
    kubernetes.io/elb.class: performance  # Load balancer type
    kubernetes.io/elb.port: '80'
spec:
  rules:
  - host: ''
    http:
      paths:
      - path: '/'
        backend:
          service:
            name: <your_service_name>  # Replace it with the name of your target Service.
            port:
              number: 8080             # Replace 8080 with the port number of your target Service.
        property:
          ingress.beta.kubernetes.io/url-match-mode: STARTS_WITH
        pathType: ImplementationSpecific
  ingressClassName: cce
```

Table 3-19 Key parameters

Parameter	Mandatory	Type	Description
kubernetes.io/elb.id	Yes	String	Load balancer ID. The value can contain 1 to 100 characters. **How to obtain**: On the management console, click **Service List**, and choose **Networking** > **Elastic Load Balance**. Click the name of the load balancer. On the **Summary** tab, find and copy the ID.
kubernetes.io/elb.ip	No	String	Service address of a load balancer. The value can be the public IP address of a public network load balancer or the private IP address of a private network load balancer.

Parameter	Mandatory	Type	Description
kubernetes.io/elb.class	Yes	String	Load balancer type. ● **performance**: dedicated load balancer **NOTE** If a LoadBalancer Ingress accesses an existing dedicated load balancer, the dedicated load balancer must be of the application load balancing (HTTP/HTTPS) type.

3.4.2 Nginx Ingresses

3.4.2.1 Creating an Nginx Ingress on the Console

Prerequisites

- A workload is available in the cluster (because an Ingress enables network access for workloads). If no workload is available, deploy a workload by referring to **Creating a Workload**.
- A ClusterIP Service has been configured for the workload. For details about how to configure the Service, see **ClusterIP**.
- To add an Nginx Ingress, ensure that the NGINX Ingress Controller add-on has been installed in the cluster. For details, see **Installing the Add-on**.

Constraints

- It is not recommended that you modify any configuration of a load balancer on the ELB console. If you modify the configuration on the ELB console, the Service will be abnormal. If you have modified the configuration, uninstall NGINX Ingress Controller and reinstall it.
- The URL specified in an Ingress forwarding policy must be the same as that used to access the backend Service. If the URL is not the same, 404 will be returned.
- The selected or created load balancer must be in the same VPC as the current cluster, and it must match the load balancer type (private or public network).
- The load balancer has at least two listeners, and ports 80 and 443 are not occupied by listeners.

Procedure

An Nginx workload is used as an example to describe how to create an Nginx Ingress.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Services & Ingresses**. On the **Ingresses** tab, click **Create Ingress** in the upper right corner.

Step 3 Configure parameters.

- **Name**: Enter a name for the Ingress, for example, **nginx-ingress-demo**.
- **Namespace**: Select the namespace to which the Ingress is to be added.
- **nginx-ingress**: This option is displayed only after the **NGINX Ingress Controller** add-on is installed in the cluster.
  - **External Protocol**: The options are **HTTP** and **HTTPS**. If NGINX Ingress Controller is installed, the default port is 80 for HTTP and 443 for HTTPS. To use HTTPS, configure a server certificate.
  - **Certificate Source**: source of a certificate for encrypting and authenticating HTTPS data transmission.
    - If you select a TLS key, you must create a key certificate of the IngressTLS or kubernetes.io/tls type beforehand. For details, see **Creating a Secret**.
    - If you select the default certificate, NGINX Ingress Controller will use its default certificate for encryption and authentication.
  - **SNI**: SNI an extended protocol of TLS. SNI allows multiple TLS-compliant domain names for external access using the same IP address and port, and different domain names can use different security certificates. If SNI is enabled, the client is allowed to submit the requested domain name when initiating a TLS handshake request. After receiving the TLS request, the load balancer searches for the certificate based on the domain name in the request. If the certificate corresponding to the domain name is found, the load balancer returns the certificate for authorization. Otherwise, the default certificate (server certificate) is returned for authorization.
- **Forwarding Policy**: When the access address of a request matches the forwarding policy (a forwarding policy consists of a domain name and URL), the request is forwarded to the corresponding target Service for processing. Click **Add Forwarding Policies** to add multiple forwarding policies.
  - **Domain Name**: Enter the domain name used for access. Ensure that the entered domain name has been registered and archived. After the Ingress is created, bind the domain name to the IP address of the automatically created load balancer (IP address of the Ingress access address). If a domain name rule is configured, the domain name must always be used for access.
  - **Path Matching Rule**
    - **Default**: Prefix match is used by default.
    - **Prefix match**: If the URL is set to **/healthz**, the URL that meets the prefix can be accessed, for example, **/healthz/v1** and **/healthz/v2**.
    - **Exact match**: The URL can be accessed only when it is fully matched. For example, if the URL is set to **/healthz**, only **/healthz** can be accessed.
  - **Path**: access path, for example, **/healthz**.

NOTE

- The access path matching rule of the Nginx Ingress is based on the path prefix separated by the slash (/) and is case-sensitive. If the subpath separated by a slash (/) matches the prefix, the access is normal. However, if the prefix is only a part of the character string in the subpath, the access is not matched. For example, if the URL is set to /healthz, /healthz/v1 is matched, but /healthzv1 is not matched.
- The access path added here must exist in the backend applications. If it does not exist, requests will fail to be forwarded.

  For example, the default access URL of the Nginx application is **/usr/share/nginx/html**. When adding **/test** to the Ingress forwarding policy, ensure the access URL of your Nginx application contains **/usr/share/nginx/html/test**. Otherwise, 404 will be returned.

  – **Destination Service**: Select an existing Service or create a Service. Services that do not meet search criteria are automatically filtered out.
  – **Destination Service Port**: Select the access port of the destination Service.
  – **Operation**: Click **Delete** to delete the configuration.

- **Annotation**: The value is in the format of key:value. You can use **annotations** to query the configurations supported by Nginx Ingress.

Step 4 Click **OK**.

After the Ingress is created, it is displayed in the Ingress list.

----End

3.4.2.2 Using kubectl to Create an Nginx Ingress

Scenario

An Nginx workload is used as an example to describe how you can create an Nginx Ingress using kubectl.

Prerequisites

- The NGINX Ingress Controller add-on has been installed in a cluster. For details about how to install the add-on, see **Installing the Add-on**.
- A workload is available in the cluster (because an Ingress enables network access for workloads). If no workload is available, deploy a workload by referring to **Creating a Workload**.
- A ClusterIP Service has been configured for the workload. For details about how to configure the Service, see **ClusterIP**.

Procedure

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a YAML file named **ingress-test.yaml**. The file name can be customized.

vi ingress-test.yaml

The following uses HTTP as an example to describe how to configure the YAML file:

```
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: ingress-test
spec:
  rules:
    - host: ''
      http:
        paths:
          - path: /
            backend:
              service:
                name: <your_service_name>  # Replace it with the name of your target Service.
                port:
                  number: <your_service_port>  # Replace it with the port number of your target Service.
            property:
              ingress.beta.kubernetes.io/url-match-mode: STARTS_WITH
            pathType: ImplementationSpecific
  ingressClassName: nginx   # Nginx Ingress is used.
```

Table 3-20 Key parameters

Parameter	Mandatory	Type	Description
ingressClassName	Yes	String	**nginx**: indicates that Nginx Ingress is used. This option cannot be used if the NGINX Ingress Controller add-on is not installed. This parameter is mandatory when an Ingress is created by calling the API.
host	No	String	Domain name for accessing the Service. By default, this parameter is left blank, and the domain name needs to be fully matched. Ensure that the domain name has been registered and licensed. Once the domain name is specified, it must be used for access.

Parameter	Mandatory	Type	Description
path	Yes	String	User-defined route path. All external access requests must match **host** and **path**. **NOTE** ● The access path matching rule of Nginx Ingress is based on the path prefix separated by the slash (/) and is case-sensitive. If the subpath separated by a slash (/) matches the prefix, the access is normal. However, if the prefix is only a part of the character string in the subpath, the access is not matched. For example, if the URL is set to /healthz, /healthz/v1 is matched, but /healthzv1 is not matched. ● The access path added here must exist in the backend application. Otherwise, the forwarding fails. For example, the default access URL of the Nginx application is **/usr/share/nginx/html**. When adding **/test** to the ingress forwarding policy, ensure the access URL of your Nginx application contains **/usr/share/nginx/html/test**. Otherwise, error 404 will be returned.
ingress.beta.kubernetes.io/url-match-mode	No	String	Route matching policy. Default: **STARTS_WITH** (prefix match) Options: ● **EQUAL_TO**: exact match ● **STARTS_WITH**: prefix match

Parameter	Mandatory	Type	Description
pathType	Yes	String	Path type. The options are as follows: ● **ImplementationSpecific**: The matching method depends on Ingress Controller. The matching method defined by **ingress.beta.kubernetes.io/url-match-mode** is used in CCE. ● **Exact**: exact matching of the URL, which is case-sensitive. ● **Prefix**: prefix matching, which is case-sensitive. With this method, the URL path is separated into multiple elements by slashes (/) and the elements are matched one by one. If each element in the URL matches the path, the subpaths of the URL can be routed normally. **NOTE** – During prefix matching, each element must be exactly matched. If the last element of the URL is the substring of the last element in the request path, no matching is performed. For example, **/foo/bar** matches **/foo/bar/baz** but does not match **/foo/barbaz**. – When elements are separated by slashes (/), if the URL or request path ends with a slash (/), the slash (/) at the end is ignored. For example, **/foo/bar** matches **/foo/bar/**. See **examples** of ingress path matching.

Step 3 Create an Ingress.

kubectl create -f ingress-test.yaml

If information similar to the following is displayed, the Ingress has been created.

```
ingress/ingress-test created
```

View the created Ingress.

kubectl get ingress

If information similar to the following is displayed, the Ingress has been created, and the workload is accessible.

```
NAME           HOSTS     ADDRESS          PORTS   AGE
ingress-test     *         121.**.**.**     80      10s
```

Step 4 Enter **http://*121.**.**.***:80** in the address box of the browser to access the workload.

*121.**.**.*** indicates the IP address of the unified load balancer.

----End

3.5 Pod Network Settings

3.5.1 Configuring an EIP for a Pod

Scenario

In CCE Autopilot clusters, pods use elastic network interfaces or supplementary network interfaces for networking so you can directly bind EIPs to pods.

To bind an EIP to a pod, simply set the value of the **yangtse.io/pod-with-eip** annotation to **true** when creating the pod. Then, the EIP is automatically allocated and bound to the pod.

Constraints

- To access a pod with an EIP bound from the Internet, you need to add security group rules to allow the Internet traffic to the pod.
- Only one EIP can be bound to a pod.
- Configure the EIP-related annotation when creating a pod. After the pod is created, the annotations related to the EIP cannot be modified.
- Do not perform operations on the EIP associated with a pod through the EIP console or API. Otherwise, the EIP may malfunction. The operations include changing the EIP name, deleting, unbinding, or binding the EIP, and changing the billing mode of the EIP.
- After an automatically allocated EIP is manually deleted, the network malfunctions. In this case, rebuild the pod.

Allocating an EIP with a Pod

When creating a pod, set the **pod-with-eip** annotation to **true**. An EIP will be automatically allocated and bound to the pod.

The following uses a Deployment named **nginx** as an example. For details about annotations, see **Table 3-22**.

- For an automatically allocated EIP with a **dedicated bandwidth** when you create a Deployment, you do not need to specify the bandwidth ID. The following shows an example:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        yangtse.io/pod-with-eip: "true" # An EIP will be automatically allocated when the pod is 
created.
        yangtse.io/eip-bandwidth-size: "5" # EIP bandwidth
        yangtse.io/eip-network-type: 5_bgp # EIP type
        yangtse.io/eip-charge-mode: bandwidth # EIP billing mode
        yangtse.io/eip-bandwidth-name: <eip_bandwidth_name> # EIP bandwidth name
    spec:
      containers:
        - name: container-0
          image: nginx:alpine
          resources:
            limits:
              cpu: 250m
              memory: 500Mi
            requests:
              cpu: 250m
              memory: 500Mi
      imagePullSecrets:
        - name: default-secret
```

Table 3-21 Annotations of an EIP with a dedicated bandwidth

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/pod-with-eip	Yes	false	Whether to bind an EIP to a pod	**false** or **true**
yangtse.io/eip-bandwidth-size	No	5	Bandwidth, in Mbit/s	The value range varies depending on the region and bandwidth billing mode. For details, see the purchase page on the EIP console. For example, in the AP-Singapore region, if an EIP is billed by bandwidth, the bandwidth ranges from 1 Mbit/s to 2,000 Mbit/s; if an EIP is billed by traffic, the bandwidth ranges from 1 Mbit/s to 300 Mbit/s.

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/eip-network-type	No	5_bgp	EIP type	The types vary by region. For details, see the EIP console. For example, the following options are available in the AP-Singapore region: ● **5_bgp**: dynamic BGP
yangtse.io/eip-charge-mode	No	None	Billed by traffic or bandwidth **You are advised to configure this parameter.** If this parameter is left blank, no billing mode is specified. In this case, the default value of the EIP API in the region is used.	● **bandwidth**: billed by bandwidth ● **traffic**: billed by traffic
yangtse.io/eip-bandwidth-name	No	Pod name	Bandwidth name	● Enter 1 to 64 characters. Only letters, digits, hyphens (-), underscores (_), and periods (.) are allowed. ● Minimum length: 1 character ● Maximum length: 64 characters

- For an automatically allocated EIP with a **shared bandwidth** when you create a Deployment, you are required to specify the bandwidth ID. The following shows an example:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx
spec:
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        yangtse.io/pod-with-eip: "true"  # An EIP will be automatically allocated when the pod is 
created.
        yangtse.io/eip-network-type: 5_bgp  # EIP type
        yangtse.io/eip-bandwidth-id: <eip_bandwidth_id>  # Shared bandwidth ID of the EIP
    spec:
      containers:
        - name: container-0
          image: nginx:alpine
          resources:
            limits:
              cpu: 250m
              memory: 500Mi
            requests:
              cpu: 250m
              memory: 500Mi
      imagePullSecrets:
        - name: default-secret
```

Table 3-22 Annotations of an EIP with a shared bandwidth

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/pod-with-eip	Yes	false	Whether to bind an EIP to a pod	**false** or **true**
yangtse.io/eip-network-type	No	5_bgp	EIP type	• 5_bgp • 5_union • 5_sbgp The types vary by region. For details, see the EIP console.

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/eip-bandwidth-id	Mandatory when a shared bandwidth is used	None	ID of an existing bandwidth ● If this parameter is not specified, the EIP with a dedicated bandwidth is used by default. For details about parameter settings for an EIP with a dedicated bandwidth, see **Table 3-21**. ● Only the **yangtse.io/eip-network-type** field can be specified concurrently, and this field is optional.	-

Checking Whether the EIP Bound to the Pod Is Available

After an EIP is allocated to a pod, the container networking controller binds the EIP to the pod and writes the allocation result back to the pod's **yangtse.io/allocated-ipv4-eip** annotation. The startup time of the pod's service containers may be earlier than the time when the EIP allocation result is written back.

You can configure an init container for the pod, associate the **yangtse.io/allocated-ipv4-eip** annotation with the init container through a downwardAPI volume, and check whether the EIP has been allocated in the init container. You can configure the init container as follows:

```
apiVersion: v1
kind: Pod
metadata:
  name: example
  annotations:
    yangtse.io/pod-with-eip: "true"
    yangtse.io/eip-bandwidth-size: "5"
    yangtse.io/eip-network-type: 5_bgp
    yangtse.io/eip-charge-mode: bandwidth
    yangtse.io/eip-bandwidth-name: "xxx"
spec:
  initContainers:
  - name: init
    image: busybox:latest
    command: ['timeout', '60', 'sh', '-c', "until grep -E '[0-9]+' /etc/eipinfo/allocated-ipv4-eip; do echo 
waiting for allocated-ipv4-eip; sleep 2; done"]
    volumeMounts:
        - name: eipinfo
          mountPath: /etc/eipinfo
  volumes:
    - name: eipinfo
```

```
      downwardAPI:
        items:
          - path: "allocated-ipv4-eip"
            fieldRef:
              fieldPath: metadata.annotations['yangtse.io/allocated-ipv4-eip']
...
```

Deleting an EIP with a Pod

When you delete a pod, the EIP automatically allocated to the pod will also be deleted.

3.5.2 Configuring a Static EIP for a Pod

Scenario

In CCE Autopilot clusters, static public IP addresses (EIPs) can be assigned to StatefulSets or pods that are created directly.

Constraints

- The static EIP function must be enabled together with the function of automatically allocating an EIP for a pod. For details, see **Configuring an EIP for a Pod**.
- Only StatefulSet pods or pods that are created directly can have static EIPs.
- After a static EIP is allocated, the EIP attributes cannot be modified through the pod within the EIP lifecycle (before the EIP expires or it is still being used by the pod).
- Do not configure a static EIP for services that do not have specific requirements on pod EIPs. Otherwise, the pod rebuilding takes a longer time.

Configuring a Static EIP for a Pod

When creating a pod to be bound with a static IP address, configure the EIP annotation. Then, an EIP will be automatically allocated and bound to the pod.

The following uses a StatefulSet named **nginx** as an example. For details about annotations, see **Table 3-23**.

- For an automatically allocated EIP with a **dedicated bandwidth** when you create a StatefulSet, you do not need to specify the bandwidth ID. The following shows an example:

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: nginx
spec:
  serviceName: nginx
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        yangtse.io/static-eip: 'true'   # Static EIP bound to the pod
        yangtse.io/static-eip-expire-no-cascading: 'false'   # Deleting the EIP with the associated workload
        yangtse.io/static-eip-expire-duration: 5m   # Interval for reclaiming expired EIPs
        yangtse.io/pod-with-eip: 'true'   # An EIP will be automatically allocated when the pod is created.
        yangtse.io/eip-bandwidth-size: '5'    # EIP bandwidth
        yangtse.io/eip-network-type: 5_bgp    # EIP type
        yangtse.io/eip-charge-mode: bandwidth    # EIP billing mode
    spec:
      containers:
        - name: container-0
          image: nginx:alpine
          resources:
            limits:
              cpu: 100m
              memory: 200Mi
            requests:
              cpu: 100m
              memory: 200Mi
      imagePullSecrets:
        - name: default-secret
```

- For an automatically allocated EIP with a **shared bandwidth** when you create a StatefulSet, you are required to specify the bandwidth ID. The following shows an example:

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: nginx
spec:
  serviceName: nginx
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
      annotations:
        yangtse.io/static-eip: 'true'   # Static EIP bound to the pod
        yangtse.io/pod-with-eip: 'true'   # An EIP will be automatically allocated when the pod is created.
        yangtse.io/eip-network-type: 5_bgp    # EIP type
        yangtse.io/eip-bandwidth-id: <eip_bandwidth_id>    # Shared bandwidth ID of the EIP
    spec:
      containers:
        - name: container-0
          image: nginx:alpine
          resources:
            limits:
              cpu: 100m
              memory: 200Mi
            requests:
              cpu: 100m
              memory: 200Mi
      imagePullSecrets:
        - name: default-secret
```

Table 3-23 Annotations of the pod's static EIP

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/static-eip	Yes	false	Specifies whether to enable the static EIP of a pod. This function is supported only for StatefulSet pods or pods without **ownerReferences**. This function is disabled by default.	**false** or **true**
yangtse.io/static-eip-expire-duration	No	5m	Specifies the interval for reclaiming the expired static EIP after the pod with a static EIP is deleted.	The time format is Go time type, for example, 1h30m and 5m. For details, see **Go time type**.
yangtse.io/static-eip-expire-no-cascading	No	false	Specifies whether to disable cascading reclamation of StatefulSet workloads. The default value is **false**, indicating that the corresponding static EIP will be deleted with the StatefulSet workload. If you want to retain the static EIP for a new StatefulSet with the same name during the interval for reclaiming the expired EIP, set the value to **true**.	**false** or **true**

Table 3-24 Annotations of an EIP with a dedicated bandwidth

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/pod-with-eip	Yes	false	Whether to allocate an EIP with a pod and bind the EIP to the pod	**false** or **true**

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/eip-bandwidth-size	No	5	Bandwidth, in Mbit/s	The value range varies depending on the region and bandwidth billing mode. For details, see the purchase page on the EIP console. For example, in the AP-Singapore region, if an EIP is billed by bandwidth, the bandwidth ranges from 1 Mbit/s to 2000 Mbit/s; if an EIP is billed by traffic, the bandwidth ranges from 1 Mbit/s to 300 Mbit/s.
yangtse.io/eip-network-type	No	5_bgp	EIP type	The type varies depending on the region. For details, see the purchase page on the EIP console. For example, the following options are available in in the AP-Singapore region: ● **5_bgp**: dynamic BGP
yangtse.io/eip-charge-mode	No	None	Billed by traffic or bandwidth **You are advised to configure this parameter.** If this parameter is left blank, no billing mode is specified. In this case, the default value of the EIP API in the region is used.	● **bandwidth**: billed by bandwidth ● **traffic**: billed by traffic

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/eip-bandwidth-name	No	Pod name	Bandwidth name	● Enter 1 to 64 characters. Only letters, digits, hyphens (-), underscores (_), and periods (.) are allowed. ● Minimum length: 1 character ● Maximum length: 64 characters

Table 3-25 Annotations of an EIP with a shared bandwidth

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/pod-with-eip	Yes	false	Whether to bind an EIP to a pod	**false** or **true**
yangtse.io/eip-network-type	No	5_bgp	EIP type	● 5_bgp ● 5_union ● 5_sbgp The types vary by region. For details, see the EIP console.

Annotation	Mandatory	Default Value	Description	Value Range
yangtse.io/eip-bandwidth-id	Mandatory when a shared bandwidth is used	None	ID of an existing bandwidth ● If this parameter is not specified, the EIP with a dedicated bandwidth is used by default. For details about parameter settings for an EIP with a dedicated bandwidth, see **Table 3-21**. ● Only the **yangtse.io/eip-network-type** field can be specified concurrently, and this field is optional.	-

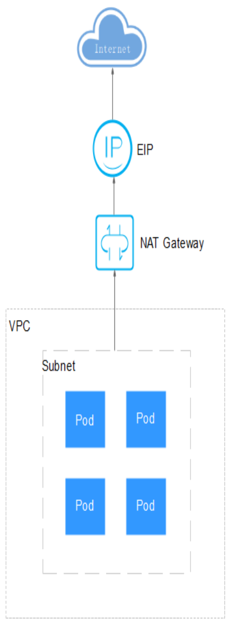
Deleting a Static EIP

After a pod is deleted, if another pod with the same name is created before the static EIP expires, the EIP can still be used. The static EIP is deleted only if there is no new pod with the name the same as that of the deleted pod before the EIP expires, or the function of deleting the EIP with the associated StatefulSet is enabled and the StatefulSet is deleted.

3.6 Accessing Public Networks from a Container

You can use NAT Gateway to enable the pods in a VPC to access public networks. NAT Gateway provides source network address translation (SNAT), which translates private IP addresses to an EIP bound to the gateway, providing secure and efficient access to the Internet. **Figure 3-23** shows the SNAT architecture. SNAT allows the pods in a VPC to access the Internet without having an EIP bound. SNAT supports a large number of concurrent connections, which makes it suitable for applications that need to handle a large number of requests.

Figure 3-23 SNAT



Procedure

To enable a container pod to access the Internet, perform the following steps:

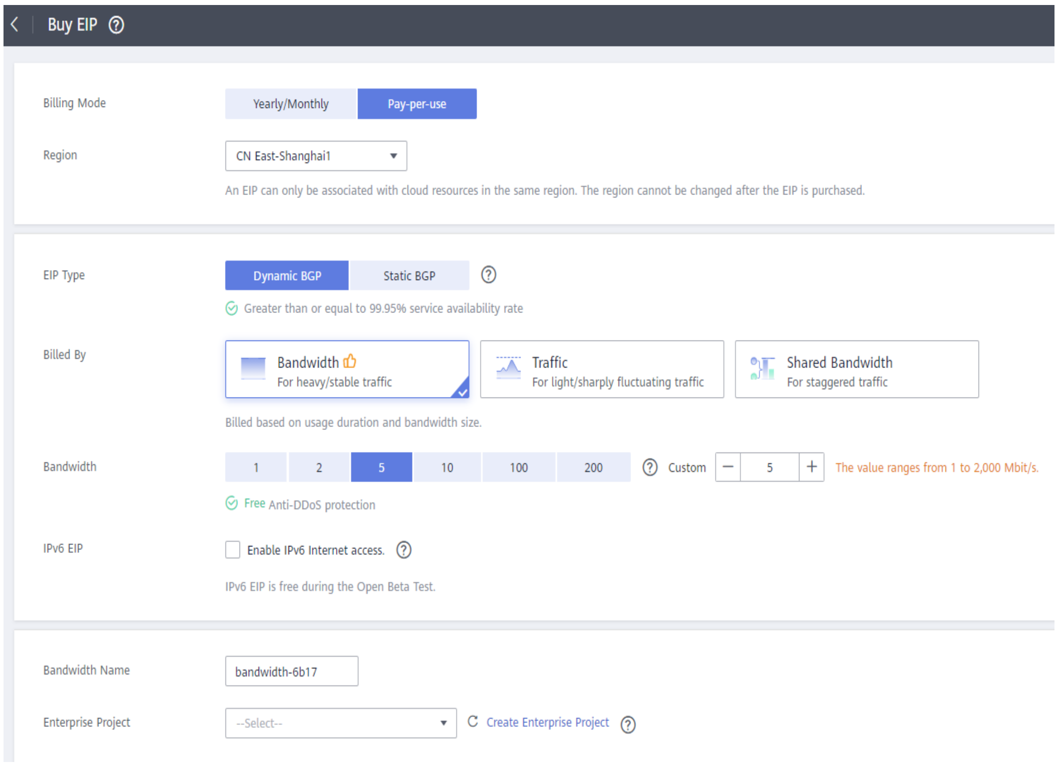
Step 1 Assign an EIP.

1. Log in to the management console.
2. Click in the upper left corner of the management console and select a region and a project.
3. Click at the upper left corner and choose **Networking** > **Elastic IP** in the expanded list.
4. On the **EIPs** page, click **Buy EIP**.
5. Configure the parameters as required.

   NOTE

   Set **Region** to the region where container pods are located.

Figure 3-24 Buying an elastic IP address



Billing Mode	Yearly/Monthly [✓ Pay-per-use]
Region	CN East-Shanghai1
	An EIP can only be associated with cloud resources in the same region. The region cannot be changed after the EIP is purchased.
EIP Type	[✓ Dynamic BGP] Static BGP
	Greater than or equal to 99.95% service availability rate
Billed By	[✓ Bandwidth For heavy/stable traffic] Traffic For light/sharply fluctuating traffic; Shared Bandwidth For staggered traffic
	Billed based on usage duration and bandwidth size.
Bandwidth	1 2 [✓ 5] 10 100 200 Custom 5 The value ranges from 1 to 2,000 Mbit/s.
	Free Anti-DDoS protection
IPv6 EIP	Enable IPv6 Internet access.
	IPv6 EIP is free during the Open Beta Test.
Bandwidth Name	bandwidth-6b17
Enterprise Project	--Select-- Create Enterprise Project

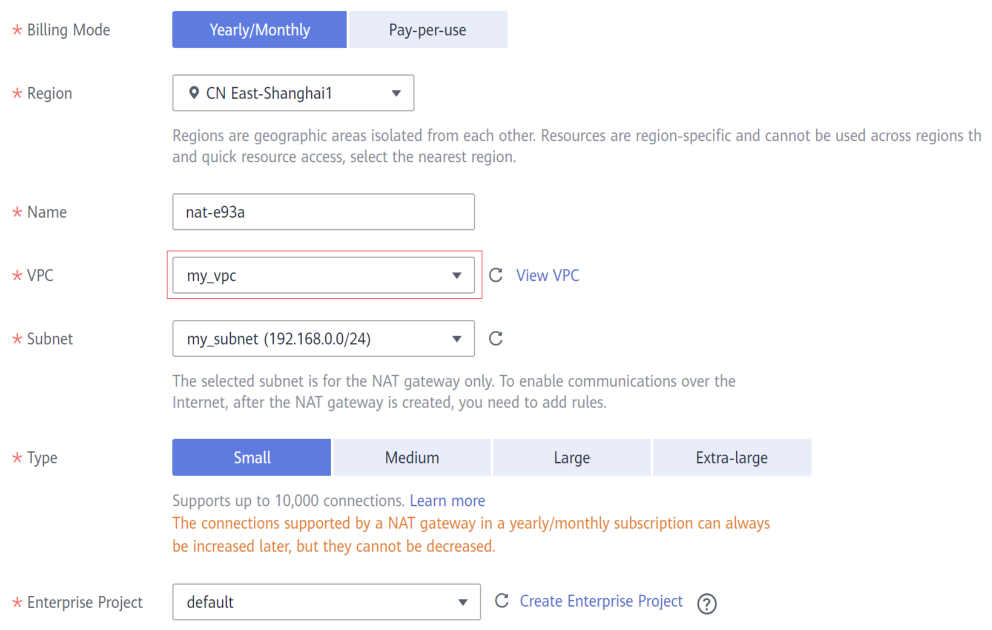
Step 2 Create a NAT gateway. For details, see **Buying a Public NAT Gateway**.

1. Log in to the management console.
2. Click in the upper left corner of the management console and select a region and a project.
3. Click at the upper left corner and choose **Networking** > **NAT Gateway** in the expanded list.
4. On the displayed page, click **Buy Public NAT Gateway** in the upper right corner.
5. Configure the parameters as required.

   NOTE

   Select the same VPC.

Figure 3-25 Buying a NAT gateway



* Billing Mode	[✓ Yearly/Monthly] Pay-per-use
* Region	CN East-Shanghai1
	Regions are geographic areas isolated from each other. Resources are region-specific and cannot be used across regions th and quick resource access, select the nearest region.
* Name	nat-e93a
* VPC	my_vpc
* Subnet	my_subnet (192.168.0.0/24)
	The selected subnet is for the NAT gateway only. To enable communications over the Internet, after the NAT gateway is created, you need to add rules.
* Type	[✓ Small] Medium Large Extra-large
	Supports up to 10,000 connections. Learn more
	The connections supported by a NAT gateway in a yearly/monthly subscription can always be increased later, but they cannot be decreased.
* Enterprise Project	default

Step 3 Configure an SNAT rule and bind the EIP to the subnet. For details, see **Adding an SNAT Rule**.

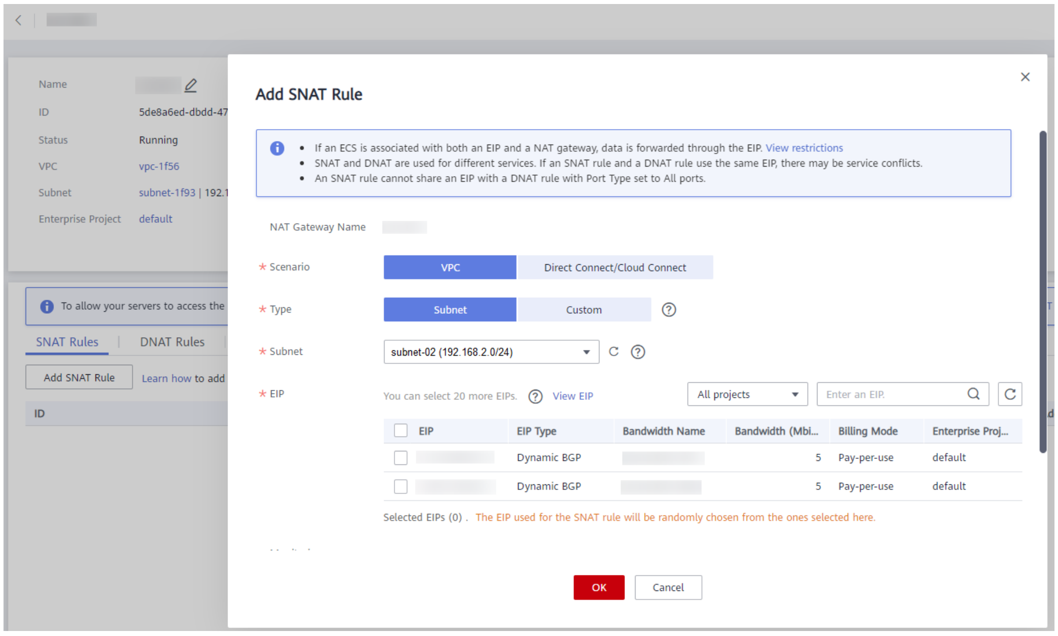
1. Log in to the management console.
2. Click in the upper left corner of the management console and select a region and a project.
3. Click at the upper left corner and choose **Networking** > **NAT Gateway** in the expanded list.
4. On the displayed page, click the name of the NAT gateway for which you want to add the SNAT rule.
5. On the **SNAT Rules** tab, click **Add SNAT Rule**.
6. Configure the parameters as required.

NOTE

SNAT rules take effect by network segment. Set **Subnet** to the subnet where the pods are located.

If there are multiple network segments, you can create multiple SNAT rules or select a user-defined network segment as long as the network segment contains the subnet where the pods are located.

Figure 3-26 Adding an SNAT rule



Name

ID 5de8a6ed-dbdd-47

Status Running

VPC vpc-1f56

Subnet subnet-1f93 | 192.1

Enterprise Project default

To allow your servers to access the

SNAT Rules | DNAT Rules

Add SNAT Rule Learn how to add

ID

Add SNAT Rule

- If an ECS is associated with both an EIP and a NAT gateway, data is forwarded through the EIP. View restrictions
- SNAT and DNAT are used for different services. If an SNAT rule and a DNAT rule use the same EIP, there may be service conflicts.
- An SNAT rule cannot share an EIP with a DNAT rule with Port Type set to All ports.

NAT Gateway Name

* Scenario: [✓ VPC] Direct Connect/Cloud Connect

* Type: [✓ Subnet] Custom

* Subnet: subnet-02 (192.168.2.0/24)

* EIP: You can select 20 more EIPs. View EIP

All projects | Enter an EIP.

EIP	EIP Type	Bandwidth Name	Bandwidth (Mbi...	Billing Mode	Enterprise Proj...
	Dynamic BGP		5	Pay-per-use	default
	Dynamic BGP		5	Pay-per-use	default

Selected EIPs (0) . The EIP used for the SNAT rule will be randomly chosen from the ones selected here.

After the SNAT rule is configured, workloads can access public networks from the container. Public networks can be pinged from the container.

----End

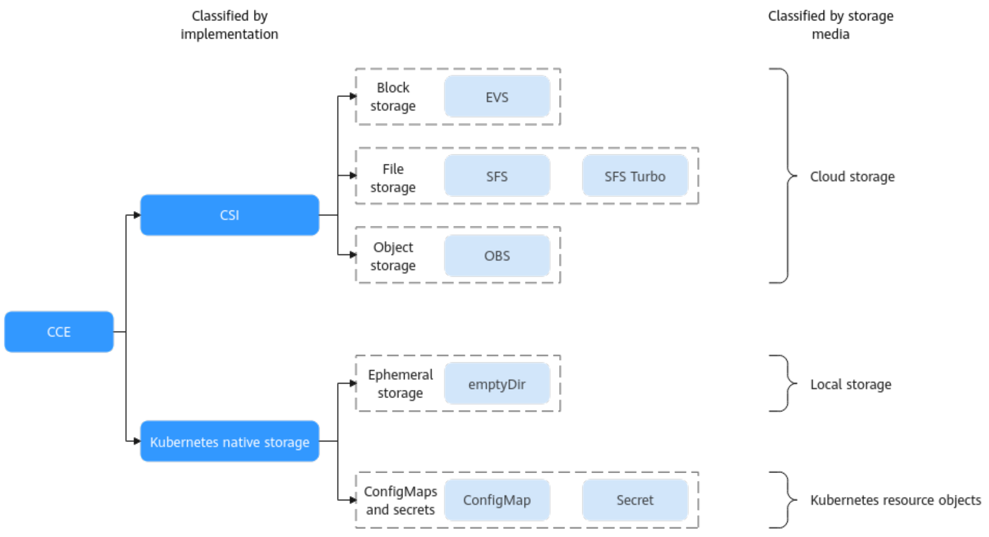
4 Storage

4.1 Storage Overview

Container Storage

CCE Autopilot container storage is implemented based on Kubernetes container storage APIs (**CSI**). CCE Autopilot integrates multiple types of cloud storage and covers different application scenarios. It is fully compatible with Kubernetes native storage services, such as emptyDir, hostPath, secret, and ConfigMap.

Figure 4-1 Container storage types



CCE Autopilot allows workload pods to use multiple types of storage:

- In terms of implementation, storage supports Container Storage Interface (CSI) and Kubernetes native storage.

Type	Description
CSI	An **out-of-tree** volume add-on, which specifies the standard container storage API and allows storage vendors to use standard custom storage plugins that are mounted using PVCs and PVs without the need to add their plugin source code to the Kubernetes repository for unified build, compilation, and release. CSI is a recommended in Kubernetes 1.13 and later versions.
Kubernetes native storage	An "in-tree" volume add-on that is built, compiled, and released with the Kubernetes repository.

- In terms of storage media, storage can be classified as cloud storage, local storage, and Kubernetes resource objects.

Type	Description	Application Scenario
Cloud storage	The storage media is provided by storage vendors. Storage volumes of this type are mounted using PVCs and PVs.	Data requires high availability or needs to be shared, for example, logs and media resources. Select a proper cloud storage type based on the application scenario. For details, see **Cloud Storage Comparison**.
Local storage	Only emptyDir is supported. The lifecycle of an emptyDir volume is the same as that of a pod. Memory can be specified as the storage medium. When the pod is deleted, the emptyDir volume is deleted, and the data is lost.	Non-HA data that requires high I/O and low latency. For details, see **Local Storage**.
Kubernetes resource objects	ConfigMaps and secrets are resources created in clusters. They are special storage types and are provided by tmpfs (RAM-based file system) on the Kubernetes API server.	ConfigMaps are used to inject configuration data to pods. Secrets are used to transmit sensitive information such as passwords to pods.

Cloud Storage Comparison

Item	EVS	SFS	SFS Turbo	OBS
Definition	EVS offers scalable block storage for cloud servers. With high reliability, high performance, and rich specifications, EVS disks can be used for distributed file systems, dev/test environments, data warehouses, and high-performance computing (HPC) applications.	Expandable to petabytes, SFS provides fully hosted shared file storage, highly available and stable to handle data- and bandwidth-intensive applications in HPC, media processing, file sharing, content management, and web services.	Expandable to 320 TB, SFS Turbo provides a fully hosted shared file storage, which is highly available and stable, to support small files and applications requiring low latency and high IOPS. You can use SFS Turbo in high-traffic websites, log storage, compression/decompression, DevOps, enterprise OA, and containerized applications.	OBS provides massive, secure, and cost-effective data storage for you to store data of any type and size. It is suitable for various scenarios, such as enterprise-level backup/archiving, video on demand (VoD), and video monitoring.
Data storage logic	Stores only binary data. To store files, you need to format the file system first.	Stores files, and sorts and displays data in the hierarchy of files and folders.	Stores files, and sorts and displays data in the hierarchy of files and folders.	Stores data as objects with metadata and unique identifiers. You can upload files directly to OBS. The system can generate metadata for files, or you can customize the metadata for files.

Item	EVS	SFS	SFS Turbo	OBS
Access method	EVS disks can only be used and accessed from applications after being attached to ECSs or BMSs and initialized.	The can be mounted to ECSs or BMSs using network protocols. A network address must be specified or mapped to a local directory for access.	SFS Turbo supports the Network File System (NFS) protocol (NFSv3 only). You can seamlessly integrate existing applications and tools with SFS Turbo.	OBS is accessible through the Internet or Direct Connect. The bucket address must be specified for access, and transfer protocols HTTP and HTTPS are used.
Static storage volumes	Supported. For details, see **Using an Existing EVS Disk Through a Static PV**.	Supported. For details, see **Using an Existing File System Through a Static PV**.	Supported. For details, see **Using an Existing SFS Turbo File System Through a Static PV**.	Supported. For details, see **Using an Existing OBS Bucket or Parallel File System Through a Static PV**.
Dynamic storage volumes	Supported. For details, see **Using an EVS Disk Through a Dynamic PV**.	Supported. For details, see **Using an SFS File System Through a Dynamic PV**.	Supported by SFS Turbo subdirectories but not by SFS Turbo. For details, see **(Recommended) Creating an SFS Turbo Subdirectory Using a Dynamic PV**.	Supported. For details, see **Using an OBS Bucket or Parallel File System Through a Dynamic PV**.
Highlights	Non-shared storage. Each volume can be mounted to only one node.	Shared storage. High-performance and high-throughput storage services are provided.	Shared storage. High-performance and high-bandwidth storage services are provided.	Shared storage and user-mode file system.

Item	EVS	SFS	SFS Turbo	OBS
Application scenarios	HPC, enterprise core cluster applications, enterprise application systems, and development and testing **NOTE** HPC apps here require high-speed and high-IOPS storage, such as industrial design and energy exploration.	HPC, media processing, content management, web services, big data, and analysis applications **NOTE** HPC apps here require high bandwidth and shared file storage, such as gene sequencing and image rendering.	High-traffic websites, log storage, DevOps, and enterprise OA	Big data analytics, static website hosting, online video on demand (VoD), gene sequencing, intelligent video surveillance, backup and archiving, and enterprise cloud boxes (web disks)
Capacity	TB	SFS 1.0: PB General purpose file system (formerly SFS 3.0): EB	General-purpose: TB	EB
Latency	1–2 ms	SFS 1.0: 3–20 ms General purpose file system (formerly SFS 3.0): 10 ms	General-purpose: 1–5 ms	10 ms
Max. IOPS	2200–256000, depending on flavors	SFS 1.0: 2,000 General purpose file system (formerly SFS 3.0): millions	General-purpose: up to 100,000	Tens of millions
Bandwidth	MB/s	SFS 1.0: GB/s General purpose file system (formerly SFS 3.0): TB/s	General-purpose: up to GB/s	TB/s

Local Storage

An emptyDir volume provides ephemeral storage for pods. Its lifecycle is the same as that of a pod. Memory can be specified as the storage medium. When the pod is deleted, the emptyDir volume is deleted, and the data is lost. For details, see **emptyDir**.

Highlights: emptyDir volumes are local ephemeral volumes. The storage space comes from the local kubelet root directory or memory.

Application scenarios:

- Scratch space, such as for a disk-based merge sort
- checkpointing a long computation for recovery from crashes
- Saving the files obtained by the content manager container when web server container data is used

Support for Enterprise Projects

- Creating storage automatically:

  When creating EVS or OBS PVCs using a StorageClass in CCE Autopilot, you can specify an enterprise project to assign the created storage resources (EVS disks and OBS) to. **This enterprise project can either be the default one or the same one as the cluster belongs to.**

  If no enterprise project is specified, the enterprise project specified in StorageClass will be used by default for creating storage resources. For the csi-disk and csi-obs storage classes provided by CCE Autopilot, the created storage resources belong to the default enterprise project.
- Using existing storage:

  When you create a PVC using a PV, ensure that **everest.io/enterprise-project-id** specified in the PVC and PV are the same because an enterprise project has been specified during storage resource creation. Otherwise, the PVC and PV cannot be bound.

Reference

- **Storage Basics**
- **EVS Volumes**
- **Scalable File Service**
- **SFS Turbo**
- **Object Storage Service**

4.2 Storage Basics

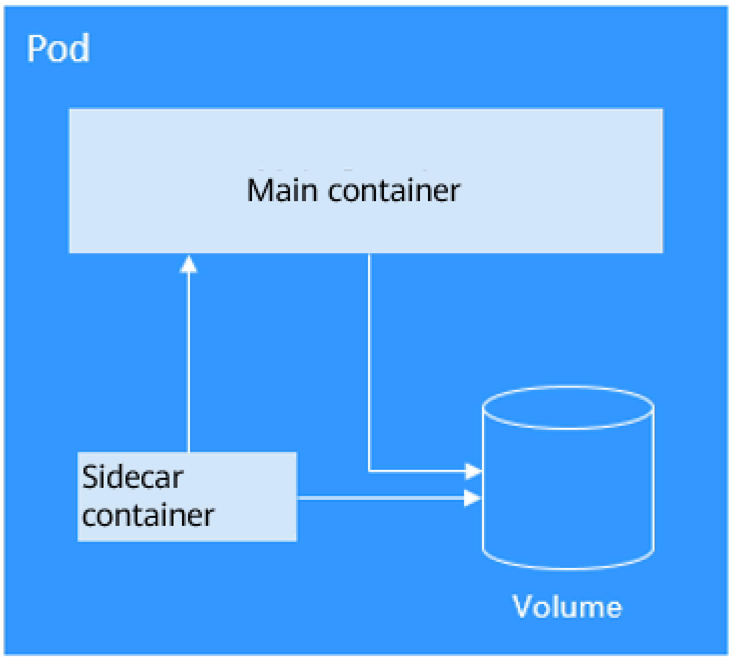
Volumes

On-disk files in a container are ephemeral, which presents the following problems to important applications running in the container:

- When a container is rebuilt, files in the container will be lost.
- When multiple containers run in a pod at the same time, files need to be shared among the containers.

Kubernetes volumes resolve both of these problems. Volumes, as part of a pod, cannot be created independently and can only be defined in pods. All containers in a pod can access its volumes, but the volumes must have been mounted to any directory in a container.

The following figure shows how a storage volume is used between containers in a pod.



The basic principles for using volumes are as follows:

- Multiple volumes can be mounted to a pod. However, do not mount too many volumes to a pod.
- Multiple types of volumes can be mounted to a pod.
- Each volume mounted to a pod can be shared among containers in the pod.
- You are advised to use PVCs and PVs to mount volumes for Kubernetes.

NOTE

The lifecycle of a volume is the same as that of the pod to which the volume is mounted. When the pod is deleted, the volume is also deleted. However, files in the volume may outlive the volume, depending on the volume type.

Kubernetes provides various volume types, which can be classified as in-tree and out-of-tree.

Volume Classification	Description
In-tree	Maintained through the Kubernetes code repository and built, edited, and released with Kubernetes binary files. Kubernetes does not accept this volume type anymore. Kubernetes-native volumes such as hostPath, emptyDir, Secret, and ConfigMap are all the in-tree type. PVCs are a special in-tree volume. Kubernetes uses this type of volume to convert from in-tree to out-of-tree. PVCs allow you to request for PVs created using the underlying storage resources provided by different storage vendors.
Out-of-tree	Out-of-tree volumes include container storage interfaces (CSIs) and FlexVolumes (deprecated). Storage vendors only need to comply with certain specifications to create custom storage add-ons and PVs that can be used by Kubernetes, without adding add-on source code to the Kubernetes code repository. Cloud storage such as SFS and OBS is used by installing storage drivers in a cluster. You need to create PVs in the cluster and mount the PVs to pods using PVCs.

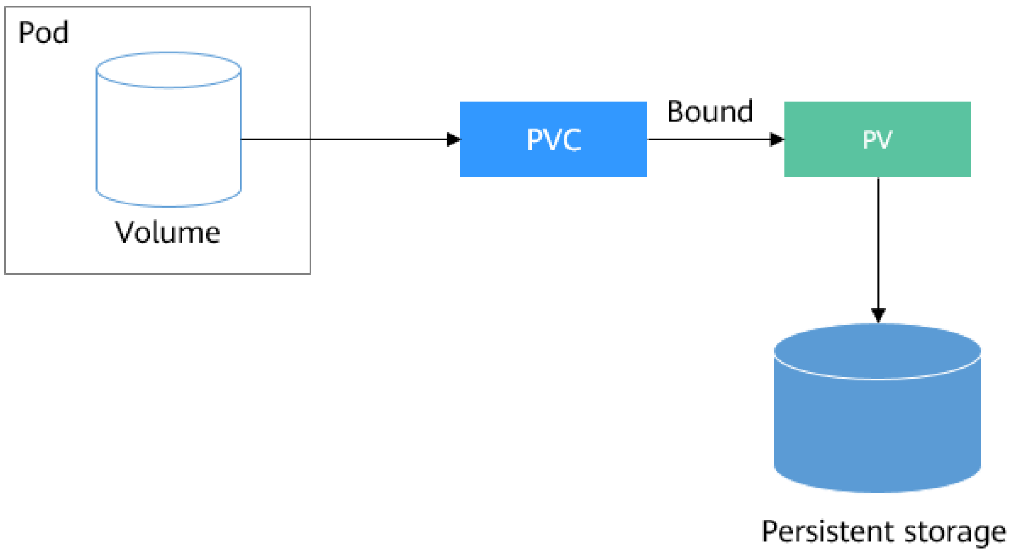
PV and PVC

Kubernetes provides PersistentVolumes (PVs) and PersistentVolumeClaims (PVCs) to abstract details of how storage is provided from how it is consumed. You can request specific size of storage when needed, just like pods can request specific levels of resources (CPU and memory).

- PV: describes a persistent storage volume in a cluster. A PV is a cluster-level resource just like a node. It applies to the entire Kubernetes cluster. A PV has a lifecycle independent of any individual Pod that uses the PV.
- PVC: describes a request for storage by a user. When configuring storage for an application, claim a storage request (that is, PVC). Kubernetes selects a PV that best meets the request and binds the PV to the PVC. A PVC to PV binding is a one-to-one mapping. When creating a PVC, describe the attributes of the requested persistent storage, such as the storage size and read/write permission.

You can bind PVCs to PVs in a pod so that the pod can use storage resources. The following figure shows the relationship between PVs and PVCs.

Figure 4-2 PVC-to-PV binding



CSI

CSI is a standard for container storage interfaces and a storage plugin implementation solution recommended by the Kubernetes community.

Volume Access Modes

Storage volumes can be mounted to the host system only in the mode supported by underlying storage resources. For example, a file storage system can be read and written by multiple pods, but an EVS volume can be read and written by only one pod.

- **ReadWriteOnce**: A storage volume can be mounted to a single pod in read-write mode.
- **ReadWriteMany**: A storage volume can be mounted to multiple pods in read-write mode.

Table 4-1 Access modes supported by storage volumes

Storage Type	ReadWriteOnce	ReadWriteMany
EVS	√	×
SFS	×	√
OBS	×	√
SFS Turbo	×	√

Mounting a Storage Volume

You can mount volumes in the following ways:

Use PVs to describe existing underlying storage resources, and then create PVCs to use the underlying storage resources in pods. You can also use the dynamic creation mode. That is, specify the StorageClass when creating a PVC and use the provisioner in the StorageClass to automatically create a PV and bind the PV to the PVC.

Table 4-2 Modes of mounting volumes

Mount Mode	Description	Supported Volume Type	Other Constraints
Statically creating storage volume (using existing storage)	You need to manually create underlying storage (such as EVS disks and SFS file systems) and PVs. Operations on the console: First, use the existing underlying storage to create a PV. Second, create a PVC and mount it to the pods for running the workload. Kubernetes binds PVCs to the matching PVs so that workloads can access storage services during PVC creation. **NOTE** If the access mode of the PV is **ReadWriteOnce**, the number of pods for running the workload can only be set to **1** when the workload is created.	All volumes	None
Dynamically creating storage volumes (automatically creating storage)	The underlying storage and PV are automatically created. Operations on the console: First, create a PVC and specify a storage class for the PVC. The Provisioner in StorageClass automatically creates the underlying storage and PV based on PVC. The automatically created PV is bound to the PVC and mounted to the pods for running the workload. Second, mount the PVC to the workload. **NOTE** If the access mode of the PV is **ReadWriteOnce**, the number of pods for running the workload can only be set to **1** when the workload is created.	EVS volumes, OBS volumes, and SFS/SFS Turbo volumes	None

Mount Mode	Description	Supported Volume Type	Other Constraints
Dynamic mounting (VolumeClaimTemplate)	The underlying storage and PV are automatically created. Achieved by using the **volumeClaimTemplates** field and depends on the dynamic PV creation capability of StorageClass. Operations on the console: Create a workload and select **VolumeClaimTemplate (VCT)** for **Data Storage**. Second, create a dynamically provisioned PVC. Specify a storage class for the PVC. The Provisioner in StorageClass automatically creates the underlying storage and PV based on PVC. **NOTE** Even if the access mode of the PV is **ReadWriteOnce**, multiple pods can be created for running the workload. This is because each pod has a unique PVC and PV. After a pod is rescheduled, the original PV can still be mounted to it based on the PVC name.	Only EVS volumes	Supported only by StatefulSets

PV Reclaim Policy

A PV reclaim policy is used to delete or reclaim underlying volumes when a PVC is deleted. The value can be **Delete** or **Retain**.

- **Delete**: Deleting a PVC will remove the PV from Kubernetes, and the associated underlying storage assets will also be removed from the external infrastructure.

  NOTE

  Yearly/Monthly resources cannot be deleted using the **Delete** reclaim policy.

- **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again.

  You can manually delete and reclaim volumes by performing the following operations:

  a. Delete the PV.
  b. Clear data on the associated underlying storage resources as required.
  c. Delete the associated underlying storage resources.

To reuse the underlying storage resources, create a PV.

CCE Autopilot also allows you to delete a PVC without deleting underlying storage resources. This function can be achieved only by using a YAML file: Set the PV reclaim policy to **Delete** and add **everest.io/reclaim-policy: retain-volume-only** to **annotations**. In this way, when the PVC is deleted, the PV is deleted, but the underlying storage resources are retained.

The following YAML file takes EVS as an example:

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: test
  namespace: default
  annotations:
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
    everest.io/disk-volume-type: SAS
  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one where the cluster is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>       # Replace the AZ with the one where the EVS disk is located.
spec:
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 10Gi
  storageClassName: csi-disk
  volumeName: pv-evs-test

---
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only
  name: pv-evs-test
  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one where the cluster is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>       # Replace the AZ with the one where the EVS disk is located.
spec:
  accessModes:
    - ReadWriteOnce
  capacity:
    storage: 10Gi
  csi:
    driver: disk.csi.everest.io
    fsType: ext4
    volumeHandle: 2af98016-6082-4ad6-bedc-1a9c673aef20
    volumeAttributes:
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/disk-mode: SCSI
      everest.io/disk-volume-type: SAS
  persistentVolumeReclaimPolicy: Delete
  storageClassName: csi-disk
```

Reference

- For more information about Kubernetes storage, see **Storage**.
- For more information about CCE Autopilot container storage, see **Storage Overview**.

4.3 EVS Volumes

4.3.1 EVS Volume Overview

For persistent storage, CCE Autopilot allows you to mount the storage volumes (PVs) created from Elastic Volume Service (EVS) disks to a path of a container. When the container is migrated within an AZ, the mounted volumes are also migrated. By using EVS volumes, you can mount the remote file directory of a storage system to a container so that data in the data volume is permanently preserved. Even if the container is deleted, the data in the data volume is still stored in the storage system.

NOTE

View the regions where EVS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where EVS volumes are supported.

EVS Disk Specifications

EVS performance metrics include:

- IOPS: the number of input/output operations performed by an EVS disk per second
- Throughput: the amount of data read from and written into an EVS disk per second
- I/O latency: the minimum interval between two consecutive I/O operations on an EVS disk

Table 4-3 EVS disk specifications

Item	Extreme SSD	General Purpose SSD	Ultra-high I/O	High I/O
Max. capacity (GiB)	• System disk: 1,024 • Data disk: 32,768	• System disk: 1,024 • Data disk: 32,768	• System disk: 1,024 • Data disk: 32,768	• System disk: 1,024 • Data disk: 32,768
Max. IOPS	128000	20000	50000	5000
Max. throughput (MiB/s)	1000	250	350	150
Burst IOPS limit	64000	8000	16000	5000
Disk IOPS	Min. (128,000, 1,800 + 50 x Capacity)	Min. (20,000, 1,800 + 12 x Capacity)	Min. (50,000, 1,800 + 50 x Capacity)	Min. (5,000, 1,800 + 8 x Capacity)

Item	Extreme SSD	General Purpose SSD	Ultra-high I/O	High I/O
Disk throughput (MiB/s)	Min. (1,000, 120 + 0.5 x Capacity)	Min. (250, 100 + 0.5 x Capacity)	Min. (350, 120 + 0.5 x Capacity)	Min. (150, 100 + 0.15 x Capacity)
Single-queue access latency (ms)	Sub-millisecond	1	1	1~ 3
API name	ESSD	GPSSD	SSD	SAS

For details about EVS disks, see **Disk Types and Performance**.

Scenario

EVS volumes can be mounted in the following modes based on application scenarios:

- **Using an existing EVS disk through a static PV**: Use an existing EVS disk to create a PV and then mount storage to the workload through a PVC. This mode applies to scenarios where the underlying storage is available or billed on a yearly/monthly basis.
- **Using an EVS disk through a dynamic PV**: You do not need to create EVS volumes beforehand. Instead, specify a StorageClass when creating a PVC. Then, an EVS volume and PV will be created automatically. This mode applies to scenarios where no underlying storage is available.
- **Dynamically mounting an EVS volume to a StatefulSet**: Each pod is associated with a unique PVC and PV. After a pod is rescheduled, the original PV can still be mounted to it based on the PVC name. This mode applies to StatefulSets with multiple pods.

Billing

- By default, the EVS disks **automatically created** when you specify the StorageClass is billed on a pay-per-use basis and cannot be changed to yearly/monthly. To use a yearly/monthly-billed EVS disk, **use an existing one**.
- For details about the EVS disk pricing, see **Billing for Disks**.

4.3.2 Using an Existing EVS Disk Through a Static PV

CCE Autopilot allows you to create a static PV using an existing EVS disk. After the PV is created, you can create a PVC and bind it to the PV. This mode applies to scenarios where the underlying storage is available or billed on a yearly/monthly basis.

Prerequisites

- You have created a cluster.
- You have created an EVS disk that meets the following requirements:

  - The EVS disk cannot be a system disk, DSS disk, or shared disk.
  - The EVS disk must be of the **SCSI** type (the default disk type is **VBD** when you purchase an EVS disk).
  - The EVS disk must not be used by other resources.
  - Encrypted EVS disks are not supported.
  - The EVS disk must be in the default enterprise project or the enterprise project to which the cluster belongs.
  - EVS disks that have been partitioned are not supported.
  - Only ext4 EVS disks are supported.
- If you want to create a cluster by running commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- Not all regions support EVS volumes. View the regions where EVS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where EVS volumes are supported.
- EVS disks cannot be used by multiple workloads, multiple pods of the same workload, or multiple tasks. If EVS disks are used when you create a Deployment, select only one pod for the Deployment.
- The cluster version must be v1.27.8-r0, v1.28.6-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- No more than 10 EVS disks can be attached to each workload in a cluster. If more than 10 EVS disks are attached, the workload may run abnormally.

Creating a Volume Using an Existing EVS Disk

Create a storage volume using an existing EVS disk and mount it to a workload on the console or using kubectl.

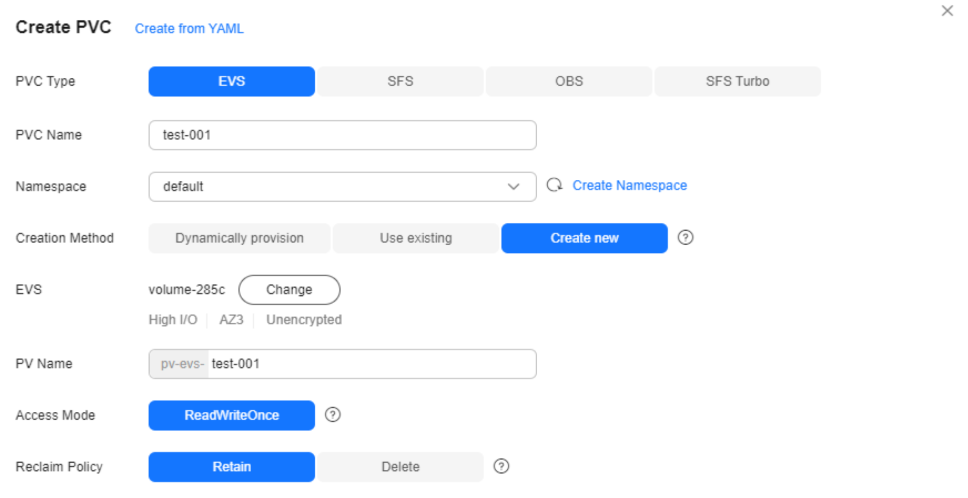
Using the CCE Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Statically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Figure 4-3 Creating a PVC



Create PVC Create from YAML

PVC Type: EVS | SFS | OBS | SFS Turbo

PVC Name: test-001

Namespace: default Create Namespace

Creation Method: Dynamically provision | Use existing | Create new

EVS: volume-285c Change
High I/O | AZ3 | Unencrypted

PV Name: pv-evs- test-001

Access Mode: ReadWriteOnce

Reclaim Policy: Retain | Delete

Parameter	Description
PVC Type	In this example, select **EVS**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Namespace	A namespace is a conceptual grouping of resources or objects. Each namespace provides isolation for data from other namespaces. After a cluster is created, the default namespace is created by default. If there is no special requirement, select the default namespace.
Creation Method	– If underlying storage is available, create a PV or use an existing PV to statically create a PVC. – If no underlying storage is available, select **Dynamically provision**. For details, see **Using an EVS Disk Through a Dynamic PV**. In this example, select **Create new** to create a PV and PVC at the same time on the console.
PV[a]	Select an existing PV in the cluster. Create a PV in advance. For details, see "Creating a storage volume" in **Related Operations**. You do not need to specify this parameter in this example.
EVS[b]	Click **Select EVS**. On the displayed page, select the EVS disk that meets your requirements and click **OK**.
PV Name[b]	Enter the PV name, which must be unique in the same cluster.

Parameter	Description
Access Mode[b]	EVS volumes support only **ReadWriteOnce**, indicating that a storage volume can only be mounted to one node in read/write mode. For details, see **Volume Access Modes**.
Reclaim Policy[b]	You can select **Delete** or **Retain** to specify the reclaim policy of the underlying storage when the PVC is deleted. For details, see **PV Reclaim Policy**. In this example, select **Retain**. **NOTE** If multiple PVs use the same underlying storage volume, use **Retain** to avoid cascading deletion of underlying volumes.

NOTE

a: The parameter is available when **Creation Method** is set to **Use existing**.

b: The parameter is available when **Creation Method** is set to **Create new**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **StatefulSets** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

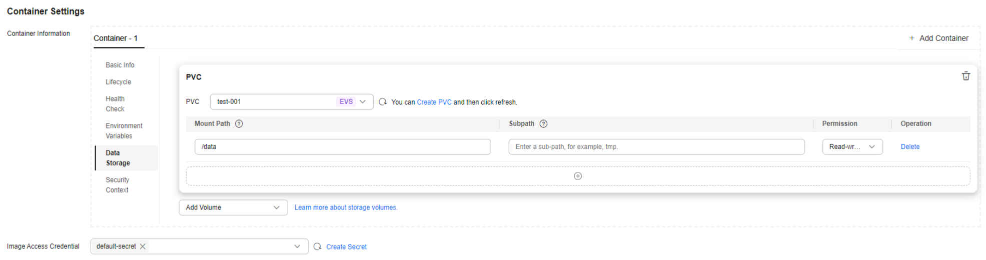
   Mount and use storage volumes. For details about the parameters, see **Table 4-4**. For other parameters, see **Creating a Workload**.

   In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the EVS disk.

   NOTE

   A non-shared EVS disk can be mounted to only one pod. If there are multiple pods, extra pods cannot start normally. Ensure that the number of pods for running a workload is 1 if an EVS disk is mounted.

Figure 4-4 Parameters for mounting a storage volume



PVC

PVC: test-001 EVS — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-wr...	Delete

Image Access Credential: default-secret — Create Secret

Table 4-4 Parameters for mounting a storage volume

Parameter	Description
PVC	Select an existing EVS volume. An EVS volume can only be mounted to one workload.
Mount Path	Enter a mount path, for example, **/data**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. **data**, for example, indicates that data in the mount path of the container is stored in the **data** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a PV. If a PV has been created in your cluster, skip this step. When using a YAML file to create a PV, do not use parameters that are not declared in the file, such as **status**, **spec.claimRef**, and **annotation.everest.io/set-disk-metadata**, or the PV may be abnormal.

1. Create the **pv-evs.yaml** file.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only         # (Optional) The underlying storage is retained when the PV is deleted.
  name: pv-evs    # PV name
  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one where the cluster is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>       # Replace the AZ with the one where the EVS disk is located.
spec:
  accessModes:
    - ReadWriteOnce     # Access mode, which must be ReadWriteOnce for EVS disks
  capacity:
    storage: 10Gi       # EVS disk capacity, in GiB. The value ranges from 1 to 32768.
  csi:
    driver: disk.csi.everest.io     # Dependent storage driver for the mounting
    fsType: ext4    # Set the file system type to ext4.
    volumeHandle: <your_volume_id>   # Volume ID of the EVS disk
    volumeAttributes:
      everest.io/disk-mode: SCSI           # Device type of the EVS disk. Only SCSI is supported.
      everest.io/disk-volume-type: SAS     # EVS disk type
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/enterprise-project-id: <your_project_id>  # (Optional) Enterprise project ID. If an enterprise project is specified, use the same enterprise project when creating a PVC. Otherwise, the PVC cannot be bound to a PV.

      everest.io/disk-iops: '3000'      # (Optional) IOPS of only a GPSSD2 EVS disk
      everest.io/disk-throughput: '125' # (Optional) Throughput of only a GPSSD2 EVS disk

  persistentVolumeReclaimPolicy: Delete    # Reclaim policy
  storageClassName: csi-disk             # StorageClass name. The value must be csi-disk for EVS disks.
```

Table 4-5 Key parameters

Parameter	Mandatory	Description
everest.io/reclaim-policy: retain-volume-only	No	This field is optional. Only **retain-volume-only** is supported. This field takes effect when the reclaim policy is **Delete**. If the reclaim policy is **Delete** and the value is **retain-volume-only**, the associated PV is deleted while the underlying storage volume is retained, when a PVC is deleted.
failure-domain.beta.kubernetes.io/region	Yes	Region of the cluster. For details, see **Regions and Endpoints**.
failure-domain.beta.kubernetes.io/zone	Yes	AZ of the EVS disk. You can obtain all supported AZs by **querying the AZ list**.
fsType	Yes	File system type. Currently, only ext4 is supported.

Parameter	Mandatory	Description
volumeHandle	Yes	Volume ID of the EVS disk. To obtain a volume ID, log in to the **Cloud Server Console**. In the navigation pane on the left, choose **Elastic Volume Service** > **Disks**. Click the name of the target EVS disk to go to its details page. On the **Summary** tab page, click the copy button after **ID**.
everest.io/disk-volume-type	Yes	EVS disk type. All letters are in uppercase. – **SAS**: high I/O – **SSD**: ultra-high I/O – **GPSSD**: general-purpose SSD – **ESSD**: extreme SSD – GPSSD2: general-purpose SSD v2. You need to specify the **everest.io/disk-iops** and **everest.io/disk-throughput** annotations.
everest.io/disk-iops	No	Preconfigured IOPS, which is supported only by general-purpose SSD v2 EVS disks. – The IOPS of general-purpose SSD EVS v2 disks ranges from 3,000 to 128,000, and the maximum value is 500 times of the capacity (GiB). If the IOPS of general-purpose SSD v2 disks is greater than 3000, extra IOPS will be billed. For details, see **Price Calculator**.
everest.io/disk-throughput	No	Preconfigured throughput, which is supported only by general-purpose SSD v2 EVS disks. The value ranges from 125 MiB/s to 1,000 MiB/s. The maximum value is a quarter of IOPS. If the throughput is greater than 125 MiB/s, extra throughput will be billed. For details, see **Price Calculator**.

Parameter	Mandatory	Description
everest.io/enterprise-project-id	No	Enterprise project ID of the EVS disk. This parameter is optional. If an enterprise project is specified, use the same enterprise project when creating a PVC, or the PVC cannot be bound to a PV. To obtain an enterprise project ID, log in to the **Cloud Server Console**. In the navigation pane on the left, choose **Elastic Volume Service** > **Disks**. Click the name of the target EVS disk to go to its details page. On the **Summary** tab page, click the enterprise project in **Management Information** to access the enterprise project console. Copy the corresponding ID to obtain the ID of the enterprise project to which the EVS disk belongs.
persistentVolumeReclaimPolicy	Yes	The **Delete** and **Retain** reclaim policies are supported. For details, see **PV Reclaim Policy**. If high data security is required, select **Retain** to prevent data from being deleted by mistake. **Delete**: – If **everest.io/reclaim-policy** is not specified, both the PV and EVS disk will be deleted when a PVC is deleted. – If **everest.io/reclaim-policy** is set to **retain-volume-only**, when a PVC is deleted, the PV will be deleted but the EVS disk will be retained. **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again.
storageClassName	Yes	The StorageClass for EVS volumes is **csi-disk**.

2. Run the following command to create a PV:
```
kubectl apply -f pv-evs.yaml
```

Step 3 Create a PVC.

1. Create the **pvc-evs.yaml** file.
```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-evs
  namespace: default
  annotations:
      everest.io/disk-volume-type: SAS    # EVS disk type
    everest.io/enterprise-project-id: <your_project_id>  # (Optional) Enterprise project ID. If an
enterprise project is specified, use the same enterprise project when creating a PVC. Otherwise, the
PVC cannot be bound to a PV.

    everest.io/disk-iops: '3000'      # (Optional) IOPS of only a GPSSD2 EVS disk
    everest.io/disk-throughput: '125' # (Optional) Throughput of only a GPSSD2 EVS disk
  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one
where the cluster is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>      # Replace the AZ with the one where
the EVS disk is located.
spec:
  accessModes:
  - ReadWriteOnce               # The value must be ReadWriteOnce for EVS disks.
  resources:
    requests:
      storage: 10Gi            # EVS disk capacity, ranging from 1 to 32768. The value must be the same
as the storage size of the existing PV.
  storageClassName: csi-disk    # The StorageClass of the EVS disk
  volumeName: pv-evs           # PV name
```

Table 4-6 Key parameters

Parameter	Mandatory	Description
failure-domain.beta.kubernetes.io/region	Yes	Region of the cluster. For example, **cn-east-3**. For details, see **Regions and Endpoints**.
failure-domain.beta.kubernetes.io/zone	Yes	AZ of the EVS disk. You can obtain all supported AZs by **querying the AZ list**.
storage	Yes	Requested capacity in the PVC, in Gi. The value must be the same as the storage size of the existing PV in **Step 2.1**.
volumeName	Yes	PV name, which must be the same as the PV name in **Step 2.1**.
storageClassName	Yes	Storage class name, which must be the same as the storage class of the PV in **Step 2.1**. The StorageClass for EVS volumes is **csi-disk**.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-evs.yaml
```

Step 4 Create a workload.

1. Create a file named **web-evs.yaml**. In this example, the EVS volume is mounted to the **/data** path.

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: web-evs
  namespace: default
spec:
 replicas: 1           # The number of workload replicas that use the EVS volume must be 1.
  selector:
    matchLabels:
      app: web-evs
  serviceName: web-evs   # Headless Service name
  template:
    metadata:
      labels:
        app: web-evs
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-disk    # Volume name, which must be the same as the volume name in the
volumes field.
          mountPath: /data  #Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-disk    # Custom volume name
          persistentVolumeClaim:
            claimName: pvc-evs    # Name of the created PVC.
---
apiVersion: v1
kind: Service
metadata:
  name: web-evs   # Headless Service name
  namespace: default
  labels:
    app: web-evs
spec:
  selector:
    app: web-evs
  clusterIP: None
  ports:
    - name: web-evs
      targetPort: 80
      nodePort: 0
      port: 80
      protocol: TCP
  type: ClusterIP
```

2. Run the following command to create a workload that the EVS volume is mounted to:

   ```
   kubectl apply -f web-evs.yaml
   ```

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Verifying Data Persistence

Step 1 View the deployed application and files stored in the EVS volume.

1. Run the following command to view the pod:

   ```
   kubectl get pod | grep web-evs
   ```

   Expected output:

   ```
   web-evs-0                1/1     Running   0              38s
   ```

2. Run the following command to verify that the EVS volume has been mounted to the **/data** path:
   ```
   kubectl exec web-evs-0 -- df | grep data
   ```
   Expected output:
   ```
   /dev/sdc              10255636     36888  10202364   0% /data
   ```
3. Run the following command to view the created file in the **/data** path:
   ```
   kubectl exec web-evs-0 -- ls /data
   ```
   Expected output:
   ```
   lost+found
   ```

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-evs-0 --  touch /data/static
```

Step 3 Run the following command to view the created file in the **/data** path:

```
kubectl exec web-evs-0 -- ls /data
```

Expected output:

```
lost+found
static
```

Step 4 Run the following command to delete the pod named **web-evs-0**:

```
kubectl delete pod web-evs-0
```

Expected output:

```
pod "web-evs-0" deleted
```

Step 5 The StatefulSet controller automatically creates a replica with the same name as the pod. Run the following command to check whether the file in the **/data** path has been modified:

```
kubectl exec web-evs-0 -- ls /data
```

Expected output:

```
lost+found
static
```

The **static** file is retained, indicating that the data can be stored persistently.

----End

Related Operations

You can also perform the operations described in **Table 4-7**.

Table 4-7 Related operations

Operation	Description	Procedure
Creating a storage volume (PV)	You can create a PV on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVs** tab. In the upper right corner, click **Create PV**. In the displayed dialog box, configure the parameters. ● **Volume Type**: Select **EVS**. ● **EVS**: Click **Select EVS**. On the displayed page, select the EVS disk that meets your requirements and click **OK**. ● **PV Name**: Enter the PV name, which must be unique in the same cluster. ● **Access Mode**: Only **ReadWriteOnce** is available. A storage volume can be mounted to one node in read/write mode. For details, see **Volume Access Modes**. ● **Reclaim Policy**: There are two options: **Retain** and **Delete**. For details, see **PV Reclaim Policy** 2. Click **Create**.
Expanding the capacity of an EVS volume	Quickly expand the capacity of a mounted EVS volume on the CCE console. Only the capacity of pay-per-use EVS volumes can be expanded on the CCE console. To expand the capacity of yearly/monthly EVS volumes, click the volume name to go to the EVS console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Scale-out** in the **Operation** column. 2. Enter the capacity to be added and click **OK**.
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).

Operation	Description	Procedure
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.3.3 Using an EVS Disk Through a Dynamic PV

CCE Autopilot allows you to dynamically create storage volumes from EVS disks by specifying a StorageClass when no EVS disks are available.

Prerequisites

- You have created a cluster.
- If you want to create a cluster by running commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- Not all regions support EVS volumes. View the regions where EVS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where EVS volumes are supported.
- EVS disks cannot be used by multiple workloads, multiple pods of the same workload, or multiple tasks. If EVS disks are used when you create a Deployment, select only one pod for the Deployment.
- The cluster version must be v1.27.8-r0, v1.28.6-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- No more than 10 EVS disks can be attached to each workload in a cluster. If more than 10 EVS disks are attached, the workload may run abnormally.
- Resource tags can be added to dynamically created EVS disks. After they are created, the resource tags cannot be updated on the CCE console. To update them, go to the EVS console. If you use an existing EVS disk to create a PV, you also need to add or update resource tags on the EVS console.

Dynamically Creating EVS Volumes

Specify StorageClass to automatically create EVS disks and storage volumes and mount the volumes to workloads for persistent storage on the console or using kubectl.

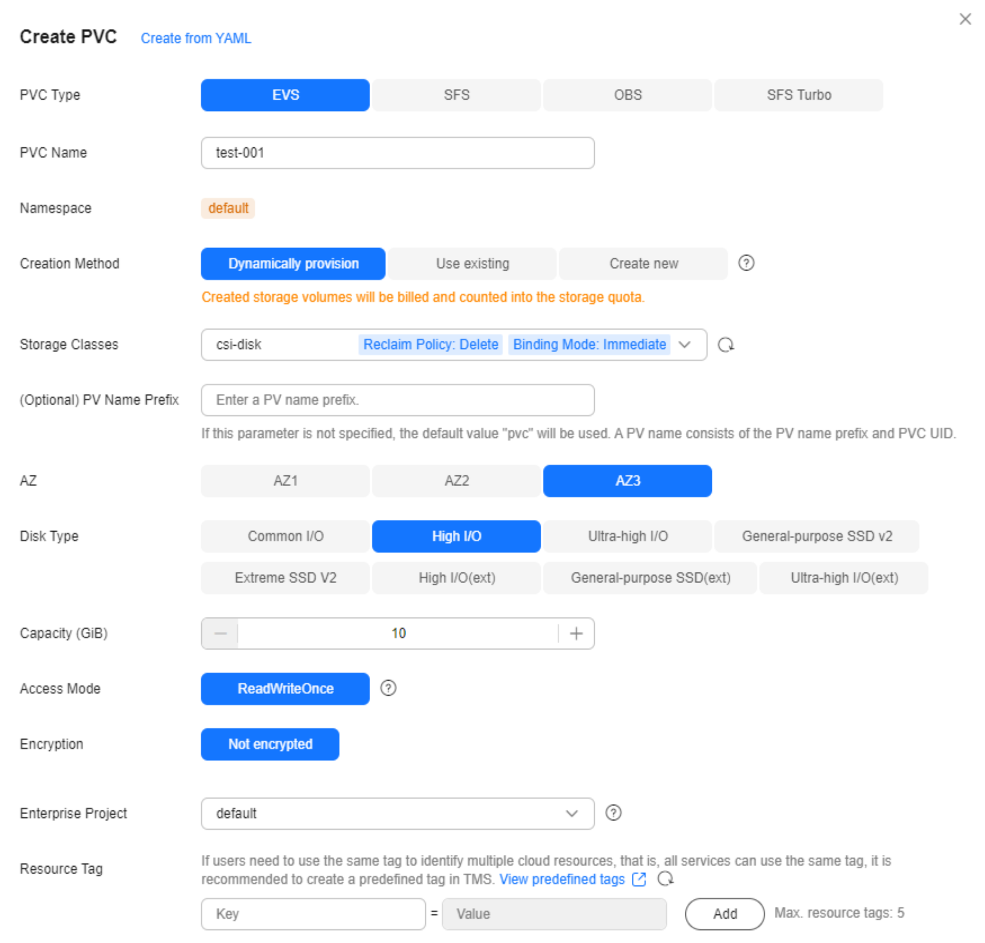
Using the CCE Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Dynamically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Figure 4-5 Creating a PVC



Create PVC
Create from YAML

PVC Type	EVS
PVC Name	test-001
Namespace	default
Creation Method	Dynamically provision
	Created storage volumes will be billed and counted into the storage quota.
Storage Classes	csi-disk Reclaim Policy: Delete Binding Mode: Immediate
(Optional) PV Name Prefix	Enter a PV name prefix.
	If this parameter is not specified, the default value "pvc" will be used. A PV name consists of the PV name prefix and PVC UID.
AZ	AZ3
Disk Type	High I/O
Capacity (GiB)	10
Access Mode	ReadWriteOnce
Encryption	Not encrypted
Enterprise Project	default
Resource Tag	If users need to use the same tag to identify multiple cloud resources, that is, all services can use the same tag, it is recommended to create a predefined tag in TMS. View predefined tags
	Key = Value Add Max. resource tags: 5

Parameter	Description
PVC Type	In this example, select **EVS**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Namespace	A namespace is a conceptual grouping of resources or objects. Each namespace provides isolation for data from other namespaces. After a cluster is created, the default namespace is created by default. If there is no special requirement, select the default namespace.

Parameter	Description
Creation Method	– If no underlying storage is available, select **Dynamically provision** to create a PVC, PV, and underlying storage on the console in cascading mode. – If underlying storage is available, select either **Use existing** or **Create new**. For details about static creation, see **Using an Existing EVS Disk Through a Static PV**. In this example, select **Dynamically provision**.
Storage Classes	The StorageClass for EVS volumes is **csi-disk**.
(Optional) PV Name Prefix	This parameter specifies the name of the underlying storage that is automatically created. The actual underlying storage name is in the format of "PV name prefix + PVC UID". If this parameter is left blank, the default prefix **pvc** will be used. For example, if the storage volume name prefix is set to **test**, the actual underlying storage name is **test-***{UID}*.
AZ	Select the AZ of the EVS disk. After the EVS disk is created, the AZ cannot be changed. In this example, select **AZ3**.
Disk Type	Select an EVS disk type. EVS disk types vary depending on regions. Obtain the available EVS types on the console. In this example, select **High I/O**. **NOTE** General-purpose SSD V2 disks allow you to specify the disk IOPS and throughput. For details, see **EVS performance data**.
Capacity (GiB)	Capacity of the requested storage volume. Set it to **10** in this example.
Access Mode	EVS volumes support only **ReadWriteOnce**, indicating that a storage volume can only be mounted to one pod in read/write mode. For details, see **Volume Access Modes**.
Encryption	Configure whether to encrypt underlying storage. If you select **Enabled (key)**, an encryption key must be configured. Currently, encryption is not supported.
Enterprise Project	The default enterprise project, the enterprise project to which the cluster belongs, or the enterprise project specified by StorageClass is available.

Parameter	Description
Resource Tag	You can add resource tags to classify resources. You can create predefined tags on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. CCE automatically creates system tags **CCE-Cluster-ID=***{Cluster ID}*, **CCE-Cluster-Name=***{Cluster name}*, and **CCE-Namespace=***{Namespace name}*. These tags cannot be modified. **NOTE** After a dynamic PV of the EVS type is created, the resource tags cannot be updated on the CCE console. To update these resource tags, go to the EVS console.

2. Click **Create**.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **StatefulSets** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

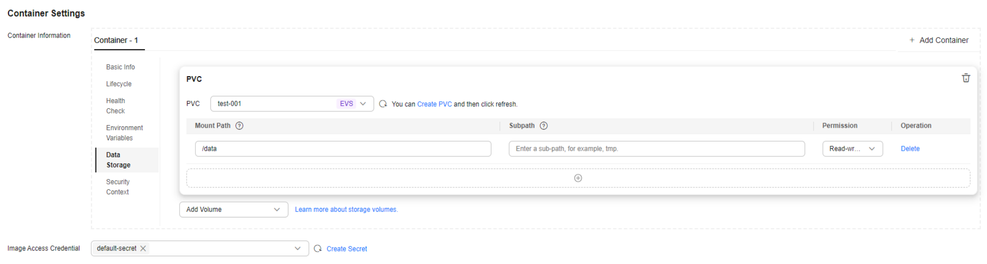
   Table 4-8 describes the parameters for mounting the volume. For details about other parameters, see **Workloads**.

   In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the EVS disk.

   NOTE

   A non-shared EVS disk can be mounted to only one pod. If there are multiple pods, extra pods cannot start normally. Ensure that the number of pods for running a workload is 1 if an EVS disk is mounted.

   Figure 4-6 Parameters for mounting a storage volume



Container Settings

Container Information

Container - 1

+ Add Container

Basic Info
Lifecycle
Health Check
Environment Variables
Data Storage
Security Context

PVC

PVC: test-001 EVS

You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-wr...	Delete

Add Volume

Learn more about storage volumes.

Image Access Credential: default-secret

Create Secret

Table 4-8 Parameters for mounting a storage volume

Parameter	Description
PVC	Select an existing EVS volume. An EVS volume can only be mounted to one workload.
Mount Path	Enter a mount path, for example, **/data**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. **data**, for example, indicates that data in the mount path of the container is stored in the **data** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Use **StorageClass** to dynamically create a PVC and PV.

1. Create the **pvc-evs-auto.yaml** file.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-evs-auto
  namespace: default
  annotations:
      everest.io/disk-volume-type: SAS    # EVS disk type
    everest.io/enterprise-project-id: <your_project_id> # (Optional) Enterprise project ID. If an
enterprise project is specified, use the same enterprise project when creating a PVC. Otherwise, the
PVC cannot be bound to a PV.

    everest.io/disk-volume-tags: '{"key1":"value1","key2":"value2"}' # (Optional) Custom resource tags

    everest.io/disk-iops: '3000'      # (Optional) IOPS of only a GPSSD2 EVS disk
    everest.io/disk-throughput: '125' # (Optional) Throughput of only a GPSSD2 EVS disk

    everest.io/csi.volume-name-prefix: test  # (Optional) Storage volume name prefix of the
automatically created underlying storage

  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one
where the cluster is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>      # Replace the AZ with the one where
the EVS disk is located.
spec:
  accessModes:
  - ReadWriteOnce              # The value must be ReadWriteOnce for EVS disks.
  resources:
    requests:
      storage: 10Gi            # EVS disk capacity, in GiB. The value ranges from 1 to 32768.
  storageClassName: csi-disk    # The StorageClass of the EVS disk
```

Table 4-9 Key parameters

Parameter	Mandatory	Description
failure-domain.beta.kubernetes.io/region	Yes	Region of the cluster. For example, **cn-east-3**. For details, see **Regions and Endpoints**.
failure-domain.beta.kubernetes.io/zone	Yes	AZ of the EVS disk. You can obtain all supported AZs by **querying the AZ list**.
everest.io/disk-volume-type	Yes	EVS disk type. All letters are in uppercase. EVS disk types vary depending on regions. Obtain the available EVS types on the console. – **SAS**: high I/O – **SSD**: ultra-high I/O – **GPSSD**: general-purpose SSD – **ESSD**: extreme SSD – GPSSD2: general-purpose SSD v2. You need to specify the **everest.io/disk-iops** and **everest.io/disk-throughput** annotations.

Parameter	Mandatory	Description
everest.io/disk-iops	No	Preconfigured IOPS, which is supported only by general-purpose SSD v2 EVS disks. – The IOPS of general-purpose SSD v2 EVS disks ranges from 3000 to 128000, and the maximum value is 500 times of the capacity (GiB). If the IOPS of general-purpose SSD v2 disks is greater than 3000, extra IOPS will be billed. For details, see **Price Calculator**.
everest.io/disk-throughput	No	Preconfigured throughput, which is supported only by general-purpose SSD v2 EVS disks. The value ranges from 125 MiB/s to 1,000 MiB/s. The maximum value is a quarter of IOPS. If the throughput is greater than 125 MiB/s, extra throughput will be billed. For details, see **Price Calculator**.
everest.io/enterprise-project-id	No	Enterprise project ID of the EVS disk. This parameter is optional. If an enterprise project is specified, use the same enterprise project when creating a PVC, or the PVC cannot be bound to a PV. To obtain an enterprise project ID, log in to the EPS console, click the name of the target enterprise project, and copy the enterprise project ID.
everest.io/disk-volume-tags	No	You can add resource tags to classify resources. This field is optional. You can create **predefined tags** on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. CCE automatically creates system tags **CCE-Cluster-ID=***{Cluster ID}*, **CCE-Cluster-Name=***{Cluster name}*, and **CCE-Namespace=***{Namespace name}*. These tags cannot be modified.

Parameter	Mandatory	Description
everest.io/csi.volume-name-prefix	No	This parameter is optional. It specifies the name of the underlying storage that is automatically created. The actual underlying storage name is in the format of "PV name prefix + PVC UID". If this parameter is left blank, the default prefix **pvc** will be used. Enter 1 to 26 characters that cannot start or end with a hyphen (-). Only lowercase letters, digits, and hyphens (-) are allowed. For example, if the storage volume name prefix is set to **test**, the actual underlying storage name is **test-***{UID}*.
storage	Yes	Requested PVC capacity, in Gi. The value ranges from **1** to **32768**.
storageClassName	Yes	The StorageClass for EVS volumes is **csi-disk**.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-evs-auto.yaml
```

Step 3 Create a workload.

1. Create a file named **web-evs-auto.yaml**. In this example, the EVS volume is mounted to the **/data** path.
```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: web-evs-auto
  namespace: default
spec:
  replicas: 1
  selector:
    matchLabels:
      app: web-evs-auto
  serviceName: web-evs-auto   # Headless Service name
  template:
    metadata:
      labels:
        app: web-evs-auto
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-disk    # Volume name, which must be the same as the volume name in the
volumes field.
          mountPath: /data  #Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-disk    # Custom volume name
          persistentVolumeClaim:
            claimName: pvc-evs-auto    # Name of the created PVC.
---
apiVersion: v1
kind: Service
```

```
metadata:
  name: web-evs-auto   # Headless Service name
  namespace: default
  labels:
    app: web-evs-auto
spec:
  selector:
    app: web-evs-auto
  clusterIP: None
  ports:
    - name: web-evs-auto
      targetPort: 80
      nodePort: 0
      port: 80
      protocol: TCP
  type: ClusterIP
```

2. Run the following command to create a workload that the EVS volume is mounted to:

```
kubectl apply -f web-evs-auto.yaml
```

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Verifying Data Persistence

Step 1 View the deployed application and files stored in the EVS volume.

1. Run the following command to view the pod:

```
kubectl get pod | grep web-evs-auto
```

   Expected output:

```
web-evs-auto-0                   1/1     Running   0               38s
```

2. Run the following command to verify that the EVS volume has been mounted to the **/data** path:

```
kubectl exec web-evs-auto-0 -- df | grep data
```

   Expected output:

```
/dev/sdc              10255636     36888  10202364   0% /data
```

3. Run the following command to view the created file in the **/data** path:

```
kubectl exec web-evs-auto-0 -- ls /data
```

   Expected output:

```
lost+found
```

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-evs-auto-0 --  touch /data/static
```

Step 3 Run the following command to view the created file in the **/data** path:

```
kubectl exec web-evs-auto-0 -- ls /data
```

Expected output:

```
lost+found
static
```

Step 4 Run the following command to delete the pod named **web-evs-auto-0**:

```
kubectl delete pod web-evs-auto-0
```

Expected output:

```
pod "web-evs-auto-0" deleted
```

Step 5 The StatefulSet controller automatically creates a replica with the same name as the pod. Run the following command to check whether the file in the **/data** path has been modified:

```
kubectl exec web-evs-auto-0 -- ls /data
```

Expected output:

```
lost+found
static
```

The **static** file is retained, indicating that the data can be stored persistently.

----End

Related Operations

You can also perform the operations described in **Table 4-10**.

Table 4-10 Related operations

Operation	Description	Procedure
Expanding the capacity of an EVS volume	Quickly expand the capacity of a mounted EVS volume on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Scale-out** in the **Operation** column. 2. Enter the capacity to be added and click **OK**.
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.3.4 Dynamically Mounting an EVS Volume to a StatefulSet

Scenario

Dynamic mounting is available only for **creating a StatefulSet**. It is implemented through a volume claim template (the **volumeClaimTemplates** field) and depends on dynamic creation of PVs through StorageClass. In this mode, each pod in a multi-pod StatefulSet is associated with a unique PVC and PV. After a pod is rescheduled, the original PV can still be mounted to it based on the PVC name. In the common mounting mode for a Deployment, if ReadWriteMany is supported, multiple pods of the Deployment will be mounted to the same underlying storage.

NOTE

When updating a StatefulSet in Kubernetes, it is not allowed to add or delete the **volumeClaimTemplates** field. This field can only be configured during the creation of the StatefulSet.

Prerequisites

- You have created a cluster.
- If you want to create a cluster by running commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- Not all regions support EVS volumes. View the regions where EVS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where EVS volumes are supported.
- The cluster version must be v1.27.8-r0, v1.28.6-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- No more than 10 EVS disks can be attached to each workload in a cluster. If more than 10 EVS disks are attached, the workload may run abnormally.

Dynamically Creating and Mounting an EVS Volume

By specifying the StorageClass, you can create a unique PVC and PV for each pod of a StatefulSet on the console or using kubectl to ensure data durability after rescheduling.

Using the CCE Console

Dynamically mount and use storage volumes. For details about other parameters, see **Creating a StatefulSet**.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

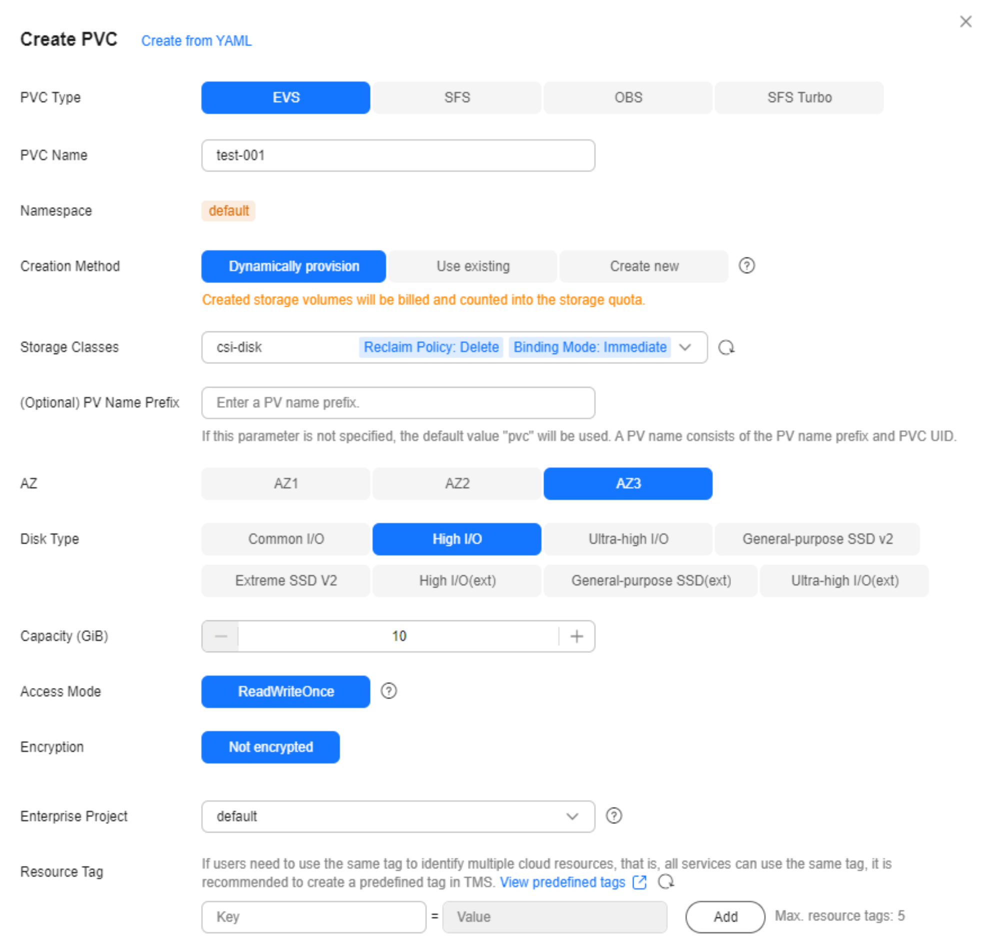
Step 2 In the navigation pane on the left, choose **Workloads**. Then click the **StatefulSets** tab.

Step 3 Click **Create Workload** in the upper right corner. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **VolumeClaimTemplate**.

Step 4 In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Click **Create**.

Figure 4-7 Creating a PVC


Create PVC Create from YAML

PVC Type: EVS | SFS | OBS | SFS Turbo (selected: EVS)

PVC Name: test-001

Namespace: default

Creation Method: Dynamically provision | Use existing | Create new (selected: Dynamically provision)

Created storage volumes will be billed and counted into the storage quota.

Storage Classes: csi-disk Reclaim Policy: Delete Binding Mode: Immediate

(Optional) PV Name Prefix: Enter a PV name prefix.

If this parameter is not specified, the default value "pvc" will be used. A PV name consists of the PV name prefix and PVC UID.

AZ: AZ1 | AZ2 | AZ3 (selected: AZ3)

Disk Type: Common I/O | High I/O | Ultra-high I/O | General-purpose SSD v2 | Extreme SSD V2 | High I/O(ext) | General-purpose SSD(ext) | Ultra-high I/O(ext) (selected: High I/O)

Capacity (GiB): 10

Access Mode: ReadWriteOnce

Encryption: Not encrypted

Enterprise Project: default

Resource Tag: If users need to use the same tag to identify multiple cloud resources, that is, all services can use the same tag, it is recommended to create a predefined tag in TMS. View predefined tags

Key = Value Add Max. resource tags: 5

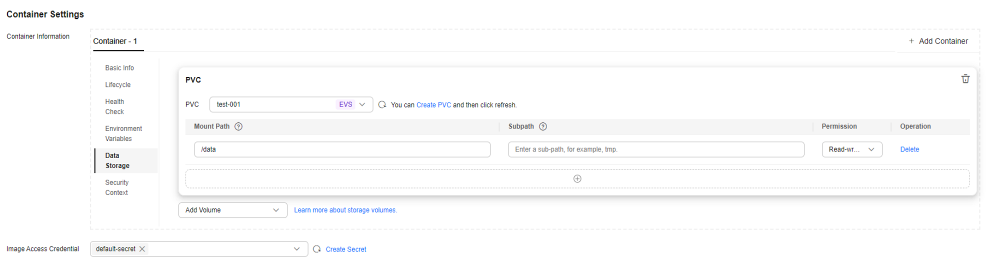
Parameter	Description
PVC Type	In this example, select **EVS**.
PVC Name	Enter the name of the PVC. After a PVC is created, a suffix is automatically added based on the number of pods. The format is <*Custom PVC name*>-<*Serial number*>, for example, example-0.
Namespace	A namespace is a conceptual grouping of resources or objects. Each namespace provides isolation for data from other namespaces. After a cluster is created, the default namespace is created by default. If there is no special requirement, select the default namespace.
Creation Method	You can select **Dynamically provision** to create a PVC, PV, and underlying storage on the console in cascading mode.
Storage Classes	The StorageClass for EVS volumes is **csi-disk**.

Parameter	Description
(Optional) PV Name Prefix	This parameter specifies the name of the underlying storage that is automatically created. The actual underlying storage name is in the format of "PV name prefix + PVC UID". If this parameter is left blank, the default prefix **pvc** will be used. For example, if the storage volume name prefix is set to **test**, the actual underlying storage name is **test-***{UID}*.
AZ	Select the AZ of the EVS disk. After the EVS disk is created, the AZ cannot be changed. In this example, select **AZ3**.
Disk Type	Select an EVS disk type. EVS disk types vary depending on regions. Obtain the available EVS types on the console. In this example, select **High I/O**. **NOTE** General-purpose SSD V2 disks allow you to specify the disk IOPS and throughput. For details, see **EVS performance data**.
Capacity (GiB)	Capacity of the requested storage volume. Set it to **10** in this example.
Access Mode	EVS volumes support only **ReadWriteOnce**, indicating that a storage volume can only be mounted to one pod in read/write mode. For details, see **Volume Access Modes**.
Encryption	Configure whether to encrypt underlying storage. If you select **Enabled (key)**, an encryption key must be configured. Currently, encryption is not supported.
Enterprise Project	The default enterprise project, the enterprise project to which the cluster belongs, or the enterprise project specified by StorageClass is available.
Resource Tag	You can add resource tags to classify resources. You can create **predefined tags** on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. CCE automatically creates system tags **CCE-Cluster-ID=***{Cluster ID}*, **CCE-Cluster-Name=***{Cluster name}*, and **CCE-Namespace=***{Namespace name}*. These tags cannot be modified. **NOTE** After a dynamic PV of the EVS type is created, the resource tags cannot be updated on the CCE console. To update these resource tags, go to the EVS console.

Step 5 Enter the path to which the volume is mounted.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the EVS disk.

Figure 4-8 Parameters for mounting a storage volume


Container Settings
Container Information: Container - 1
PVC: test-001 (EVS). You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-wr...	Delete

Image Access Credential: default-secret

Table 4-11 Parameters for mounting a storage volume

Parameter	Description
Mount Path	Enter a mount path, for example, **/data**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. **data**, for example, indicates that data in the mount path of the container is stored in the **data** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	● **Read-only**: You can only read the data in the mounted volume. ● **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

Step 6 Configure other parameters and click **Create Workload**. After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a file named **statefulset-evs.yaml**. In this example, the EVS volume is mounted to the **/data** path.

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: statefulset-evs
  namespace: default
spec:
  selector:
    matchLabels:
      app: statefulset-evs
  template:
    metadata:
      labels:
        app: statefulset-evs
    spec:
      containers:
        - name: container-1
          image: nginx:latest
          volumeMounts:
            - name: pvc-disk          # The value must be the same as that in the volumeClaimTemplates field.
              mountPath: /data         # Location where the storage volume is mounted
      imagePullSecrets:
        - name: default-secret
  serviceName: statefulset-evs         # Headless Service name
  replicas: 2
  volumeClaimTemplates:
    - apiVersion: v1
      kind: PersistentVolumeClaim
      metadata:
        name: pvc-disk
        namespace: default
        annotations:
          everest.io/disk-volume-type: SAS    # EVS disk type
          everest.io/enterprise-project-id: <your_project_id>  # (Optional) Enterprise project ID. If an enterprise
project is specified, use the same enterprise project when creating a PVC. Otherwise, the PVC cannot be
bound to a PV.
          everest.io/disk-iops: '3000'      # (Optional) IOPS of only a GPSSD2 EVS disk
          everest.io/disk-throughput: '125' # (Optional) Throughput of only a GPSSD2 EVS disk
          everest.io/csi.volume-name-prefix: test  # (Optional) Storage volume name prefix of the
automatically created underlying storage
        labels:
          failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one
where the cluster is located.
          failure-domain.beta.kubernetes.io/zone: <your_zone>       # Replace the AZ with the one where
the EVS disk is located.
      spec:
        accessModes:
          - ReadWriteOnce               # The value must be ReadWriteOnce for EVS disks.
        resources:
          requests:
            storage: 10Gi             # EVS disk capacity, in GiB. The value ranges from 1 to 32768.
        storageClassName: csi-disk    # The StorageClass of the EVS disk
---
apiVersion: v1
kind: Service
metadata:
  name: statefulset-evs   # Headless Service name
  namespace: default
  labels:
    app: statefulset-evs
spec:
  selector:
    app: statefulset-evs
```

```
clusterIP: None
ports:
  - name: statefulset-evs
    targetPort: 80
    nodePort: 0
    port: 80
    protocol: TCP
type: ClusterIP
```

Table 4-12 Key parameters

Parameter	Mandatory	Description
failure-domain.beta.kubernetes.io/region	Yes	Region of the cluster. For example, **cn-east-3**. For details, see **Regions and Endpoints**.
failure-domain.beta.kubernetes.io/zone	Yes	AZ of the EVS disk. You can obtain all supported AZs by **querying the AZ list**.
everest.io/disk-volume-type	Yes	EVS disk type. All letters are in uppercase. EVS disk types vary depending on regions. Obtain the available EVS types on the console. ● **SAS**: high I/O ● **SSD**: ultra-high I/O ● **GPSSD**: general-purpose SSD ● **ESSD**: extreme SSD ● GPSSD2: general-purpose SSD v2. You need to specify the **everest.io/disk-iops** and **everest.io/disk-throughput** annotations.
everest.io/disk-iops	No	Preconfigured IOPS, which is supported only by general-purpose SSD v2 EVS disks. ● The IOPS of general-purpose SSD v2 EVS disks ranges from 3000 to 128000, and the maximum value is 500 times of the capacity (GiB). If the IOPS of general-purpose SSD v2 disks is greater than 3000, extra IOPS will be billed. For details, see **Price Calculator**.
everest.io/disk-throughput	No	Preconfigured throughput, which is supported only by general-purpose SSD v2 EVS disks. The value ranges from 125 MiB/s to 1,000 MiB/s. The maximum value is a quarter of IOPS. If the throughput is greater than 125 MiB/s, extra throughput will be billed. For details, see **Price Calculator**.

Parameter	Mandatory	Description
everest.io/enterprise-project-id	No	Enterprise project ID of the EVS disk. This parameter is optional. If an enterprise project is specified, use the same enterprise project when creating a PVC, or the PVC cannot be bound to a PV. To obtain an enterprise project ID, log in to the EPS console, click the name of the target enterprise project, and copy the enterprise project ID.
everest.io/disk-volume-tags	No	You can add resource tags to classify resources. This field is optional. You can create **predefined tags** on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. CCE automatically creates system tags **CCE-Cluster-ID=***{Cluster ID}*, **CCE-Cluster-Name=***{Cluster name}*, and **CCE-Namespace=***{Namespace name}*. These tags cannot be modified.
everest.io/csi.volume-name-prefix	No	This parameter is optional. It specifies the name of the underlying storage that is automatically created. The actual underlying storage name is in the format of "PV name prefix + PVC UID". If this parameter is left blank, the default prefix **pvc** will be used. Enter 1 to 26 characters that cannot start or end with a hyphen (-). Only lowercase letters, digits, and hyphens (-) are allowed. For example, if the storage volume name prefix is set to **test**, the actual underlying storage name is **test-***{UID}*.
storage	Yes	Requested PVC capacity, in Gi. The value ranges from **1** to **32768**.
storageClassName	Yes	The StorageClass for EVS volumes is **csi-disk**.

Step 3 Run the following command to create a workload that the EVS volume is mounted to:

```
kubectl apply -f statefulset-evs.yaml
```

After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence**.

----End

Verifying Data Persistence

Step 1 View the deployed application and files stored in the EVS volume.

1. Run the following command to view the pod:
   ```
   kubectl get pod | grep statefulset-evs
   ```
   Expected output:
   ```
   statefulset-evs-0          1/1     Running   0             45s
   statefulset-evs-1          1/1     Running   0             28s
   ```
2. Run the following command to verify that the EVS volume has been mounted to the **/data** path:
   ```
   kubectl exec statefulset-evs-0 -- df | grep data
   ```
   Expected output:
   ```
   /dev/sdd              10255636     36888  10202364   0% /data
   ```
3. Run the following command to view the created file in the **/data** path:
   ```
   kubectl exec statefulset-evs-0 -- ls /data
   ```
   Expected output:
   ```
   lost+found
   ```

Step 2 Run the following command to create a file named **static** in the **/data** path:
```
kubectl exec statefulset-evs-0 --  touch /data/static
```

Step 3 Run the following command to view the created file in the **/data** path:
```
kubectl exec statefulset-evs-0 -- ls /data
```
Expected output:
```
lost+found
static
```

Step 4 Run the following command to delete the pod named **web-evs-auto-0**:
```
kubectl delete pod statefulset-evs-0
```
Expected output:
```
pod "statefulset-evs-0" deleted
```

Step 5 The StatefulSet controller automatically creates a replica with the same name as the pod. Run the following command to check whether the file in the **/data** path has been modified:
```
kubectl exec statefulset-evs-0 -- ls /data
```
Expected output:
```
lost+found
static
```
The **static** file is retained, indicating that the data can be stored persistently.

----End

Related Operations

You can also perform the operations described in **Table 4-13**.

Table 4-13 Related operations

Operation	Description	Procedure
Expanding the capacity of an EVS volume	Quickly expand the capacity of a mounted EVS volume on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Scale-out** in the **Operation** column. 2. Enter the capacity to be added and click **OK**.
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.3.5 Expanding the Capacity of an EVS Volume

If the EVS volume mounted to a workload does not have enough space, you can increase its capacity by expanding it. This section describes how to expand the capacity of an EVS volume through the console.

Prerequisites

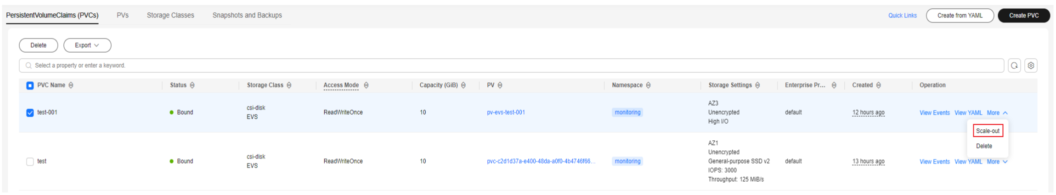
You have created a cluster of v1.27.8-r0, v1.28.6-r0, or later.

Procedure

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Scale-out** in the **Operation** column.

Figure 4-9 Expanding the capacity of an EVS volume



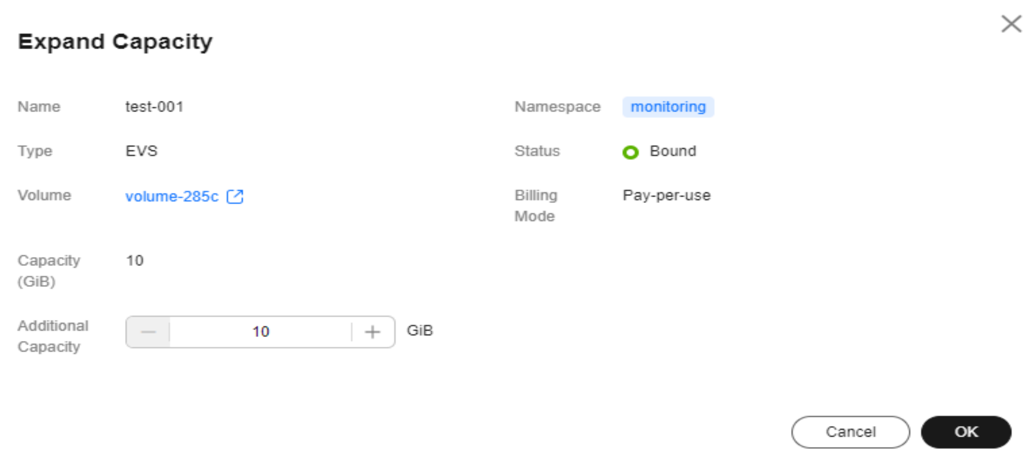
PVC Name	Status	Storage Class	Access Mode	Capacity (GiB)	PV	Namespace	Storage Settings	Enterprise Pr...	Created	Operation
test-001	Bound	csi-disk EVS	ReadWriteOnce	10	pv-evs-test-001	monitoring	AZ3 Unencrypted High I/O	default	12 hours ago	View Events View YAML More
test	Bound	csi-disk EVS	ReadWriteOnce	10	pvc-c2d1d37a-e400-48da-a0f0-4b4746f66...	monitoring	AZ1 Unencrypted General-purpose SSD v2 IOPS: 3000 Throughput: 125 MiB/s	default	13 hours ago	View Events View YAML More

Step 3 Enter the capacity to be added and click **OK**. In this example, 10 GiB is added.

NOTE

The capacity can only be increased.

Figure 4-10 Configuring the capacity to be expanded



Expand Capacity

Name	test-001	Namespace	monitoring
Type	EVS	Status	Bound
Volume	volume-285c	Billing Mode	Pay-per-use
Capacity (GiB)	10		
Additional Capacity	10 GiB		

Step 4 Verify that the expansion is successful. On the **PVCs** tab, you can see that the PVC capacity is expanded from 10 GiB to 20 GiB.

----End

4.3.6 Creating a Snapshot and Backup

CCE Autopilot clusters allow you to create snapshots for storage volumes of the EVS type. A snapshot is a complete copy or image of the disk data at a specific point in time, which can be used for data disaster recovery. If data is lost, you can roll back the disk data to the state when a snapshot was created.

You can create snapshots to rapidly save the disk data at a certain time. In addition, you can use snapshots to create disks so that the created disks will contain the snapshot data in the beginning.

Prerequisites

- You have created a cluster.
- If you want to create a cluster by running commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- Not all regions support EVS volumes. View the regions where EVS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where EVS volumes are supported.
- The cluster version must be v1.27.8-r0, v1.28.6-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- The subtype (such as common I/O, high I/O, or ultra-high I/O), disk mode (SCSI or VBD), data encryption, sharing status, and capacity of an EVS disk

created from a snapshot must be the same as those of the disk associated with the snapshot. These attributes cannot be modified after being checked or configured.

- Snapshots can be created only for EVS disks that are available or in use, and a maximum of seven snapshots can be created for a single EVS disk.
- The snapshots of encrypted disks are stored encrypted, and those of non-encrypted disks are stored non-encrypted.

Scenario

Snapshots can help meet your following needs:

- **Routine data backup**

  You can create snapshots for EVS disks regularly and use snapshots to recover your data in case that data loss or data inconsistency occurred due to misoperations, viruses, or attacks.

- **Rapid data restoration**

  You can create a snapshot or multiple snapshots before an OS change, application software upgrade, or a service data migration. If an exception occurs during the upgrade or migration, service data can be rapidly restored to the time point when the snapshot was created.

  For example, a fault occurred on system disk A of ECS A, and therefore ECS A cannot be started. Because system disk A is already faulty, the data on system disk A cannot be restored by rolling back snapshots. In this case, you can use an existing snapshot of system disk A to create EVS disk B and attach it to ECS B that is running properly. Then, ECS B can read data from system disk A using EVS disk B.

  NOTE

  The snapshots provided by CCE Autopilot are the same as the CSI snapshots provided by the Kubernetes community. They can only be used to create EVS disks, but cannot be used to roll back to source EVS disks.

- **Rapid deployment of multiple services**

  You can use a snapshot to create multiple EVS disks containing the same initial data, and these disks can be used as data resources for various services, for example, data mining, report query, and development and testing. This method protects the initial data and creates disks rapidly, meeting the diversified service data requirements.

Creating a Snapshot

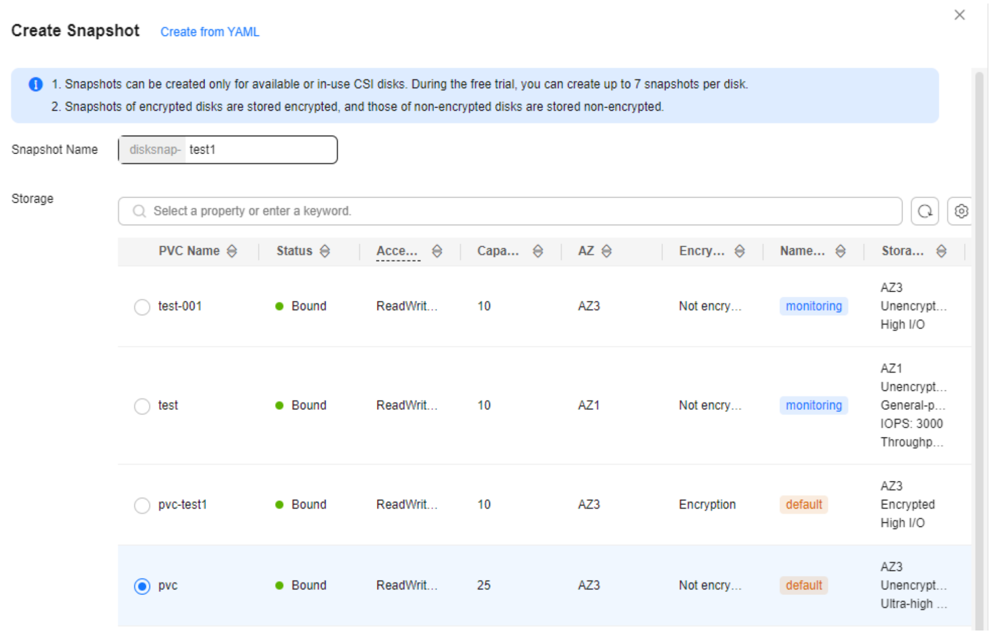
Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console. In the navigation pane on the left, choose **Storage**. Then click the **Snapshots and Backups** tab.

Step 3 In the upper right corner, click **Create Snapshot**. In the displayed dialog box, configure the parameters.

Figure 4-11 Creating a snapshot



1. Snapshots can be created only for available or in-use CSI disks. During the free trial, you can create up to 7 snapshots per disk.
2. Snapshots of encrypted disks are stored encrypted, and those of non-encrypted disks are stored non-encrypted.

Snapshot Name: disksnap- test1

Storage: Select a property or enter a keyword.

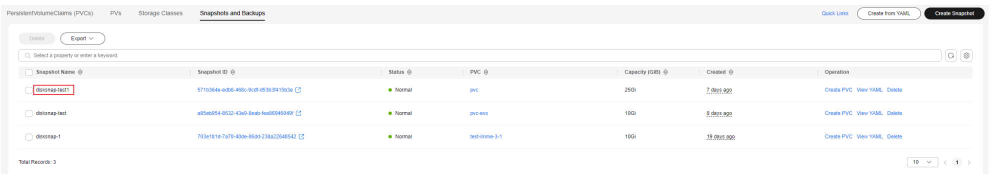
	PVC Name	Status	Acce...	Capa...	AZ	Encry...	Name...	Stora...
○	test-001	● Bound	ReadWrit...	10	AZ3	Not encry...	monitoring	AZ3 Unencrypt... High I/O
○	test	● Bound	ReadWrit...	10	AZ1	Not encry...	monitoring	AZ1 Unencrypt... General-p... IOPS: 3000 Throughp...
○	pvc-test1	● Bound	ReadWrit...	10	AZ3	Encryption	default	AZ3 Encrypted High I/O
◉	pvc	● Bound	ReadWrit...	25	AZ3	Not encry...	default	AZ3 Unencrypt... Ultra-high ...

Table 4-14 Parameters for creating a snapshot

Parameter	Description
Snapshot Name	Enter the snapshot name, for example, **disksnap-test1**.
Storage	Select the PVC for which you want to create a snapshot. Only EVS PVCs can be selected. In this example, select **pvc**.

Step 4 Click **Create**. On the **Snapshots and Backups** tab, view details about the snapshot.

Figure 4-12 Snapshot details



Snapshot Name	Snapshot ID	Status	PVC	Capacity (GiB)	Created	Operation
disksnap-test1	571b364e-edb8-468c-9cdf-d53b3f415b3e	● Normal	pvc	25Gi	7 days ago	Create PVC View YAML Delete
disksnap-test	a85eb954-8632-43e9-8eab-fea86946049f	● Normal	pvc-evs	10Gi	8 days ago	Create PVC View YAML Delete
disksnap-1	703e181d-7a70-40de-86dd-238a22648542	● Normal	test-imme-3-1	10Gi	19 days ago	Create PVC View YAML Delete

Total Records: 3

----End

Creating from YAML

```
kind: VolumeSnapshot
apiVersion: snapshot.storage.k8s.io/v1beta1
metadata:
  name: disksnap-test1 # Snapshot name
  namespace: default
spec:
  source:
    persistentVolumeClaimName: pvc     # PVC name. Only an EVS PVC can be selected.
  volumeSnapshotClassName: csi-disk-snapclass
```

Using a Snapshot to Create a PVC

The disk type, encryption setting, and disk mode of the created EVS PVC are consistent with those of the snapshot's source EVS disk.

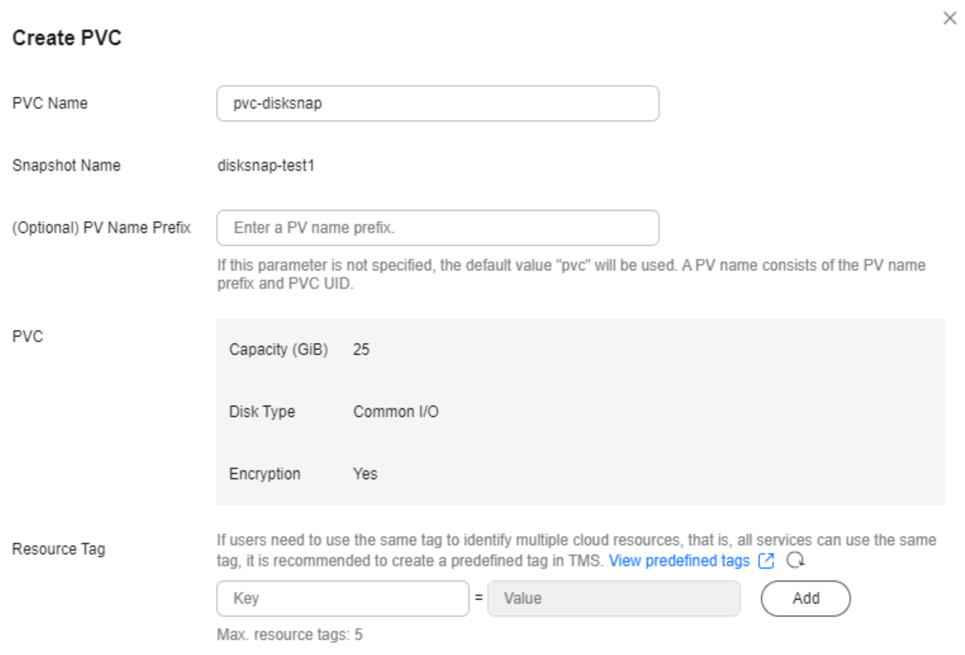
Using the CCE Console

Step 1 Log in to the CCE console.

Step 2 Click the cluster name to go to the cluster console. In the navigation pane on the left, choose **Storage**. Then click the **Snapshots and Backups** tab.

Step 3 Locate the snapshot that you want to use for creating a PVC, click **Create PVC**, and configure PVC parameters in the displayed dialog box.

Figure 4-13 Using a snapshot to create a PVC



Create PVC

PVC Name: pvc-disksnap

Snapshot Name: disksnap-test1

(Optional) PV Name Prefix: Enter a PV name prefix.

If this parameter is not specified, the default value "pvc" will be used. A PV name consists of the PV name prefix and PVC UID.

PVC
Capacity (GiB): 25
Disk Type: Common I/O
Encryption: Yes

Resource Tag: If users need to use the same tag to identify multiple cloud resources, that is, all services can use the same tag, it is recommended to create a predefined tag in TMS. View predefined tags

Key = Value Add

Max. resource tags: 5

Table 4-15 Parameters for creating a PVC using a snapshot

Parameter	Description
PVC Name	Enter a PVC name, for example, **pvc-disksnap**.
(Optional) PV Name Prefix	This parameter specifies the name of the underlying storage that is automatically created. The actual underlying storage name is in the format of "PV name prefix + PVC UID". If this parameter is left blank, the default prefix **pvc** will be used. For example, if the storage volume name prefix is set to **test**, the actual underlying storage name is **test-***{UID}*.

Parameter	Description
Resource Tag	You can add resource tags to classify resources. You can create **predefined tags** on the TMS console. The predefined tags are available to all resources that support tags. You can use predefined tags to improve the tag creation and resource migration efficiency. For details, see **Creating Predefined Tags**. CCE automatically creates system tags **CCE-Cluster-ID=***{Cluster ID}*, **CCE-Cluster-Name=***{Cluster name}*, and **CCE-Namespace=***{Namespace name}*. These tags cannot be modified.

Step 4 Click **Create**. Click the **PVCs** tab to view the created PVC.

----End

Creating from YAML

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-disksnap # PVC name
  namespace: default
  annotations:
    everest.io/disk-volume-type: SSD     # EVS disk type, which must be the same as that of the snapshot's
source EVS disk.
    everest.io/disk-iops: '3000'      # (Optional) IOPS of only a GPSSD2 EVS disk
    everest.io/disk-throughput: '125' # (Optional) Throughput of only a GPSSD2 EVS disk
  labels:
    failure-domain.beta.kubernetes.io/region: <your_region>   # Replace the region with the one where
the EVS disk is located.
    failure-domain.beta.kubernetes.io/zone: <your_zone>       # Replace the AZ with the one where the
EVS disk is located.
spec:
  accessModes:
  - ReadWriteOnce
  resources:
    requests:
      storage: 10Gi
  storageClassName: csi-disk
  dataSource:
    name: disksnap-test1            # Snapshot name
    kind: VolumeSnapshot
    apiGroup: snapshot.storage.k8s.io
```

4.4 Scalable File Service

4.4.1 SFS Volume Overview

Introduction

CCE Autopilot allows you to mount a volume that is created from a Scalable File Service (SFS) file system to a container for persistent data storage. SFS volumes are commonly used in ReadWriteMany scenarios for large-capacity expansion and cost-sensitive services, such as media processing, content management, big data analysis, and workload analysis. For services with a massive number of small files, SFS Turbo file systems are recommended.

NOTE

View the regions where SFS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS volumes are supported.

Expandable to petabytes, SFS provides fully hosted shared file storage, highly available and stable to handle data- and bandwidth-intensive applications

- **Standard file protocols**: You can mount file systems as volumes to servers, the same as using local directories.
- **Data sharing**: The same file system can be mounted to multiple servers, so that data can be shared.
- **Private network**: Users can access data only in private networks of data centers.
- **Capacity and performance**: The capacity of a single file system is high (PB level) and the performance is excellent (ms-level read/write I/O latency).
- **Use cases**: Deployments and StatefulSets in the ReadWriteMany mode, and jobs created for high-performance computing (HPC), media processing, content management, web services, big data analysis, and workload process analysis

Performance

SFS provides two types of file systems: General Purpose File System (formerly SFS 3.0) and SFS Capacity-Oriented. CCE Autopilot clusters support only General Purpose File System (formerly SFS 3.0). For details about the performance of General Purpose File System, see **File System Types**.

Table 4-16 Performance

Parameter	General Purpose File System (formerly SFS 3.0)
Maximum bandwidth	1.25 TB/s
Maximum IOPS	Million
Latency	10 ms
Maximum capacity	EB

Application Scenarios

SFS supports the following mounting modes based on application scenarios:

- **Using an Existing File System Through a Static PV**: static provisioning. You use an existing file system to create a PV and then mount the PV to the workload through a PVC. This mode applies to scenarios where the underlying storage is available or billed on a yearly/monthly basis.
- **Using an SFS File System Through a Dynamic PV**: dynamic provisioning. You do not need to create volumes in advance. Instead, you need to specify a StorageClass during PVC creation. A file system and a PV will be automatically created. This mode applies to scenarios where no underlying storage is available.

Billing

- For SFS volumes **automatically created** using StorageClass, the default billing mode is **Pay-per-use**. You are billed based on the storage capacity and duration. For SFS pricing details, see **Billing**.
- If you want to be billed on a yearly/monthly basis, **use existing SFS volumes**.

4.4.2 Using an Existing File System Through a Static PV

SFS is a network-attached storage (NAS) that provides shared, scalable, and high-performance file storage. It applies to large-capacity expansion and cost-sensitive services. You can mount the PVs created from general-purpose file systems (formerly SFS 3.0) to pods in CCE Autopilot clusters for file storage. This section describes how to use an existing file system to statically create PVs and PVCs and implement data persistence and sharing in workloads.

Prerequisites

- If you want to create a cluster using commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.
- You have created a general purpose file system (formerly SFS 3.0) that is in the same VPC as the cluster.
- You have configured a VPC endpoint required by the general purpose file system (formerly SFS 3.0). For details, see **Configuring a VPC Endpoint**.

Constraints

- Not all regions support volumes created from file systems. View the regions where SFS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS volumes are supported.
- Multiple PVs can use the same general purpose file system (formerly SFS 3.0) if the following restrictions are meet:
  - All PVCs or PVs that use the same file system cannot be mounted to a pod. This will result in a pod startup failure because not all PVCs can be mounted to the pod due to the same **volumeHandle** values of these PVs.
  - The **persistentVolumeReclaimPolicy** parameter in the PVs should be set to **Retain**. If any other value is used, when a PV is deleted, the associated

underlying SFS file system may be deleted. In this case, other PVs associated with the underlying file system malfunction.

  – When a file system is repeatedly used, enable isolation and protection for ReadWriteMany at the application layer to prevent data overwriting and loss.

- If a general-purpose file system (formerly SFS 3.0) is used, the owner group and permission of the mount point cannot be modified.
- If a general-purpose file system (formerly SFS 3.0) is used, there may be a latency when the PVCs or PVs are created or deleted. The billing duration depends on the time when the file systems are created or deleted on the SFS console.
- If the reclamation policy of the volumes created from general purpose file systems (formerly SFS 3.0) is set to **Delete**, the volumes cannot be reclaimed automatically. Before deleting a PV or PVC, you are required to delete all files in the file systems manually.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Statically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **SFS**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Creation Method	– If underlying storage is available, create a PV or use an existing PV to statically create a PVC. – If no underlying storage is available, select **Dynamically provision**. For details, see **Using an SFS File System Through a Dynamic PV**. In this example, select **Create new** to create a PV and PVC at the same time on the console.
PV[a]	Select an existing PV in the cluster. Create a PV in advance. For details, see "Creating a storage volume" in **Related Operations**. You do not need to specify this parameter in this example.
SFS[b]	Click **Select SFS**. On the displayed page, select the SFS file system that meets your requirements and click **OK**. **NOTE** Only general purpose file systems (formerly SFS 3.0) are supported.

Parameter	Description
PV Name[b]	Enter the PV name, which must be unique in the same cluster.
Access Mode[b]	SFS volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.
Reclaim Policy[b]	You can select **Delete** or **Retain** to specify the reclaim policy of the underlying storage when the PVC is deleted. For details, see **PV Reclaim Policy**. **NOTE** If multiple PVs use the same underlying storage volume, use **Retain** to avoid cascading deletion of underlying volumes.
Mount Options[b]	Enter the mounting parameter key-value pairs. For details, see **Configuring SFS Volume Mount Options**.

NOTE

a: The parameter is available when **Creation Method** is set to **Use existing**.

b: The parameter is available when **Creation Method** is set to **Create new**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

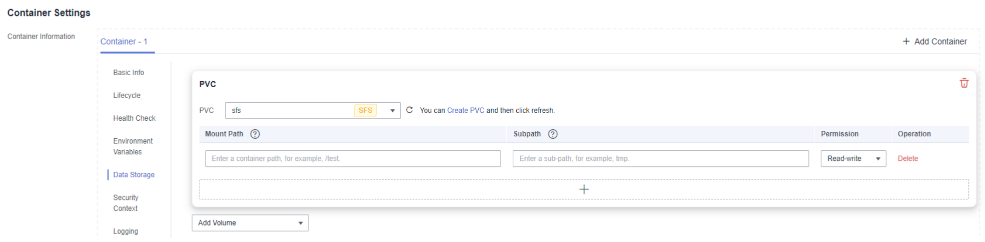
   Mount and use storage volumes, as shown in **Table 4-17**. For details about other parameters, see **Workloads**.

Table 4-17 Mounting a storage volume

Parameter	Description
PVC	Select an existing SFS volume.

Parameter	Description
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter a subpath, for example, **tmp**, indicating that data in the mount path of the container is stored in the **tmp** directory of the storage volume. A subpath is used to mount a local volume so that the same volume is used in a single pod. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the SFS file system.


Container Settings

Container Information

Container - 1 + Add Container

Basic Info
Lifecycle
Health Check
Environment Variables
Data Storage
Security Context
Logging

PVC

PVC: sfs SFS — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
Enter a container path, for example, /test.	Enter a sub-path, for example, tmp.	Read-write	Delete

Add Volume

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a PV.

1. Create the **pv-sfs.yaml** file.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only      # (Optional) The underlying storage is retained when the PV is deleted.
  name: pv-sfs    # PV name.
spec:
  accessModes:
  - ReadWriteMany      # Access mode. The value must be ReadWriteMany for SFS.
  capacity:
    storage: 1Gi     # SFS volume capacity.
  csi:
    driver: nas.csi.everest.io    # Dependent storage driver for the mounting
    fsType: nfs
    volumeHandle: <your_volume_id>   # Enter the name of the general purpose file system (formerly SFS 3.0) is used.
    volumeAttributes:
      everest.io/share-export-location: <your_location>  # Shared path of the SFS volume.
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/sfs-version: sfs3.0       # A general-purpose file system (formerly SFS 3.0) is used for storage.
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-sfs              # StorageClass name, where csi-sfs indicates general purpose file systems (formerly SFS 3.0).
  mountOptions: []                         # Mount options.
```

Table 4-18 Key parameters

Parameter	Mandatory	Description
everest.io/reclaim-policy: retain-volume-only	No	Optional. Currently, only **retain-volume-only** is supported. This parameter is valid only when the reclaim policy is set to **Delete**. If the reclaim policy is **Delete** and the value is **retain-volume-only**, the associated PV is deleted while the underlying storage volume is retained, when a PVC is deleted.
volumeHandle	Yes	If a general purpose file system (formerly SFS 3.0) is used, enter the name of the file system.
everest.io/share-export-location	Yes	Shared path of the file system. A shared path is in the following format: `{your_sfs30_name}.sfs3.{region}.myhuaweicloud.com:/{your_sfs30_name}`

Parameter	Mandatory	Description

mountOptions	Yes	Mount options.
		If this parameter is not specified, the following configurations are used by default. For details, see **Configuring SFS Volume Mount Options**.
		`mountOptions:`
		`- vers=3`
		`- timeo=600`
		`- nolock`
		`- hard`

persistentVolumeReclaimPolicy	Yes	The **Delete** and **Retain** reclaim policies are supported. For details, see **PV Reclaim Policy**. If multiple PVs use the same SFS volume, use **Retain** to prevent the underlying volume from being deleted with a PV. **Delete**: – If **everest.io/reclaim-policy** is not specified, both the PV and SFS file system will be deleted when a PVC is deleted. – If **everest.io/reclaim-policy** is set to **retain-volume-only**, when a PVC is deleted, the PV will be deleted but the SFS file system will be retained. **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again.
storage	Yes	Requested capacity in the PVC, in Gi. For SFS, this field is used only for verification (cannot be empty or **0**). Its value is fixed at **1**, and any value you set does not take effect for SFS file systems.

2. Run the following command to create a PV:
```
kubectl apply -f pv-sfs.yaml
```

Step 3 Create a PVC.

1. Create the **pvc-sfs.yaml** file.
```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-sfs
  namespace: default
  annotations:
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
```

```
spec:
  accessModes:
  - ReadWriteMany              # The value must be ReadWriteMany for SFS.
  resources:
    requests:
      storage: 1Gi             # SFS volume capacity.
  storageClassName: csi-sfs       # Storage class name, which must be the same as that of the PV.
  volumeName: pv-sfs    # PV name.
```

Table 4-19 Key parameters

Parameter	Mandatory	Description
storage	Yes	Requested capacity in the PVC, in Gi. The value must be the same as the storage size of the existing PV.
volumeName	Yes	PV name, which must be the same as the PV name in **Step 2.1**.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-sfs.yaml
```

Step 4 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the SFS volume is mounted to the **/data** path.
```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-sfs-volume    # Volume name, which must be the same as the volume name in the
volumes field.
          mountPath: /data  # Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-sfs-volume    # Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-sfs    # PVC name.
```

2. Run the following command to create a workload that the SFS volume is mounted to:
```
kubectl apply -f web-demo.yaml
```

After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-mjhm9   1/1     Running   0          46s
   web-demo-846b489584-wvv5s   1/1     Running   0          46s
   ```
2. Run the following commands in sequence to view the files in the **/data** path of the pods:
   ```
   kubectl exec web-demo-846b489584-mjhm9 -- ls /data
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:
```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to view the created file in the **/data** path:
```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```
Expected output:
```
static
```

Step 4 Verify data persistence.

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:
   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```
   Expected output:
   ```
   pod "web-demo-846b489584-mjhm9" deleted
   ```
   After the deletion, the Deployment controller automatically creates a replica.
2. Run the following command to view the pod:
   ```
   kubectl get pod | grep web-demo
   ```
   The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:
   ```
   web-demo-846b489584-d4d4j   1/1     Running   0          110s
   web-demo-846b489584-wvv5s   1/1     Running   0          7m50s
   ```
3. Run the following command to check whether the file in the **/data** path of the new pod has been modified:
   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```
   Expected output:
   ```
   static
   ```
   The **static** file is retained, indicating that the data can be stored persistently.

Step 5 **Verify data sharing.**

1. Run the following command to view the pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:

```
web-demo-846b489584-d4d4j   1/1     Running   0            7m
web-demo-846b489584-wvv5s   1/1     Running   0            13m
```

2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.

   ```
   kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
   ```

   Check the files in the **/data** path of the pod.

   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```

   Expected output:

   ```
   share
   static
   ```

3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.

   ```
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```

   Expected output:

   ```
   share
   static
   ```

   After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-20**.

Table 4-20 Related operations

Operation	Description	Procedure
Creating a storage volume (PV)	Create a PV on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVs** tab. In the upper right corner, click **Create PersistentVolume**. In the displayed dialog box, configure the parameters. ● **Volume Type**: Select **SFS**. ● **SFS**: Click **Select SFS**. On the displayed page, select the SFS file system that meets your requirements and click **OK**. ● **PV Name**: Enter the PV name, which must be unique in the same cluster. ● **Access Mode**: Only **ReadWriteMany** is available. A storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**. ● **Reclaim Policy**: There are two options: **Retain** and **Delete**. For details, see **PV Reclaim Policy** NOTE If multiple PVs use the same underlying storage volume, use **Retain** to prevent the underlying volume from being deleted with a PV. ● **Mount Options**: Enter the mounting parameter key-value pairs. For details, see **Configuring SFS Volume Mount Options**. 2. Click **Create**.
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.4.3 Using an SFS File System Through a Dynamic PV

You can mount the PVs created from general-purpose file systems (formerly SFS 3.0) to pods in CCE Autopilot clusters for file storage. This section describes how to use storage classes to dynamically create PVs and PVCs for data persistence and sharing in workloads.

Prerequisites

- If you want to create a cluster using commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.
- You have configured a VPC endpoint required by the general-purpose file system (formerly SFS 3.0). For details, see **Configuring a VPC Endpoint**.

Constraints

- Not all regions support volumes created from file systems. View the regions where SFS volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS volumes are supported.
- If a general-purpose file system (formerly SFS 3.0) is used, the owner group and permission of the mount point cannot be modified.
- If a general-purpose file system (formerly SFS 3.0) is used, there may be a latency when the PVCs or PVs are created or deleted. The billing duration depends on the time when the file systems are created or deleted on the SFS console.
- If the reclamation policy of the volumes created from general-purpose file systems (formerly SFS 3.0) is set to **Delete**, PVs and PVCs can only be deleted after all files in the file systems are deleted manually.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Dynamically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **SFS**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.

Parameter	Description
Creation Method	– If no underlying storage is available, select **Dynamically provision** to create a PVC, PV, and underlying storage on the console in cascading mode. – If underlying storage is available, select either **Use existing** or **Create new**. For details about static creation, see **Using an Existing File System Through a Static PV**. In this example, select **Dynamically provision**.
Storage Classes	The storage class of SFS volumes is **csi-sfs**.
Access Mode	SFS volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

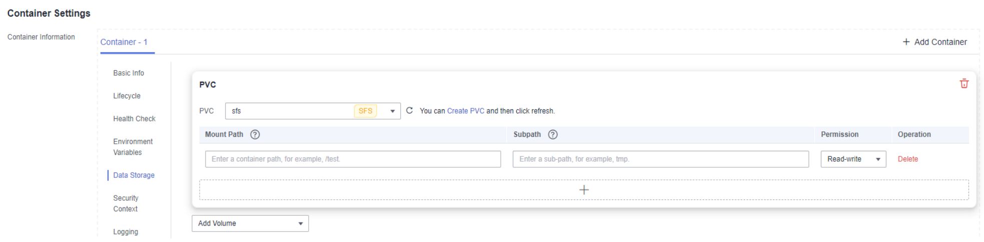
   Table 4-21 describes the parameters for mounting the volume. For details about other parameters, see **Creating a Workload**.

Table 4-21 Parameters for mounting a storage volume

Parameter	Description
PVC	Select an existing SFS volume.
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.

Parameter	Description
Subpath	Enter a subpath, for example, **tmp**, indicating that data in the mount path of the container is stored in the **tmp** directory of the storage volume. A subpath is used to mount a local volume so that the same volume is used in a single pod. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the SFS file system.



PVC

PVC: sfs SFS — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
Enter a container path, for example, /test.	Enter a sub-path, for example, tmp.	Read-write	Delete

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Use **StorageClass** to dynamically create a PVC and PV.

1. Create the **pvc-sfs-auto.yaml** file.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-sfs-auto
  namespace: default
  annotations:
    everest.io/crypt-key-id: <your_key_id>      # (Optional) ID of the key for encrypting file systems

    everest.io/crypt-alias: sfs/default        # (Optional) Key name. Mandatory for encrypting volumes.

    everest.io/crypt-domain-id: <your_domain_id>   # (Optional) ID of the tenant to which an 
encrypted volume belongs. Mandatory for encrypting volumes.

spec:
  accessModes:
    - ReadWriteMany            # The value must be ReadWriteMany for SFS.
  resources:
    requests:
      storage: 1Gi           # SFS volume capacity.
  storageClassName: csi-sfs    # The StorageClass of the SFS file system
```

Table 4-22 Key parameters

Parameter	Mandatory	Description
storage	Yes	Requested capacity in the PVC, in Gi. For SFS, this field is used only for verification (cannot be empty or **0**). Its value is fixed at **1**, and any value you set does not take effect for SFS file systems.
everest.io/crypt-key-id	No	This parameter is mandatory when an SFS system is encrypted. Enter the encryption key ID selected during SFS system creation. You can use a custom key or the default key named **sfs/default**. To obtain a key ID, log in to the DEW console, locate the key to be encrypted, and copy the key ID.
everest.io/crypt-alias	No	Key name, which is mandatory when you create an encrypted volume. To obtain a key name, log in to the DEW console, locate the key to be encrypted, and copy the key name.
everest.io/crypt-domain-id	No	ID of the tenant to which the encrypted volume belongs. This parameter is mandatory for creating an encrypted volume. To obtain a tenant ID, hover the cursor over the username in the upper right corner of the ECS console, choose **My Credentials**, and copy the account ID.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-sfs-auto.yaml
```

Step 3 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the SFS volume is mounted to the **/data** path.
```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
```

```
    app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-sfs-volume    # Volume name, which must be the same as the volume name in the
volumes field.
          mountPath: /data  # Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-sfs-volume    # Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-sfs-auto    # PVC name.
```

2. Run the following command to create a workload that the SFS volume is mounted to:
   ```
   kubectl apply -f web-demo.yaml
   ```
   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-mjhm9   1/1     Running   0             46s
   web-demo-846b489584-wvv5s   1/1     Running   0             46s
   ```
2. Run the following commands in sequence to view the files in the **/data** path of the pods:
   ```
   kubectl exec web-demo-846b489584-mjhm9 -- ls /data
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:
```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to check the files in the **/data** path:
```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```
Expected output:
```
static
```

Step 4 **Verify data persistence.**

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:
   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```
   Expected output:
   ```
   pod "web-demo-846b489584-mjhm9" deleted
   ```
   After the deletion, the Deployment controller automatically creates a replica.

2. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:
   ```
   web-demo-846b489584-d4d4j   1/1     Running   0             110s
   web-demo-846b489584-wvv5s   1/1     Running   0             7m50s
   ```
3. Run the following command to check whether the files in the **/data** path of the new pod have been modified:
   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```
   Expected output:
   ```
   static
   ```
   The **static** file is retained, indicating that the data in the file system can be stored persistently.

Step 5 **Verify data sharing.**

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-d4d4j   1/1     Running   0             7m
   web-demo-846b489584-wvv5s   1/1     Running   0             13m
   ```
2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.
   ```
   kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
   ```
   Check the files in the **/data** path of the pod.
   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```
   Expected output:
   ```
   share
   static
   ```
3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.
   ```
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   Expected output:
   ```
   share
   static
   ```
   After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-23**.

Table 4-23 Related operations

Operation	Description	Procedure
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.4.4 Configuring SFS Volume Mount Options

This section describes how to configure SFS volume mount options. You can configure mount options in a PV and bind the PV to a PVC. Alternatively, configure mount options in a StorageClass and use the StorageClass to create a PVC. In this way, PVs can be dynamically created and inherit mount options configured in the StorageClass by default.

SFS Volume Mount Options

CCE Autopilot presets the options described in **Table 4-24** for mounting SFS volumes.

Table 4-24 SFS volume mount options

Parameter	Value	Description
keep-original-ownership	Leave it blank.	Whether to retain the ownership of the file mount point. ● By default, this option is not added, and the mount point ownership is **root:root** when SFS is mounted. ● If this option is added, the original ownership of the file system is retained when SFS is mounted.
vers	3	File system version. Currently, only NFSv3 is supported. Value: **3**

Parameter	Value	Description
nolock	Leave it blank.	Whether to lock files on the server using the NLM protocol. If **nolock** is selected, the lock is valid for applications on one host. For applications on another host, the lock is invalid.
timeo	600	Waiting time before the NFS client retransmits a request. The unit is 0.1 seconds. Recommended value: **600**
hard/soft	Leave it blank.	Mount mode. ● **hard**: If the NFS request times out, the client keeps resending the request until the request is successful. ● **soft**: If the NFS request times out, the client returns an error to the invoking program. The default value is **hard**.
sharecache/nosharecache	Leave it blank.	How the data cache and attribute cache are shared when one file system is concurrently mounted to different clients. If this parameter is set to **sharecache**, the caches are shared between the mountings. If this parameter is set to **nosharecache**, the caches are not shared, and one cache is configured for each client mounting. The default value is **sharecache**. **NOTE** The **nosharecache** setting will affect the performance. The mounting information must be obtained for each mounting, which increases the communication overhead with the NFS server and the memory consumption of the NFS clients. In addition, the **nosharecache** setting on the NFS clients may lead to inconsistent caches. Determine whether to use **nosharecache** based on site requirements.

You can set other mount options if needed. For details, see **Mounting an NFS File System to ECSs (Linux)**.

Configuring Mount Options in a PV

You can use the **mountOptions** field to configure mount options in a PV. The options you can configure in **mountOptions** are listed in **SFS Volume Mount Options**.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Configure mount options in a PV.

The following is an example:

```
apiVersion: v1
kind: PersistentVolume
```

```
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only      # (Optional) The underlying storage is retained when
the PV is deleted.
  name: pv-sfs
spec:
  accessModes:
  - ReadWriteMany      # Access mode. The value must be ReadWriteMany for SFS.
  capacity:
    storage: 1Gi     # SFS volume capacity.
  csi:
    driver: nas.csi.everest.io    # Dependent storage driver for the mounting
    fsType: nfs
    volumeHandle: <your_volume_id>   # Enter the name of the general purpose file system (formerly SFS
3.0) is used.
    volumeAttributes:
      everest.io/share-export-location: <your_location>  # Shared path of the SFS volume.
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/sfs-version: sfs3.0       # A general-purpose file system (formerly SFS 3.0) is used for storage.
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-sfs              # StorageClass name, where csi-sfs indicates general purpose file
systems (formerly SFS 3.0).
  mountOptions:                          # Mount options.
  - vers=3
  - nolock
  - timeo=600
  - hard
```

Step 3 After a PV is created, you can create a PVC and bind it to the PV, and then mount the PV to the container in the workload. For details, see **Using an Existing File System Through a Static PV**.

Step 4 Check whether the mount options take effect.

In this example, the PVC is mounted to the workload that uses the **nginx:latest** image. You can run the **mount -l** command to check whether the mount options take effect.

1. View the pod to which the SFS volume has been mounted. In this example, the workload name is **web-sfs**.

```
kubectl get pod | grep web-sfs
```

   Command output:

```
web-sfs-***   1/1     Running   0             23m
```

2. Run the following command to check the mount options (**web-sfs-*** ** is an example pod):

```
kubectl exec -it web-sfs-*** -- mount -l | grep nfs
```

   If the mounting information in the command output is consistent with the configured mount options, the mount options have been configured.

```
<Your shared path> on /data type nfs
(rw,relatime,vers=3,rsize=1048576,wsize=1048576,namlen=255,hard,nolock,noresvport,proto=tcp,
timeo=600,retrans=2,sec=sys,mountaddr=**.**.**.**,mountvers=3,mountport=2050,mountproto=tcp
,local_lock=all,addr=**.**.**.**)
```

----End

Configuring Mount Options in a StorageClass

You can use the **mountOptions** field to configure mount options in a StorageClass. The options you can configure in **mountOptions** are listed in **SFS Volume Mount Options**.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a customized StorageClass. Example:

```
apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
  name: csi-sfs-mount-option
provisioner: everest-csi-provisioner
parameters:
  csi.storage.k8s.io/csi-driver-name: nas.csi.everest.io
  csi.storage.k8s.io/fstype: nfs
everest.io/share-access-to: <your_vpc_id> # VPC ID of the cluster.
  everest.io/sfs-version: sfs3.0              # A general-purpose file system (formerly SFS 3.0) is used for storage.
reclaimPolicy: Delete
volumeBindingMode: Immediate
mountOptions:                            # Mount options
- vers=3
- nolock
- timeo=600
- hard
```

Step 3 After the StorageClass is configured, you can use it to create a PVC. By default, the dynamically created PVs inherit the mount options configured in the StorageClass. For details, see **Using an SFS File System Through a Dynamic PV**.

Step 4 Check whether the mount options take effect.

In this example, the PVC is mounted to the workload that uses the **nginx:latest** image. You can run the **mount -l** command to check whether the mount options take effect.

1. View the pod to which the SFS volume has been mounted. In this example, the workload name is **web-sfs**.

```
kubectl get pod | grep web-sfs
```

Command output:

```
web-sfs-***   1/1     Running   0            23m
```

2. Run the following command to check the mount options (**web-sfs-***** is an example pod):

```
kubectl exec -it web-sfs-*** -- mount -l | grep nfs
```

If the mounting information in the command output is consistent with the configured mount options, the mount options have been configured.

```
<Your shared path> on /data type nfs
(rw,relatime,vers=3,rsize=1048576,wsize=1048576,namlen=255,hard,nolock,noresvport,proto=tcp,
timeo=600,retrans=2,sec=sys,mountaddr=**.**.**.**,mountvers=3,mountport=2050,mountproto=tcp
,local_lock=all,addr=**.**.**.**)
```

----End

4.5 SFS Turbo

4.5.1 SFS Turbo Volume Overview

Introduction

CCE Autopilot allows you to mount storage volumes created from SFS Turbo file systems and their subdirectories to a path of a container for persistent data storage. SFS Turbo file systems are fast, on-demand, and scalable, which are suitable for scenarios with a massive number of small files, such as DevOps, containerized microservices, and enterprise office applications.

NOTE

Not all regions support SFS Turbo volumes. View the regions where SFS Turbo volumes are supported on the console. You can also view **Function Overview** to learn about all regions where these volumes are supported.

Expandable to 320 TB, SFS Turbo provides a fully hosted shared file storage, which is highly available and stable, to support small files and applications requiring low latency and high IOPS.

- **Standard file protocols**: You can mount file systems as volumes to servers, the same as using local directories.
- **Data sharing**: The same file system can be mounted to multiple servers, so that data can be shared.
- **Private network**: Users can access data only in private networks of data centers.
- **Data isolation**: The on-cloud storage service provides exclusive cloud file storage, which delivers data isolation and ensures IOPS performance.
- **Use cases**: Deployments and StatefulSets in the ReadWriteMany mode, and jobs created for high-traffic websites, log storage, DevOps, and enterprise OA applications

SFS Turbo Performance

For details about the performance parameters of SFS Turbo, see **File System Types**.

Application Scenarios

SFS Turbo supports the following mounting modes:

- **Using an Existing SFS Turbo File System Through a Static PV**: static provisioning. You use an existing file system to create a PV and then mount the PV to the workload through a PVC.
- **(Recommended) Creating an SFS Turbo Subdirectory Using a Dynamic PV**: CCE Autopilot allows you to create subdirectories for PVs that are dynamically provisioned from SFS Turbo file systems when creating PVCs and to configure the capacity of the subdirectories so that different workloads can share the PVs.

Billing

SFS Turbo does not support dynamic creation. Only created SFS Turbo volumes can be mounted. You can select the pay-per-use billing mode or yearly/monthly package as required. For pricing details about SFS Turbo, see **Billing**.

4.5.2 Using an Existing SFS Turbo File System Through a Static PV

SFS Turbo is a shared file system with high availability and durability. It is suitable for applications that contain massive small files and require low latency, and high IOPS. You can use a manually created SFS Turbo file system to create a PV, which can be bound to a PVC to allow pods share this PV for persistent data storage. This section describes how to use an existing SFS Turbo file system to statically create PVs and PVCs and implement data persistence and sharing in workloads. You can choose to create new subdirectories when creating a PV.

Prerequisites

- You have created an available SFS Turbo file system that is in the same VPC as the cluster.
- If you want to create a cluster using commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- Not all regions support SFS Turbo volumes. View the regions where SFS Turbo volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS Turbo volumes are supported.
- Multiple PVs can use the same SFS Turbo file system with the following restrictions:
  - All PVCs or PVs that use the same SFS Turbo file system cannot be mounted to a pod. This will result in a pod startup failure because not all PVCs can be mounted to the pod due to the same **volumeHandle** values of these PVs.
  - The **persistentVolumeReclaimPolicy** parameter in the PVs should be set to **Retain**. If any other value is used, when a PV is deleted, the associated underlying SFS Turbo file system may be deleted. In this case, other PVs associated with the underlying file system malfunction.
  - When a file system is repeatedly used, enable isolation and protection for ReadWriteMany at the application layer to prevent data overwriting and loss.
- To create subdirectories for an SFS Turbo volume, the cluster version must be v1.27.9-r0, v1.28.7-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- For SFS Turbo storage, the yearly/monthly SFS Turbo resources will not be reclaimed when the cluster or PVC is deleted. Reclaim the resources on the SFS Turbo console.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Statically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **SFS Turbo**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Namespace	Namespace of the PVC. – In the navigation pane, choose **Storage**. If **Namespace** in the upper part of the page is **All namespaces** or **Non-system namespace**, you can select an appropriate namespace. – If you choose **Storage** in the navigation pane and a non-system namespace (for example, **default**) has been specified for **Namespace** in the upper part of the page, the namespace is displayed here.
Creation Method	You can create a storage volume or use an existing storage volume to statically create a PVC based on whether a PV is available. In this example, select **Create new** to create a PV and PVC at the same time on the console.
PV[a]	Select an existing PV in the cluster. Create a PV in advance. For details, see "Creating a storage volume" in **Related Operations**. You do not need to specify this parameter in this example.
SFS Turbo[b]	Click **Select SFS Turbo**. On the displayed page, select the SFS Turbo file system that meets your requirements and click **OK**.
Subdirectory[b]	Determine whether to use subdirectories to create PVs. If this parameter is set to **Yes**, enter the absolute path of the subdirectory, for example, **/a/b**. In this case, the underlying storage associated with the PV is a subdirectory in the selected SFS Turbo file system. Ensure that the subdirectory is available.
PV Name[b]	Enter the PV name, which must be unique in the same cluster.
Access Mode[b]	SFS Turbo volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.

Parameter	Description
Reclaim Policy[b]	Only **Retain** is available if you do not use subdirectories to create PVs. This indicates that the PV is not deleted when the PVC is deleted. For details, see **PV Reclaim Policy**. If you choose to create a subdirectory, this parameter can only be set to **Delete**.
Subdirectory Reclaim Policy[b]	Determine whether to retain subdirectories when a PVC is deleted. This parameter must be used with **PV Reclaim Policy** and can be configured when **PV Reclaim Policy** is set to **Delete**. – **Retain**: If a PVC is deleted, the PV will be deleted, but **the subdirectories associated with the PV will be retained**. – **Delete**: After a PVC is deleted, **the PV and its associated subdirectories will also be deleted**.
Mount Options[b]	Enter the mounting parameter key-value pairs. For details, see **Configuring SFS Turbo Mount Options**.

NOTE

a: The parameter is available when **Creation Method** is set to **Use existing**.

b: The parameter is available when **Creation Method** is set to **Create new**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

   Table 4-25 describes the parameters for mounting the volume. For details about other parameters, see **Workloads**.

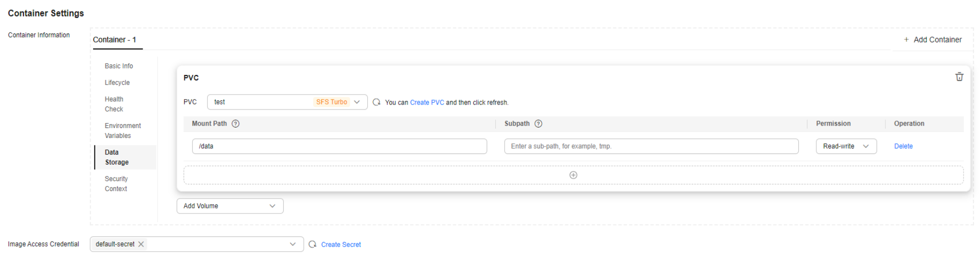
Table 4-25 Mounting a storage volume

Parameter	Description
PVC	Select an existing SFS Turbo volume.

Parameter	Description
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter a subdirectory of the PV. This subdirectory will be mounted to a specified path of a container. In this way, different subdirectories of the same PV can be used in a single pod. Enter a subpath, for example, **tmp**, indicating that data in the mount path of the container is stored in the **tmp** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the SFS Turbo file system.

Figure 4-14 Mounting a PV


Container Settings

Container Information: Container - 1

PVC: test (SFS Turbo). You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-write	Delete

Image Access Credential: default-secret

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

You can determine whether to use subdirectories based on your service requirements.

Using an Existing SFS Turbo File System Without Subdirectories

An entire SFS Turbo file system is used, and there are no subdirectories.

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a PV.

1. Create the **pv-sfsturbo.yaml** file.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
  name: pv-sfsturbo    # PV name.
spec:
  accessModes:
  - ReadWriteMany      # Access mode. The value must be ReadWriteMany for SFS Turbo.
  capacity:
    storage: 500Gi       # SFS Turbo volume capacity.
  csi:
    driver: sfsturbo.csi.everest.io    # Dependent storage driver for the mounting.
    fsType: nfs
    volumeHandle: <your_volume_id>  # SFS Turbo volume ID.
    volumeAttributes:
      everest.io/share-export-location: <your_location>   # Shared path of the SFS Turbo volume.
      everest.io/enterprise-project-id: <your_project_id>  # Project ID of the SFS Turbo volume.

      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-sfsturbo          # Storage class name of the SFS Turbo volume.
  mountOptions: []                        # Mount options.
```

Table 4-26 Key parameters

Parameter	Mandatory	Description
volumeHandle	Yes	SFS Turbo volume ID. How to obtain: Log in to the console, choose **Service List** > **Storage** > **Scalable File Service**, and select **SFS Turbo**. In the list, click the name of the target SFS Turbo file system. On the details page, copy the content following **ID**.
everest.io/share-export-location	Yes	Shared path of the SFS Turbo volume. Log in to the console, choose **Service List** > **Storage** > **Scalable File Service**, and select **SFS Turbo**. You can obtain the shared path of the file system from the **Mount Address** column.

Parameter	Mandatory	Description
everest.io/enterprise-project-id	No	Project ID of the SFS Turbo volume. If you use the default enterprise project, set this field to **0**. How to obtain: On the SFS console, click **SFS Turbo** in the left navigation pane. Click the name of the SFS Turbo file system to interconnect. On the **Basic Info** tab, find and click the enterprise project to go to the console, and copy the ID.

mountOptions	No	Mount options.
		If this parameter is not specified, the following configurations are used by default. For details, see **Configuring SFS Turbo Mount Options**.
		`mountOptions:`
		`- vers=3`
		`- timeo=600`
		`- nolock`
		`- hard`

persistentVolumeReclaimPolicy	Yes	Only the **Retain** reclaim policy is supported. For details, see **PV Reclaim Policy**. **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again.
storage	Yes	Requested capacity in the PVC, in Gi.
storageClassName	Yes	The storage class name of SFS Turbo volumes is **csi-sfsturbo**.

2. Run the following command to create a PV:
```
kubectl apply -f pv-sfsturbo.yaml
```

Step 3 Create a PVC.

1. Create the **pvc-sfsturbo.yaml** file.
```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-sfsturbo
  namespace: default
  annotations:
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
    everest.io/enterprise-project-id: <your_project_id>  # Project ID of the SFS Turbo volume.

spec:
  accessModes:
  - ReadWriteMany                  # The value must be ReadWriteMany for SFS Turbo.
  resources:
    requests:
      storage: 500Gi              # SFS Turbo volume capacity.
  storageClassName: csi-sfsturbo       # Storage class of the SFS Turbo volume. It must be the same
as that of the PV.
  volumeName: pv-sfsturbo    # PV name.
```

Table 4-27 Key parameters

Parameter	Mandatory	Description
everest.io/enterprise-project-id	No	Project ID of the SFS Turbo volume. If you use the default enterprise project, set this field to **0**. How to obtain: On the SFS console, click **SFS Turbo** in the left navigation pane. Click the name of the SFS Turbo file system to interconnect. On the **Basic Info** tab, find and click the enterprise project to go to the console, and copy the ID.
storage	Yes	Requested capacity in the PVC, in Gi. The value must be the same as the storage size of the existing PV.
storageClassName	Yes	Storage class name, which must be the same as the storage class of the PV in **Step 2.1**. The storage class name of SFS Turbo volumes is **csi-sfsturbo**.
volumeName	Yes	PV name, which must be the same as the PV name in **Step 2.1**.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-sfsturbo.yaml
```

Step 4 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the SFS Turbo volume is mounted to the **/data** path.
```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
```

```
        image: nginx:latest
        volumeMounts:
        - name: pvc-sfsturbo-volume    #Volume name, which must be the same as the volume name in 
the volumes field.
          mountPath: /data  #Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-sfsturbo-volume    #Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-sfsturbo    #Name of the created PVC.
```

2. Run the following command to create a workload that the SFS Turbo volume is mounted to:

```
kubectl apply -f web-demo.yaml
```

After the workload is created, you can try **Verifying Data Persistence and Sharing**.

----End

Using an Existing SFS Turbo File System with Subdirectories

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a PV.

1. Create the **pv-sfsturbo.yaml** file.

   Example:

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only    # When a PVC is deleted, the PV will be deleted but 
the subdirectories associated with the PV will be retained. This parameter is available only when 
subdirectories are used and the reclaim policy is Delete.
  name: pv-sfsturbo    # PV name.
spec:
  accessModes:
  - ReadWriteMany      # Access mode. The value must be ReadWriteMany for SFS Turbo.
  capacity:
    storage: 500Gi       # SFS Turbo volume capacity.
  csi:
    driver: sfsturbo.csi.everest.io    # Dependent storage driver for the mounting.
    fsType: nfs
    volumeHandle: pv-sfsturbo   # PV name when subdirectories are used.
    volumeAttributes:
      everest.io/share-export-location: <sfsturbo_path>:/<absolute_path>   # Shared path and 
subdirectory of the SFS Turbo file system.
      everest.io/enterprise-project-id: <your_project_id>  # Project ID of the SFS Turbo volume.
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/volume-as: absolute-path   # (Optional) An SFS Turbo subdirectory is used.
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy, which can be set to Delete when 
subdirectories are automatically created.
  storageClassName: csi-sfsturbo         # Storage class name of the SFS Turbo volume.
  mountOptions: []                        # Mount options.
```

Table 4-28 Key parameters

Parameter	Mandatory	Description
volumeHandle	Yes	PV name when an SFS Turbo subdirectory is used to create the PV.
everest.io/share-export-location	Yes	Shared path of the SFS Turbo subdirectory. Format: `{sfsturbo_path}:/{absolute_path}` Log in to the console, choose **Service List** > **Storage** > **Scalable File Service**, and select **SFS Turbo**. You can obtain the shared path of the file system.
everest.io/enterprise-project-id	No	Project ID of the SFS Turbo volume. If you use the default enterprise project, set this field to **0**. How to obtain: On the SFS console, click **SFS Turbo** in the left navigation pane. Click the name of the SFS Turbo file system to interconnect. On the **Basic Info** tab, find and click the enterprise project to go to the console, and copy the ID.

mountOptions	No	Mount options.
		If this parameter is not specified, the following configurations are used by default. For details, see **Configuring SFS Turbo Mount Options**.
		`mountOptions:`
		`- vers=3`
		`- timeo=600`
		`- nolock`
		`- hard`

persistentVolumeReclaimPolicy	Yes	Reclamation policy. For details, see **PV Reclaim Policy**. – **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again. – **Delete**: This parameter can be configured when subdirectories are automatically created, indicating that the PV is deleted when a PVC is deleted.

Parameter	Mandatory	Description
everest.io/reclaim-policy	No	Whether to retain the subdirectory when deleting the PVC. This parameter is used in conjunction with **PV Reclamation Policy**. This parameter is available only when the PV reclaim policy is **Delete**. – **retain-volume-only**: If a PVC is deleted, the PV will be deleted, but **the subdirectories associated with the PV will be retained**. – **delete**: After a PVC is deleted, **the PV and its associated subdirectories will also be deleted**. **NOTE** When a subdirectory is deleted, only the absolute path of the subdirectory configured in the PVC will be deleted. The upper-layer directory will not be deleted.
everest.io/volume-as	No	The value is fixed at **absolute-path**, indicating that a dynamically created SFS Turbo subdirectory is used.
storage	Yes	Requested capacity in the PVC, in Gi. If a subdirectory is used, this parameter serves no purpose other than for verification and must have a non-empty, non-zero value.
storageClassName	Yes	The storage class name of SFS Turbo volumes is **csi-sfsturbo**.

2. Run the following command to create a PV:
```
kubectl apply -f pv-sfsturbo.yaml
```

Step 3 Create a PVC.

1. Create the **pvc-sfsturbo.yaml** file.
```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-sfsturbo
  namespace: default
  annotations:
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
      everest.io/enterprise-project-id: <your_project_id>  # Project ID of the SFS Turbo volume.
spec:
  accessModes:
  - ReadWriteMany                  # The value must be ReadWriteMany for SFS Turbo.
  resources:
    requests:
      storage: 500Gi               # SFS Turbo volume capacity.
  storageClassName: csi-sfsturbo       # Storage class of the SFS Turbo volume. It must be the same
as that of the PV.
  volumeName: pv-sfsturbo    # PV name.
```

Table 4-29 Key parameters

Parameter	Mandatory	Description
everest.io/enterprise-project-id	No	Project ID of the SFS Turbo volume. If you use the default enterprise project, set this field to **0**. How to obtain: On the SFS console, click **SFS Turbo** in the left navigation pane. Click the name of the SFS Turbo file system to interconnect. On the **Basic Info** tab, find and click the enterprise project to go to the console, and copy the ID.
storage	Yes	Requested capacity in the PVC, in Gi. The value must be the same as the storage size of the existing PV.
storageClassName	Yes	Storage class name, which must be the same as the storage class of the PV in **Step 2.1**. The storage class name of SFS Turbo volumes is **csi-sfsturbo**.
volumeName	Yes	PV name, which must be the same as the PV name in **Step 2.1**.

2. Run the following command to create a PVC:
```
kubectl apply -f pvc-sfsturbo.yaml
```

Step 4 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the SFS Turbo volume is mounted to the **/data** path.
```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-sfsturbo-volume    #Volume name, which must be the same as the volume name in
the volumes field.
          mountPath: /data  #Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-sfsturbo-volume    # Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-sfsturbo    #Name of the created PVC.
```

2. Run the following command to create a workload that the SFS Turbo volume is mounted to:
   ```
   kubectl apply -f web-demo.yaml
   ```
   After the workload is created, you can try **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-mjhm9   1/1     Running   0             46s
   web-demo-846b489584-wvv5s   1/1     Running   0             46s
   ```
2. Run the following commands in sequence to view the files in the **/data** path of the pods:
   ```
   kubectl exec web-demo-846b489584-mjhm9 -- ls /data
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to check the files in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```

Expected output:

```
static
```

Step 4 **Verify data persistence.**

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:
   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```
   Expected output:
   ```
   pod "web-demo-846b489584-mjhm9" deleted
   ```
   After the deletion, the Deployment controller automatically creates a replica.
2. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:
   ```
   web-demo-846b489584-d4d4j   1/1     Running   0             110s
   web-demo-846b489584-wvv5s    1/1     Running   0             7m50s
   ```
3. Run the following command to check whether the files in the **/data** path of the new pod have been modified:
   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```
   Expected output:
   ```
   static
   ```

The **static** file is retained, indicating that the data in the file system can be stored persistently.

Step 5 **Verify data sharing.**

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-d4d4j   1/1     Running   0            7m
   web-demo-846b489584-wvv5s   1/1     Running   0            13m
   ```
2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.
   ```
   kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
   ```
   Check the files in the **/data** path of the pod.
   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```
   Expected output:
   ```
   share
   static
   ```
3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.
   ```
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   Expected output:
   ```
   share
   static
   ```
   After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-30**.

Table 4-30 Related operations

Operation	Description	Procedure
Creating a storage volume (PV)	Create a PV on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVs** tab. In the upper right corner, click **Create PersistentVolume**. In the displayed dialog box, configure the parameters. ● **Volume Type**: Select **SFS Turbo**. ● **SFS Turbo**: Click **Select SFS Turbo**. On the page displayed, select the SFS Turbo file system that meets your requirements and click **OK**. ● **Subdirectory**: Determine whether to use subdirectories to create PVs. Enter the absolute path of a subdirectory, for example, **/a/b**. Ensure that the subdirectory is available. ● **PV Name**: Enter the PV name, which must be unique in the same cluster. ● **Access Mode**: SFS volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**. ● **Reclaim Policy**: Only **Retain** is supported if you do not use subdirectories. For details, see **PV Reclaim Policy**. If you choose to create a subdirectory, this parameter can only be set to **Delete**. ● **Mount Options**: Enter the mounting parameter key-value pairs. For details, see **Configuring SFS Turbo Mount Options**. 2. Click **Create**.
Expanding the capacity of an SFS Turbo volume	Quickly expand the capacity of a mounted SFS Turbo volume on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Scale-out** in the **Operation** column. 2. Enter the capacity to be added and click **OK**.

Operati on	Description	Procedure
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.5.3 Configuring SFS Turbo Mount Options

This section describes how to configure SFS Turbo volume mount options. For SFS Turbo, you can only set mount options in a PV and bind the PV by creating a PVC.

SFS Turbo Mount Options

CCE Autopilot presets the options described in **Table 4-31** for mounting SFS Turbo volumes.

Table 4-31 SFS Turbo mount options

Parameter	Value	Description
vers	3	File system version. Currently, only NFS v3 is supported. Value: **3**
nolock	Leave it blank.	Whether to lock files on the server using the NLM protocol. If **nolock** is selected, the lock is valid for applications on one host. For applications on another host, the lock is invalid.
timeo	600	Waiting time before the NFS client retransmits a request. The unit is 0.1 seconds. Recommended value: **600**

Parameter	Value	Description
hard/soft	Leave it blank.	Mount mode. ● **hard**: If the NFS request times out, the client keeps resending the request until the request is successful. ● **soft**: If the NFS request times out, the client returns an error to the invoking program. The default value is **hard**.

You can set other mount options if needed. For details, see **Mounting an NFS File System to ECSs (Linux)**.

Configuring Mount Options in a PV

You can use the **mountOptions** field to configure mount options in a PV. The options you can configure in **mountOptions** are listed in **SFS Turbo Mount Options**.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Configure mount options in a PV. The following is an example:

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
  name: pv-sfsturbo    # PV name.
spec:
  accessModes:
  - ReadWriteMany      # Access mode. The value must be ReadWriteMany for SFS Turbo.
  capacity:
    storage: 500Gi       # SFS Turbo volume capacity.
  csi:
    driver: sfsturbo.csi.everest.io    # Dependent storage driver for the mounting.
    fsType: nfs
    volumeHandle: {your_volume_id}  # SFS Turbo volume ID.
    volumeAttributes:
      everest.io/share-export-location: {your_location}  # Shared path of the SFS Turbo volume.
      everest.io/enterprise-project-id: {your_project_id}  # Project ID of the SFS Turbo volume.
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-sfsturbo          # Storage class name of the SFS Turbo volume.
  mountOptions:                         # Mount options.
  - vers=3
  - nolock
  - timeo=600
  - hard
```

Step 3 After a PV is created, you can create a PVC and bind it to the PV, and then mount the PV to the container in the workload. For details, see **Using an Existing SFS Turbo File System Through a Static PV**.

Step 4 Check whether the mount options take effect.

In this example, the PVC is mounted to the workload that uses the **nginx:latest** image. You can run the **mount -l** command to check whether the mount options take effect.

1. View the pod to which the SFS Turbo volume has been mounted. In this example, the workload name is **web-sfsturbo**.
   ```
   kubectl get pod | grep web-sfsturbo
   ```
   Command output:
   ```
   web-sfsturbo-***   1/1     Running   0             23m
   ```
2. Run the following command to check the mount options (**web-sfsturbo-******* is an example pod):
   ```
   kubectl exec -it web-sfsturbo-*** -- mount -l | grep nfs
   ```
   If the mounting information in the command output is consistent with the configured mount options, the mount options have been configured.
   ```
   <Your mount path> on /data type nfs
   (rw,relatime,vers=3,rsize=1048576,wsize=1048576,namlen=255,hard,nolock,noresvport,proto=tcp,
   timeo=600,retrans=2,sec=sys,mountaddr=**.**.**.**,mountvers=3,mountport=20048,mountproto=tc
   p,local_lock=all,addr=**.**.**.**)
   ```

----End

4.5.4 (Recommended) Creating an SFS Turbo Subdirectory Using a Dynamic PV

When an SFS Turbo volume is mounted to a workload container, the root directory is mounted to the container by default. However, the minimum capacity of an SFS Turbo volume is 500 GiB, which exceeds the capacity required by most workloads, leading to a waste of storage resources. CCE Autopilot enables efficient utilization of storage capacity by creating SFS Turbo subdirectories dynamically when you create a PVC. You can configure the capacities of these subdirectories. This allows multiple workloads to share the SFS Turbo file system.

Prerequisites

- If you want to create a cluster using commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.
- You have created an available SFS Turbo file system that is in the same VPC as the cluster.

Constraints

- Not all regions support SFS Turbo volumes. View the regions where SFS Turbo volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS Turbo volumes are supported.
- To create subdirectories for an SFS Turbo volume, the cluster version must be v1.27.9-r0, v1.28.7-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.
- Constraints on the subdirectory quota:
  - If the subdirectory quota is less than 1 GiB, it is automatically set to 1 GiB. The maximum quota cannot exceed the capacity of the target SFS Turbo file system. The minimum scale-out step is 1 GiB, and capacity reduction is not allowed.

- Once a subdirectory quota is configured, it cannot be canceled.
- The number of files or subdirectories in a file system is determined by the quota capacity (in KB) divided by 16. The upper limit is set at 1 billion and cannot be changed by users.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **SFS Turbo**.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Namespace	Namespace of the PVC. ● In the navigation pane, choose **Storage**. If **Namespace** in the upper part of the page is **All namespaces** or **Non-system namespace**, you can select an appropriate namespace. ● If you choose **Storage** in the navigation pane and a non-system namespace (for example, **default**) has been specified for **Namespace** in the upper part of the page, the namespace is displayed here.
Creation Method	Select **New subdirectory**.
Storage Class	Choose **csi-sfsturbo**.
Access Mode	SFS Turbo volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.
SFS Turbo	Click **Select SFS Turbo**. On the displayed page, select the SFS Turbo file system that meets your requirements and click **OK**.
Subdirectory	Enter the absolute path of a subdirectory, for example, **/a/b**. If the subdirectory does not exist, a subdirectory is automatically created.

Parameter	Description
Subdirectory Reclaim Policy	Whether to retain subdirectories when a PVC is deleted. ● **Retain**: If a PVC is deleted, the PV will be deleted, but **the subdirectories associated with the PV will be retained**. ● **Delete**: After a PVC is deleted, **the PV and its associated subdirectories will also be deleted**. **NOTE** When a subdirectory is deleted, only the absolute path of the subdirectory configured in the PVC will be deleted. The upper-layer directory will not be deleted.
Subdirectory Capacity	● **Not limited**: The subdirectory capacity is not limited. ● **Limited**: The subdirectory capacity is limited.
Capacity	The maximum capacity of the subdirectory, in GiB. This parameter is available only when the subdirectory capacity limit is enabled.

Step 3 Click **Create** to create a PVC and a PV.

In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 4 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

   Table 4-32 describes the parameters for mounting the volume. For details about other parameters, see Workloads.

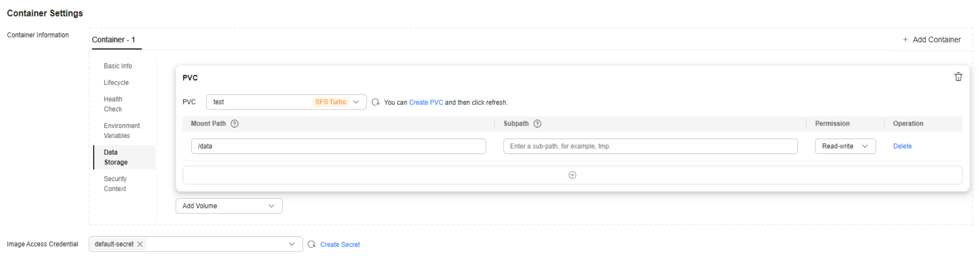
Table 4-32 Parameters for mounting a storage volume

Parameter	Description
PVC	Select an existing SFS Turbo volume.

Parameter	Description
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter a subdirectory of the PV. This subdirectory will be mounted to a specified path of a container. In this way, different subdirectories of the same PV can be used in a single pod. **tmp**, for example, indicates that data in the mount path of the container is stored in the **tmp** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the SFS Turbo file system.

Figure 4-15 Mounting a PV


Container Settings

Container Information: Container - 1

PVC: test (SFS Turbo) — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-write	Delete

Image Access Credential: default-secret

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Create the **pvc-sfsturbo-subpath.yaml** file.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-sfsturbo-subpath    # PVC name.
  namespace: default
  annotations:
    everest.io/volume-as: absolute-path                # An SFS Turbo subdirectory is used.
    everest.io/sfsturbo-share-id: <sfsturbo_id>        # SFS Turbo ID.
    everest.io/path: /a                               # Subdirectory that is automatically created, which must be an absolute path.
    everest.io/reclaim-policy: retain-volume-only      # When a PVC is deleted, the PV will be deleted, but the subdirectories associated with the PV will be retained.
spec:
  accessModes:
    - ReadWriteMany      # ReadWriteMany must be selected for SFS Turbo.
  resources:
    requests:
      storage: 10Gi      # For SFS Turbo subdirectory PVCs, this configuration is meaningless and only used for verification (the value cannot be empty or 0) (the value cannot be empty or 0).
  storageClassName: csi-sfsturbo     # Storage class name of the SFS Turbo volume.
```

Table 4-33 Key parameters

Parameter	Mandatory	Description
everest.io/volume-as	No	The value is fixed at **absolute-path**, indicating that a dynamically created SFS Turbo subdirectory is used.
everest.io/sfsturbo-share-id	No	SFS Turbo ID. How to obtain: Log in to the console, choose **Service List** > **Storage** > **Scalable File Service**, and select **SFS Turbo**. In the list, click the name of the target SFS Turbo file system. On the details page, copy the content following **ID**.
everest.io/path	No	Subdirectory that is automatically created, which must be an absolute path.

Parameter	Mandatory	Description
everest.io/reclaim-policy	No	Whether to retain the subdirectory when deleting the PVC. This parameter is used in conjunction with **PV Reclamation Policy**. This parameter is available only when the PV reclaim policy is **Delete**. ● **retain-volume-only**: If a PVC is deleted, the PV will be deleted, but **the subdirectories associated with the PV will be retained**. ● **delete**: After a PVC is deleted, **the PV and its associated subdirectories will also be deleted**. **NOTE** When a subdirectory is deleted, only the absolute path of the subdirectory configured in the PVC will be deleted. The upper-layer directory will not be deleted.
storage	Yes	Requested capacity in the PVC, in Gi. For dynamically created subdirectories using a PVC: ● You can specify the subdirectory capacity. ● If subdirectory capacity is not limited, this parameter is meaningless and is used only for verification. You can set this parameter to a fixed value **10Gi**.
everest.io/csi.enable-sfsturbo-dir-quota	No	Whether to limit the subdirectory capacity. If the value is set to **true**, the subdirectory capacity is limited. If the value is empty or set to any other value, the subdirectory capacity is not limited.

Step 3 Run the following command to create a PVC:

```
kubectl apply -f pvc-sfsturbo-subpath.yaml
```

Step 4 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the SFS Turbo volume is mounted to the **/data** path.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
```

```
   labels:
     app: web-demo
  spec:
    containers:
    - name: container-1
      image: nginx:latest
      volumeMounts:
      - name: pvc-sfsturbo-volume    #Volume name, which must be the same as the volume name in
the volumes field.
        mountPath: /data  #Location where the storage volume is mounted.
    imagePullSecrets:
      - name: default-secret
    volumes:
      - name: pvc-sfsturbo-volume    # Volume name, which can be changed as needed.
        persistentVolumeClaim:
          claimName: pvc-sfsturbo-subpath    # PVC name.
```

2. Run the following command to create a workload that the SFS Turbo volume is mounted to:
   ```
   kubectl apply -f web-demo.yaml
   ```
   After the workload is created, you can verify it by referring to the procedure in **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the pods:
   ```
   kubectl get pod | grep web-demo
   ```
   Expected output:
   ```
   web-demo-846b489584-mjhm9   1/1     Running   0             46s
   web-demo-846b489584-wvv5s   1/1     Running   0             46s
   ```
2. Run the following commands in sequence to view the files in the **/data** path of the pods:
   ```
   kubectl exec web-demo-846b489584-mjhm9 -- ls /data
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```
   If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to view the created file in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```

Expected output:

```
static
```

Step 4 Verify data persistence.

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:
   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```
   Expected output:
   ```
   pod "web-demo-846b489584-mjhm9" deleted
   ```
   After the deletion, the Deployment controller automatically creates a replica.
2. Run the following command to view the pod:
   ```
   kubectl get pod | grep web-demo
   ```

The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:

```
web-demo-846b489584-d4d4j   1/1     Running   0            110s
web-demo-846b489584-wvv5s    1/1     Running   0            7m50s
```

3. Run the following command to check whether the file in the **/data** path of the new pod has been modified:

   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```

   Expected output:

   ```
   static
   ```

   The **static** file is retained, indicating that the data can be stored persistently.

Step 5 Verify data sharing.

1. Run the following command to view the pod:

   ```
   kubectl get pod | grep web-demo
   ```

   Expected output:

   ```
   web-demo-846b489584-d4d4j   1/1     Running   0            7m
   web-demo-846b489584-wvv5s   1/1     Running   0            13m
   ```

2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.

   ```
   kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
   ```

   Check the files in the **/data** path of the pod.

   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```

   Expected output:

   ```
   share
   static
   ```

3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.

   ```
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```

   Expected output:

   ```
   share
   static
   ```

   After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-34**.

Table 4-34 Related operations

Operation	Description	Steps
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.5.5 Dynamically Creating and Mounting Subdirectories of an SFS Turbo Volume

Background

The minimum capacity of an SFS Turbo file system is 500 GiB, and the SFS Turbo file system cannot be billed by usage. By default, the root directory of an SFS Turbo file system is mounted to a container which, in most case, does not require such a large capacity.

CCE Autopilot allows you to dynamically create subdirectories in an SFS Turbo volume and mount these subdirectories to containers so that this volume can be shared and the storage capacity can be used more economically.

Not all regions support SFS Turbo volumes. View the regions where SFS Turbo volumes are supported on the console. You can also view **Function Overview** to learn about all regions where SFS Turbo volumes are supported.

Creating an SFS Turbo Volume of the subpath Type

NOTE

Compared with CCE standard or CCE Turbo clusters, CCE Autopilot clusters do not support archiving of SFS Turbo subdirectories when PVCs are deleted. When deleting a PVC, if **persistentVolumeReclaimPolicy** of the PV bound to the PVC is set to **Delete**, the corresponding SFS Turbo subdirectory will be deleted.

Step 1 Create an SFS Turbo file system in the same VPC and subnet as the cluster.

Step 2 Create a YAML file of StorageClass, for example, **sfsturbo-subpath-sc.yaml**.

The following is an example:

```
apiVersion: storage.k8s.io/v1
allowVolumeExpansion: true
kind: StorageClass
metadata:
  name: sfsturbo-subpath-sc
mountOptions:
- lock
parameters:
  csi.storage.k8s.io/csi-driver-name: sfsturbo.csi.everest.io
  csi.storage.k8s.io/fstype: nfs
  everest.io/share-access-to: 7ca2dba2-1234-1234-1234-626371a8fb3a
  everest.io/share-expand-type: bandwidth
  everest.io/share-export-location: 192.168.1.1:/sfsturbo/
  everest.io/share-source: sfs-turbo
  everest.io/share-volume-type: STANDARD
  everest.io/volume-as: subpath
  everest.io/volume-id: 0d773f2e-1234-1234-1234-de6a35074696
provisioner: everest-csi-provisioner
reclaimPolicy: Delete
volumeBindingMode: Immediate
```

In this example:

- **name**: indicates the name of the StorageClass.
- **mountOptions**: indicates the mount options. This field is optional.

  The following configuration is used by default. For details, see **Setting Mount Options**. Do not set **nolock** to **true**. If you do this, the mount operation will fail.
  ```
  mountOptions:
  - vers=3
  - timeo=600
  - nolock
  - hard
  ```
- **everest.io/volume-as**: This parameter is set to **subpath** to use the **subpath** volume.
- **everest.io/share-access-to**: This parameter is optional. In a subpath volume, set this parameter to the ID of the VPC where the SFS Turbo file system is located.
- **everest.io/share-expand-type**: This parameter is optional. If the type of the SFS Turbo file system is SFS Turbo Standard – Enhanced or SFS Turbo Performance – Enhanced, set this parameter to **bandwidth**.
- **everest.io/share-export-location**: This parameter indicates the mount directory. It consists of the SFS Turbo shared path and subdirectory. The shared path can be obtained on the SFS Turbo console. The subdirectory is user-defined. The PVCs created using the StorageClass are located in this subdirectory.
- **everest.io/share-volume-type**: This parameter is optional. It specifies the SFS Turbo file system type. The value can be **STANDARD** or **PERFORMANCE**. For enhanced types, this parameter must be used together with **everest.io/share-expand-type** (whose value should be **bandwidth**).
- **everest.io/zone**: This parameter is optional. Set it to the AZ where the SFS Turbo file system is located.
- **everest.io/volume-id**: This parameter indicates the ID of the SFS Turbo volume. You can obtain the volume ID on the SFS Turbo page.

Step 3 Run **kubectl create -f sfsturbo-subpath-sc.yaml**.

Step 4 Create a PVC YAML file named **sfs-turbo-test.yaml**.

The following is an example:

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: sfs-turbo-test
  namespace: default
spec:
  accessModes:
  - ReadWriteMany
  resources:
    requests:
      storage: 50Gi
  storageClassName: sfsturbo-subpath-sc
  volumeMode: Filesystem
```

In this example:

- **name**: indicates the name of the PVC.
- **storageClassName**: specifies the name of the StorageClass.
- **storage**: In the subpath mode, it is useless to specify this parameter. The storage capacity is limited by the total capacity of the SFS Turbo file system. If the total capacity of the SFS Turbo file system is insufficient, expand the capacity on the SFS Turbo page in a timely manner.

Step 5 Run the **kubectl create -f sfs-turbo-test.yaml** command to create a PVC.

----End

NOTE

It is meaningless to conduct capacity expansion on an SFS Turbo volume created in the subpath mode. This operation does not expand the capacity of the SFS Turbo file system. Ensure that the total capacity of the SFS Turbo file system is not used up.

Creating a Deployment and Mounting an Existing Volume

Step 1 Create a YAML file for the Deployment, for example, **deployment-test.yaml**.

The following is an example:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: test-turbo-subpath-example
  namespace: default
  generation: 1
  labels:
    appgroup: ''
spec:
  replicas: 1
  selector:
    matchLabels:
      app: test-turbo-subpath-example
  template:
    metadata:
      labels:
        app: test-turbo-subpath-example
    spec:
      containers:
      - image: nginx:latest
        name: container-0
        volumeMounts:
```

```
        - mountPath: /tmp
          name: pvc-sfs-turbo-example
      restartPolicy: Always
      imagePullSecrets:
      - name: default-secret
      volumes:
      - name: pvc-sfs-turbo-example
        persistentVolumeClaim:
          claimName: sfs-turbo-test
```

In this example:

- **name**: indicates the name of the Deployment.
- **image**: specifies the image used by the Deployment.
- **mountPath**: indicates the mount path of the container. In this example, the volume is mounted to the **/tmp** directory.
- **claimName**: indicates the name of an existing PVC.

Step 2 Create the Deployment.

kubectl create -f deployment-test.yaml

----End

Dynamically Creating a subpath Volume for a StatefulSet

Step 1 Create a YAML file for a StatefulSet, for example, **statefulset-test.yaml**.

The following is an example:

```
apiVersion: apps/v1
kind: StatefulSet
metadata:
  name: test-turbo-subpath
  namespace: default
  generation: 1
  labels:
    appgroup: ''
spec:
  replicas: 2
  selector:
    matchLabels:
      app: test-turbo-subpath
  template:
    metadata:
      labels:
        app: test-turbo-subpath
      annotations:
        metrics.alpha.kubernetes.io/custom-endpoints: '[{"api":"","path":"","port":"","names":""}]'
        pod.alpha.kubernetes.io/initialized: 'true'
    spec:
      containers:
        - name: container-0
          image: 'nginx:latest'
          resources: {}
          volumeMounts:
            - name: sfs-turbo-160024548582479676
              mountPath: /tmp
          terminationMessagePath: /dev/termination-log
          terminationMessagePolicy: File
          imagePullPolicy: IfNotPresent
      restartPolicy: Always
      terminationGracePeriodSeconds: 30
      dnsPolicy: ClusterFirst
      securityContext: {}
      imagePullSecrets:
        - name: default-secret
      affinity: {}
      schedulerName: default-scheduler
  volumeClaimTemplates:
    - metadata:
        name: sfs-turbo-160024548582479676
        namespace: default
        annotations: {}
      spec:
        accessModes:
          - ReadWriteOnce
        resources:
          requests:
            storage: 10Gi
        storageClassName: sfsturbo-subpath-sc
  serviceName: wwww
  podManagementPolicy: OrderedReady
  updateStrategy:
    type: RollingUpdate
  revisionHistoryLimit: 10
```

In this example:

- **name**: indicates the name of the StatefulSet.
- **image**: specifies the image used by the StatefulSet.
- **mountPath**: indicates the mount path of the container. In this example, the volume is mounted to the **/tmp** directory.
- **spec.template.spec.containers.volumeMounts.name** and **spec.volumeClaimTemplates.metadata.name**: must be consistent because they have a mapping relationship.
- **storageClassName**: specifies the name of the StorageClass.

Step 2 Create the StatefulSet.

kubectl create -f statefulset-test.yaml

----End

4.6 Object Storage Service

4.6.1 OBS Volume Overview

Introduction

Object Storage Service (OBS) provides massive, secure, and cost-effective data storage for you to store data of any type and size. You can use it in enterprise backup/archiving, video on demand (VoD), video surveillance, and many other scenarios.

- **Standard APIs**: With HTTP RESTful APIs, OBS allows you to use client tools or third-party tools to access object storage.
- **Data sharing**: Servers, embedded devices, and IoT devices can use the same path to access shared object data in OBS.
- **Public/Private networks**: OBS allows data to be accessed from public networks to meet Internet application requirements.

- **Capacity and performance**: No capacity limit; high performance (read/write I/O latency within 10 ms).
- **Use cases**: Deployments/StatefulSets in the **ReadOnlyMany** mode and jobs created for big data analysis, static website hosting, online VOD, gene sequencing, intelligent video surveillance, backup and archiving, and enterprise cloud boxes (web disks). You can create object storage by using the OBS console, tools, and SDKs.

OBS Specifications

OBS provides multiple storage classes to meet customers' requirements on storage performance and costs.

- Object buckets provide reliable, high-performance, secure, and budget-friendly storage for data. They have no restrictions on the quantity of files or storage capacity.
  - Standard: features low latency and high throughput. It is therefore good for storing frequently (multiple times per month) accessed files or small files (less than 1 MB). Its application scenarios include big data analytics, mobile apps, hot videos, and social apps.
  - OBS Infrequent Access: applicable to storing semi-frequently accessed (less than 12 times a year) data requiring quick response. Its application scenarios include file synchronization or sharing, and enterprise-level backup. This storage class has the same durability, low latency, and high throughput as the Standard storage class, with a lower cost, but its availability is slightly lower than the Standard storage class.
- A parallel file system is a high-performance file system provided by OBS. It is designed to provide high-performance file semantics for big data scenarios. For details, see **About PFS**.

For details about OBS storage classes, see **Storage Classes**.

Performance

Every time an OBS volume is mounted to a container workload, a resident process is created in the backend. When a workload uses too many OBS volumes or reads and writes a large number of object storage files, resident processes will consume a significant amount of memory. The amount of memory required in these scenarios is listed **Table 4-35**. To ensure stable running of the workload, make sure that the number of OBS volumes used does not exceed the requested memory. For example, if the workload requests for 4 GiB of memory, the number of OBS volumes should be **no more than** 4.

Table 4-35 Memory required by a single object storage resident process

Test Item	Memory Usage
Long-term stable running	About 50 MiB
Concurrent write to a 10 MB file from two processes	About 110 MiB

Test Item	Memory Usage
Concurrent write to a 10 MB file from four processes	About 220 MiB
Write to a 100 GB file from a single process	About 300 MiB

Application Scenarios

OBS supports the following mounting modes based on application scenarios:

- **Using an Existing OBS Bucket or Parallel File System Through a Static PV**: static provisioning. You use an existing OBS bucket or high-performance file system to create a PV and then mount the PV to the workload through a PVC. This mode applies to scenarios where the underlying storage is available or billed on a yearly/monthly basis.
- **Using an OBS Bucket or Parallel File System Through a Dynamic PV**: dynamic provisioning. You do not need to create OBS volumes in advance. Instead, specify a StorageClass during PVC creation. An OBS volume and a PV will be automatically created. This mode applies to scenarios where no underlying storage is available.

Billing

- When an OBS volume is mounted, the billing mode of OBS **automatically created** using StorageClass is **pay-per-use** by default. For OBS pricing details, see **OBS Pricing Details**.
- If you want to be billed on a yearly/monthly basis, **use existing OBS volumes**.

4.6.2 Using an Existing OBS Bucket or Parallel File System Through a Static PV

This section describes how to use an existing OBS bucket or parallel file system to statically create PVs and PVCs for implement data persistence and sharing in workloads.

Prerequisites

If you want to create a cluster using commands, kubectl has been used to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Constraints

- If the volumes are created from an OBS bucket or parallel file system, the owner group and permission of the mount point cannot be modified.
- Hard links are not supported when common buckets are mounted.
- Multiple PVs can use the same OBS bucket or parallel file system with the following restrictions:

- All PVCs or PVs that use the same bucket or parallel file system cannot be mounted to a pod. This will result in a pod startup failure because not all PVCs can be mounted to the pod due to the same **volumeHandle** values of these PVs.
- The **persistentVolumeReclaimPolicy** parameter in the PVs should be set to **Retain**. If any other value is used, when a PV is deleted, the associated bucket or parallel file system may be deleted. In this case, other PVs associated with the bucket or parallel file system malfunction.
- If the bucket or parallel file system is repeatedly used, you are required to maintain data consistency. Enable isolation and protection for ReadWriteMany at the application layer and prevent multiple clients from writing the same file to prevent data overwriting and loss.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Statically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **OBS**.
OBS Endpoint	To access an OBS volume in a CCE Autopilot cluster, you need to create a VPC endpoint for OBS.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Creation Method	– If underlying storage is available, create a PV or use an existing PV to statically create a PVC. – If no underlying storage is available, select **Dynamically provision**. For details, see **Using an OBS Bucket or Parallel File System Through a Dynamic PV**. In this example, select **Create new** to create a PV and PVC at the same time on the console.
PV[a]	Select an existing PV in the cluster. Create a PV in advance. For details, see "Creating a storage volume" in **Related Operations**. You do not need to specify this parameter in this example.
OBS[b]	Click **Select OBS**. In the displayed dialog box, select the OBS storage that meets your requirements and click **OK**.
PV Name[b]	Enter the PV name, which must be unique in the same cluster.

Parameter	Description
Access Mode[b]	OBS volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.
Reclaim Policy[b]	You can select **Delete** or **Retain** to specify the reclaim policy of the underlying storage when the PVC is deleted. For details, see **PV Reclaim Policy**. **NOTE** If multiple PVs use the same OBS volume, use **Retain** to avoid cascading deletion of underlying volumes.
Access Key (AK/SK)[b]	**Custom**: Customize a secret if you want to assign different user permissions to different OBS storage devices. For details, see **Using a Custom Access Key (AK/SK) to Mount an OBS Volume**. Only secrets with the **secret.kubernetes.io/used-by = csi** label can be selected. The secret type is cfe/secure-opaque. If no secret is available, click **Create Secret** to create one. – **Name**: Enter a secret name. – **Namespace**: Select the namespace where the secret is located. – **Access Key (AK/SK)**: Upload a key file in .csv format. For details, see **Obtaining an Access Key**.
Mount Options[b]	Enter the mounting parameter key-value pairs. For details, see **Configuring OBS Mount Options**.

NOTE

a: The parameter is available when **Creation Method** is set to **Use existing**.

b: The parameter is available when **Creation Method** is set to **Create new**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

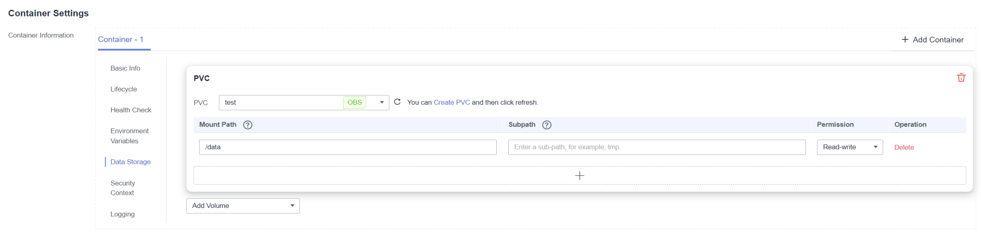
1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

   Mount and use storage volumes, as shown in **Table 4-36**. For details about other parameters, see **Creating a Workload**.

Table 4-36 Mounting a storage volume

Parameter	Description
PVC	Select an existing OBS volume.
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. **tmp**, for example, indicates that data in the mount path of the container is stored in the **tmp** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the OBS bucket or parallel file system.



PVC

PVC: test OBS — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-write	Delete

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Create a PV.

1. Create the **pv-obs.yaml** file.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only      # (Optional) The underlying storage is retained 
when the PV is deleted.
  name: pv-obs       # PV name.
spec:
  accessModes:
  - ReadWriteMany    # Access mode. The value must be ReadWriteMany for OBS.
  capacity:
    storage: 1Gi     # OBS volume capacity.
  csi:
    driver: obs.csi.everest.io        # Dependent storage driver for the mounting.
    fsType: obsfs                     # Instance type.
    volumeHandle: <your_volume_id>    # Name of the OBS volume.
    volumeAttributes:
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/obs-volume-type: STANDARD
      everest.io/region: <your_region>                        # Region where the OBS volume is located.
    nodePublishSecretRef:            # Custom secret of the OBS volume.
      name: <your_secret_name>       # Custom secret name.
      namespace: <your_namespace>    # Namespace of the custom secret.
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-obs               # Storage class name.
  mountOptions: []                         # Mount options.
```

Table 4-37 Key parameters

Parameter	Mandatory	Description
everest.io/reclaim-policy: retain-volume-only	No	Optional. Currently, only **retain-volume-only** is supported. If the reclaim policy is **Delete** and the value is **retain-volume-only**, the associated PV is deleted while the underlying storage volume is retained, when a PVC is deleted.
fsType	Yes	Instance type. The value can be **obsfs** or **s3fs**. – **obsfs**: Parallel file system, which is mounted using **obsfs**. – **s3fs**: Object bucket, which is mounted using s3fs.
volumeHandle	Yes	OBS volume name.

Parameter	Mandatory	Description
everest.io/obs-volume-type	Yes	OBS storage class. – If **fsType** is set to **s3fs**, **STANDARD** (standard bucket) and **WARM** (infrequent access bucket) are supported. – This parameter is invalid when **fsType** is set to **obsfs**.
everest.io/region	Yes	Region where the OBS bucket is deployed. For details, see **Regions and Endpoints**.

nodePublishSecretRef	No	Access key (AK/SK) used for mounting the OBS volume. You can use the AK/SK to create a secret and mount it to the PV. For details, see **Using a Custom Access Key (AK/SK) to Mount an OBS Volume**.
		An example is as follows:
		`nodePublishSecretRef:`
		`  name: secret-demo`
		`  namespace: default`

mountOptions	No	Mount options. For details, see **Configuring OBS Mount Options**.
persistentVolumeReclaimPolicy	Yes	The **Delete** and **Retain** reclaim policies are supported. For details, see **PV Reclaim Policy**. If multiple PVs use the same OBS volume, use **Retain** to avoid cascading deletion of underlying volumes. **Delete**: – If **everest.io/reclaim-policy** is not specified, both the PV and OBS bucket will be deleted when a PVC is deleted. – If **everest.io/reclaim-policy** is set to **retain-volume-only**, when a PVC is deleted, the PV will be deleted but the OBS bucket or parallel file system will be retained. **Retain**: When a PVC is deleted, both the PV and underlying storage are retained. You need to manually delete these resources. After the PVC is deleted, the PV is in the **Released** state and cannot be bound to a PVC again.
storage	Yes	Storage capacity, in Gi. For OBS, this field is used only for verification (cannot be empty or 0). Its value is fixed at **1**, and any value you set does not take effect for OBS.

Parameter	Mandatory	Description
storageClassName	Yes	The StorageClass for OBS volumes is **csi-obs**.

2. Run the following command to create a PV:
```
kubectl apply -f pv-obs.yaml
```

Step 3 Create a PVC.

1. Create the **pvc-obs.yaml** file.
```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-obs
  namespace: default
  annotations:
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
    everest.io/obs-volume-type: STANDARD
    csi.storage.k8s.io/fstype: obsfs
    csi.storage.k8s.io/node-publish-secret-name: <your_secret_name>  # Custom secret name.
    csi.storage.k8s.io/node-publish-secret-namespace: <your_namespace>        # Namespace of the 
custom secret.
spec:
  accessModes:
  - ReadWriteMany                  # The value must be ReadWriteMany for OBS.
  resources:
    requests:
      storage: 1Gi
  storageClassName: csi-obs       # Storage class name, which must be the same as that of the PV.
  volumeName: pv-obs    # PV name.
```

Table 4-38 Key parameters

Parameter	Mandatory	Description
csi.storage.k8s.io/node-publish-secret-name	No	Name of the custom secret specified in the PV.
csi.storage.k8s.io/node-publish-secret-namespace	No	Namespace of the custom secret specified in the PV.
storage	Yes	Requested capacity in the PVC, in Gi. For OBS, this field is used only for verification (cannot be empty or **0**). Its value is fixed at **1**, and any value you set does not take effect for OBS.
storageClassName	Yes	Storage class name, which must be the same as the storage class of the PV in **Step 2.1**. The StorageClass for OBS volumes is **csi-obs**.

Parameter	Mandatory	Description
volumeName	Yes	PV name, which must be the same as the PV name in **Step 2.1**.

2. Run the following command to create a PVC:

```
kubectl apply -f pvc-obs.yaml
```

Step 4 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the OBS volume is mounted to the **/data** path.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-obs-volume    # Volume name, which must be the same as the volume name in the volumes field.
          mountPath: /data  # Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-obs-volume    # Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-obs    # PVC name.
```

2. Run the following command to create a workload that the OBS volume is mounted to:

```
kubectl apply -f web-demo.yaml
```

After the workload is created, you can try **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the created pod:

```
kubectl get pod | grep web-demo
```

Expected output:

```
web-demo-846b489584-mjhm9   1/1     Running   0             46s
web-demo-846b489584-wvv5s   1/1     Running   0             46s
```

2. Run the following commands in sequence to view the files in the **/data** path of the pods:

```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
kubectl exec web-demo-846b489584-wvv5s -- ls /data
```

If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to check the files in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```

Expected output:

```
static
```

Step 4 **Verify data persistence.**

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:

   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```

   Expected output:

   ```
   pod "web-demo-846b489584-mjhm9" deleted
   ```

   After the deletion, the Deployment controller automatically creates a replica.

2. Run the following command to view the created pod:

   ```
   kubectl get pod | grep web-demo
   ```

   The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:

   ```
   web-demo-846b489584-d4d4j   1/1     Running   0             110s
   web-demo-846b489584-wvv5s    1/1     Running   0             7m50s
   ```

3. Run the following command to check whether the files in the **/data** path of the new pod have been modified:

   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```

   Expected output:

   ```
   static
   ```

   The **static** file is retained, indicating that the data in the file system can be stored persistently.

Step 5 **Verify data sharing.**

1. Run the following command to view the created pod:

   ```
   kubectl get pod | grep web-demo
   ```

   Expected output:

   ```
   web-demo-846b489584-d4d4j   1/1     Running   0             7m
   web-demo-846b489584-wvv5s   1/1     Running   0             13m
   ```

2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.

   ```
   kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
   ```

   Check the files in the **/data** path of the pod.

   ```
   kubectl exec web-demo-846b489584-d4d4j -- ls /data
   ```

   Expected output:

   ```
   share
   static
   ```

3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.

   ```
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```

   Expected output:

```
share
static
```

After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-39**.

Table 4-39 Related operations

Operation	Description	Procedure
Creating a storage volume (PV)	You can create a PV on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVs** tab. In the upper right corner, click **Create PersistentVolume**. In the displayed dialog box, configure the parameters. ● **Volume Type**: Select **OBS**. ● **OBS**: Click **Select OBS**. In the dialog box displayed, select the OBS storage that meets your requirements and click **OK**. ● **PV Name**: Enter the PV name, which must be unique in the same cluster. ● **Access Mode**: Only **ReadWriteMany** is available. A storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**. ● **Reclaim Policy**: There are two options: **Retain** and **Delete**. For details, see **PV Reclaim Policy** NOTE If multiple PVs use the same underlying storage volume, use **Retain** to avoid cascading deletion of underlying volumes. ● **Access Key (AK/SK)**: Customize a secret if you want to assign different user permissions to different OBS storage devices. For details, see **Using a Custom Access Key (AK/SK) to Mount an OBS Volume**. Only secrets with the **secret.kubernetes.io/used-by = csi** label can be selected. The secret type is cfe/secure-opaque. If no secret is available, click **Create Secret** to create one. ● **Mount Options**: Enter the mounting parameter key-value pairs. For details, see **Configuring OBS Mount Options**. 2. Click **Create**.
Updating an access key	You can update the access key of object storage on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Update Access Key** in the **Operation** column. 2. Upload a key file in .csv format. For details, see **Obtaining an Access Key**. Click **OK**.

Operatio n	Description	Procedure
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.6.3 Using an OBS Bucket or Parallel File System Through a Dynamic PV

This section describes how to create an OBS volume dynamically. It is applicable when no underlying storage is available.

Constraints

- If the volumes are created from an OBS bucket or parallel file system, the owner group and permission of the mount point cannot be modified.
- Hard links are not supported when common buckets are mounted.
- OBS allows a single user to create a maximum of 100 buckets. If a large number of dynamic PVCs are created, the number of buckets may exceed the upper limit, and no more OBS buckets can be created. In this case, use OBS by calling its API or SDK and do not mount OBS buckets to workloads.

Using the Console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Dynamically create a PVC and PV.

1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. In the upper right corner, click **Create PVC**. In the displayed dialog box, configure the parameters.

Parameter	Description
PVC Type	In this example, select **OBS**.

Parameter	Description
OBS Endpoint	To access an OBS volume in a CCE Autopilot cluster, you need to create a VPC endpoint for OBS.
PVC Name	Enter the PVC name, which must be unique in the same namespace.
Creation Method	– If no underlying storage is available, select **Dynamically provision** to create a PVC, PV, and underlying storage on the console in cascading mode. – If underlying storage is available, select either **Use existing** or **Create new**. For details about static creation, see **Using an Existing OBS Bucket or Parallel File System Through a Static PV**. In this example, select **Dynamically provision**.
Storage Classes	The storage class of OBS volumes is **csi-obs**.
Instance Type	– **Parallel file system**: a high-performance file system provided by OBS. It provides millisecond-level access latency, TB/s-level bandwidth, and million-level IOPS. **Parallel file systems are recommended.** – **Object bucket**: a container that stores objects in OBS. All objects in a bucket are at the same logical level.
OBS Class	You can select the following object bucket types: – **Standard**: Applicable when a large number of hotspot files or small-sized files need to be accessed frequently (multiple times per month on average) and require fast access response. – **Infrequent access**: Applicable when data is not frequently accessed (fewer than 12 times per year on average) but requires fast access response.
Access Mode	OBS volumes support only **ReadWriteMany**, indicating that a storage volume can be mounted to multiple pods in read/write mode. For details, see **Volume Access Modes**.

Parameter	Description
Access Key (AK/SK)	**Custom**: Customize a secret if you want to assign different user permissions to different OBS storage devices. For details, see **Using a Custom Access Key (AK/SK) to Mount an OBS Volume**. Only secrets with the **secret.kubernetes.io/used-by = csi** label can be selected. The secret type is cfe/secure-opaque. If no secret is available, click **Create Secret** to create one. – **Name**: Enter a secret name. – **Namespace**: Select the namespace where the secret is located. – **Access Key (AK/SK)**: Upload a key file in .csv format. For details, see **Obtaining an Access Key**.

2. Click **Create** to create a PVC and a PV.

   In the navigation pane on the left, choose **Storage**. View the created PVC and PV on the **PVCs** and **PVs** tabs, respectively.

Step 3 Create a workload.

1. In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.
2. In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **PVC**.

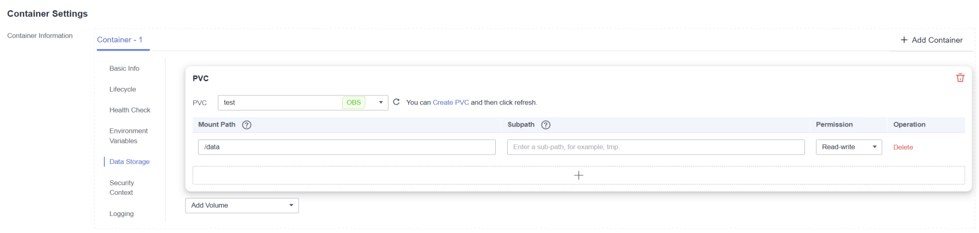
   Mount and use storage volumes, as shown in **Table 4-40**. For details about other parameters, see **Creating a Workload**.

Table 4-40 Mounting a storage volume

Parameter	Description
PVC	Select an existing OBS volume.
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.

Parameter	Description
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. **tmp**, for example, indicates that data in the mount path of the container is stored in the **tmp** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	– **Read-only**: You can only read the data in the mounted volume. – **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

In this example, the volume is mounted to the **/data** path of the container. The container data generated in this path is stored in the OBS bucket or parallel file system.


Container Settings

Container Information

Container - 1 + Add Container

Basic Info
Lifecycle
Health Check
Environment Variables
Data Storage
Security Context
Logging

PVC

PVC: test (OBS) — You can Create PVC and then click refresh.

Mount Path	Subpath	Permission	Operation
/data	Enter a sub-path, for example, tmp.	Read-write	Delete

Add Volume

3. Configure other parameters and click **Create Workload**.

   After the workload is created, the data in the container mount directory will be persistently stored. Verify the storage by referring to **Verifying Data Persistence and Sharing**.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster.

Step 2 Use **StorageClass** to dynamically create a PVC and PV.

1. Create the **pvc-obs-auto.yaml** file.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: pvc-obs-auto
  namespace: default
  annotations:
    everest.io/obs-volume-type: STANDARD    # Object storage type.
    csi.storage.k8s.io/fstype: obsfs        # Instance type.
    csi.storage.k8s.io/node-publish-secret-name: <your_secret_name>       # Custom secret name.
    csi.storage.k8s.io/node-publish-secret-namespace: <your_namespace>    # Namespace of the
custom secret.
spec:
  accessModes:
    - ReadWriteMany             # The value must be ReadWriteMany for OBS.
  resources:
```

```
  requests:
    storage: 1Gi             # OBS volume capacity.
storageClassName: csi-obs    # The StorageClass type of OBS
```

Table 4-41 Key parameters

Parameter	Mandatory	Description
everest.io/obs-volume-type	Yes	OBS storage class. – If **fsType** is set to **s3fs**, **STANDARD** (standard bucket) and **WARM** (infrequent access bucket) are supported. – This parameter is invalid when **fsType** is set to **obsfs**.
csi.storage.k8s.io/fstype	Yes	Instance type. The value can be **obsfs** or **s3fs**. – **obsfs**: Parallel file system, which is mounted using **obsfs**. – **s3fs**: Object bucket, which is mounted using s3fs.
csi.storage.k8s.io/node-publish-secret-name	No	Custom secret name. (Recommended) Select this option if you want to assign different user permissions to different OBS storage devices. For details, see **Using a Custom Access Key (AK/SK) to Mount an OBS Volume**.
csi.storage.k8s.io/node-publish-secret-namespace	No	Namespace of a custom secret.
storage	Yes	Requested capacity in the PVC, in Gi. For OBS, this field is used only for verification (cannot be empty or 0). Its value is fixed at **1**, and any value you set does not take effect for OBS.
storageClassName	Yes	The StorageClass for OBS volumes is **csi-obs**.

2. Run the following command to create a PVC:

```
kubectl apply -f pvc-obs-auto.yaml
```

Step 3 Create a workload.

1. Create a file named **web-demo.yaml**. In this example, the OBS volume is mounted to the **/data** path.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: web-demo
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: web-demo
  template:
    metadata:
      labels:
        app: web-demo
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: pvc-obs-volume    #Volume name, which must be the same as the volume name in the volumes field.
          mountPath: /data  # Location where the storage volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: pvc-obs-volume    # Volume name, which can be changed as needed.
          persistentVolumeClaim:
            claimName: pvc-obs-auto    # PVC name.
```

2. Run the following command to create a workload that the OBS volume is mounted to:

   ```
   kubectl apply -f web-demo.yaml
   ```

   After the workload is created, you can try **Verifying Data Persistence and Sharing**.

----End

Verifying Data Persistence and Sharing

Step 1 View the deployed application and files.

1. Run the following command to view the created pod:

   ```
   kubectl get pod | grep web-demo
   ```

   Expected output:

   ```
   web-demo-846b489584-mjhm9   1/1     Running   0             46s
   web-demo-846b489584-wvv5s   1/1     Running   0             46s
   ```

2. Run the following commands in sequence to view the files in the **/data** path of the pods:

   ```
   kubectl exec web-demo-846b489584-mjhm9 -- ls /data
   kubectl exec web-demo-846b489584-wvv5s -- ls /data
   ```

   If no result is returned for both pods, no file exists in the **/data** path.

Step 2 Run the following command to create a file named **static** in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 --  touch /data/static
```

Step 3 Run the following command to check the files in the **/data** path:

```
kubectl exec web-demo-846b489584-mjhm9 -- ls /data
```

Expected output:

```
static
```

Step 4 **Verify data persistence.**

1. Run the following command to delete the pod named **web-demo-846b489584-mjhm9**:

   ```
   kubectl delete pod web-demo-846b489584-mjhm9
   ```

Expected output:

```
pod "web-demo-846b489584-mjhm9" deleted
```

After the deletion, the Deployment controller automatically creates a replica.

2. Run the following command to view the created pod:

```
kubectl get pod | grep web-demo
```

The expected output is as follows, in which **web-demo-846b489584-d4d4j** is the newly created pod:

```
web-demo-846b489584-d4d4j   1/1     Running   0            110s
web-demo-846b489584-wvv5s    1/1     Running   0            7m50s
```

3. Run the following command to check whether the files in the **/data** path of the new pod have been modified:

```
kubectl exec web-demo-846b489584-d4d4j -- ls /data
```

Expected output:

```
static
```

The **static** file is retained, indicating that the data in the file system can be stored persistently.

Step 5 **Verify data sharing.**

1. Run the following command to view the created pod:

```
kubectl get pod | grep web-demo
```

Expected output:

```
web-demo-846b489584-d4d4j   1/1     Running   0            7m
web-demo-846b489584-wvv5s   1/1     Running   0            13m
```

2. Run the following command to create a file named **share** in the **/data** path of either pod: In this example, select the pod named **web-demo-846b489584-d4d4j**.

```
kubectl exec web-demo-846b489584-d4d4j --  touch /data/share
```

Check the files in the **/data** path of the pod.

```
kubectl exec web-demo-846b489584-d4d4j -- ls /data
```

Expected output:

```
share
static
```

3. Check whether the **share** file exists in the **/data** path of another pod (**web-demo-846b489584-wvv5s**) as well to verify data sharing.

```
kubectl exec web-demo-846b489584-wvv5s -- ls /data
```

Expected output:

```
share
static
```

After you create a file in the **/data** path of a pod, if the file is also created in the **/data** path of the other pod, the two pods share the same volume.

----End

Related Operations

You can also perform the operations described in **Table 4-42**.

Table 4-42 Related operations

Operation	Description	Procedure
Updating an access key	Update the access key of object storage on the CCE console.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** tab. Locate the target PVC and click **More** > **Update Access Key** in the **Operation** column. 2. Upload a key file in .csv format. For details, see **Obtaining an Access Key**. Click **OK**.
Viewing events	You can view event names, event types, number of occurrences, Kubernetes events, first occurrence time, and last occurrence time of the PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View Events** in the **Operation** column to view events generated within one hour (events are retained for one hour).
Viewing a YAML file	You can view, copy, and download the YAML files of a PVC or PV.	1. In the navigation pane on the left, choose **Storage**. Then click the **PVCs** or **PVs** tab. 2. Locate the target PVC or PV, click **View YAML** in the **Operation** column to view or download the YAML.

4.6.4 Configuring OBS Mount Options

This section describes how to configure OBS volume mount options. You can configure mount options in a PV and bind the PV to a PVC. Alternatively, configure mount options in a StorageClass and use the StorageClass to create a PVC. In this way, PVs can be dynamically created and inherit mount options configured in the StorageClass by default.

OBS Mount Options

When mounting an OBS volume, the Everest add-on presets the options described in **Table 4-43** and **Table 4-44** by default. The options in **Table 4-43** are mandatory.

Table 4-43 Mandatory mount options configured by default

Parameter	Value	Description
use_ino	Leave it blank.	If enabled, obsfs allocates the **inode** number. Enabled by default in read/write mode.

Parameter	Value	Description
big_writes	Leave it blank.	If configured, the maximum size of the cache can be modified.
nonempty	Leave it blank.	Allows non-empty mount paths.
allow_other	Leave it blank.	Allows other users to access the parallel file system.
no_check_certificate	Leave it blank.	Disables server certificate verification.
sigv2	Leave it blank.	Specifies the signature version. Used by default in object buckets.
public_bucket	1	If this parameter is set to **1**, public buckets are mounted anonymously. Enabled by default in object bucket read-only mode.

Table 4-44 Optional mount options configured by default

Parameter	Value	Description
max_write	131072	This parameter is valid only when **big_writes** is configured. The recommended value is **128 KB**.
ssl_verify_hostname	0	Disables SSL certificate verification based on the host name.
max_background	100	Allows setting the maximum number of waiting requests in the background. Used by default in parallel file systems.
umask	0	Mask of the configuration file permission. For example, if the **umask** value is **022**, the directory permission (the maximum permission is **777**) is **755** (777 – 022 = 755, rwxr-xr-x).

Configuring Mount Options in a PV

You can use the **mountOptions** field to configure mount options in a PV. The options you can configure in **mountOptions** are listed in **OBS Mount Options**.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Configure mount options in a PV. The following is an example:

```
apiVersion: v1
kind: PersistentVolume
metadata:
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
    everest.io/reclaim-policy: retain-volume-only      # (Optional) The underlying storage is retained when
the PV is deleted.
  name: pv-obs       # PV name.
spec:
  accessModes:
  - ReadWriteMany    # Access mode. The value must be ReadWriteMany for OBS.
  capacity:
    storage: 1Gi     # OBS volume capacity.
  csi:
    driver: obs.csi.everest.io        # Dependent storage driver for the mounting.
    fsType: obsfs                    # Instance type.
    volumeHandle: <your_volume_id>    # Name of the OBS volume.
    volumeAttributes:
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
      everest.io/obs-volume-type: STANDARD
      everest.io/region: <your_region>                        # Region where the OBS volume is located.
      everest.io/enterprise-project-id: <your_project_id>     # (Optional) Enterprise project ID. If an enterprise
project is specified, use the same enterprise project when creating a PVC. Otherwise, the PVC cannot be
bound to a PV.

    nodePublishSecretRef:            # Custom secret of the OBS volume.
      name: <your_secret_name>       # Custom secret name.
      namespace: <your_namespace>    # Namespace of the custom secret.
  persistentVolumeReclaimPolicy: Retain    # Reclaim policy.
  storageClassName: csi-obs               # Storage class name.
  mountOptions:                            # Mount options.
  - umask=027
```

Step 3 After a PV is created, you can create a PVC and bind it to the PV, and then mount the PV to the container in the workload. For details, see **Using an Existing OBS Bucket or Parallel File System Through a Static PV**.

----End

Configuring Mount Options in a StorageClass

You can use the **mountOptions** field to configure mount options in a StorageClass. The options you can configure in **mountOptions** are listed in **OBS Mount Options**.

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a customized StorageClass. Example:

```
kind: StorageClass
apiVersion: storage.k8s.io/v1
metadata:
  name: csi-obs-mount-option
provisioner: everest-csi-provisioner
parameters:
  csi.storage.k8s.io/csi-driver-name: obs.csi.everest.io
  csi.storage.k8s.io/fstype: s3fs
  everest.io/obs-volume-type: STANDARD
reclaimPolicy: Delete
volumeBindingMode: Immediate
mountOptions:                          # Mount options.
- umask=027
```

Step 3 After the StorageClass is configured, you can use it to create a PVC. By default, the dynamically created PVs inherit the mount options configured in the StorageClass. For details, see **Using an Existing OBS Bucket or Parallel File System Through a Static PV**.

----End

4.6.5 Using a Custom Access Key (AK/SK) to Mount an OBS Volume

Scenario

CCE supports custom access keys so IAM users can use their own custom access keys to mount an OBS volume. For details, see **How Can I Control Access to OBS?**

Constraints

When an OBS volume is mounted using a custom access key (AK/SK), the access key cannot be deleted or disabled. Otherwise, the service container cannot access the mounted OBS volume.

Obtaining an Access Key

Step 1 Log in to the management console.

Step 2 Hover the cursor over the username in the upper right corner and choose **My Credentials** from the drop-down list.

Step 3 In the navigation pane on the left, choose **Access Keys**.

Step 4 Click **Create Access Key**. The **Create Access Key** dialog box is displayed.

Step 5 Click **OK** to download the access key.

----End

Creating a Secret Using an Access Key

Step 1 Obtain an access key.

Step 2 Encode the keys using Base64. (Assume that the AK is xxx and the SK is yyy.)

echo -n xxx|base64

echo -n yyy|base64

Record the encoded AK and SK.

Step 3 Create a YAML file for the secret, for example, **test-user.yaml**.

```
apiVersion: v1
data:
  access.key: WE5WWVhVNU*****
  secret.key: Nnk4emJyZ0*****
kind: Secret
metadata:
  name: test-user
  namespace: default
  labels:
    secret.kubernetes.io/used-by: csi
type: cfe/secure-opaque
```

Specifically:

Parameter	Description
access.key	Base64-encoded AK.
secret.key	Base64-encoded SK.
name	Secret name.
namespace	Namespace of the secret.
secret.kubernetes.io/used-by: csi	Add this label in the YAML file if you want to make it available on the CCE console when you create an OBS PV/PVC.
type	Secret type. The value must be **cfe/secure-opaque**. When this type is used, the data entered by users is automatically encrypted.

Step 4 Create the secret.

kubectl create -f test-user.yaml

----End

Mounting a Secret When Statically Creating an OBS Volume

After a secret is created using the AK/SK, you can associate the secret with the PV to be created and then use the AK/SK in the secret to mount an OBS volume.

Step 1 Log in to the OBS console, create an OBS bucket, and record the bucket name and storage class. The parallel file system is used as an example.

Step 2 Create a YAML file for the PV, for example, **pv-example.yaml**.

```
apiVersion: v1
kind: PersistentVolume
metadata:
  name: pv-obs-example
  annotations:
    pv.kubernetes.io/provisioned-by: everest-csi-provisioner
spec:
  accessModes:
  - ReadWriteMany
  capacity:
    storage: 1Gi
  csi:
    nodePublishSecretRef:
      name: test-user
      namespace: default
    driver: obs.csi.everest.io
    fsType: obsfs
    volumeAttributes:
      everest.io/obs-volume-type: STANDARD
      everest.io/region: ap-southeast-1
      storage.kubernetes.io/csiProvisionerIdentity: everest-csi-provisioner
    volumeHandle: obs-normal-static-pv
  persistentVolumeReclaimPolicy: Delete
  storageClassName: csi-obs
```

Parameter	Description
nodePublishSecre-tRef	Secret specified during the mounting. ● **name**: name of the secret ● **namespace**: namespace of the secret
fsType	File type. The value can be **obsfs** or **s3fs**. If the value is **s3fs**, an OBS bucket is created and mounted. If the value is **obsfs**, an OBS parallel file system is created and mounted.
volumeHandle	OBS bucket name.

Step 3 Create a PV.

kubectl create -f pv-example.yaml

After a PV is created, you can create a PVC and associate it with the PV.

Step 4 Create a YAML file for the PVC, for example, **pvc-example.yaml**.

Example YAML file for the PVC:

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  annotations:
    csi.storage.k8s.io/node-publish-secret-name: test-user
    csi.storage.k8s.io/node-publish-secret-namespace: default
    volume.beta.kubernetes.io/storage-provisioner: everest-csi-provisioner
    everest.io/obs-volume-type: STANDARD
    csi.storage.k8s.io/fstype: obsfs
  name: obs-secret
  namespace: default
spec:
  accessModes:
  - ReadWriteMany
  resources:
    requests:
      storage: 1Gi
  storageClassName: csi-obs
  volumeName: pv-obs-example
```

Parameter	Description
csi.storage.k8s.io/node-publish-secret-name	Secret name
csi.storage.k8s.io/node-publish-secret-namespace	Namespace of the secret

Step 5 Create a PVC.

kubectl create -f pvc-example.yaml

After the PVC is created, you can create a workload and associate it with the PVC to create volumes.

----End

Mounting a Secret When Dynamically Creating an OBS Volume

When dynamically creating an OBS volume, you can use the following method to specify a secret:

Step 1 Create a YAML file for the PVC, for example, **pvc-example.yaml**.

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  annotations:
    csi.storage.k8s.io/node-publish-secret-name: test-user
    csi.storage.k8s.io/node-publish-secret-namespace: default
    everest.io/obs-volume-type: STANDARD
    csi.storage.k8s.io/fstype: obsfs
  name: obs-secret
  namespace: default
spec:
  accessModes:
  - ReadWriteMany
  resources:
    requests:
      storage: 1Gi
  storageClassName: csi-obs
```

Parameter	Description
csi.storage.k8s.io/node-publish-secret-name	Secret name
csi.storage.k8s.io/node-publish-secret-namespace	Namespace of the secret

Step 2 Create a PVC.

kubectl create -f pvc-example.yaml

After the PVC is created, you can create a workload and associate it with the PVC to create volumes.

----End

Verification

You can use a secret of an IAM user to mount an OBS volume. Assume that a workload named **obs-secret** is created, the mount path in the container is **/temp**, and the IAM user has the CCE **ReadOnlyAccess** and **Tenant Guest** permissions.

1. Query the name of the workload pod.

   kubectl get po | grep obs-secret

   Expected outputs:

   ```
   obs-secret-5cd558f76f-vxslv          1/1     Running   0          3m22s
   ```

2. Query the objects in the mount path. In this example, the query is successful.

   kubectl exec obs-secret-5cd558f76f-vxslv -- ls -l /temp/

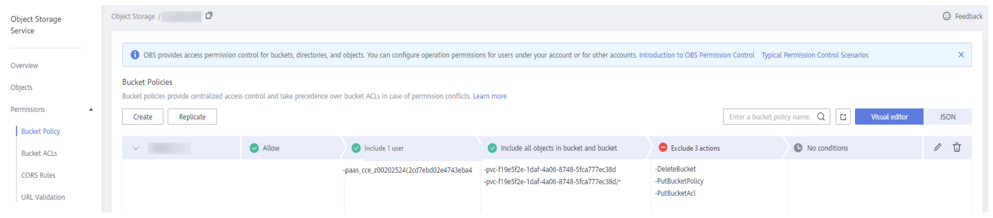
3. Write data into the mount path. In this example, the write operation failed.

   kubectl exec obs-secret-5cd558f76f-vxslv -- touch /temp/test

   Expected outputs:

```
touch: setting times of '/temp/test': No such file or directory
command terminated with exit code 1
```

4. Set the read/write permissions for the IAM user who mounted the OBS volume by referring to the bucket policy configuration.



OBS provides access permission control for buckets, directories, and objects. You can configure operation permissions for users under your account or for other accounts. Introduction to OBS Permission Control Typical Permission Control Scenarios

Bucket Policies

Bucket policies provide centralized access control and take precedence over bucket ACLs in case of permission conflicts. Learn more

[redacted]	Allow	Include 1 user	Include all objects in bucket and bucket	Exclude 3 actions	No conditions
		-paas_cce_z00202524(2cd7ebd02e4743eba4	-pvc-f19e5f2e-1daf-4a06-8748-5fca777ec38d -pvc-f19e5f2e-1daf-4a06-8748-5fca777ec38d/*	-DeleteBucket -PutBucketPolicy -PutBucketAcl	

5. Write data into the mount path again. In this example, the write operation succeeded.

   kubectl exec obs-secret-5cd558f76f-vxslv -- touch /temp/test

6. Check the mount path in the container to see whether the data is successfully written.

   kubectl exec obs-secret-5cd558f76f-vxslv -- ls -l /temp/

   Expected outputs:

```
-rwxrwxrwx 1 root root 0 Jun  7 01:52 test
```

4.7 emptyDir

An emptyDir volume provides ephemeral storage for pods. Its lifecycle is the same as that of a pod. Memory can be specified as the storage medium. When the pod is deleted, the emptyDir volume is deleted, and the data is lost.

Using the Console to Use a Temporary Path

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab.

Step 3 In the upper right corner, click **Create Workload**. On the displayed page, click **Data Storage** in the **Container Settings** area and click **Add Volume** to select **emptyDir**.

Table 4-45 describes the parameters for mounting the volume. For details about other parameters, see **Creating a Workload**.

Table 4-45 Parameters for mounting an emptyDir volume

Parameter	Description
Storage Medium	**Memory**: ● You can select this option to improve the running speed, but the storage capacity is subject to the memory size. This option is suitable when data volume is small and efficient read and write is required. ● If this option is not selected, data is stored in local disks. This fits to the scenarios where a large amount of data needs to be stored but read/write efficiency is not high. **NOTE** ● If **Memory** is selected, pay attention to the memory size. If the storage capacity exceeds the memory size, OOM will occur. ● If **Memory** is selected, the size of an emptyDir volume is the same as the pod specifications. ● If **Memory** is not selected, emptyDir volumes will not occupy the system memory.
Mount Path	Enter a mount path, for example, **/tmp**. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter the subpath of the storage volume and mount a path in the storage volume to the container. In this way, different folders of the same storage volume can be used in a single pod. Enter a subpath, for example, **tmp**, indicating that data in the mount path of the container is stored in the **tmp** directory of the storage volume. If this parameter is left blank, the root path is used by default.
Permission	● **Read-only**: You can only read the data in the mounted volume. ● **Read-write**: You can modify the volume mounted to the path. Newly written data will not be migrated if the container is migrated, which may cause data loss.

Step 4 Configure other parameters and click **Create Workload**.

----End

Using kubectl to Use a Temporary Path

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-emptydir.yaml** and edit it.

vi nginx-emptydir.yaml

The content of the YAML file is as follows:

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-emptydir
  namespace: default
spec:
  replicas: 2
  selector:
    matchLabels:
      app: nginx-emptydir
  template:
    metadata:
      labels:
        app: nginx-emptydir
    spec:
      containers:
        - name: container-1
          image: nginx:latest
          volumeMounts:
            - name: vol-emptydir     # Volume name, which must be the same as the volume name in the 
volumes field.
              mountPath: /tmp       # Path to which an emptyDir volume is mounted.
      imagePullSecrets:
        - name: default-secret
      volumes:
        - name: vol-emptydir        # Volume name, which can be customized.
          emptyDir:
            medium: Memory          # emptyDir volume medium: If this parameter is set to Memory, the 
memory is enabled. If this parameter is left blank, the native default storage medium is used.
            sizeLimit: 1Gi          # Volume capacity.
```

Step 3 Create a workload.

kubectl apply -f nginx-emptydir.yaml

----End

4.8 Increasing the Ephemeral Storage Space of a Pod

Each pod in a workload is provided with 30-GiB free storage capacity by default (maximum IOPS of 2,500 and maximum burst IOPS of 16,000). The OS and CCE platform components will occupy some storage space, and the total available capacity for images, containers, and ephemeral storage is about 20 GiB. If the storage capacity cannot meet service requirements, you can add ephemeral storage for a pod on the console or by running kubectl commands. The ephemeral storage is billed by capacity and usage duration. For details, see **Pricing Details**.

Constraints

The cluster version must be v1.27.8-r0, v1.28.6-r0, or later. If the cluster version does not meet the requirements, you need to upgrade the cluster.

Adding Ephemeral Storage for a Pod

Increase the ephemeral storage capacity for a pod using the console or kubectl.

Using the CCE Console

The following are operations for you to add ephemeral storage capacity for a pod on the console. For details about workload parameters, see **Creating a Workload**.

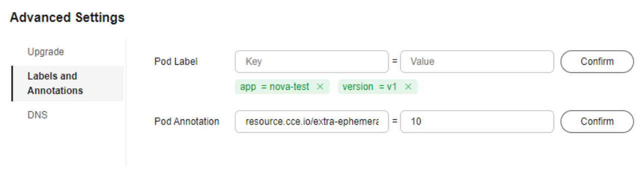
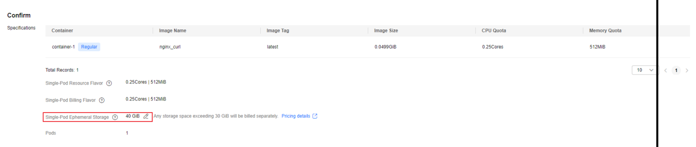
Step 1 Log in to the **CCE console** and click the cluster name to access the cluster console.

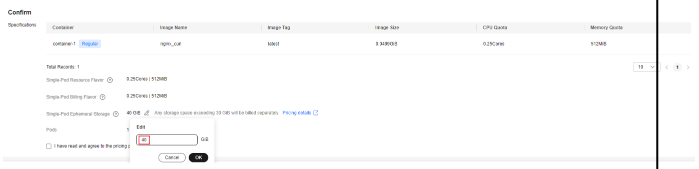
Step 2 In the navigation pane on the left, choose **Workloads**.

- To add ephemeral storage for a pod during workload creation, select the workload type and click **Create Workload** in the upper right corner of the page.
- To add ephemeral storage for a pod of an existing workload, locate the workload and click **Upgrade** in the **Operation** column.

Step 3 Add ephemeral storage for a pod in either of the following methods.

Table 4-46 Two methods of adding ephemeral storage for a pod

Method	Example	Procedure
Adding an annotation	Key: resource.cce.io/extra-ephemeral-storage-in-GiB Value: 10	1. In the **Advanced Settings** area, choose **Labels and Annotations**. 2. Enter **resource.cce.io/extra-ephemeral-storage-in-GiB** as the key of the pod annotation. 3. In the **Value** text box of the pod annotation, enter the storage capacity (a value from **0** to **994**) to be added and click **Confirm**. **NOTE** After this annotation is added to the pod, the ephemeral storage per pod in the area for confirming the specifications changes from 30 GiB to 40 GiB (the default pod storage plus the added ephemeral storage). **Figure 4-16** Adding ephemeral storage for a pod  Advanced Settings Upgrade / Labels and Annotations / DNS Pod Label: Key = Value Confirm app = nova-test × version = v1 × Pod Annotation: resource.cce.io/extra-ephemera = 10 Confirm **Figure 4-17** Single-Pod Ephemeral Storage  Confirm Specifications Container: container-1 Regular; Image Name: nginx_curl; Image Tag: latest; Image Size: 0.0499GiB; CPU Quota: 0.25Cores; Memory Quota: 512MiB Total Records: 1 Single-Pod Resource Flavor: 0.25Cores | 512MiB Single-Pod Billing Flavor: 0.25Cores | 512MiB Single-Pod Ephemeral Storage: 40 GiB Any storage space exceeding 30 GiB will be billed separately. Pricing details Pods: 1

Method	Example	Procedure
Modifying the ephemeral storage for a pod	40 GiB	1. In the area for confirming the specifications, click next to **Single-Pod Ephemeral Storage**. 2. In the displayed dialog box, enter the required capacity. The required capacity is the sum of the default 30-GiB capacity and the ephemeral storage to be added. The value ranges from **30** to **1024**. **NOTICE** If you are not sure how much ephemeral storage needs to be added for a pod, retain the recommended value, which is the value of **Single-Pod Ephemeral Storage**. After an image is selected, the recommended value will be calculated based on the size of the selected image and the storage capacity required by the OS. The recommended ephemeral storage for a pod is either the 30-GiB free storage or the sum of system storage and 2-fold image size, whichever is greater. (Formula: Recommended ephemeral storage for a pod = Max. (30-GiB free storage, Sum of system storage and image size x 2)) **Figure 4-18** Modifying the ephemeral storage for a pod  Confirm Specifications Container: container-1 Regular; Image Name: nginx_curl; Image Tag: latest; Image Size: 0.0499GiB; CPU Quota: 0.25Cores; Memory Quota: 512MiB Total Records: 1 Single-Pod Resource Flavor: 0.25Cores | 512MiB Single-Pod Billing Flavor: 0.25Cores | 512MiB Single-Pod Ephemeral Storage: 40 GiB Any storage space exceeding 30 GiB will be billed separately. Pricing details Pods: 1 I have read and agree to the pricing p Edit: 40 GiB Cancel OK 3. Click **OK**.

Step 4 Select **I have read and agree to the pricing policy** and click **Create Workload** or **Upgrade Workload**.

----End

Using kubectl

You can run kubectl commands to add ephemeral storage capacity for a pod. You can add ephemeral storage capacity when creating a workload or modify the ephemeral storage for a pod of an existing workload. For details about workload parameters, see **Creating a Workload**.

NOTE

Command line operations are required. You can perform related operations in either of the following ways:

- Use the command line tool in the cluster. The kubectl commands have been configured for the command line tool, which has been connected to the cluster.
- Log in to an ECS that is in the same VPC as the cluster, and access the cluster using kubectl by following the instructions in **Connecting to a Cluster Using kubectl**.

Create a workload and add ephemeral storage for a pod.

1. Create a YAML file for creating a workload. The file name can be customized. In this example, the file name is **extra-ephemeral-storage-test.yaml**.
   ```
   vim extra-ephemeral-storage-test.yaml
   ```
   Example file content:
   ```
   apiVersion: apps/v1
   kind: Deployment
   metadata:
     name: extra-ephemeral-storage-test # Workload name
   spec:
     replicas: 1 # Number of pods
     selector:
       matchLabels: # Selector for selecting resources with specific labels
         app: nginx
     template:
       metadata:
         labels:     # Labels
           app: nginx
         annotations:
           resource.cce.io/extra-ephemeral-storage-in-GiB: '10'
       spec:
         containers:
         - image: nginx:latest   # {Image name}:{Image tag}
           name: nginx
         imagePullSecrets:
         - name: default-secret
   ```
   resource.cce.io/extra-ephemeral-storage-in-GiB indicates the ephemeral storage capacity added to the pod. The value ranges from **0** to **994**. In this example, the value is **10**. The total ephemeral storage capacity of a pod is the sum of the 30-GiB default capacity and the value of this parameter.

   Press **Esc** to exit editing mode and enter **:wq** to save the file.
2. Create a workload.
   ```
   kubectl create -f extra-ephemeral-storage-test.yaml
   ```
3. View the workload status.
   ```
   kubectl get deployment
   ```
   If **READY** is **1/1** in the command output, the workload has been created. The ephemeral storage capacity of the pod in the workload has changed to the value you set.
   ```
   NAME         READY   UP-TO-DATE   AVAILABLE   AGE
   nova-test    1/1     1            1           59s
   ```
4. View the ephemeral storage capacity of the pod.

   Run the following command to query the pod name (**extra-ephemeral-storage-test** indicates the workload name):
   ```
   kubectl get pod | grep extra-ephemeral-storage-test
   ```
   Information similar to the following is displayed:
   ```
   extra-ephemeral-storage-test-85dbdb8c5d-4vftc      1/1     Running   0          2m24s
   ```
   Run the following command to access the pod:
   ```
   kubectl exec -it extra-ephemeral-storage-test-85dbdb8c5d-4vftc /bin/sh
   ```
   Query the ephemeral storage capacity of the pod.
   ```
   lsblk
   ```
   If the total storage capacity of the pod is 40 GiB in the command output, the ephemeral storage is added for the pod. Enter **exit** and press **Enter** to exit the pod.
   ```
   NAME    MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
   vda      254:0    0   40G  0 disk
   ```

```
├─vda1 254:1    0   59M  0 part
├─vda2 254:2    0  1.5G  0 part
├─vda3 254:3    0  1.5G  0 part
└─vda4 254:4    0   37G  0 part /etc/resolv.conf
```

Add ephemeral storage capacity for a pod of an existing workload.

1. Modify the YAML file of the target workload and add the **resource.cce.io/extra-ephemeral-storage-in-GiB** annotation. The following uses a Deployment named **extra-ephemeral-storage-test** as an example.

   ```
   kubectl edit deployment extra-ephemeral-storage-test
   ```

   Add the **resource.cce.io/extra-ephemeral-storage-in-GiB** annotation in the lower part of the file. Press **Esc** to exit editing mode and enter **:wq** to save the file.

   ```
   template:
     metadata:
       creationTimestamp: null
       labels:
         app: nginx
       annotations:
         resource.cce.io/extra-ephemeral-storage-in-GiB: '10'
   ```

   resource.cce.io/extra-ephemeral-storage-in-GiB indicates the ephemeral storage capacity added to the pod. The value ranges from **0** to **994**. In this example, the value is **10**. The total ephemeral storage capacity of a pod is the sum of the 30-GiB default capacity and the value of this parameter.

2. View the workload status.

   ```
   kubectl get deployment
   ```

   If **READY** is **1/1**, the workload has been upgraded. The ephemeral storage capacity of the pod in the workload has changed to the value you set.

   ```
   NAME        READY   UP-TO-DATE   AVAILABLE   AGE
   nova-test   1/1     1            1           59m
   ```

3. View the ephemeral storage capacity of the pod.

   Run the following command to query the pod name (**extra-ephemeral-storage-test** indicates the workload name):

   ```
   kubectl get pod | grep extra-ephemeral-storage-test
   ```

   Information similar to the following is displayed:

   ```
   extra-ephemeral-storage-test-85dbdb8c5d-4vftc     1/1     Running   0          60m24s
   ```

   Run the following command to access the pod:

   ```
   kubectl exec -it extra-ephemeral-storage-test-85dbdb8c5d-4vftc /bin/sh
   ```

   Query the ephemeral storage capacity of the pod.

   ```
   lsblk
   ```

   If the total storage capacity of the pod is 40 GiB in the command output, the ephemeral storage is added for the pod. Enter **exit** and press **Enter** to exit the pod.

   ```
   NAME    MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
   vda     254:0    0   40G  0 disk
   ├─vda1 254:1    0   59M  0 part
   ├─vda2 254:2    0  1.5G  0 part
   ├─vda3 254:3    0  1.5G  0 part
   └─vda4 254:4    0   37G  0 part /etc/resolv.conf
   ```

5 O&M

5.1 Monitoring Center

5.1.1 Enabling Cluster Monitoring

To enable monitoring for a cluster, you need to install the Cloud Native Cluster Monitoring add-on for metric collection. After cluster monitoring is enabled, cluster metrics are collected and reported to AOM instances. This section describes how to enable cluster monitoring.

NOTICE

- After cluster monitoring is enabled, cluster metrics are reported to the selected AOM instance. Basic metrics are free, but custom metrics are billed based on the standard pricing of AOM. For details, see **AOM Pricing Details**.
- Running the Cloud Native Monitoring add-on in a cluster consumes cluster resources. Ensure that there are required cluster resources for installing the add-on. To view resource consumption, go to the add-on details page.

Prerequisites

You have an account in the **admin** user group to delegate CCE and its dependent services.

The authorization dialog box is automatically displayed on the **Monitoring Center** page. After you confirm the authorization, the system automatically completes the authorization.

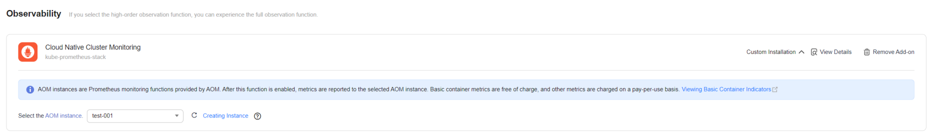
Constraints

Before using Monitoring Center, you need to use an account in the **admin** user group to delegate CCE and its dependent services. After the authorization is complete, users with the CCE Administrator role or CCE FullAccess permission can perform all operations on Monitoring Center. Users with the CCE ReadOnlyAccess permission can view all resource information but cannot perform any operations.

Enabling Cluster Monitoring

- **Enabling cluster monitoring during cluster purchase**
  a. Log in to the CCE console and purchase a cluster.
  b. On the **Select Add-on** page, select the **Cloud Native Cluster Monitoring** add-on.
  c. On the **Add-on Configuration** page, select the AOM instance to be interconnected with the add-on. If there is no access code, create one first.

Figure 5-1 Enabling cluster monitoring



Observability If you select the high-order observation function, you can experience the full observation function.

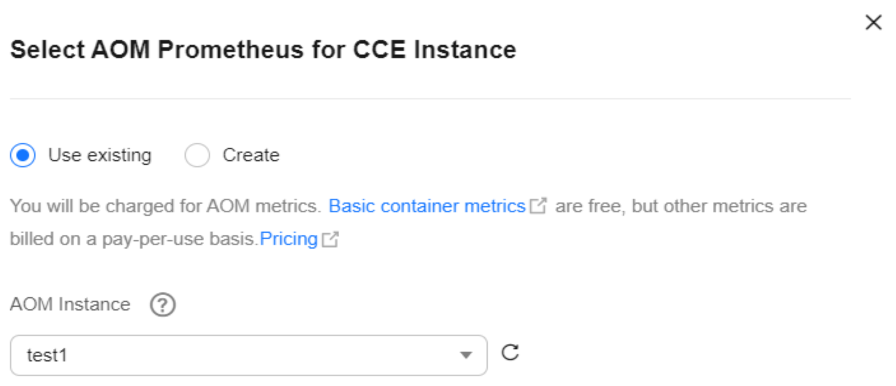
Cloud Native Cluster Monitoring
kube-prometheus-stack

AOM instances are Prometheus monitoring functions provided by AOM. After this function is enabled, metrics are reported to the selected AOM instance. Basic container metrics are free of charge, and other metrics are charged on a pay-per-use basis. Viewing Basic Container Indicators

Select the AOM instance. test-001

- **Enabling cluster monitoring on the Monitoring Center page**
  a. Click the cluster name to access the cluster console. In the navigation pane on the left, choose **Monitoring Center**.
  b. Click **Enable** and then select the AOM instance that metrics are reported to.

Figure 5-2 Enabling cluster monitoring



Select AOM Prometheus for CCE Instance

[✓ Use existing] Create

You will be charged for AOM metrics. Basic container metrics are free, but other metrics are billed on a pay-per-use basis. Pricing

AOM Instance

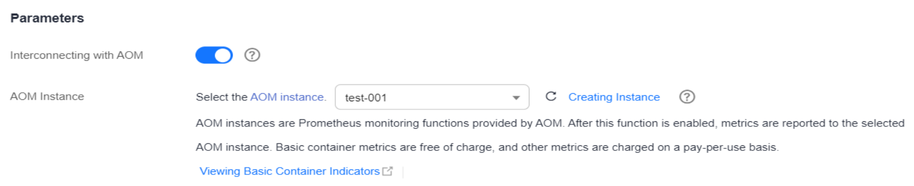
test1

  c. Wait for 3 to 5 minutes until the monitoring data is reported to the AOM instance.

     The functions of Monitoring Center are available.

- **Enabling cluster monitoring on the Add-ons page**
  a. Click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**.
  b. Select the Cloud Native Cluster Monitoring add-on and click **Install**.
  c. Enable the option for interconnecting with AOM to report the metrics to the selected AOM instance.

Figure 5-3 Installing the Cloud Native Cluster Monitoring add-on



Parameters

Interconnecting with AOM	●
AOM Instance	Select the AOM instance. test-001 ⟳ Creating Instance

AOM instances are Prometheus monitoring functions provided by AOM. After this function is enabled, metrics are reported to the selected AOM instance. Basic container metrics are free of charge, and other metrics are charged on a pay-per-use basis.

Viewing Basic Container Indicators

d. Wait for 3 to 5 minutes until the monitoring data is reported to the AOM instance.

The functions of Monitoring Center are available.

NOTE

To disable cluster monitoring, uninstall the Cloud Native Cluster Monitoring add-on on the **Add-ons** page or disable the option for interconnecting with AOM.

5.1.2 Cluster Monitoring

On the **Clusters** tab, you can view the monitoring data of each cluster from multiple dimensions, as described in **Resource Overview**,**Top Resource Consumption Statistics**, and **Data Plane Monitoring**.

Navigation Path

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

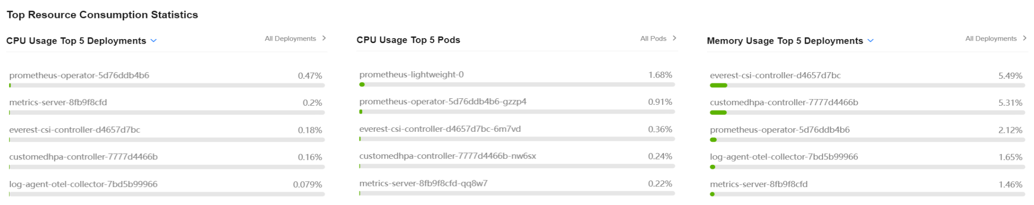
Step 2 In the navigation pane on the left, choose **Monitoring Center**. Then click the **Clusters** tab.

----End

Resource Overview

Resource Overview displays the percentages of abnormal resources in workloads and pods and the total number of namespaces.

Figure 5-4 Resource Overview



Top Resource Consumption Statistics

CPU Usage Top 5 Deployments	All Deployments
prometheus-operator-5d76ddb4b6	0.47%
metrics-server-8fb9f8cfd	0.2%
everest-csi-controller-d4657d7bc	0.18%
customedhpa-controller-7777d4466b	0.16%
log-agent-otel-collector-7bd5b99966	0.079%

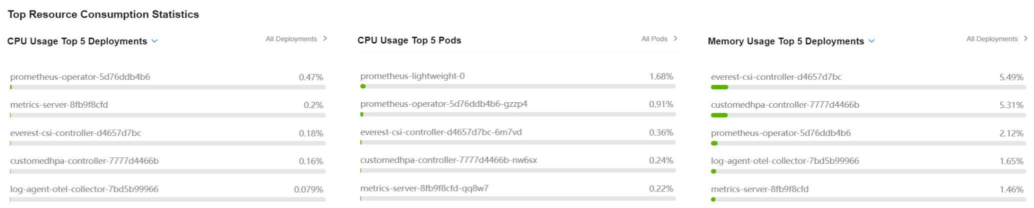
CPU Usage Top 5 Pods	All Pods
prometheus-lightweight-0	1.68%
prometheus-operator-5d76ddb4b6-gzzp4	0.91%
everest-csi-controller-d4657d7bc-6m7vd	0.36%
customedhpa-controller-7777d4466b-nw6sx	0.24%
metrics-server-8fb9f8cfd-qq8w7	0.22%

Memory Usage Top 5 Deployments	All Deployments
everest-csi-controller-d4657d7bc	5.49%
customedhpa-controller-7777d4466b	5.31%
prometheus-operator-5d76ddb4b6	2.12%
log-agent-otel-collector-7bd5b99966	1.65%
metrics-server-8fb9f8cfd	1.46%

Top Resource Consumption Statistics

CCE collects statistics on top 5 Deployments, StatefulSets, and pods by CPU and memory usages, helping you identify workloads with high resource consumption. To view all data, click the **Workloads** or **Pods** tab.

Figure 5-5 Top Resource Consumption Statistics



Top Resource Consumption Statistics

CPU Usage Top 5 Deployments	
prometheus-operator-5d76ddb4b6	0.47%
metrics-server-8fb9f8cfd	0.2%
everest-csi-controller-d4657d7bc	0.18%
customedhpa-controller-7777d4466b	0.16%
log-agent-otel-collector-7bd5b99966	0.079%

CPU Usage Top 5 Pods	
prometheus-lightweight-0	1.68%
prometheus-operator-5d76ddb4b6-gzzp4	0.91%
everest-csi-controller-d4657d7bc-6m7vd	0.36%
customedhpa-controller-7777d4466b-nw6sx	0.24%
metrics-server-8fb9f8cfd-qq8w7	0.22%

Memory Usage Top 5 Deployments	
everest-csi-controller-d4657d7bc	5.49%
customedhpa-controller-7777d4466b	5.31%
prometheus-operator-5d76ddb4b6	2.12%
log-agent-otel-collector-7bd5b99966	1.65%
metrics-server-8fb9f8cfd	1.46%

Monitoring metrics

- CPU Usage

  Workload CPU usage = Average CPU usage in each pod of the workload

  Pod CPU usage = CPU cores used by a pod/CPU limits of service containers

- Memory Usage

  Workload memory usage = Average memory usage in each pod of the workload

  Pod memory usage = Memory used by a pod/Sum of workload container memory limits

Data Plane Monitoring

By default, resource usages are collected from each dimension in the last hour, last 8 hours, and last 24 hours. To view more monitoring data, click **View All Metrics** to access the **Dashboard** page. For details, see **Using Dashboard**.

NOTE

You can hover over a chart to view the monitoring data in each minute.

- **Pod Status and Quantity**: real-time status and number of pods in a cluster.
- **Trend of Total Pod Restarts**: the total number of pod restarts in the cluster in the last 5 minutes.

5.1.3 Workload Monitoring

On the **Workloads** tab, you can view the resource usages of workloads. This page provides information about all workloads in a cluster and monitoring data of a single workload, such as the CPU/memory usage and network inbound/outbound rate.

Navigation Path

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Monitoring Center**. Then click the **Workloads** tab.

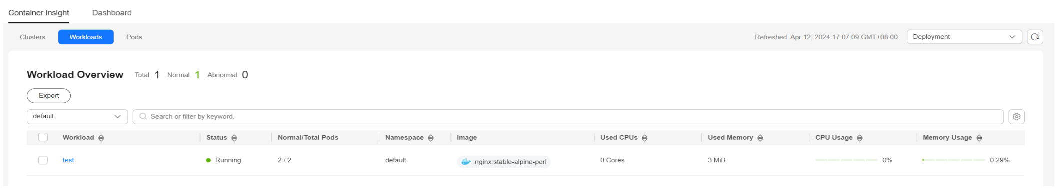
Information about all workloads is displayed. To view the monitoring data of a workload, click the workload name to go to the **Overview** tab. You can also click **Pods** or **Monitoring** to view corresponding information.

----End

Workload List

You can view the name, status, number of normal pods, number of total pods, namespace, image, used CPUs, used memory, CPU usage, and memory usage of each workload.

Figure 5-6 Workloads



Refreshed: Apr 12, 2024 17:07:09 GMT+08:00

Deployment

Workload Overview Total 1 Normal 1 Abnormal 0

default

Search or filter by keyword.

	Workload	Status	Normal/Total Pods	Namespace	Image	Used CPUs	Used Memory	CPU Usage	Memory Usage
	test	● Running	2 / 2	default	nginx:stable-alpine-perl	0 Cores	3 MiB	0%	0.29%

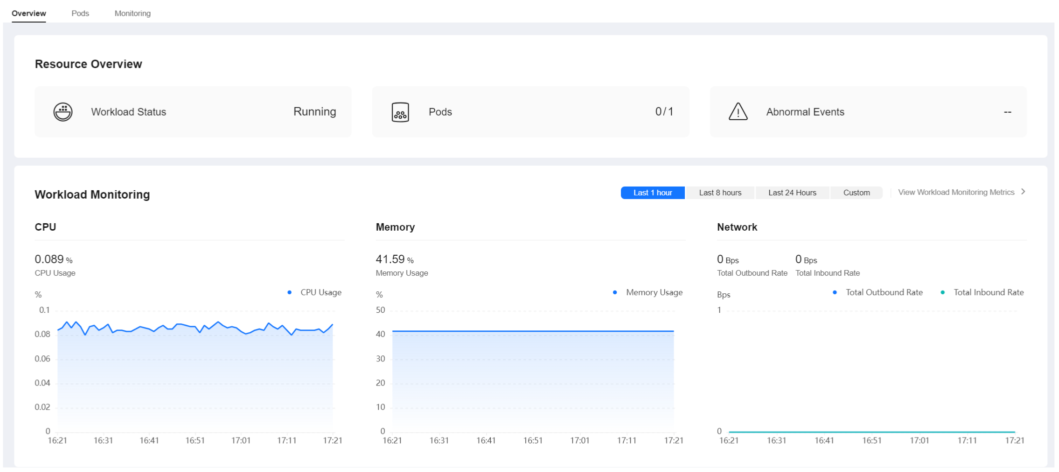
You can select a workload type in the upper right corner, or select **Workload name**, **Status**, and **Namespace** above the list to quickly locate the desired workload.

You can click **Export** in the upper right corner of the list to export details of all workloads or the selected workloads. The exported file is in .xlsx format, and the file name contains the timestamp.

Overview

You can click the workload name to view the resource overview, such as the workload status, number of normal pods, number of total pods, and abnormal events. You can also view the monitoring overview in the last hour, including the CPU usage, memory usage, and network inbound/outbound rate.

Figure 5-7 Resource overview and monitoring overview



Resource Overview

Workload Status Running

Pods 0 / 1

Abnormal Events --

Workload Monitoring

CPU

0.089 %
CPU Usage

Memory

41.59 %
Memory Usage

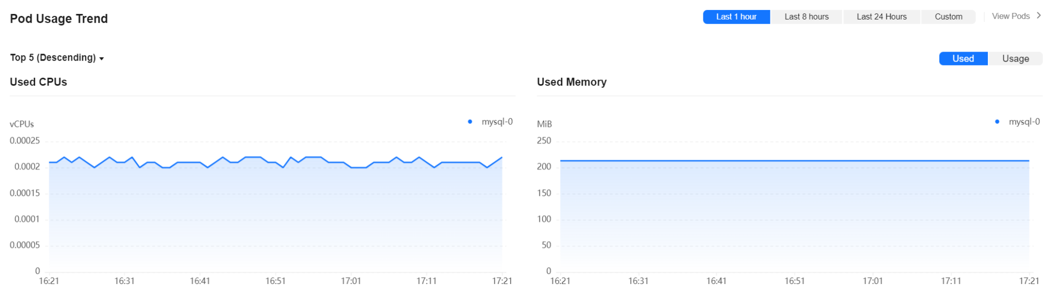
Network

0 Bps
Total Outbound Rate

0 Bps
Total Inbound Rate

The **Overview** tab also shows the pod usage trend. You can switch the metrics in the upper right corner of the chart to view the CPU usage, used CPUs, memory usage, and used memory of each pod of the workload. You can also click **Top 5 (Descending)** or **Top 5 (Ascending)** in the upper left corner to view the top 5 data in descending or ascending order.

Figure 5-8 Pod usage trend

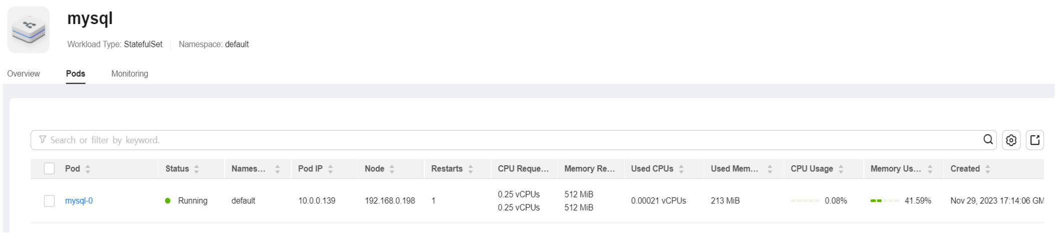


For more metrics, go to the **Monitoring** tab.

Pods

You can view the name, status, namespace, IP address, node, number of restarts, CPU request, CPU limit, memory request, memory limit, used CPUs, used memory, CPU usage, and memory usage of each pod.

Figure 5-9 Pods



Pod	Status	Names...	Pod IP	Node	Restarts	CPU Reque...	Memory Re...	Used CPUs	Used Mem...	CPU Usage	Memory Us...	Created
mysql-0	● Running	default	10.0.0.139	192.168.0.198	1	0.25 vCPUs 0.25 vCPUs	512 MiB 512 MiB	0.00021 vCPUs	213 MiB	0.08%	41.59%	Nov 29, 2023 17:14:06 GM

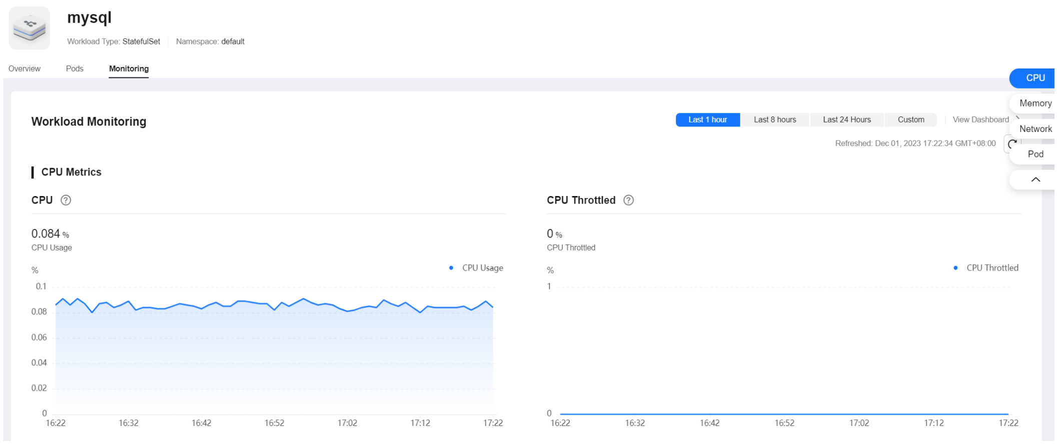
You can find the desired pod by name, status, namespace, IP address, or node. You can click **Export** in the upper right corner of the list to export details of all pods or the selected pods. The exported file is in .xlsx format, and the file name contains the timestamp.

You can click the name of a pod to view its monitoring data. For more information, see **Pod Monitoring**.

Monitoring

This tab shows the resource usage of the workload in each dimension in the last 1 hour, last 8 hours, last 24 hours, or a custom period. To view more monitoring data, click **View Dashboard** to access the **Dashboard** page. For details, see **Using Dashboard**.

Figure 5-10 Workload monitoring



- **CPU Metrics**
  - CPU usage: the percentage of the CPU used by containers in all pods of the workload in different time periods with respect to the total CPU limit for all containers.
  - CPU throttled: the average percentage of time duration containers have been throttled in all pods of the workload in different time periods.
- **Memory Metrics**
  - Memory usage: the percentage of memory used by containers in all pods of the workload in different time periods with respect to the total memory limit for all containers.
- **Networking Metrics**
  - Total outbound rate: the total number of bytes transmitted by containers in all pods of the workload per second in different time periods.
  - Total inbound rate: the total number of bytes received by containers in all pods of the workload per second in different time periods.
  - Packet loss rate (transmit): the percentage of packets not received by the recipient to packets sent from containers in all pods of the workload in different time periods.
  - Packet loss rate (receive): the percentage of packets not received by containers in all pods of the workload to packets sent to the containers in different time periods.
- **Pod Metrics**
  - Pod CPU usage: the percentage of CPU used by each pod of the workload in different time periods with respect to the CPU limit for each pod.
  - Pod memory usage: the percentage of memory used by each pod of the workload in different time periods with respect to the memory limit for each pod.
  - Pod status and quantity: the total number of pods in the **Unavailable**, **Unready**, **Running**, **Completed**, or **Other** state of the workload in different time periods.
  - Pod quantity trend: the number of pods (replicas) of the workload in different time periods.

5.1.4 Pod Monitoring

To view the resource usages of pods, go to the **Pods** tab, where you can view information about all pods in a cluster and monitoring data of each pod, such as the CPU usage, memory usage, inbound rate, and outbound rate.

Navigation Path

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Monitoring Center**. Then click the **Pods** tab.

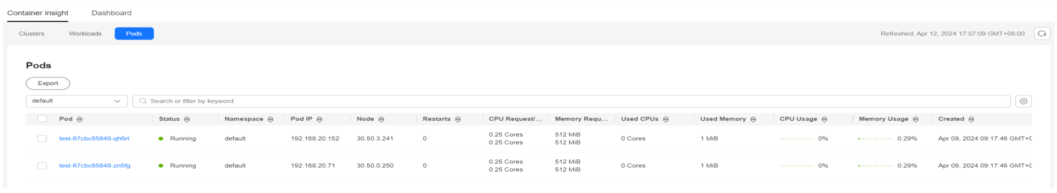
Information about all pods is displayed. To view the monitoring data of a pod, click the pod name to go to the **Overview** tab. You can also click the **Containers** or **Monitoring** tab to view the corresponding information.

----End

Pods

You can view the name, status, namespace, IP address, node, number of restarts, CPU request, CPU limit, memory request, memory limit, used CPUs, used memory, CPU usage, and memory usage of each pod.

Figure 5-11 Pods


Container insight
Dashboard
Clusters
Workloads
Pods
Refreshed: Apr 12, 2024 17:07:09 GMT+08:00
Pods
Export
default
Search or filter by keyword.

Pod	Status	Namespace	Pod IP	Node	Restarts	CPU Request/...	Memory Requ...	Used CPUs	Used Memory	CPU Usage	Memory Usage	Created
test-67cbc85848-qh6rt	● Running	default	192.168.20.152	30.50.3.241	0	0.25 Cores 0.25 Cores	512 MiB 512 MiB	0 Cores	1 MiB	0%	0.29%	Apr 09, 2024 09:17:46 GMT+C
test-67cbc85848-zn5fg	● Running	default	192.168.20.71	30.50.0.250	0	0.25 Cores 0.25 Cores	512 MiB 512 MiB	0 Cores	1 MiB	0%	0.29%	Apr 09, 2024 09:17:46 GMT+C

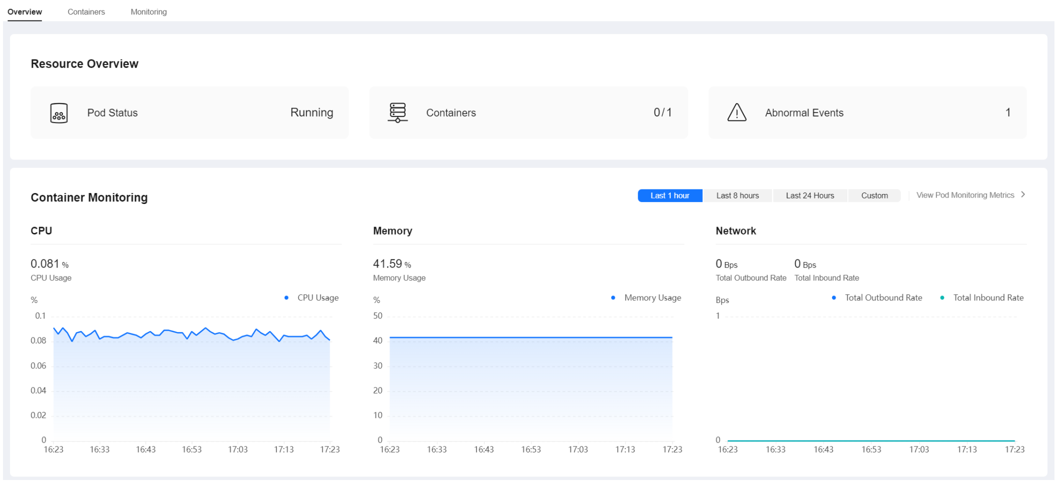
You can select a namespace in the upper right corner, or select **Pod**, **Status**, **Namespace**, **Pod IP**, and **Node** above the list to quickly locate the required pod.

You can click **Export** in the upper right corner of the list to export details of all pods or the selected pods. The exported file is in .xlsx format, and the file name contains the timestamp.

Overview

You can click the pod name to view the resource overview, including the pod status, number of containers (abnormal/total), and abnormal events. You can also view the monitoring overview of the pod in the last hour, including the CPU usage, memory usage, and network inbound/outbound rate.

Figure 5-12 Resource overview and monitoring overview



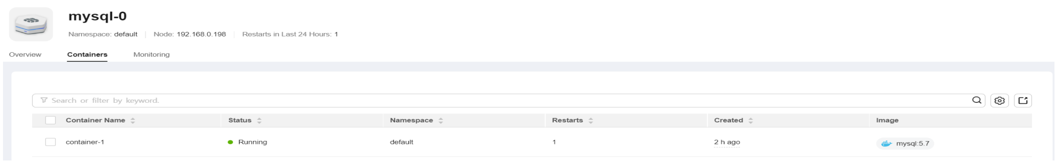
The **Overview** tab also shows the container usage trend. You can switch the metrics in the upper right corner of the chart to view the CPU usage, used CPUs, memory usage, and used memory of each container in the pod. You can also click **Top 5 (Descending)** or **Top 5 (Ascending)** in the upper left corner to view the top 5 data in descending or ascending order.

For more metrics, go to the **Monitoring** tab.

Containers

This tab contains details such as the name, status, namespace, number of restarts, and image of each container.

Figure 5-13 Containers



mysql-0

Namespace: default | Node: 192.168.0.198 | Restarts in Last 24 Hours: 1

Overview Containers Monitoring

Search or filter by keyword.

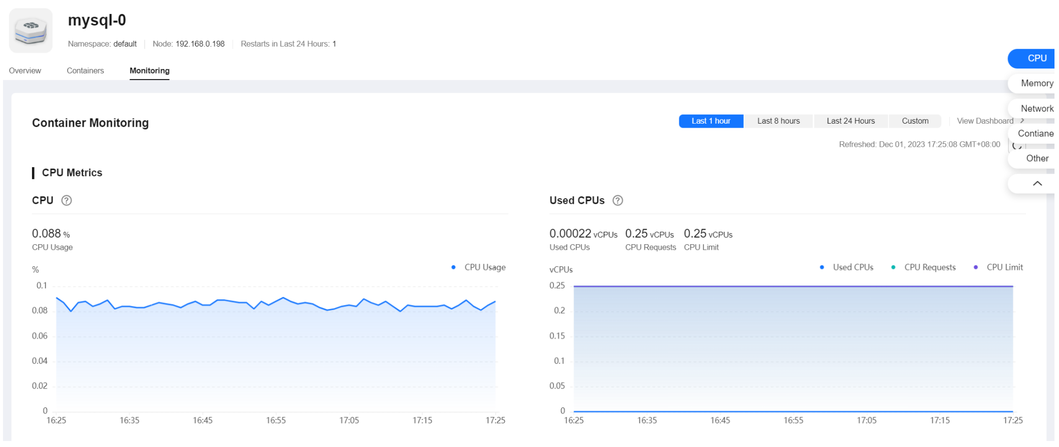
Container Name	Status	Namespace	Restarts	Created	Image
container-1	● Running	default	1	2 h ago	mysql:5.7

You can find the desired container by name, status, or namespace. You can click **Export** in the upper right corner of the list to export details of all containers or the selected containers. The exported file is in .xlsx format, and the file name contains the timestamp.

Monitoring

This tab shows the resource usage of the pod in each dimension in the last 1 hour, last 8 hours, last 24 hours, or a custom period. To view more monitoring data, click **View Dashboard** to access the **Dashboard** page. For details, see **Using Dashboard**.

Figure 5-14 Pod monitoring



- **CPU Metrics**
  - CPU usage: the percentage of CPU used by all containers in a pod in different time periods with respect to the total CPU limit for all containers.
  - Used CPU: the CPU that the pod is using.
  - CPU request: the CPU requested for the pod.
  - CPU limit: the CPU limit configured for the pod. When the used CPU is close to this limit, the CPU usage of the containers will be limited, affecting container performance.
- **Memory Metrics**
  - Memory usage: the percentage of memory used by all containers in the pod in different time periods with respect to the total memory limit for all containers.
  - Used memory: the memory that the pod is using.
  - Memory request: the memory requested for the pod.
  - Memory limit: the memory limit configured for the pod. When the used memory is close to this limit, OOM will occur.
- **Networking Metrics**
  - Total outbound rate: the total number of bytes transmitted by all containers in the pod per second.
  - Total inbound rate: the total number of bytes received by all containers in the pod per second.
- **Container Metrics**
  - Container CPU usage: the percentage of CPU used by each container in the pod in different time periods with respect to the CPU limit for each container.
  - Container memory usage: the percentage of memory used by each container in the pod in different time periods with respect to the memory limit for each container.
  - Container CPU throttled: the percentage of time duration each container has been throttled in different time periods.

- Container network packet loss rate: the percentage of packets not received by each container of the pod to packets sent to the container in different time periods.

- **Other Metrics**
  - Historical pod status: the status of the pod in different periods.
  - Historical container status: the status of each container in the pod in different time periods.

5.1.5 Dashboard

5.1.5.1 Using Dashboard

A dashboard integrates high-frequency monitoring metrics of different components from different perspectives. Different metrics are displayed on the same screen in charts, helping you monitor the cluster running in real time.

The dashboard displays monitoring metrics in various views, such as the cluster view and pod view.

Prerequisites

- The cluster is in the **Running** state.
- Monitoring Center has been enabled for the cluster.

Checking and Switching Views

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Monitoring Center**. Then click the **Dashboard** tab.

The cluster view is displayed by default.

Step 3 The dashboard provides preset views. You can click the **Switch View** button next to the view name to select monitoring data to view.

Step 4 Configure the parameters.

Step 5 Specify the view window.

Select or customize time periods in the upper right corner of the page, and click ⟳ to refresh the page.

----End

5.2 Logging

5.2.1 Collecting Logs

Cloud Native Log Collection is an add-on based on Fluent Bit and OpenTelemetry for collecting logs and Kubernetes events. This add-on supports CRD-based log collection policies. It collects and forwards standard output logs, container file logs, and Kubernetes events in a cluster. It also reports all abnormal Kubernetes events and some normal Kubernetes events to AOM.

Constraints

- A maximum of 50 log collection rules can be configured for each cluster.
- This add-on cannot collect .gz, .tar, and .zip logs or access symbolic links of logs.
- In each cluster, up to 10,000 single-line logs can be collected per second, and up to 2,000 multi-line logs can be collected per second.
- If a volume is attached to the data directory of a service container, this add-on cannot collect data from the parent directory. In this case, you need to configure a complete data directory.

Billing

LTS does not charge you for creating log groups and offers a free quota for log collection every month. You pay only for log volume that exceeds the quota. For details, see **Price Calculator**.

Log Collection

Step 1 Enable log collection.

Enabling log collection during cluster creation

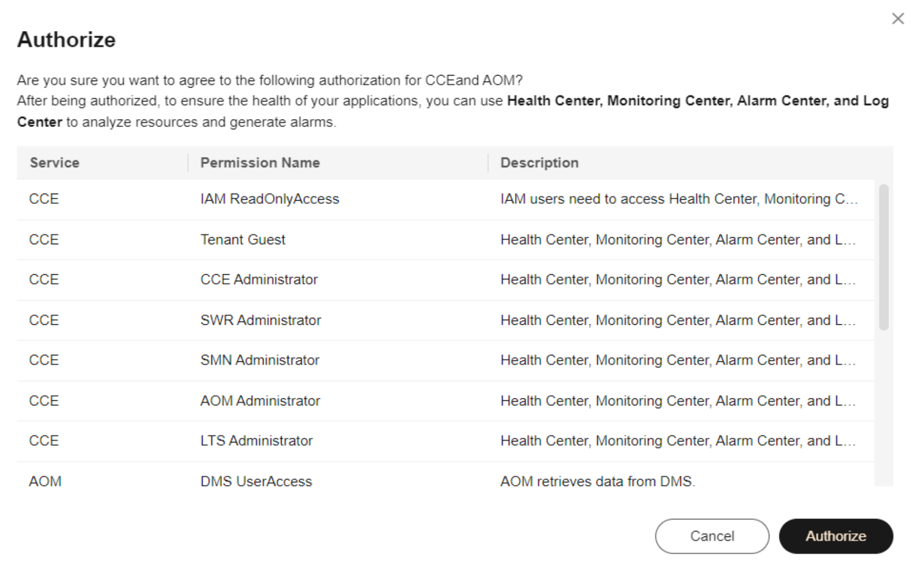
1. Log in to the CCE console.
2. In the upper right corner, click **Buy Cluster**.
3. In the **Select Add-on** step, select **Cloud Native Log Collection**.
4. Click **Next: Add-on Configuration** in the lower right corner and select the required logs.
   - Container logs: A log collection policy named **default-stdout** will be created, and standard output logs in all namespaces will be reported to LTS.
   - Kubernetes events: A log collection policy named **default-event** will be created, and Kubernetes events in all namespaces will be reported to LTS.
5. Click **Next: Confirm Configuration**. On the displayed page, click **Submit**.

Enabling log collection for an existing cluster

1. Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Logging**.
2. (Optional) If you are not authorized, obtain required permissions first.

   In the displayed dialog box, click **Authorize**.

Figure 5-15 Adding authorization



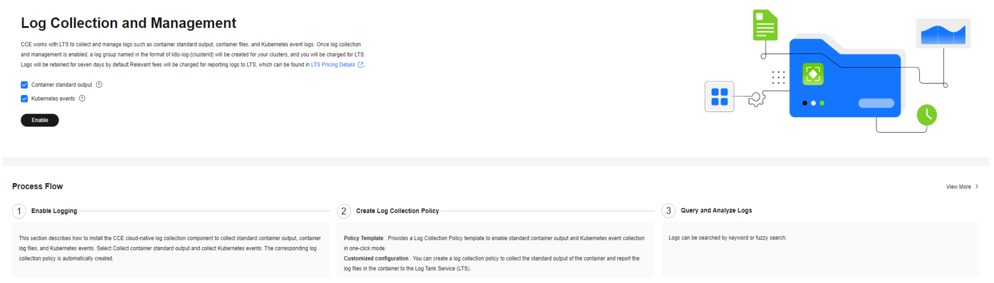
Authorize

Are you sure you want to agree to the following authorization for CCEand AOM?
After being authorized, to ensure the health of your applications, you can use **Health Center, Monitoring Center, Alarm Center, and Log Center** to analyze resources and generate alarms.

Service	Permission Name	Description
CCE	IAM ReadOnlyAccess	IAM users need to access Health Center, Monitoring C...
CCE	Tenant Guest	Health Center, Monitoring Center, Alarm Center, and L...
CCE	CCE Administrator	Health Center, Monitoring Center, Alarm Center, and L...
CCE	SWR Administrator	Health Center, Monitoring Center, Alarm Center, and L...
CCE	SMN Administrator	Health Center, Monitoring Center, Alarm Center, and L...
CCE	AOM Administrator	Health Center, Monitoring Center, Alarm Center, and L...
CCE	LTS Administrator	Health Center, Monitoring Center, Alarm Center, and L...
AOM	DMS UserAccess	AOM retrieves data from DMS.

3. Click **Enable** and wait for about 30 seconds until the log page is automatically displayed.
   - Standard output logs: A log collection policy named **default-stdout** will be created, and standard output logs in all namespaces will be reported to LTS.
   - Kubernetes events: A log collection policy named **default-event** will be created, and Kubernetes events in all namespaces will be reported to LTS.

Figure 5-16 Enabling log collection



Log Collection and Management

CCE works with LTS to collect and manage logs such as container standard output, container files, and Kubernetes event logs. Once log collection and management is enabled, a log group named in the format of k8s-log-{clusterId} will be created for your clusters, and you will be charged for LTS. Logs will be retained for seven days by default.Relevant fees will be charged for reporting logs to LTS, which can be found in LTS Pricing Details.

[✓] Container standard output
[✓] Kubernetes events

Process Flow

1 **Enable Logging**

This section describes how to install the CCE cloud-native log collection component to collect standard container output, container log files, and Kubernetes events. Select Collect container standard output and collect Kubernetes events. The corresponding log collection policy is automatically created.

2 **Create Log Collection Policy**

Policy Template : Provides a Log Collection Policy template to enable standard container output and Kubernetes event collection in one-click mode.
Customized configuration : You can create a log collection policy to collect the standard output of the container and report the log files in the container to the Log Tank Service (LTS).

3 **Query and Analyze Logs**

Logs can be searched by keyword or fuzzy search.

Step 2 View and configure log collection policies.

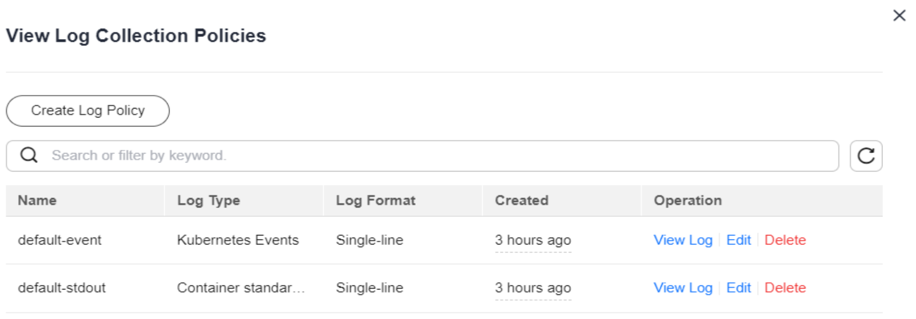
1. On the CCE console, click the cluster name to access the cluster console. In the navigation pane on the left, choose **Logging**.
2. Click **View Log Collection Policies** in the upper right corner.

   All log collection policies reported to LTS are displayed.

   If **Container standard output** and **Kubernetes events** are selected during the add-on installation, two log collection policies will be created, and the collected logs will be reported to the default log group and log streams.
   - **default-stdout**: collects standard output. Default log group: **k8s-logs-{*Cluster ID*}**; Default log stream: **stdout-{*Cluster ID*}**

– **default-event**: collects Kubernetes events. Default log group: **k8s-logs-{*Cluster ID*}**; Default log stream: **event-{*Cluster ID*}**

Figure 5-17 Viewing log collection policies

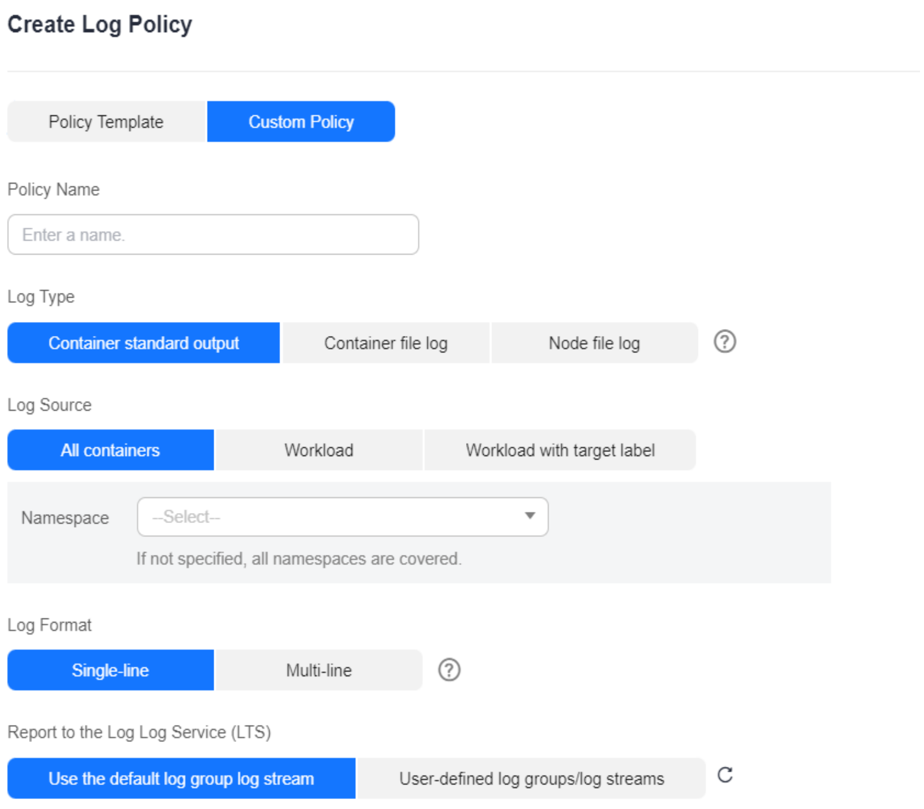

View Log Collection Policies

Create Log Policy

Name	Log Type	Log Format	Created	Operation
default-event	Kubernetes Events	Single-line	3 hours ago	View Log | Edit | Delete
default-stdout	Container standar...	Single-line	3 hours ago	View Log | Edit | Delete

3. Click **Create Log Policy** and configure parameters as required.
   – **Policy Template**: If **Container standard output** and **Kubernetes events** are not selected during add-on installation or their log collection policies are deleted, you can use this option to create a default log collection policy.
   – **Custom Policy**: You can use this option to create a log collection policy.

Figure 5-18 Custom policy


Create Log Policy

Policy Template | Custom Policy

Policy Name: Enter a name.

Log Type: Container standard output | Container file log | Node file log

Log Source: All containers | Workload | Workload with target label

Namespace: --Select--
If not specified, all namespaces are covered.

Log Format: Single-line | Multi-line

Report to the Log Log Service (LTS): Use the default log group log stream | User-defined log groups/log streams

Table 5-1 Custom policy parameters

Parameter	Description
Log Type	Type of logs to be collected. – **Container standard output**: used to collect container standard output logs. You can create a log collection policy by namespace, workload name, or instance label. – **Container file log**: used to collect text logs. You can create a log collection policy by workload or instance label.
Log Source	Containers whose logs are to be collected. – **All containers**: You can specify all containers in a namespace. If this parameter is not specified, logs of containers in all namespaces will be collected. – **Workload**: You can specify a workload and its containers. If this parameter is not specified, logs of all containers running the workload will be collected. – **Workload with target label**: You can specify a workload by label and its containers. If this parameter is not specified, logs of all containers running the workload will be collected.

Log Format	– **Single-line**
	Each log contains only one line of text. The newline character \n denotes the start of a new log.
	– **Multi-line**
	Some programs (for example, Java program) print a log that occupies multiple lines. By default, logs are collected by line. If you want to display logs as a single message, you can enable multi-line logging and use the regular pattern. If you select the multi-line text, you need to enter the log matching format.
	Example:
	If logs need to be collected by line, enter **\d{4}-\d{2}-\d{2} \d{2}\:\d{2}\:\d{2}.***.
	The following three lines starting with the date are regarded as a log.
	`2022-01-01 00:00:00 Exception in thread "main"`
	`java.lang.RuntimeException: Something has gone wrong, aborting!`
	`at com.myproject.module.MyProject.badMethod(MyProject.java:22)`
	`at com.myproject.module.MyProject.oneMoreMethod(MyProject.java:18)`

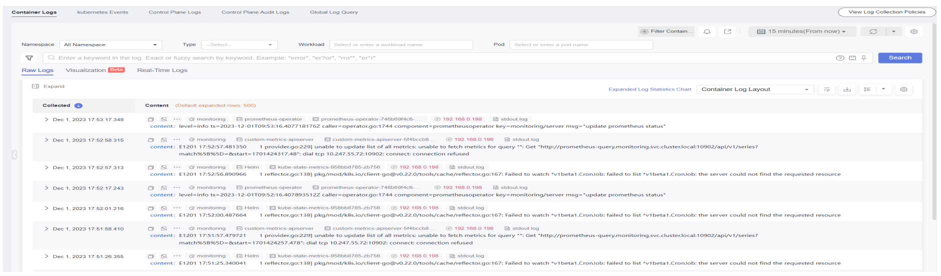
Parameter	Description
Report to the Log Tank Service (LTS)	This parameter is used to configure the log group and log stream for log reporting. – Default log groups/log streams: The default log group (**k8s-log-***{Cluster ID}*) and default log stream (**stdout-***{Cluster ID}*) are automatically selected. – Custom log groups/log streams: You can select any log group and log stream. ▪ A log group is the basic unit for LTS to manage logs. If you do not have a log group, CCE prompts you to create one. The default name is **k8s-log-***{Cluster ID}*, for example, **k8s-log-bb7eaa87-07dd-11ed-ab6c-0255ac1001b3**. ▪ A log stream is the basic unit for log read and write. You can create log streams in a log group to store different types of logs for finer log management. When you install the add-on or create a log policy based on a template, the following log streams are automatically created: **stdout-***{Cluster ID}* for standard output logs, for example, **stdout-bb7eaa87-07dd-11ed-ab6c-0255ac1001b3** **event-***{Cluster ID}* for Kubernetes events, for example, **event-bb7eaa87-07dd-11ed-ab6c-0255ac1001b3**

4. Click **Edit** to modify an existing log collection policy.
5. Click **Delete** to delete an existing log collection policy.

Step 3 View the logs.

1. On the CCE console, click the cluster name to access the cluster console. In the navigation pane on the left, choose **Logging**.
2. View different types of logs:
   – **Container Logs**: displays all logs in the default log stream **stdout-***{Cluster ID}* of the default log group **k8s-log-***{Cluster ID}*. You can search for logs by workload.

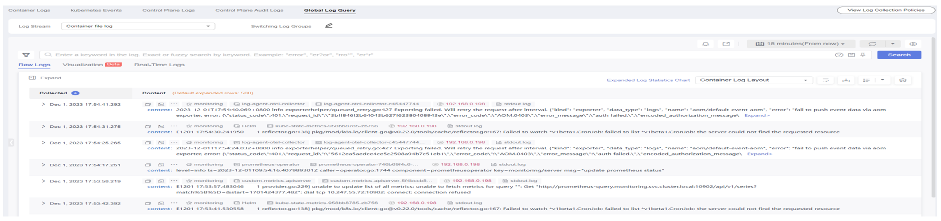
Figure 5-19 Querying container logs



```
Dec 1, 2023 17:53:17.348   monitoring  prometheus-operator  prometheus-operator-746b69f4c6-...  192.168.0.198  stdout.log
  content: level=info ts=2023-12-01T09:53:16.407718176Z caller=operator.go:1744 component=prometheusoperator key=monitoring/server msg="update prometheus status"
Dec 1, 2023 17:52:58.315   monitoring  custom-metrics-apiserver  custom-metrics-apiserver-5f4bccb8...  192.168.0.198  stdout.log
  content: E1201 17:52:57.481350       1 provider.go:229] unable to update list of all metrics: unable to fetch metrics for query "": Get "http://prometheus-query.monitoring.svc.cluster.local:10902/api/v1/series?
           match%5B%5D=&start=1701424317.48": dial tcp 10.247.55.72:10902: connect: connection refused
Dec 1, 2023 17:52:57.313   monitoring  Helm  kube-state-metrics-958bb8785-zb756  192.168.0.198  stdout.log
  content: E1201 17:52:56.890966       1 reflector.go:138] pkg/mod/k8s.io/client-go@v0.22.0/tools/cache/reflector.go:167: Failed to watch *v1beta1.CronJob: failed to list *v1beta1.CronJob: the server could not find the requested resource
Dec 1, 2023 17:52:17.243   monitoring  prometheus-operator  prometheus-operator-746b69f4c6-...  192.168.0.198  stdout.log
  content: level=info ts=2023-12-01T09:52:16.407893512Z caller=operator.go:1744 component=prometheusoperator key=monitoring/server msg="update prometheus status"
Dec 1, 2023 17:52:01.216   monitoring  Helm  kube-state-metrics-958bb8785-zb756  192.168.0.198  stdout.log
  content: E1201 17:52:00.487664       1 reflector.go:138] pkg/mod/k8s.io/client-go@v0.22.0/tools/cache/reflector.go:167: Failed to watch *v1beta1.CronJob: failed to list *v1beta1.CronJob: the server could not find the requested resource
Dec 1, 2023 17:51:58.410   monitoring  custom-metrics-apiserver  custom-metrics-apiserver-5f4bccb8...  192.168.0.198  stdout.log
  content: E1201 17:51:57.479721       1 provider.go:229] unable to update list of all metrics: unable to fetch metrics for query "": Get "http://prometheus-query.monitoring.svc.cluster.local:10902/api/v1/series?
           match%5B%5D=&start=1701424257.478": dial tcp 10.247.55.72:10902: connect: connection refused
Dec 1, 2023 17:51:26.355   monitoring  Helm  kube-state-metrics-958bb8785-zb756  192.168.0.198  stdout.log
  content: E1201 17:51:25.340041       1 reflector.go:138] pkg/mod/k8s.io/client-go@v0.22.0/tools/cache/reflector.go:167: Failed to watch *v1beta1.CronJob: failed to list *v1beta1.CronJob: the server could not find the requested resource
```

– **Kubernetes Events**: displays all Kubernetes events in the default log stream **event-*{Cluster ID}*** of the default log group **k8s-log-*{Cluster ID}***.
– **Global Log Query**: You can view logs in the log streams of all log groups. You can specify a log stream to view the logs. By default, the default log group **k8s-log-*{Cluster ID}*** is selected. You can click the edit icon on the right of **Switching Log Groups** to switch to another log group.

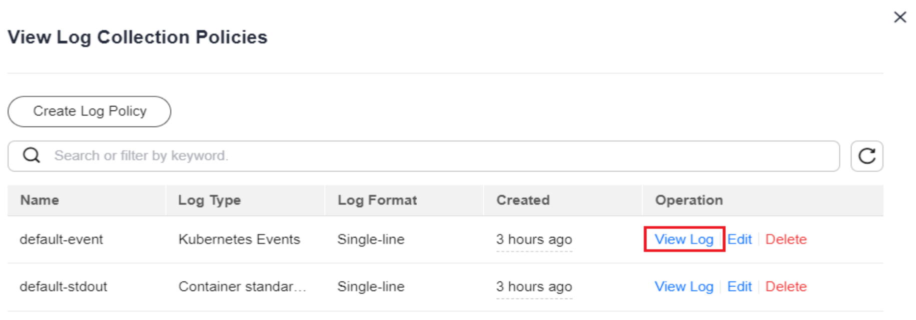
Figure 5-20 Global log query



```
Dec 1, 2023 17:54:41.292   monitoring   log-agent-otel-collector   log-agent-otel-collector-c45447744...   192.168.0.198   stdout.log
content: 2023-12-01T17:54:40.069+0800 info exporterhelper/queued_retry.go:427 Exporting failed. Will retry the request after interval. {"kind": "exporter", "data_type": "logs", "name": "aom/default-event-aom", "error": "fail to push event data via aom exporter, error: {\"status_code\":401,\"request_id\":\"3bff846f2b64043b627f62380408943e\",\"error_code\":\"AOM.0403\",\"error_message\":\"auth failed.\",\"encoded_authorization_message\  Expand>
Dec 1, 2023 17:54:31.275   monitoring   Helm   kube-state-metrics-958bb8785-zb756   192.168.0.198   stdout.log
content: E1201 17:54:30.241950       1 reflector.go:138] pkg/mod/k8s.io/client-go@v0.22.0/tools/cache/reflector.go:167: Failed to watch *v1beta1.CronJob: failed to list *v1beta1.CronJob: the server could not find the requested resource
Dec 1, 2023 17:54:25.265   monitoring   log-agent-otel-collector   log-agent-otel-collector-c45447744...   192.168.0.198   stdout.log
content: 2023-12-01T17:54:24.032+0800 info exporterhelper/queued_retry.go:427 Exporting failed. Will retry the request after interval. {"kind": "exporter", "data_type": "logs", "name": "aom/default-event-aom", "error": "fail to push event data via aom exporter, error: {\"status_code\":401,\"request_id\":\"5612ea5aedce4ce5c2508a94b7c51eb1\",\"error_code\":\"AOM.0403\",\"error_message\":\"auth failed.\",\"encoded_authorization_message\  Expand>
Dec 1, 2023 17:54:17.251   monitoring   prometheus-operator   prometheus-operator-746b69f4c6-...   192.168.0.198   stdout.log
content: level=info ts=2023-12-01T09:54:16.407989301Z caller=operator.go:1744 component=prometheusoperator key=monitoring/server msg="update prometheus status"
Dec 1, 2023 17:53:58.219   monitoring   custom-metrics-apiserver   custom-metrics-apiserver-5f4bccb8...   192.168.0.198   stdout.log
content: E1201 17:53:57.483046       1 provider.go:229] unable to update list of all metrics: unable to fetch metrics for query "": Get "http://prometheus-query.monitoring.svc.cluster.local:10902/api/v1/series?match%5B%5D=&start=1701424377.482": dial tcp 10.247.55.72:10902: connect: connection refused
Dec 1, 2023 17:53:42.392   monitoring   Helm   kube-state-metrics-958bb8785-zb756   192.168.0.198   stdout.log
content: E1201 17:53:41.530558       1 reflector.go:138] pkg/mod/k8s.io/client-go@v0.22.0/tools/cache/reflector.go:167: Failed to watch *v1beta1.CronJob: failed to list *v1beta1.CronJob: the server could not find the requested resource
```

3. Click **View Log Collection Policies** in the upper right corner. Locate the log collection policy and click **View Log** to go to the log list.

Figure 5-21 Viewing logs



View Log Collection Policies

Create Log Policy

Search or filter by keyword.

Name	Log Type	Log Format	Created	Operation
default-event	Kubernetes Events	Single-line	3 hours ago	View Log Edit Delete
default-stdout	Container standar...	Single-line	3 hours ago	View Log Edit Delete

----End

Troubleshooting

- **All components except log-operator are not ready, and the volume failed to be attached to the node.**

  Solution: Check the logs of log-operator. During add-on installation, the configuration files required by other components are generated by log-operator. If the configuration files are invalid, all components cannot be started.

  The log information is as follows:

  ```
  MountVolume.SetUp failed for volume "otel-collector-config-vol":configmap "log-agent-otel-collector-config" not found
  ```

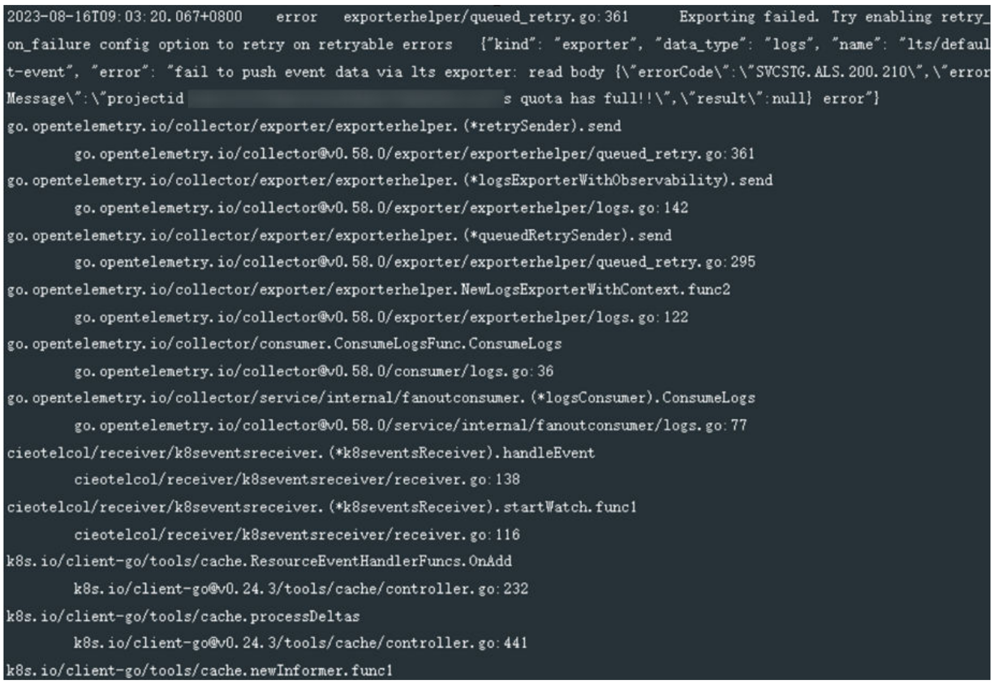
- **"Failed to create log group, the number of log groups exceeds the quota" is reported in the standard output log of log-operator.**

  Example:

  ```
  2023/05/05 12:17:20.799 [E] call 3 times failed, resion: create group failed, projectID: xxx, groupName: k8s-log-xxx, err: create groups status code: 400, response: {"error_code":"LTS.0104","error_msg":"Failed to create log group, the number of log groups exceeds the quota"}, url: https://lts.cn-north-4.myhuaweicloud.com/v2/xxx/groups, process will retry after 45s
  ```

  Solution: On the LTS console, delete unnecessary log groups. For details about the log group quota, see **Log Groups**.

- **Logs cannot be reported, and "log's quota has full" is reported in the standard output log of the OTel component.**

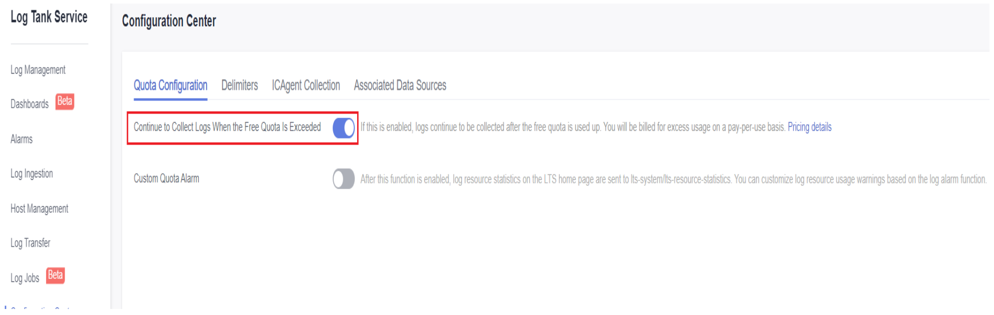


```
2023-08-16T09:03:20.067+0800    error   exporterhelper/queued_retry.go:361      Exporting failed. Try enabling retry_
on_failure config option to retry on retryable errors   {"kind": "exporter", "data_type": "logs", "name": "lts/defaul
t-event", "error": "fail to push event data via lts exporter: read body {\"errorCode\":\"SVCSTG.ALS.200.210\",\"error
Message\":\"projectid [redacted] s quota has full!!\",\"result\":null} error"}
go.opentelemetry.io/collector/exporter/exporterhelper.(*retrySender).send
        go.opentelemetry.io/collector@v0.58.0/exporter/exporterhelper/queued_retry.go:361
go.opentelemetry.io/collector/exporter/exporterhelper.(*logsExporterWithObservability).send
        go.opentelemetry.io/collector@v0.58.0/exporter/exporterhelper/logs.go:142
go.opentelemetry.io/collector/exporter/exporterhelper.(*queuedRetrySender).send
        go.opentelemetry.io/collector@v0.58.0/exporter/exporterhelper/queued_retry.go:295
go.opentelemetry.io/collector/exporter/exporterhelper.NewLogsExporterWithContext.func2
        go.opentelemetry.io/collector@v0.58.0/exporter/exporterhelper/logs.go:122
go.opentelemetry.io/collector/consumer.ConsumeLogsFunc.ConsumeLogs
        go.opentelemetry.io/collector@v0.58.0/consumer/logs.go:36
go.opentelemetry.io/collector/service/internal/fanoutconsumer.(*logsConsumer).ConsumeLogs
        go.opentelemetry.io/collector@v0.58.0/service/internal/fanoutconsumer/logs.go:77
cieotelcol/receiver/k8seventsreceiver.(*k8seventsReceiver).handleEvent
        cieotelcol/receiver/k8seventsreceiver/receiver.go:138
cieotelcol/receiver/k8seventsreceiver.(*k8seventsReceiver).startWatch.func1
        cieotelcol/receiver/k8seventsreceiver/receiver.go:116
k8s.io/client-go/tools/cache.ResourceEventHandlerFuncs.OnAdd
        k8s.io/client-go@v0.24.3/tools/cache/controller.go:232
k8s.io/client-go/tools/cache.processDeltas
        k8s.io/client-go@v0.24.3/tools/cache/controller.go:441
k8s.io/client-go/tools/cache.newInformer.func1
```

Solution:

LTS provides a free log quota. If the quota is used up, you will be charged for the excess log usage. If an error message is displayed, the free quota has been used up. To continue collecting logs, log in to the LTS console, choose **Configuration Center** in the navigation pane on the left, and enable **Continue to Collect Logs When the Free Quota Is Exceeded**.

Figure 5-22 Quota configuration



Continue to Collect Logs When the Free Quota Is Exceeded — If this is enabled, logs continue to be collected after the free quota is used up. You will be billed for excess usage on a pay-per-use basis. Pricing details

Custom Quota Alarm — After this function is enabled, log resource statistics on the LTS home page are sent to lts-system/lts-resource-statistics. You can customize log resource usage warnings based on the log alarm function.

- **Text logs cannot be collected because wildcards are configured for the collection directory.**

  Troubleshooting: Check the volume mounting status in the workload configuration. If a volume is attached to the data directory of a service container, this add-on cannot collect data from the parent directory. In this case, you need to set the collection directory to a complete data directory. For example, if the data volume is attached to the **/var/log/service** directory, logs cannot be collected from the **/var/log** or **/var/log/*** directory. In this case, you need to set the collection directory to **/var/log/service**.

  Solution: If the log generation directory is **/application/logs/***{Application name}***/*.log**, attach the data volume to the **/application/logs** directory and set the collection directory in the log collection policy to **/application/logs/*/*.log**.

5.2.2 Collecting Kubernetes Events

The Cloud Native Log Collection add-on works with LTS to collect and store Kubernetes events and works with AOM to generate alarms.

Billing

LTS does not charge you for creating log groups and offers a free quota for log collection every month. You pay only for log volume that exceeds the quota. For details, see **Price Calculator**.

Reporting Kubernetes Events to LTS

The Cloud Native Log Collection add-on has not been installed in a cluster.

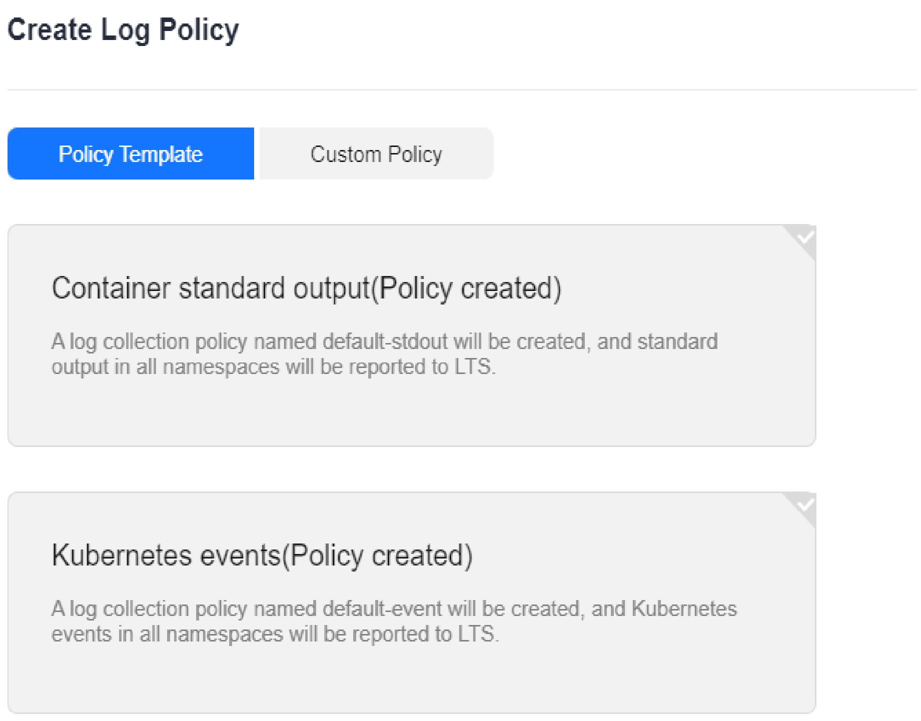
You can select **Kubernetes events** when installing the Cloud Native Log Collection add-on. A default log collection policy will be created, and all events collected will be reported to LTS. For details about how to install the add-on, see **Collecting Logs**.

The Cloud Native Log Collection add-on has been installed in a cluster.

1. Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Logging**.
2. Click **View Log Collection Policies** in the upper right corner.

   All log collection policies reported to LTS are displayed.
3. Click **Create Log Policy** and configure parameters as required.

   Policy Template: If **Kubernetes events** is not selected during add-on installation or the log collection policy is deleted, you can use this option to create a default log collection policy.



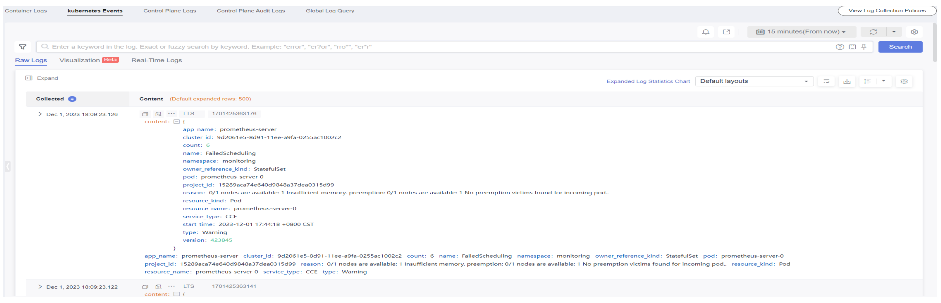
Container standard output(Policy created)

A log collection policy named default-stdout will be created, and standard output in all namespaces will be reported to LTS.

Kubernetes events(Policy created)

A log collection policy named default-event will be created, and Kubernetes events in all namespaces will be reported to LTS.

4. On the **Logging** page, select the log stream configured in the log collection policy to view the events reported to LTS.


Container Logs
kubernetes Events
Control Plane Logs
Control Plane Audit Logs
Global Log Query
View Log Collection Policies
15 minutes(From now)
Enter a keyword in the log. Exact or fuzzy search by keyword. Example: "error", "er?or", "rro*", "er*r"
Search
Raw Logs
Visualization Beta
Real-Time Logs
Expand
Expanded Log Statistics Chart
Default layouts
Collected
Content (Default expanded rows: 500)

```
Dec 1, 2023 18:09:23.126    LTS  1701425363176
content: {
    app_name: prometheus-server
    cluster_id: 9d2061e5-8d91-11ee-a9fa-0255ac1002c2
    count: 6
    name: FailedScheduling
    namespace: monitoring
    owner_reference_kind: StatefulSet
    pod: prometheus-server-0
    project_id: 15289aca74e640d9848a37dea0315d99
    reason: 0/1 nodes are available: 1 Insufficient memory. preemption: 0/1 nodes are available: 1 No preemption victims found for incoming pod..
    resource_kind: Pod
    resource_name: prometheus-server-0
    service_type: CCE
    start_time: 2023-12-01 17:44:18 +0800 CST
    type: Warning
    version: 423845
}
app_name: prometheus-server  cluster_id: 9d2061e5-8d91-11ee-a9fa-0255ac1002c2  count: 6  name: FailedScheduling  namespace: monitoring  owner_reference_kind: StatefulSet  pod: prometheus-server-0
project_id: 15289aca74e640d9848a37dea0315d99  reason: 0/1 nodes are available: 1 Insufficient memory. preemption: 0/1 nodes are available: 1 No preemption victims found for incoming pod..  resource_kind: Pod
resource_name: prometheus-server-0  service_type: CCE  type: Warning
Dec 1, 2023 18:09:23.122    LTS  1701425363141
content: {
```

Reporting Kubernetes Events to AOM

After the Cloud Native Log Collection add-on is installed, all Warning events and some Normal events will be reported to AOM by default. The reported events can be used to configure alarms.

Custom Event Reporting

If the reported events cannot meet requirements, you can modify the settings for the events.

Step 1 Run the following command on the cluster to edit the event collection settings:

kubectl edit logconfig -n kube-system default-event-aom

Step 2 Modify the event collection settings as required.

```
apiVersion: logging.openvessel.io/v1
kind: LogConfig
metadata:
  annotations:
    helm.sh/resource-policy: keep
  name: default-event-aom
  namespace: kube-system
spec:
  inputDetail:    # Settings on CCE from which events are collected
    type: event    # Type of logs to be collected from CCE. Do not change the value.
    event:
      normalEvents:    # Used to configure Normal events
        enable: true    # Whether to enable Normal event collection
        includeNames:    # Name of the Normal event to be collected. If this parameter is not specified, all Normal events will be collected.
        - NotTriggerScaleUp
        excludeNames:    # Name of the Normal event that is not collected. If this parameter is not specified, all Normal events will be collected.
        - ScaleDown
      warningEvents:    # Used to configure Warning events
        enable: true    # Whether to enable Warning event collection
        includeNames:    # Name of the Warning event to be collected. If this parameter is not specified, all Warning events will be collected.
        - NotTriggerScaleUp
        excludeNames:    # Name of the Warning event that is not collected. If this parameter is not specified, all Warning events will be collected.
        - ScaleDown
  outputDetail:
    type: AOM    # Type of the system that receives the events. Do not change the value.
    AOM:
      events:
      - name: DeleteNodeWithNoServer    # Event name. This parameter is mandatory.
        resourceType: Namespace    # Type of the resource that operations are performed on.
        severity: Major    # Event severity after an event is reported to AOM, which can be Critical, Major, Minor, or Info. The default value is Major.
```

----End

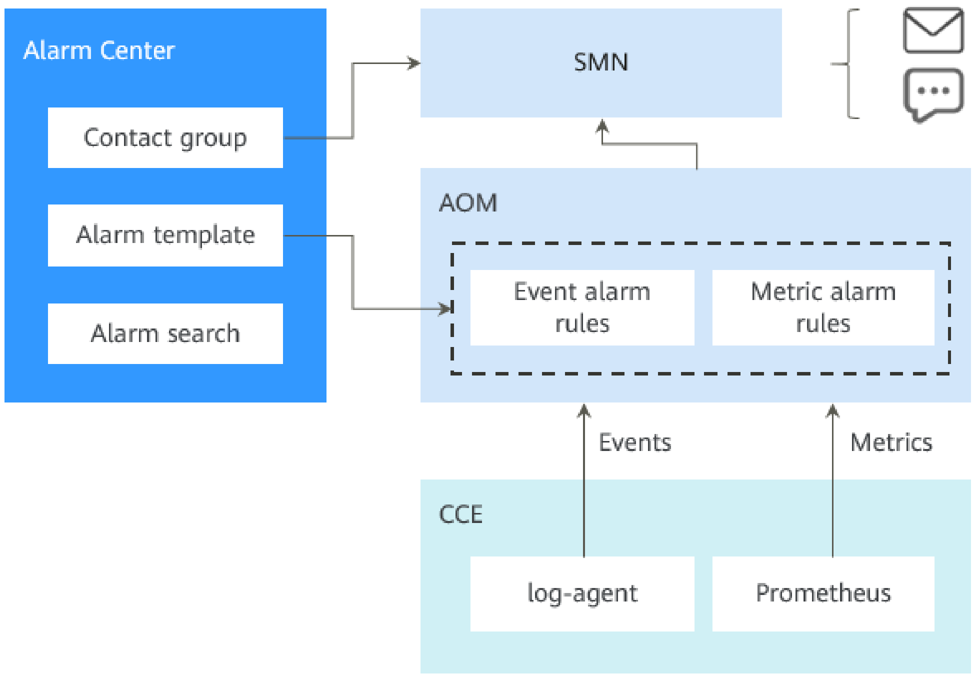
5.3 Alarm Center

5.3.1 Overview

Alarm reporting is an essential aspect of observability. In addition to traditional resource usage alarms (such as CPU and memory usage alarms), there are custom monitoring metric alarms (such as container restart alarms and application access failure alarms).

CCE works with AOM for metric and event alarm reporting and provides Alarm Center that allows you to quickly configure and view common alarms (such as resource usage alarms).

Figure 5-23 Alarm Center architecture



- Alarm Center

  Based on the alarm capabilities of AOM, Alarm Center provides quick alarm search and configuration for clusters. You can use Alarm Center to configure common alarm rules with just a few clicks.
- AOM

  AOM is a one-stop, multi-dimensional O&M management platform for monitoring and alarm reporting of cloud applications.
- Simple Message Notification

  Simple Message Notification (SMN) connects cloud applications. After events or alarms are triggered, SMN will send notifications. In cloud native scenarios, alarms triggered by AOM are sent by SMS message, email, or HTTP message configured on SMN.

5.3.2 Configuring Alarms in Alarm Center

By using AOM, Alarm Center can promptly detect cluster faults and generate alarms for service stability. Alarm Center provides built-in alarm rules, which can free you from manually configuring alarm rules on AOM. These rules are established based on the extensive cluster O&M experience of our Huawei Cloud container team and can cover container service exceptions, key metric alarms of basic cluster resources, and metric alarms of applications in a cluster to meet your routine O&M requirements.

Constraints

Only Huawei Cloud accounts, HUAWEI IDs, or IAM users with CCE administrator or FullAccess permissions can perform all operations using Alarm Center. IAM users with the CCE ReadOnlyAccess permission can only view all resources.

Enabling Alarm Center

Step 1 Click the cluster name to access the cluster console. In the navigation pane on the left, choose **Alarm Center**.

Step 2 On the **Alarm Rules** tab, click **Enable Alarm Center**. In the window that slides out from the right, select one or more contact groups to manage subscription endpoints and receive alarm messages by group. If no contact group is available, create one by referring to **Binding Contact Groups**.

Step 3 Click **OK**.

NOTE

Metric alarm rules can be created in Alarm Center only after the Cloud Native Cluster Monitoring add-on is installed and the AOM Prometheus instance is interconnected. For details about how to enable Monitoring Center, see **Enabling Cluster Monitoring**.

Event alarms in **Table 5-2** can be reported only when Kubernetes event collection is enabled in Logging. For details, see **Collecting Kubernetes Events**.

----End

Configuring Alarm Rules

After Alarm Center is enabled for clusters, you can configure and manage alarm rules.

Step 1 Log in to the CCE console.

Step 2 On the cluster list page, click the name of the target cluster to go to the details page.

Step 3 In the navigation pane on the left, choose **Alarm Center**. Then, click the **Alarm Rules** tab and configure and manage alarm rules.

By default, Alarm Center generates alarm rules for containers. The rules are intended for alarms including event alarms and metric alarms for exceptions. Alarm rules are classified into several sets. You can associate an alarm rule set with multiple contact groups and enable or disable alarm items. An alarm rule set consists of multiple alarm rules. An alarm rule corresponds to the check items for a single exception. **Table 5-2** lists default alarm rules.

----End

Table 5-2 Default alarm rules

Rule Type	Alarm Item	Description	Alarm Type	Dependency Item	PromQL/Event Name
Load rule set	Abnormal pod	Check whether the pod is running normally.	Metric	Cloud Native Cluster Monitoring	sum(min_over_time(kube_pod_status_phase{phase=~"Pending|Unknown|Failed"}[10m]) and count_over_time(kube_pod_status_phase{phase=~"Pending|Unknown|Failed"}[10m]) > 18)by (namespace,pod, phase, cluster_name, cluster) > 0
	Frequent pod restarts	Check whether the pod frequently restarts.	Metric	Cloud Native Cluster Monitoring	increase(kube_pod_container_status_restarts_total[5m]) > 3
	Unexpected number of Deployment replicas	Check whether the number of Deployment replicas is the same as the expected value.	Metric	Cloud Native Cluster Monitoring	(kube_deployment_spec_replicas != kube_deployment_status_replicas_available) and (changes(kube_deployment_status_replicas_updated[5m]) == 0)
	Unexpected number of StatefulSet replicas	Check whether the number of StatefulSet replicas is the same as the expected value.	Metric	Cloud Native Cluster Monitoring	(kube_statefulset_status_replicas_ready != kube_statefulset_status_replicas) and (changes(kube_statefulset_status_replicas_updated[5m]) == 0)

Rule Type	Alarm Item	Description	Alarm Type	Dependency Item	PromQL/Event Name
	Container CPU usage higher than 80%	Check whether the container CPU usage is higher than 80%.	Metric	Cloud Native Cluster Monitoring	100 * (sum(rate(container_cpu_usage_seconds_total{image!="", container!="POD"}[1m])) by (cluster_name,pod,node,namespace,container, cluster) / sum(kube_pod_container_resource_limits{resource="cpu"}) by (cluster_name,pod,node,namespace,container, cluster)) > 80
	Container memory usage higher than 80%	Check whether the container memory usage is higher than 80%.	Metric	Cloud Native Cluster Monitoring	(sum(container_memory_working_set_bytes{image!="", container!="POD"}) BY (cluster_name, node,container, pod , namespace, cluster) / sum(container_spec_memory_limit_bytes > 0) BY (cluster_name, node, container, pod , namespace, cluster) * 100) > 80
	Abnormal container	Check whether the container is running normally.	Metric	Cloud Native Cluster Monitoring	sum by (namespace, pod, container, cluster_name, cluster) (kube_pod_container_status_waiting_reason) > 0
	UpdateLoadBalancerFailed	Check whether a load balancer is updated.	Event	Cloud Native Log Collection	N/A

Rule Type	Alarm Item	Description	Alarm Type	Dependency Item	PromQL/Event Name
	Pod OOM	Check whether an OOM occurs in the pod.	Event	CCE Node Problem Detector Cloud Native Log Collection	PodOOMKilling
Cluster status rule set	Unavailable cluster	Check whether the cluster is available.	Event	Cloud Native Log Collection	N/A

Binding Contact Groups

NOTE

An alarm rule set can be bound to a maximum of five contact groups.

A contact group, backed on **Simple Message Notification**, enables message publishers and subscribers to contact each other. A contact group contains one or more terminals. You can bind an alarm rule to a contact group to manage terminals that have subscribed to alarm messages.

Step 1 Log in to the CCE console.

Step 2 On the cluster list page, click the name of the target cluster to go to the details page.

Step 3 In the navigation pane on the left, choose **Alarm Center**. Then, click the **Default Contact Groups** tab.

Step 4 Click **Bind Contact Group**. You can select a contact group created in SMN or create a contact group. The parameters for creating a contact group are described as follows:

- **Contact Group Name**: Enter the name of the contact group, which cannot be changed after the contact group is created. The name can contain 1 to 255 characters and must start with a letter or digit. Only letters, digits, hyphens (-), and underscores (_) are allowed.
- **Alarm Message Display Name**: Enter the title of the message received by the specified subscription endpoint. For example, if you set **endpoint Type** to **Email** and specify a display name, the name you specified will be displayed as the alarm message sender. If no alarm message display name is specified, the sender will be **username@example.com**. The alarm message display name can be changed after a contact group is created.
- **Add Subscription endpoint**: Add one or more endpoints to receive alarm messages. The endpoint type can be **SMS** or **Email**. If you select **SMS**, enter a valid mobile number. If you select **Email**, enter a valid email address.

Step 5 Click **OK**.

You will be redirected to the contact group list. The subscription endpoint is in the **Unconfirmed** state. Send a subscription request to the endpoint to verify its validity.

Step 6 Click **Request Confirmation** in the **Operation** column to send a subscription request to the endpoint. After the endpoint receives and confirms the request, the subscription endpoint status changes to **Confirmed**.

----End

Viewing Alarms

You can view the latest historical alarms on the **Alarm list** tab.

Step 1 Log in to the CCE console.

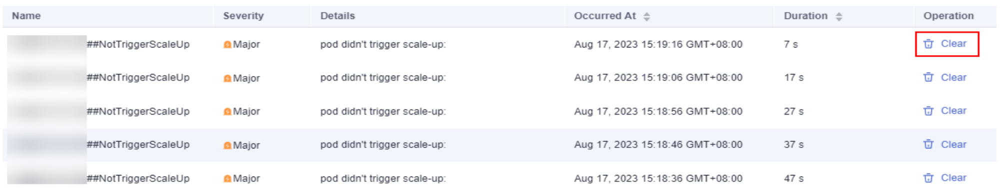
Step 2 On the cluster list page, click the name of the target cluster to go to the details page.

Step 3 In the navigation pane on the left, choose **Alarm Center**. Then, click the **Alarms** tab.

By default, all alarms to be cleared are displayed in the list. You can query alarms by alarm keyword, alarm severity, or alarm time. In addition, you can view the distribution of alarms that meet the specified criteria in different periods.

If you confirm that an alarm has been handled, click **Clear** in the **Operation** column. After the alarm is cleared, you can view it in the historical alarm list.

Figure 5-24 Querying alarms



Name	Severity	Details	Occurred At	Duration	Operation
[redacted]##NotTriggerScaleUp	Major	pod didn't trigger scale-up:	Aug 17, 2023 15:19:16 GMT+08:00	7 s	Clear
[redacted]##NotTriggerScaleUp	Major	pod didn't trigger scale-up:	Aug 17, 2023 15:19:06 GMT+08:00	17 s	Clear
[redacted]##NotTriggerScaleUp	Major	pod didn't trigger scale-up:	Aug 17, 2023 15:18:56 GMT+08:00	27 s	Clear
[redacted]##NotTriggerScaleUp	Major	pod didn't trigger scale-up:	Aug 17, 2023 15:18:46 GMT+08:00	37 s	Clear
[redacted]##NotTriggerScaleUp	Major	pod didn't trigger scale-up:	Aug 17, 2023 15:18:36 GMT+08:00	47 s	Clear

----End

5.3.3 Configuring Custom Alarms on CCE

If the default alarm rules cannot meet your requirements, you can create alarm rules on CCE. Based on the alarm rules, you can check whether resources in clusters are normal in a timely manner.

Adding Metric Alarms

NOTE

To create Prometheus metric threshold-crossing alarm rules and metric alarm rules, you need to enable Monitoring Center. For details, see **Enabling Cluster Monitoring**.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Alarm Center**. Then, choose **Alarm Rules** > **Custom Alarm Rules** and click **Create Alarm Rule**.

Step 3 Configure the alarm rule parameters.

- **Rule Type**: Select **Metric alarm**.
- **Alarm Template**: If you select **No template**, you need to configure the parameters in **Rule Details**. You can also set this parameter to **Use template** to quickly define a PromQL-based alarm rule or modify an existing template.
- **Rule Details**: Configure the parameters listed in the following table.

Parameter	Description	Example Value
Rule Name	Enter the name of the alarm rule.	CoreDNS memory usage higher than 80%
(Optional) Description	Describe the alarm rule.	Check whether the memory usage of CoreDNS is higher than 80%.
Alarm Rule (PromQL)	Enter a Prometheus query statement. For details about how to compile Prometheus query statements, see **Query Examples**.	The following is an example statement for generating an alarm when the maximum memory usage of CoreDNS is higher than 80%: `(sum(container_memory_working_set_bytes{image!="", container!="POD",namespace="kube-system",container="coredns"}) BY (cluster_name, node,container, pod, namespace, cluster) / sum(container_spec_memory_limit_bytes{namespace="kube-system", container="coredns"} > 0) BY (cluster_name, node, container, pod , namespace, cluster) * 100) > 80`
Severity	Select **Critical**, **Major**, **Minor**, or **Warning**.	Critical
Duration	Select an alarm duration from the drop-down list. The default value is **1 minute**.	1 minute
Alarm Content	Define the content in the alarm notification. Variables in Prometheus can be obtained in the form of ${variable}.	Example: `Cluster: ${cluster_name}, Namespace: ${namespace}, Pod: ${pod}, Container: ${container} memory usage is higher than 80%. The current value is ${value} %.`
Contact Group	Select an existing contact group. You can also click **Create Contact Group** to create a contact group. For details about the parameters, see **Binding Contact Groups**.	CCEGroup

In the preceding example, an alarm rule named **CoreDNS memory usage higher than 80%** is set for CoreDNS in the **kube-system** namespace, and its severity is **Critical**. When the maximum memory usage is higher than 80% for 1 minute, a notification is sent to all alarm contacts in the **CCEGroup** contact group by SMS message or email. The notification contains the cluster name, namespace, pod name, container name, and current memory usage.

- (Optional) Advanced Settings
  - **Alarm Tag**: An attribute for identifying and grouping alarms to reduce noise. In the message template, the tag value is referenced as *$event.metadata.* A maximum of 10 alarm tags can be added.
  - **Alarm Annotation**: An attribute that is not used for alarm identification. In the message template, the annotation value is referenced as *$event.annotations.* A maximum of 10 alarm annotations can be added.

Step 4 Click **OK**. Then, go to the **Custom Alarm Rules** page to check whether the rule is successfully created.

----End

Adding Event Alarms

NOTE

To create event-triggered alarm rules, you need to enable Logging and Kubernetes event collection. For details, see **Collecting Logs**.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Alarm Center**. Then, choose **Alarm Rules** > **Custom Alarm Rules** and click **Create Alarm Rule**.

Step 3 Configure the alarm rule parameters.

- Rule Type: Select **Event alarm**. Common events include Kubernetes events and cloud service events.
- **Rule Details**: Configure the parameters listed in the following table.

Parameter	Description	Example Value
Rule Name	Enter the name of the alarm rule.	ReplicaSet quantity change
(Optional) Description	Describe the alarm rule.	The number of ReplicaSets changes more than three times within 5 minutes.
Event Name	Enter the event name based on the actual Kubernetes event or cloud service event.	ScalingReplicaSet

Parameter	Description	Example Value
Triggering Mode	– **Immediate trigger**: An alarm is generated as long as the event occurs. – **Accumulative trigger**: An alarm is generated only after the event is triggered for a preset number of times within the triggering period.	Select **Accumulative trigger**, and set **Monitoring Interval** to **5 minutes** and **Occurrences** to **> 3**.
Severity	Select **Critical**, **Major**, **Minor**, or **Warning**.	Minor
Contact Group	Select an existing contact group. You can also click **Create Contact Group** to create a contact group. For details about the parameters, see **Binding Contact Groups**.	CCEGroup

In the preceding example, an alarm named **ReplicaSet quantity change** is set for the **ScalingReplicaSet** event, and its severity is **Minor**. When the number of ReplicaSet changes more than three times within 5 minutes, a notification is sent to all alarm contacts in the **CCEGroup** by SMS or email.

Step 4 Click **OK**. Then, go to the **Custom Alarm Rules** page to check whether the rule is successfully created.

----End

5.3.4 CCE Autopilot Cluster Events

CCE Autopilot can report a range of events in a running cluster to AOM. You can add event alarms as required to monitor the health of cluster data plane and control plane components. This helps you quickly identify and resolve problems, ensuring cluster stability and reliability.

- **Data Plane Events**: user operation events, such as workload, network, storage, and auto scaling events.
  – **Workload events**
  – **Network events**
  – **Storage events**
  – **Auto scaling events**

- **Control Plane Events**: master node events, which are usually caused by faults or upgrades of control plane components.

Data Plane Events

Table 5-3 Workload events

Object	Event Name	Severity	Description
Pod	PodOOMKilling	Major	Check whether the pod exits due to OOM. This event is reported by CCE Node Problem Detector (1.18.41 or later) and Cloud Native Log Collection (1.3.2 or later).
Pod	FailedStart	Major	Check whether the pod is started.
Pod	FailedPullImage	Major	Check whether the pod has pulled an image.
Pod	BackOffStart	Major	Check whether the pod fails to be restarted.
Pod	FailedScheduling	Major	Check whether the pod is scheduled.
Pod	BackOffPullImage	Major	Check whether the pod has pulled an image after a retry.
Pod	FailedCreate	Major	Check whether the pod is created.
Pod	Unhealthy	Minor	Check whether the pod health check is successful.
Pod	FailedDelete	Minor	Check whether the workload is deleted.
Pod	ErrImageNeverPull	Minor	Check whether the workload has pulled an image.
Pod	FailedScaleOut	Minor	Check whether the workload is scaled out using replicas.
Pod	FailedReconfig	Minor	Check whether the pod configuration is updated.
Pod	FailedActive	Minor	Check whether the pod is activated.
Pod	FailedRollback	Minor	Check whether the pod is rolled back.
Pod	FailedUpdate	Minor	Check whether the pod is updated.
Pod	FailedScaleIn	Minor	Check whether the pod scale-in failed.
Pod	FailedRestart	Minor	Check whether the pod is restarted.

Object	Event Name	Severity	Description
Deployment	SelectorOverlap	Minor	Check whether label selectors in the cluster conflict.
Deployment	ReplicaSetCreateError	Minor	Check whether a workload ReplicaSet can be created.
Deployment	DeploymentRollbackRevisionNotFound	Minor	Check whether the Deployment rollback version is available.
Job	TooManyActivePods	Minor	Check whether there are still active pods after the number of pods in a job reaches the preset value.
Job	TooManySucceededPods	Minor	Check whether there are extra running pods after the number of pods in a job reaches the preset value.
CronJob	FailedGet	Minor	Check whether the CronJob can be obtained.
CronJob	FailedList	Minor	Check whether the list of pods can be obtained.
CronJob	UnexpectedJob	Minor	Check whether there are any unknown CronJobs.

Table 5-4 Network events

Type	Event Name	Severity	Description
Service	CreatingLoadBalancerFailed	Minor	Check whether a load balancer is created.
Service	DeletingLoadBalancerFailed	Minor	Check whether the load balancer is deleted.
Service	UpdateLoadBalancerFailed	Minor	Check whether the load balancer is updated.

Table 5-5 Storage events

Type	Event Name	Severity	Description
PV	DetachVolumeFailed	Minor	Check whether the block storage is detached.
PV	VolumeUnknownReclaimPolicy	Minor	Check whether a volume reclamation policy is specified.

Type	Event Name	Severity	Description
PV	SetUpAtVolumeFailed	Minor	Check whether the volume is mounted.
PV	VolumeFailedRecycle	Minor	Check whether the volume is reclaimed.
PV	WaitForAttachVolumeFailed	Minor	Check whether the block storage is attached to the node.
PV	VolumeFailedDelete	Minor	Check whether the volume is deleted.
PV	MountDeviceFailed	Minor	Check whether the device is mounted.
PV	TearDownAtVolumeFailed	Minor	Check whether the volume is unmounted.
PV	UnmountDeviceFailed	Minor	Check whether the device is unmounted.
PV	AttachVolumeFailed	Minor	Check whether the block storage is detached from the node.
PVC	VolumeResizeFailed	Minor	Check whether the volume capacity is expanded.
PVC	ClaimLost	Minor	Check whether the PVC is normal.
PVC	ProvisioningFailed	Minor	Check whether the volume is created.
PVC	ProvisioningClea-nupFailed	Minor	Check whether the volume has been cleared.
PVC	ClaimMisbound	Minor	Check whether the PVC is bound to an incorrect volume.

Table 5-6 Auto scaling events

Type	Event Name	Severity	Description
HPA	InvalidTargetRange	Major	● Invalid **extendedhpa.metrics** is configured in **annotations** of HPA. ● The metric type in **spec** of HPA is incorrect.
HPA	FailedGetScale	Major	HPA failed to obtain the resource object to be scaled.

Type	Event Name	Severity	Description
HPA	FailedComputeMetricsReplicas	Major	An error occurs when the number of copies to be adjusted for resources is calculated. For example, metric-server is unavailable, resource metric collection fails, or the CPU usage is incorrectly set. You can run the following command to view details: `kubectl describe horizontalpodautoscaler <hpa-name>`
HPA	FailedGetObjectMetric	Major	Failed to obtain the metrics of the specified object (such as PVC and ConfigMaps).
HPA	FailedGetPodsMetric	Major	Failed to obtain pod resource metrics (resource usages of a pod).
HPA	FailedGetResourceMetric	Major	Failed to obtain cluster resource metrics (resource usages of a cluster).
HPA	FailedGetContainerResourceMetric	Major	Failed to obtain the resource metrics of a container.
HPA	FailedGetExternalMetric	Major	Failed to obtain external metrics.
HPA	FailedRescale	Major	Failed to update the desired number of copies of the resource object to be scaled.
HPA	SuccessfulRescale	Minor	The desired number of copies of the resource object to be scaled is updated.
CronHPA	ScaleFailed	Major	CronHPA failed to update the desired number of copies of the resource object to be scaled.
CronHPA	FailedGetHorizontalPodAutoscaler	Major	CronHPA failed to query the associated HPA object. (Generally, kube-apiserver cannot respond.)
CronHPA	FailedGetHpaScale	Major	CronHPA failed to obtain the resource object to be scaled.
CronHPA	UpdateHPAFailed	Major	CronHPA failed to update the associated HPA object.
CronHPA	UpdateHPASuccess	Minor	CronHPA successfully updates the associated HPA object.

Type	Event Name	Severity	Description
CronHPA	SkipUpdateHPA	Minor	CronHPA skips updating the associated HPA object.
CronHPA	SkipUpdateTarget	Minor	CronHPA skips updating the number of copies of the resource object to be scaled.
CronHPA	UpdateTargetSuccess	Minor	CronHPA successfully updates the number of copies of the resource object to be scaled.
CustomedHPA	FailedSetPolicy-Settings	Major	Failed to parse the cooldown period of CustomedHPA.
CustomedHPA	FailedSubmitRule	Major	CustomedHPA failed to process schedule rules or metric rules.
CustomedHPA	FailedComputeReplicas	Major	CustomedHPA failed to trigger resource scaling based on the compute metrics.
CustomedHPA	FailedScale	Major	CustomedHPA failed to update the desired number of copies of the resource object to be scaled. (Generally, kube-apiserver cannot respond).
CustomedHPA	MetricScaleSuccess	Minor	CustomedHPA triggers resource scaling based on the metric rule.
CustomedHPA	CronScaleSuccess	Minor	CustomedHPA triggers resource scaling based on the periodic rule.

Control Plane Events

Table 5-7 Control plane events

Event ID	Severity	Description
Internal error	Major	Check whether an internal error occurs in the cluster.
Failed to check component status or components are abnormal	Major	Check whether the statuses of cluster components can be obtained or whether the components malfunction.
Cluster status is Unavailable	Major	Check whether the cluster is available.
Cluster status is Error	Major	Check whether the cluster is faulty.

Event ID	Severity	Description
Cluster status is not updated for a long time	Major	Check whether the cluster is stuck in a state for a long time.
Failed to update cluster status	Major	Check whether the cluster status is updated.
Failed to delete the unavailable connection of the Kubernetes cluster	Major	Check whether unavailable Kubernetes connections are deleted.
Failed to sync the cluster cert	Major	Check whether the cluster certificate is synchronized.

6 Namespaces

6.1 Creating a Namespace

When to Use Namespaces

A namespace is a collection of resources and objects. Multiple namespaces can be created inside a cluster and isolated from each other. This enables namespaces to share the same cluster Services without affecting each other.

For example, you can deploy workloads in a development environment into one namespace, and deploy workloads in a testing environment into another namespace.

Prerequisites

At least one cluster has been created.

Constraints

A maximum of 6,000 Services can be created in each namespace. The Services mentioned here indicate the Kubernetes Service resources added for workloads.

Namespace Types

Namespaces can be created in either of the following ways:

- Created automatically: When a cluster is up, the **default**, **kube-public**, **kube-system**, and **kube-node-lease** namespaces are created by default.
  - **default**: All objects for which no namespace is specified are allocated to this namespace.
  - **kube-public**: Resources in this namespace can be accessed by all users (including unauthenticated users), such as public add-ons and container charts.
  - **kube-system**: All resources created by Kubernetes are in this namespace.
  - **kube-node-lease**: Each node has an associated Lease object in this namespace. The object is periodically updated by the node. Both

NodeStatus and NodeLease are considered as heartbeats from a node. In versions earlier than v1.13, only NodeStatus is available. The NodeLease feature is introduced in v1.13. NodeLease is more lightweight than NodeStatus. This feature significantly improves the cluster scalability and performance.

- Created manually: You can create namespaces to serve separate purposes. For example, you can create three namespaces, one for a development environment, one for joint debugging environment, and one for test environment. You can also create one namespace for login services and one for game services.

Creating a Namespace

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Namespaces**. Then click **Create Namespace** in the upper right corner.

Step 3 Configure the parameters based on **Table 6-1**.

Table 6-1 Parameters for creating a namespace

Parameter	Description
Name	Unique name of the created namespace.
Description	Description about the namespace.
Quota Management	Resource quotas can limit the number of resources available in namespaces, for resource allocation by namespace. **NOTICE** **You are advised to set resource quotas in the namespace as required to prevent cluster or node exceptions caused by resource overload.** Enter an integer. If the quota of a resource is not specified, no limit is posed on the resource. If you want to limit the CPU or memory quota, you must specify the CPU or memory request value when creating a workload.

Step 4 Click **OK**.

----End

Using kubectl to Create a Namespace

Define a namespace.

```
apiVersion: v1
kind: Namespace
metadata:
  name: custom-namespace
```

Run the **kubectl** command to create it.

```
$ kubectl create -f custom-namespace.yaml
namespace/custom-namespace created
```

You can also run the **kubectl create namespace** command to create a namespace.

```
$ kubectl create namespace custom-namespace
namespace/custom-namespace created
```

6.2 Managing Namespaces

Using Namespaces

- When creating a workload, you can select a namespace to isolate resources or users.
- When querying workloads, you can select a namespace to view all workloads in the namespace.

Isolating Namespaces

- **Isolating namespaces by environment**

  An application generally goes through the development, joint debugging, and testing stages before it is launched. In this process, the workloads deployed in each environment (stage) are the same, but are logically defined. There are two ways to define them:

  – Group them in different clusters for different environments.

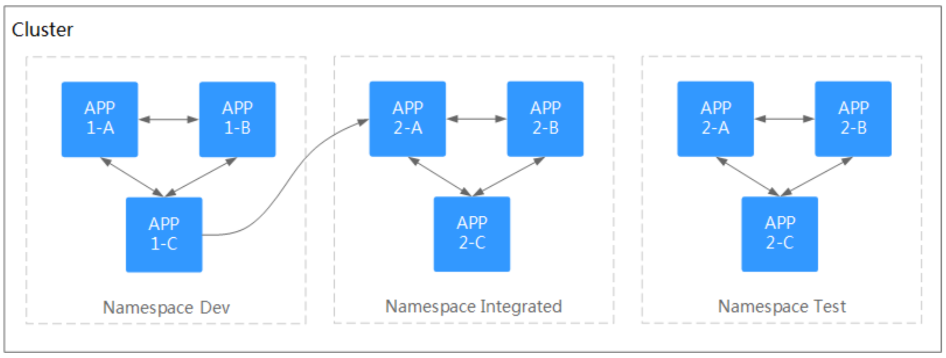
    Resources cannot be shared among different clusters. In addition, services in different environments can access each other only through load balancing.

  – Group them in different namespaces for different environments.

    Workloads in the same namespace can be mutually accessed by using the Service name. Cross-namespace access can be implemented by using the Service name or namespace name.

    The following figure shows namespaces created for the development, joint debugging, and testing environments, respectively.

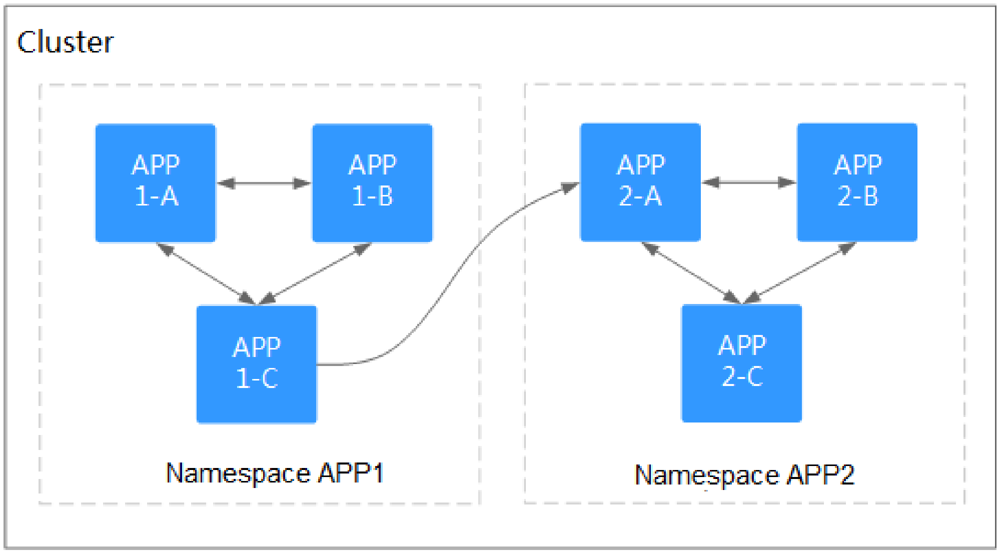
Figure 6-1 One namespace for one environment



- **Isolating namespaces by application**

  You are advised to use this method if a large number of workloads are deployed in the same environment. For example, in the following figure, different namespaces (APP1 and APP2) are created to logically manage workloads as different groups. Workloads in the same namespace access each other using the Service name, and workloads in different namespaces access each other using the Service name or namespace name.

Figure 6-2 Grouping workloads into different namespaces



Managing Namespace Labels

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Namespaces**.

Step 2 Locate the row containing the target namespace and choose **More** > **Manage Label** in the **Operation** column.

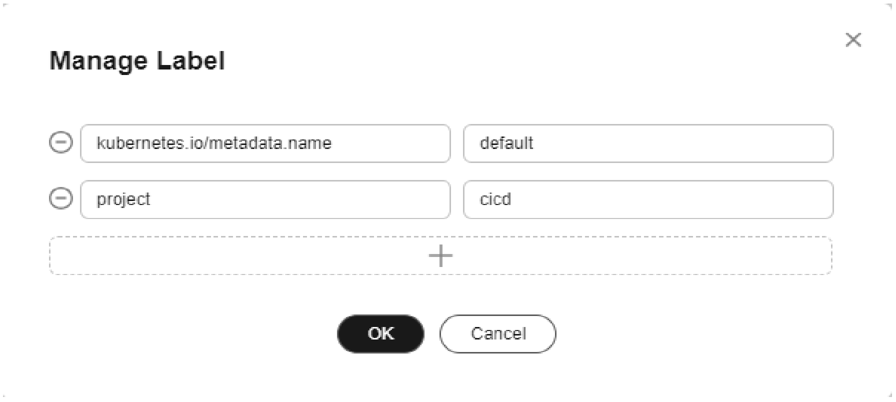
Step 3 In the dialog box that is displayed, the existing labels of the namespace are displayed. Modify the labels as needed.

- Adding a label: Click the add icon, enter the key and value of the label to be added, and click **OK**.

  For example, the key is **project** and the value is **cicd**, indicating that the namespace is used to deploy CICD.

- Deleting a label: Click ⊖ next the label to be deleted and then **OK**.

Figure 6-3 Adding or deleting a namespace label



kubernetes.io/metadata.name	default
project	cicd

Step 4 Switch to the **Manage Label** dialog box again and check the modified labels.

----End

Deleting a Namespace

If a namespace is deleted, all resources (such as workloads, jobs, and ConfigMaps) in this namespace will also be deleted. Exercise caution when deleting a namespace.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Namespaces**. Locate the target namespace and choose **More** > **Delete** in the **Operation** column.

Follow the prompts to delete the namespace. The default namespaces cannot be deleted.

----End

6.3 Setting Resource Quotas

Namespace-level resource quotas limit the amount of resources available to teams or users when these teams or users use the same cluster. The quotas include the total number of a type of objects and the total amount of compute resources (CPU and memory) consumed by the objects.

Usage

By default, running pods can use the CPUs and memory of a node without restrictions. This means the pods in a namespace may exhaust all resources of the cluster.

Kubernetes provides namespaces for you to group workloads in a cluster. By setting resource quotas for each namespace, you can prevent resource exhaustion and ensure cluster reliability.

You can configure quotas for resources such as CPU, memory, and the number of pods in a namespace. For more information, see **Resource Quotas**.

Constraints

Kubernetes provides optimistic concurrency control (OCC), also known as optimistic locking, for frequent data updates. You can use optimistic locking by defining the **resourceVersion** field. This field is in the object metadata. This field identifies the internal version number of the object. When the object is modified, this field is modified accordingly. You can use kube-apiserver to check whether an object has been modified. When the API server receives an update request containing the **resourceVersion** field, the server compares the requested data with the resource version number of the server. If they are different, the object on the server has been modified when the update is submitted. In this case, the API server returns a conflict error (409). Obtain the server data, modify the data, and submit the data to the server again. The resource quota limits the total resource consumption of each namespace and records the resource information in the cluster. Therefore, after the **enable-resource-quota** option is enabled, the probability of resource creation conflicts increases in large-scale concurrency scenarios, affecting the performance of batch resource creation.

Procedure

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Namespaces**.

Step 3 Click **Quota Management** next to the target namespace.

This operation cannot be performed on system namespaces **kube-system** and **kube-public**.

Step 4 Set the resource quotas and click **OK**.

NOTICE

- After setting CPU and memory quotas for a namespace, you must specify the request and limit values of CPU and memory resources when creating a workload. Otherwise, the workload cannot be created. If the quota of a resource is set to **0**, the resource usage is not limited.
- Accumulated quota usage includes the resources used by CCE to create default components, such as the Kubernetes Services (which can be viewed using kubectl) created under the **default** namespace. Therefore, you are advised to set a resource quota greater than expected to reserve resource for creating default components.

----End

7 ConfigMaps and Secrets

7.1 Creating a ConfigMap

Scenario

A ConfigMap is a type of resource that stores configuration information required by a workload. Its content is user-defined. After creating ConfigMaps, you can use them as files or environment variables in a containerized workload.

ConfigMaps allow you to decouple configuration files from container images to enhance the portability of workloads.

Benefits of ConfigMaps:

- Manage configurations of different environments and services.
- Deploy workloads in different environments. Multiple versions are supported for configuration files so that you can update and roll back workloads easily.
- Quickly import configurations in the form of files to containers.

Notes and Constraints

- The size of a ConfigMap resource file cannot exceed 1 MB.
- ConfigMaps cannot be used in **static pods**.

Procedure

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Choose **ConfigMaps and Secrets** in the navigation pane and click **Create ConfigMap** in the upper right corner.

Step 3 Configure parameters.

Table 7-1 Parameters for creating a ConfigMap

Parameter	Description
Name	Name of the ConfigMap you create, which must be unique in a namespace.
Namespace	Namespace to which the ConfigMap belongs. If you do not specify this parameter, the value **default** is used by default.
Description	Description of the ConfigMap.
Data	Data of a ConfigMap, in the key-value pair format. Click + to add data. The value can be in string, JSON, or YAML format.
Label	Label of the ConfigMap. Enter a key-value pair and click **Confirm**.

Step 4 Click **OK**.

The new ConfigMap is displayed in the ConfigMap list.

----End

Creating a ConfigMap Using kubectl

Step 1 Use kubectl to access the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **cce-configmap.yaml** and edit it.

vi cce-configmap.yaml

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: cce-configmap
data:
  SPECIAL_LEVEL: Hello
  SPECIAL_TYPE: CCE
```

Table 7-2 Key parameters

Parameter	Description
apiVersion	The value is fixed at **v1**.
kind	The value is fixed at **ConfigMap**.
metadata.name	ConfigMap name, which can be customized.
data	ConfigMap data. The value must be key-value pairs.

Step 3 Run the following commands to create a ConfigMap.

kubectl create -f cce-configmap.yaml

Run the following commands to view the created ConfigMap:

kubectl get cm

```
NAME            DATA         AGE
cce-configmap    3           7m
```

----End

Related Operations

After creating a ConfigMap, you can update or delete it as described in **Table 7-3**.

Table 7-3 Related operations

Operation	Description
Editing a YAML file	Click **Edit YAML** in the row where the target ConfigMap resides to edit its YAML file.
Updating a ConfigMap	1. Select the name of the ConfigMap to be updated and click **Update**. 2. Modify the secret data. For more information, see **Table 7-1**. 3. Click **OK**.
Deleting a ConfigMap	Select the configuration you want to delete and click **Delete**. Follow the prompts to delete the ConfigMap.

7.2 Using a ConfigMap

You can use a ConfigMap to set environment variables or command line parameters, or mount a ConfigMap as a volume.

- **Configuring Environment Variables for a Workload**
- **Configuring Command Line Parameters**
- **Mounting a ConfigMap Volume to a Workload**

The following example shows how to use a ConfigMap.

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: cce-configmap
data:
  SPECIAL_LEVEL: Hello
  SPECIAL_TYPE: CCE
```

NOTICE

- When a ConfigMap is used, it must be in the same cluster and namespace as the workload.
- When a ConfigMap mounted as a volume is updated, Kubernetes updates the data in the volume at the same time.

  If a ConfigMap is mounted as a volume using **subPath**, data in the volume cannot be automatically updated when the ConfigMap is updated.
- When a ConfigMap is used as an environment variable, data in the environment variable cannot be automatically updated when the ConfigMap is updated. To update the data, restart the pod.

Configuring Environment Variables for a Workload

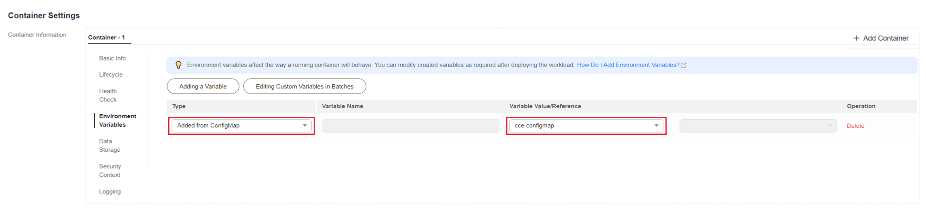
Using the console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. In the upper right corner of the displayed page, click **Create Workload**.

When creating a workload, click **Environment Variables** in the **Container Settings** area, and click **Add Variable**.

- **Added from ConfigMap**: Select a ConfigMap to import all of its keys as environment variables.

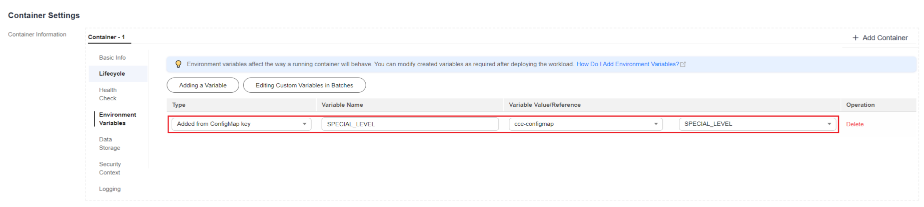


Environment variables affect the way a running container will behave. You can modify created variables as required after deploying the workload. How Do I Add Environment Variables?

Type	Variable Name	Variable Value/Reference		Operation
Added from ConfigMap		cce-configmap		Delete

- **Added from ConfigMap key**: Import the key in a ConfigMap as the value of an environment variable.
  - **Variable Name**: Name of the environment variable. The name can be customized and is set to the key name selected in the ConfigMap by default.
  - **Variable Value/Reference**: Select a ConfigMap and the key to be imported. The corresponding value is imported as a workload environment variable.

  For example, after you import the value **Hello** of **SPECIAL_LEVEL** in ConfigMap **cce-configmap** as the value of workload environment variable **SPECIAL_LEVEL**, an environment variable named **SPECIAL_LEVEL** with its value **Hello** exists in the container.



Environment variables affect the way a running container will behave. You can modify created variables as required after deploying the workload. How Do I Add Environment Variables?

Type	Variable Name	Variable Value/Reference		Operation
Added from ConfigMap key	SPECIAL_LEVEL	cce-configmap	SPECIAL_LEVEL	Delete

Step 3 Configure other workload parameters and click **Create Workload**.

When the workload runs normally, **log in to the container** and run the following statement to check whether the ConfigMap has been set as an environment variable for the workload:

```
printenv SPECIAL_LEVEL
```

The following is an example output:

```
Hello
```

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-configmap.yaml** and edit it.

vi nginx-configmap.yaml

The content of the YAML file is as follows:

- **Added from ConfigMap**: To add all data in a ConfigMap to environment variables, use the **envFrom** parameter. The keys in the ConfigMap will become names of environment variables in the workload.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-configmap
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-configmap
  template:
    metadata:
      labels:
        app: nginx-configmap
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        envFrom:                    # Use envFrom to specify a ConfigMap to be referenced by
environment variables.
        - configMapRef:
            name: cce-configmap       # Name of the referenced ConfigMap.
      imagePullSecrets:
      - name: default-secret
```

- **Added from ConfigMap key**: When creating a workload, you can use a ConfigMap to set environment variables and use the **valueFrom** parameter to reference the key-value pair in the ConfigMap separately.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-configmap
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-configmap
  template:
    metadata:
      labels:
```

```
      app: nginx-configmap
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        env:                          # Set an environment variable for the workload.
        - name: SPECIAL_LEVEL           # Name of the environment variable for the workload.
          valueFrom:                  # Specify a ConfigMap to be referenced by the environment variable.
            configMapKeyRef:
              name: cce-configmap       # Name of the referenced ConfigMap.
              key: SPECIAL_LEVEL        # Key in the referenced ConfigMap.
        - name: SPECIAL_TYPE            # Add multiple environment variables to import them at the 
same time.
          valueFrom:
            configMapKeyRef:
              name: cce-configmap
              key: SPECIAL_TYPE
      imagePullSecrets:
      - name: default-secret
```

Step 3 Create a workload.

kubectl apply -f nginx-configmap.yaml

Step 4 View the environment variables in the pod.

1. Run the following command to view the created pod:
```
kubectl get pod | grep nginx-configmap
```
Expected output:
```
nginx-configmap-***   1/1     Running   0              2m18s
```
2. Run the following command to view the environment variables in the pod:
```
kubectl exec nginx-configmap-*** -- printenv SPECIAL_LEVEL SPECIAL_TYPE
```
Expected output:
```
Hello
CCE
```
The ConfigMap has been set as environment variables of the workload.

----End

Configuring Command Line Parameters

You can use a ConfigMap as an environment variable to set commands or parameter values for a container by using the environment variable substitution syntax $(VAR_NAME).

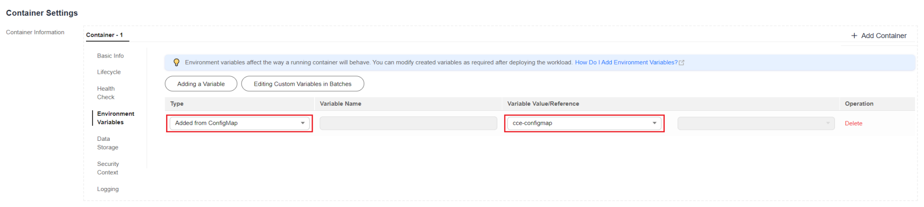
Using the console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. In the upper right corner of the displayed page, click **Create Workload**.

When creating a workload, click **Environment Variables** in the **Container Settings** area, and click **Add Variable**. In this example, select **Added from ConfigMap**.

- **Added from ConfigMap**: Select a ConfigMap to import all of its keys as environment variables.


Container Settings

Container Information

Container - 1 + Add Container

Basic Info | Lifecycle | Health Check | Environment Variables | Data Storage | Security Context | Logging

Environment variables affect the way a running container will behave. You can modify created variables as required after deploying the workload. How Do I Add Environment Variables?

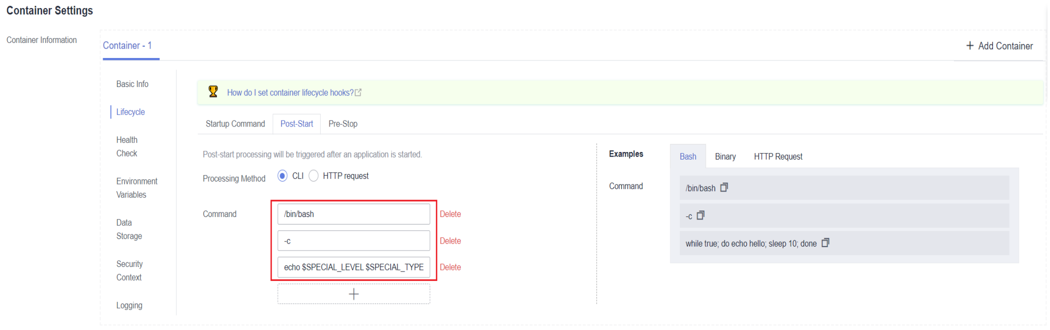
Adding a Variable | Editing Custom Variables in Batches

Type	Variable Name	Variable Value/Reference		Operation
Added from ConfigMap		cce-configmap		Delete

Step 3 Click **Lifecycle** in the **Container Settings** area, click the **Post-Start** tab on the right, and configure the following parameters:

- **Processing Method**: **CLI**
- **Command**: Enter the following three command lines. *SPECIAL_LEVEL* and *SPECIAL_TYPE* are the environment variable names in the workload, that is, the key names in the **cce-configmap** ConfigMap.

```
/bin/bash
-c
echo $SPECIAL_LEVEL $SPECIAL_TYPE > /usr/share/nginx/html/index.html
```


Container Settings

Container Information

Container - 1 + Add Container

Basic Info | Lifecycle | Health Check | Environment Variables | Data Storage | Security Context | Logging

How do I set container lifecycle hooks?

Startup Command | Post-Start | Pre-Stop

Post-start processing will be triggered after an application is started.

Processing Method [✓ CLI] HTTP request

Command

```
/bin/bash
-c
echo $SPECIAL_LEVEL $SPECIAL_TYPE
```

Examples: Bash | Binary | HTTP Request

Command

```
/bin/bash
-c
while true; do echo hello; sleep 10; done
```

Step 4 Configure other workload parameters and click **Create Workload**.

When the workload runs normally, **log in to the container** and run the following statement to check whether the ConfigMap has been set as an environment variable for the workload:

```
cat /usr/share/nginx/html/index.html
```

The following is an example output:

```
Hello CCE
```

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-configmap.yaml** and edit it.

vi nginx-configmap.yaml

In this example, the **cce-configmap** ConfigMap is imported to the workload. *SPECIAL_LEVEL* and *SPECIAL_TYPE* are the environment variable names in the workload, that is, the key names in the **cce-configmap** ConfigMap.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-configmap
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-configmap
  template:
    metadata:
      labels:
        app: nginx-configmap
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        lifecycle:
          postStart:
            exec:
              command: [ "/bin/sh", "-c", "echo $SPECIAL_LEVEL $SPECIAL_TYPE > /usr/share/nginx/html/
index.html" ]
        envFrom:                    # Use envFrom to specify a ConfigMap to be referenced by environment
variables.
        - configMapRef:
            name: cce-configmap       # Name of the referenced ConfigMap.
      imagePullSecrets:
        - name: default-secret
```

Step 3 Create a workload.

kubectl apply -f nginx-configmap.yaml

Step 4 Wait until the workload runs normally. The following content will be added to the **/usr/share/nginx/html/index.html** file of the container:

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep nginx-configmap
   ```
   Expected output:
   ```
   nginx-configmap-***   1/1     Running   0              2m18s
   ```
2. Run the following command to view the environment variables in the pod:
   ```
   kubectl exec nginx-configmap-*** -- cat /usr/share/nginx/html/index.html
   ```
   Expected output:
   ```
   Hello CCE
   ```

----End

Mounting a ConfigMap Volume to a Workload

A ConfigMap can be mounted as a volume to a specified container path. The workload code is separated from the configuration files. ConfigMap volumes are used to store workload configuration parameters. You need to create ConfigMaps in advance by following the instructions in **Creating a ConfigMap**.

Using the console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. In the upper right corner of the displayed page, click **Create Workload**.

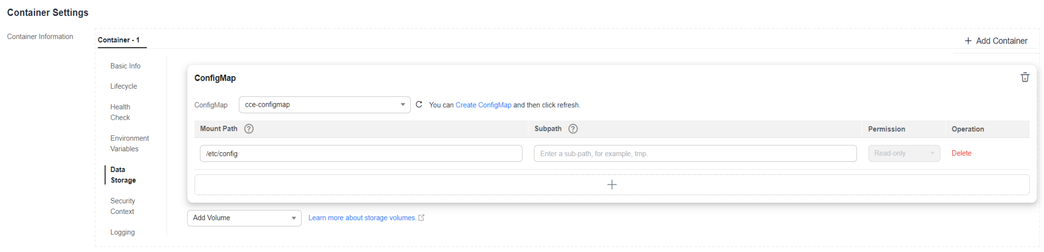
When creating a workload, click **Data Storage** in the **Container Settings** area. Click **Add Volume** and select **ConfigMap** from the drop-down list.

Step 3 Configure the parameters for mounting a ConfigMap volume based on **Table 7-4**.

Table 7-4 Parameters for mounting a ConfigMap volume

Parameter	Description
ConfigMap	Select the desired ConfigMap. A ConfigMap must be created beforehand. For details, see **Creating a ConfigMap**.
Mount Path	Enter a mount point. After the ConfigMap volume is mounted, a configuration file with the key as the file name and value as the file content is generated in the mount path of the container. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.
Subpath	Enter a subpath of the mount path. ● A subpath is used to mount a local volume so that the same volume is used in a single pod. If this parameter is left blank, the root path is used by default. ● The subpath can be the key and value of a ConfigMap. If the subpath is a key-value pair that does not exist, the data import does not take effect. ● The data imported using a subpath will not be updated along with the ConfigMap.
Permission	Read-only permission. The volume in the path is read-only.

Figure 7-1 Mounting a ConfigMap volume



Container Settings

Container Information

Container - 1 | + Add Container

Basic Info / Lifecycle / Health Check / Environment Variables / Data Storage / Security Context / Logging

ConfigMap

ConfigMap: cce-configmap — You can Create ConfigMap and then click refresh.

Mount Path	Subpath	Permission	Operation
/etc/config	Enter a sub-path, for example, tmp.	Read-only	Delete

Add Volume — Learn more about storage volumes.

Step 4 Configure other parameters and click **Create Workload**.

When the workload runs normally, the **SPECIAL_LEVEL** and **SPECIAL_TYPE** files will be generated in the **/etc/config** directory in this example. The content of file **SPECIAL_LEVEL** is **Hello**, and that of **SPECIAL_TYPE** is **CCE**.

Access the container and run the following statement to view the **SPECIAL_LEVEL** or **SPECIAL_TYPE** file in the container:

```
cat /etc/config/SPECIAL_LEVEL
```

Expected output:

```
Hello
```

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-configmap.yaml** and edit it.

vi nginx-configmap.yaml

In this example, after the ConfigMap volume is mounted, a configuration file with the key as the file name and value as the file content is generated in the **/etc/config** directory of the container.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-configmap
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-configmap
  template:
    metadata:
      labels:
        app: nginx-configmap
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
        - name: config-volume
          mountPath: /etc/config           # Mount the volume to the /etc/config directory.
          readOnly: true
      volumes:
      - name: config-volume
        configMap:
          name: cce-configmap              # Name of the referenced ConfigMap.
```

Step 3 Create a workload.

kubectl apply -f nginx-configmap.yaml

Step 4 When the workload runs normally, the **SPECIAL_LEVEL** and **SPECIAL_TYPE** files will be generated in the **/etc/config** directory in this example. The content of file **SPECIAL_LEVEL** is **Hello**, and that of **SPECIAL_TYPE** is **CCE**.

1. Run the following command to view the created pod:
   ```
   kubectl get pod | grep nginx-configmap
   ```
   Expected output:
   ```
   nginx-configmap-***   1/1     Running   0              2m18s
   ```
2. Run the following command to view the **SPECIAL_LEVEL** or **SPECIAL_TYPE** file in the pod:
   ```
   kubectl exec nginx-configmap-*** -- cat /etc/config/SPECIAL_LEVEL
   ```

Expected output:

```
Hello
```

----End

7.3 Creating a Secret

Scenario

A secret is a type of resource that holds sensitive data, such as authentication and key information. Its content is user-defined. After creating secrets, you can use them as files or environment variables in a containerized workload.

Notes and Constraints

Secrets cannot be used in **static pods**.

Procedure

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Choose **ConfigMaps and Secrets** in the navigation pane, click the **Secrets** tab, and click **Create Secret** in the upper right corner.

Step 3 Configure parameters.

Table 7-5 Parameters for creating a secret

Parameter	Description
Name	Name of the secret you create, which must be unique.
Namespace	Namespace to which the secret belongs. If you do not specify this parameter, the value **default** is used by default.
Description	Description of a secret.
Type	Type of the secret you create. • Opaque: common secret. • kubernetes.io/dockerconfigjson: a secret that stores the authentication information required for pulling images from a private repository. • **kubernetes.io/tls**: Kubernetes TLS secret, which is used to store the certificate required by layer-7 load balancing Services. For details about examples of the kubernetes.io/tls secret and its description, see **TLS secrets**. • **IngressTLS**: TLS secret provided by CCE to store the certificate required by layer-7 load balancing Services. • Other: another type of secret, which is specified manually.

Parameter	Description
Secret Data	Workload secret data can be used in containers. ● If **Secret Type** is **Opaque**, click +. In the dialog box displayed, enter a key-value pair and select **Auto Base64 Encoding**. ● If **Secret Type** is **kubernetes.io/dockerconfigjson**, enter the account and password for logging in to the private image repository. ● If **Secret Type** is **kubernetes.io/tls** or **IngressTLS**, upload the certificate file and private key file. **NOTE** – A certificate is a self-signed or CA-signed credential used for identity authentication. – A certificate request is a request for a signature with a private key.
Secret Label	Label of the secret. Enter a key-value pair and click **Confirm**.

Step 4 Click **OK**.

The new secret is displayed in the key list.

----End

Secret Resource File Configuration Example

This section describes configuration examples of secret resource description files.

- Opaque type

  The **secret.yaml** file is defined as shown below. The **data** field is filled in as a key-value pair, and the **value** field must be encoded using Base64. For details about the Base64 encoding method, see **Base64 Encoding**.

```
apiVersion: v1
kind: Secret
metadata:
  name: mysecret           #Secret name
  namespace: default       #Namespace. The default value is default.
data:
  <your_key>: <your_value>  # Enter a key-value pair. The value must be encoded using Base64.
type: Opaque
```

- kubernetes.io/dockerconfigjson type

  The **secret.yaml** file is defined as shown below. The value of **.dockerconfigjson** must be encoded using Base64. For details, see **Base64 Encoding**.

```
apiVersion: v1
kind: Secret
metadata:
  name: mysecret           #Secret name
  namespace: default       #Namespace. The default value is default.
data:
  .dockerconfigjson: eyJh*****    # Content encoded using Base64.
type: kubernetes.io/dockerconfigjson
```

To obtain the **.dockerconfigjson** content, perform the following steps:

a. Obtain the following login information of the image repository.

   - Image repository address: The section uses *address* as an example. Replace it with the actual address.
   - Username: The section uses *username* as an example. Replace it with the actual username.
   - Password: The section uses *password* as an example. Replace it with the actual password.

b. Use Base64 to encode the key-value pair *username:password* and fill the encoded content in 3.

```
echo -n "username:password" | base64
```

Command output:

```
dXNlcm5hbWU6cGFzc3dvcmQ=
```

c. Use Base64 to encode the following JSON content:

```
echo -n '{"auths":{"address":
{"username":"username","password":"password","auth":"dXNlcm5hbWU6cGFzc3dvcmQ="}}}'
| base64
```

Command output:

```
eyJhdXRocyI6eyJhZGRyZXNzIjp7InVzZXJuYW1lIjoidXNlcm5hbWUiLCJwYXNzd29yZCI6InBhc3N3b3J
kIiwiYXV0aCI6ImRYTmxjbTVoYldVNmNHRnpjM2R2Y21RPSJ9fX0=
```

The encoded content is the **.dockerconfigjson** content.

- kubernetes.io/tls type

  The value of **tls.crt** and **tls.key** must be encoded using Base64. For details, see **Base64 Encoding**.

```
kind: Secret
apiVersion: v1
metadata:
  name: mysecret          #Secret name
  namespace: default      #Namespace. The default value is default.
data:
  tls.crt: LS0tLS1CRU*****FURS0tLS0t  # Certificate content, which must be encoded using Base64.
  tls.key: LS0tLS1CRU*****VZLS0tLS0=  # Private key content, which must be encoded using Base64.
type: kubernetes.io/tls
```

- IngressTLS type

  The value of **tls.crt** and **tls.key** must be encoded using Base64. For details, see **Base64 Encoding**.

```
kind: Secret
apiVersion: v1
metadata:
  name: mysecret          #Secret name
  namespace: default      #Namespace. The default value is default.
data:
  tls.crt: LS0tLS1CRU*****FURS0tLS0t  # Certificate content, which must be encoded using Base64.
  tls.key: LS0tLS1CRU*****VZLS0tLS0=  # Private key content, which must be encoded using Base64.
type: IngressTLS
```

Creating a Secret Using kubectl

Step 1 Use kubectl to access the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create and edit the Base64-encoded **cce-secret.yaml** file.

```
# echo -n "content to be encoded" | base64
******
```

vi cce-secret.yaml

The following YAML file uses the Opaque type as an example. For details about other types, see **Secret Resource File Configuration Example**.

```
apiVersion: v1
kind: Secret
metadata:
  name: mysecret
type: Opaque
data:
  <your_key>: <your_value>  # Enter a key-value pair. The value must be encoded using Base64.
```

Step 3 Create a secret.

kubectl create -f cce-secret.yaml

You can query the secret after creation.

kubectl get secret -n default

----End

Related Operations

After creating a secret, you can update or delete it as described in **Table 7-6**.

NOTE

The secret list contains system secret resources that can be queried only. The system secret resources cannot be updated or deleted.

Table 7-6 Related Operations

Operation	Description
Editing a YAML file	Click **Edit YAML** in the row where the target secret resides to edit its YAML file.
Updating a secret	1. Select the name of the secret to be updated and click **Update**. 2. Modify the secret data. For more information, see **Table 7-5**. 3. Click **OK**.
Deleting a secret	Select the secret you want to delete and click **Delete**. Follow the prompts to delete the secret.
Deleting secrets in batches	1. Select the secrets to be deleted. 2. Click **Delete** above the secret list. 3. Follow the prompts to delete the secrets.

Base64 Encoding

To Base64-encode a string, run the **echo -n** ***content to be encoded*** **| base64** command. The following is an example:

```
root@ubuntu:~# echo -n "content to be encoded" | base64
******
```

7.4 Using a Secret

After secrets are created, they can be mounted as volumes or exposed as environment variables for a container.

NOTICE

Do not perform any operation on the following secrets (for details, see **Cluster Secrets**):

- Secrets under kube-system
- default-secret and paas.elb in any namespace default-secret is used to pull private images from SWR, and paas.elb is used to connect the services in the namespace to ELB.

- **Configuring Environment Variables for a Workload**
- **Configuring a Data Volume for a Workload**

The following example shows how to use a secret.

```
apiVersion: v1
kind: Secret
metadata:
  name: mysecret
type: Opaque
data:
  username: ******  #The value must be encoded using Base64.
  password: ******  #The value must be encoded using Base64.
```

NOTICE

- When a secret is used in a pod, it must be in the same cluster and namespace as the pod.
- When a secret mounted as a volume is updated, Kubernetes updates the data in the volume at the same time.

  However, when a secret mounted using **subPath** is updated, Kubernetes cannot automatically update the data in the volume.

Configuring Environment Variables for a Workload

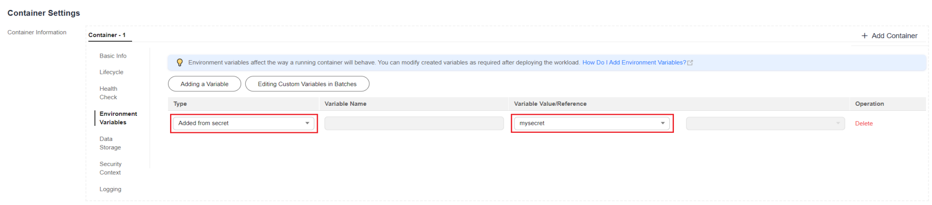
Using the console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. In the upper right corner of the displayed page, click **Create Workload**.

When creating a workload, click **Environment Variables** in the **Container Settings** area, and click **Add Variable**.

- **Added from secret**: Select a secret and import all keys in the secret as environment variables.

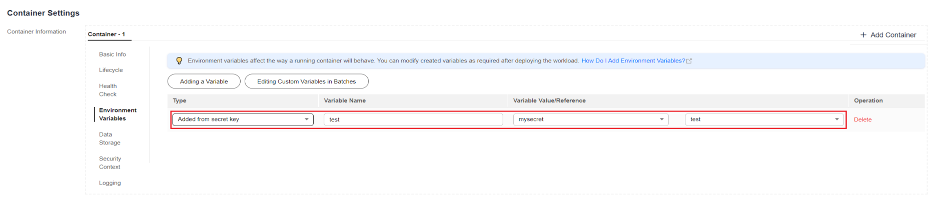


Environment variables affect the way a running container will behave. You can modify created variables as required after deploying the workload. How Do I Add Environment Variables?

Type	Variable Name	Variable Value/Reference		Operation
Added from secret		mysecret		Delete

- **Added from secret key**: Import the key in a secret as the value of an environment variable.
  - **Variable Name**: Name of the environment variable. The name can be customized and is set to the key name selected in the secret by default.
  - **Variable Value/Reference**: Select a secret and the key to be imported. The corresponding value is imported as a workload environment variable.

  For example, after you import the value of **username** in secret **mysecret** as the value of workload environment variable **username**, an environment variable named **username** exists in the container.



Environment variables affect the way a running container will behave. You can modify created variables as required after deploying the workload. How Do I Add Environment Variables?

Type	Variable Name	Variable Value/Reference		Operation
Added from secret key	test	mysecret	test	Delete

Step 3 Configure other workload parameters and click **Create Workload**.

When the workload runs normally, log in to the container and run the following statement to check whether the secret has been set as an environment variable for the workload:

```
printenv username
```

If the output is the same as the content in the secret, the secret has been set as an environment variable for the workload.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-secret.yaml** and edit it.

vi nginx-secret.yaml

The content of the YAML file is as follows:

- **Added from secret**: To add all data in a secret to environment variables, use the **envFrom** parameter. The keys in the secret will become names of environment variables in the workload.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-secret
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-secret
  template:
    metadata:
      labels:
        app: nginx-secret
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        envFrom:                 # Use envFrom to specify a secret to be referenced by environment
variables.
        - secretRef:
            name: mysecret       # Name of the referenced secret.
      imagePullSecrets:
      - name: default-secret
```

- **Added from secret key**: When creating a workload, you can use a secret to set environment variables and use the **valueFrom** parameter to reference the key-value pair in the secret separately.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-secret
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-secret
  template:
    metadata:
      labels:
        app: nginx-secret
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        env:                             # Set an environment variable for the workload.
        - name: SECRET_USERNAME          # Name of the environment variable for the workload.
          valueFrom:                     # Use valueFrom to specify a secret to be referenced by environment
variables.
            secretKeyRef:
              name: mysecret       # Name of the referenced secret.
              key: username        # Key in the referenced secret.
        - name: SECRET_PASSWORD          # Add multiple environment variables to import them at
the same time.
          valueFrom:
            secretKeyRef:
              name: mysecret
              key: password
      imagePullSecrets:
      - name: default-secret
```

Step 3 Create a workload.

kubectl apply -f nginx-secret.yaml

Step 4 View the environment variables in the pod.

1. Run the following command to view the created pod:

```
kubectl get pod | grep nginx-secret
```

Expected output:

```
nginx-secret-***   1/1     Running   0              2m18s
```

2. Run the following command to view the environment variables in the pod:

```
kubectl exec nginx-secret-*** -- printenv SPECIAL_USERNAME SPECIAL_PASSWORD
```

If the output is the same as the content in the secret, the secret has been set as an environment variable for the workload.

----End

Configuring a Data Volume for a Workload

You can mount a secret as a volume to the specified container path. The content of a secret is user-defined. You need to create a secret in advance. For details, see **Creating a Secret**.

Using the console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Workloads**. Then click the **Deployments** tab. Click **Create Workload** in the upper right corner.

When creating a workload, click **Data Storage** in the **Container Settings** area. Click **Add Volume** and select **Secret** from the drop-down list.

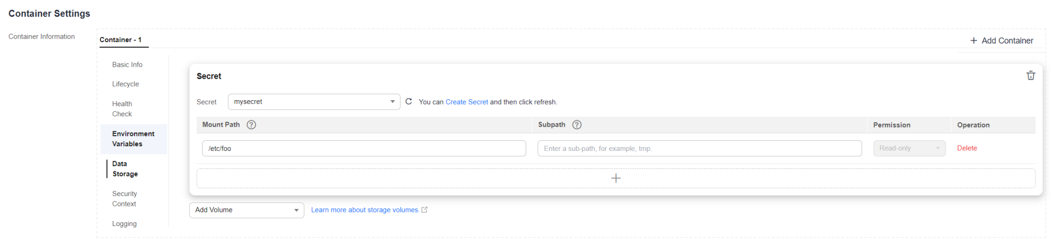
Step 3 Configure parameters for mounting a secret volume based on **Table 7-7**.

Table 7-7 Parameters for mounting a secret volume

Parameter	Description
Secret	Select the desired secret. A secret must be created beforehand. For details, see **Creating a Secret**.
Mount Path	Enter a mount point. After the secret volume is mounted, a secret file with the key as the file name and value as the file content is generated in the mount path of the container. This parameter indicates the container path that the volume will be mounted to. Do not mount the volume to a system directory such as **/** or **/var/run**. This may cause container errors. Mount the volume to an empty directory. If the directory is not empty, ensure that there are no files that affect container startup. If there are such files, they will be replaced, which will lead to a container startup or workload creation failure. **NOTICE** If a volume is mounted to a high-risk directory, use an account with minimum permissions to start the container, or high-risk files on the host may be damaged.

Parameter	Description
Subpath	Enter a subpath of the mount path. ● A subpath is used to mount a local volume so that the same volume is used in a single pod. If this parameter is left blank, the root path is used by default. ● The subpath can be the key and value of a secret. If the subpath is a key-value pair that does not exist, the data import does not take effect. ● The data imported using a subpath will not be updated along with the secret.
Permission	Read-only permission. The volume in the path is read-only.

Figure 7-2 Mounting a secret volume


Container Settings

Container Information

Container - 1 | + Add Container

Basic Info | Lifecycle | Health Check | Environment Variables | Data Storage | Security Context | Logging

Secret

Secret: mysecret — You can Create Secret and then click refresh.

Mount Path	Subpath	Permission	Operation
/etc/foo	Enter a sub-path, for example, tmp.	Read-only	Delete

Add Volume — Learn more about storage volumes.

Step 4 Configure other parameters and click **Create Workload**.

When the workload runs normally, the **username** and **password** files will be generated in the **/etc/foo** directory in this example. The file content is the secret values.

Access the container and run the following statement to view the **username** or **password** file in the container:

```
cat /etc/foo/username
```

The expected output is the same as the content in the secret.

----End

Using kubectl

Step 1 Use kubectl to connect to the cluster. For details, see **Connecting to a Cluster Using kubectl**.

Step 2 Create a file named **nginx-secret.yaml** and edit it.

vi nginx-secret.yaml

In the following example, the username and password in the **mysecret** secret are saved in the **/etc/foo** directory as files.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-secret
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-secret
  template:
    metadata:
      labels:
        app: nginx-secret
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
       - name: foo
         mountPath: /etc/foo          # Mount the secret volume to the /etc/foo directory.
         readOnly: true
    volumes:
    - name: foo
      secret:
        secretName: mysecret      # Name of the referenced secret.
```

You can also use the **items** field to control the mapping path of secret keys. For example, store username in the **/etc/foo/my-group/my-username** directory in the container.

NOTE

- If you use the **items** field to specify the mapping path of the secret keys, the keys that are not specified will not be created as files. In the following example, if the **password** key is not specified, the file will not be created.
- If you want to use all keys in a secret, you must list all keys in the **items** field.
- All keys listed in the **items** field must exist in the corresponding secret, or the volume will not be created.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-secret
spec:
  replicas: 1
  selector:
    matchLabels:
      app: nginx-secret
  template:
    metadata:
      labels:
        app: nginx-secret
    spec:
      containers:
      - name: container-1
        image: nginx:latest
        volumeMounts:
       - name: foo
         mountPath: /etc/foo          # Mount the secret volume to the /etc/foo directory.
         readOnly: true
    volumes:
    - name: foo
      secret:
        secretName: mysecret      # Name of the referenced secret.
        items:
        - key: username      # Name of the referenced key.
          path: my-group/my-username    # Mapping path of the secret key.
```

Step 3 Create a workload.

kubectl apply -f nginx-secret.yaml

Step 4 Wait until the workload runs normally. The **username** and **password** files will be generated in the **/etc/foo** directory.

1. Run the following command to view the created pod:

   ```
   kubectl get pod | grep nginx-secret
   ```

   Expected output:

   ```
   nginx-secret-***    1/1     Running   0              2m18s
   ```

2. Run the following command to view the **username** or **password** file in the pod:

   ```
   kubectl exec nginx-secret-*** -- cat /etc/foo/username
   ```

   The expected output is the same as the content in the secret.

----End

7.5 Cluster Secrets

By default, CCE Autopilot creates the following secrets in each namespace:

- default-secret
- paas.elb

The functions of these secrets are described as follows.

default-secret

The type of **default-secret** is **kubernetes.io/dockerconfigjson**. The data is the credential for logging in to the SWR image repository and is used to pull images from SWR. To pull an image from SWR when creating a workload on CCE Autopilot, set **imagePullSecrets** to **default-secret**.

```
apiVersion: v1
kind: Pod
metadata:
  name: nginx
spec:
  containers:
  - image: nginx:alpine
    name: container-0
    resources:
      limits:
        cpu: 100m
        memory: 200Mi
      requests:
        cpu: 100m
        memory: 200Mi
  imagePullSecrets:
  - name: default-secret
```

The data of **default-secret** is updated periodically, and the current data will expire after a certain period of time. You can run the **describe** command to view the expiration time in of default-secret.

NOTICE

Use default-secret directly instead of copying the secret content to create a new one. The credential in the copied secret will expire and the image cannot be pulled.

```
$ kubectl describe secret default-secret
Name:         default-secret
```

```
Namespace:    default
Labels:       secret-generated-by=cce
Annotations:  temporary-ak-sk-expires-at: 2021-11-26 20:55:31.380909 +0000 UTC

Type:  kubernetes.io/dockerconfigjson

Data
====
.dockerconfigjson:  347 bytes
```

paas.elb

The **paas.elb** data stores a temporary AK/SK that is used when a pod is created or a load balancer is automatically created. The **paas.elb** data is updated periodically and has a specific time limit before it expires.

In practice, you will not directly use **paas.elb**. Do not delete it, as doing so will result in the failure of creating a pod or load balancer.

8 Auto Scaling

8.1 Scaling a Workload

8.1.1 How a Workload Is Scaled

How HPA Works

Horizontal Pod Autoscaler (HPA) is a controller that controls horizontal pod scaling. HPA periodically checks the pod metrics, calculates the number of replicas required to meet the target values configured for HPA resources, and then adjusts the value of the **replicas** field in the target resource object (such as a Deployment).

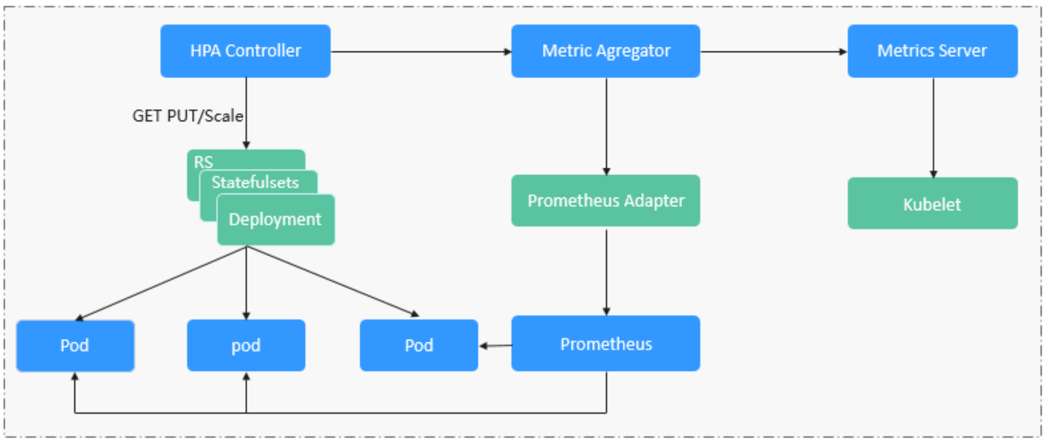
A prerequisite for auto scaling is that your container running data can be collected, such as the numbers of cluster nodes and pods, and CPU and memory usages of containers. Kubernetes does not have built-in monitoring capabilities, but you can use extensions like **Prometheus** and **Metrics Server** to monitor and collect data.

- **Prometheus** is an open-source monitoring and alarming framework that can collect multiple types of metrics. Prometheus has been a standard monitoring solution of Kubernetes.
- **Metrics Server** is a cluster-wide aggregator of resource utilization data. Metrics Server collects metrics from the Summary API exposed by kubelet. These metrics are set for core Kubernetes resources, such as pods, nodes, containers, and Services. Metrics Server provides a set of standard APIs for external systems to collect these metrics.

Horizontal Pod Autoscaler (HPA) can work with Metrics Server to implement auto scaling based on CPU and memory usages. It can also work with Prometheus to implement auto scaling based on custom metrics.

Figure 8-1 shows how HPA works.

Figure 8-1 HPA working process



Two core modules of HPA:

- Data Source Monitoring

  The community provided only CPU- and memory-based HPA at the early stage. With the population of Kubernetes and Prometheus, developers need more custom metrics or monitoring information at the access layer for their own applications, for example, the QPS of the load balancer and the number of online users of the website. In response, the community defines a set of standard metric APIs to provide services externally through these aggregated APIs.

  – **metrics.k8s.io** provides monitoring metrics related to the CPU and memory of pods and nodes.

  – **custom.metrics.k8s.io** provides custom monitoring metrics related to Kubernetes objects.

  – **external.metrics.k8s.io** provides metrics that come from external systems and are irrelevant to any Kubernetes resource metrics.

- Scaling Decision-Making Algorithms

  The HPA controller calculates the scaling ratio based on the current metric values and desired metric values using the following formula:

  desiredReplicas = ceil[currentReplicas x (currentMetricValue/ desiredMetricValue)]

  For example, if the current metric value is 200m and the target value is 100m, the desired number of pods will be doubled according to the formula. In practice, pods may be constantly added or reduced. To ensure stability, the HPA controller is optimized from the following aspects:

  – Cooldown interval: In v1.11 and earlier versions, Kubernetes introduced the startup parameters **horizontal-pod-autoscaler-downscale-stabilization-window** and **horizontal-pod-autoScaler-upscale-stabilization-window** to indicate the cooldown intervals after a scale-in and scale-out, respectively, in which no scaling operation will not be performed. In versions later than v1.14, the scheduling queue is introduced to store all decision-making suggestions detected within a period of time. Then, the system makes decisions based on all valid decision-making suggestions to minimize changes of the desired number of replicas to ensure stability.

– Tolerance: It can be considered as a buffer zone. If the pod number changes can be tolerated, the number of pods remains unchanged.

  Use the formula: ratio = currentMetricValue/desiredMetricValue

  When |ratio – 1.0| ≤ tolerance, scaling will not be performed.

  When |ratio – 1.0| > tolerance, the desired value is calculated using the formula mentioned above.

  The default value is **0.1** in the current community version.

The HPA performs scaling based on metric thresholds. Common metrics include the CPU and memory usage. You can also set custom metrics, such as the QPS and number of connections, to trigger scaling. However, metric-based scaling brings in latency of minutes generated during data collection, determination, and scaling phases. Such latency may cause high CPU usage and slow response. To solve this problem, CCE allows you to configure scheduled policies to scale resources regularly for applications with periodic changes.

8.1.2 HPA Policies

Horizontal Pod Autoscaling (HPA) in Kubernetes implements horizontal scaling of pods. In a CCE HPA policy, you can configure different cooldown time windows and scaling thresholds for different applications based on the Kubernetes HPA.

Prerequisites

The following add-on has been installed in the cluster:

- **Kubernetes Metrics Server**: provides basic resource usage metrics, such as container CPU and memory usage.
- **Creating an HPA Policy Using Custom Metrics**: Custom metrics need to be aggregated to the Kubernetes API server for auto scaling.

Creating an HPA Policy

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **Scaling Policies**, click the **HPA Policies** tab and then **Create HPA Policy** in the upper right corner.

Step 3 Configure basic information.

- **Policy Name**: Enter a name for the policy.
- **Namespace**: Select the namespace that the workload belongs to.
- **Associated Workload**: Select the workload that the policy is configured for.

Step 4 Configure other parameters.

Table 8-1 HPA policy parameters

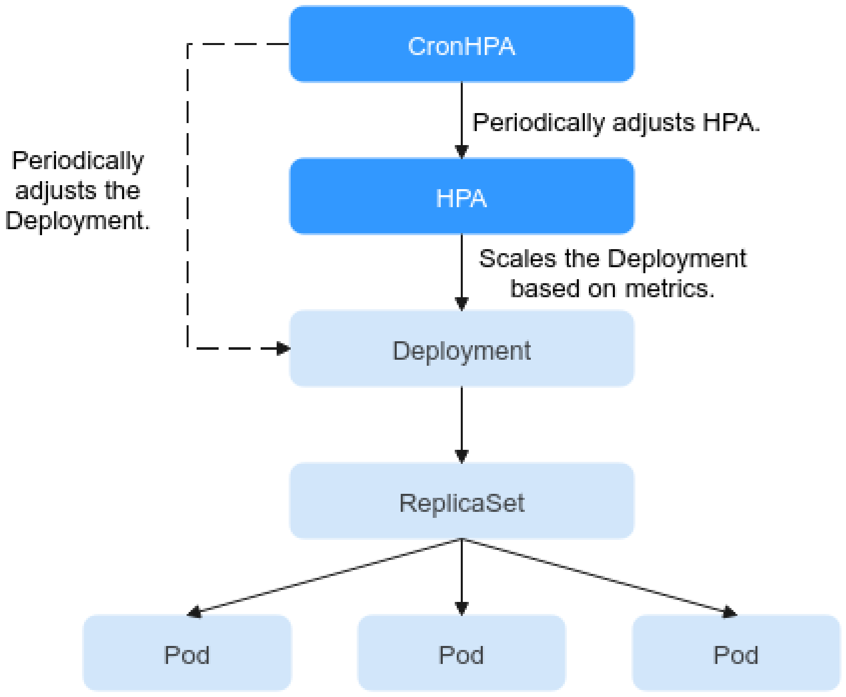
Parameter	Description
Pod Range	Minimum and maximum numbers of pods. When a policy is triggered, the pods are scaled within this range. **NOTICE** In CCE Autopilot clusters, if you use a dedicated load balancer for a workload, the number of pods cannot exceed the load balancer's backend server group quota, which is 500 by default. If this limit is exceeded, no more pods can be added as the load balancer backend.
Scaling Behavior	● **Default**: Scale workloads using the Kubernetes default behavior. For details, see **Default Behavior**. ● **Custom**: Scale workloads using custom policies such as stabilization window, steps, and priorities. Unspecified parameters use the values recommended by Kubernetes. – **Disable scale-out/scale-in**: Select whether to disable scale-out or scale-in. – **Stabilization Window**: a period during which CCE continuously checks whether the metrics used for scaling keep fluctuating. CCE triggers scaling if the desired state is not maintained for the entire window. This window restricts the unwanted flapping of pod count due to metric changes. – **Step**: specifies the scaling step. You can set the number or percentage of pods to be scaled in or out within a specified period. If there are multiple policies, you can select the policy that maximizes or minimizes the number of pods.
System Policy	● **Metric**: You can select **CPU usage** or **Memory usage**. **NOTE** Usage = CPUs or memory used by pods/Requested CPUs or memory ● **Desired Value**: Enter the desired average resource usage. This parameter indicates the desired value of the selected metric. Number of pods to be scaled (rounded up) = (Current metric value/Desired value) x Number of current pods **NOTE** When calculating the number of pods to be added or reduced, the HPA policy uses the maximum number of pods in the last 5 minutes. ● **Tolerance Range**: Scaling is not triggered when the metric value is within the tolerance range. The desired value must be within the tolerance range. If the metric value is greater than the scale-in threshold and less than the scale-out threshold, no scaling is triggered. **This parameter is only supported in clusters v1.15 or later.**

Step 5 Click **Create**.

----End

8.1.3 CronHPA Policies

There are predictable and unpredictable traffic peaks for some services. For such services, CCE CronHPA allows you to scale resources in fixed periods. It can work with HPA policies to periodically adjust the HPA scaling scope, implementing workload scaling.


CronHPA
Periodically adjusts HPA.
HPA
Scales the Deployment based on metrics.
Periodically adjusts the Deployment.
Deployment
ReplicaSet
Pod
Pod
Pod

CronHPA can periodically adjust the maximum and minimum numbers of pods in the HPA policy or directly adjust the number of pods of a Deployment.

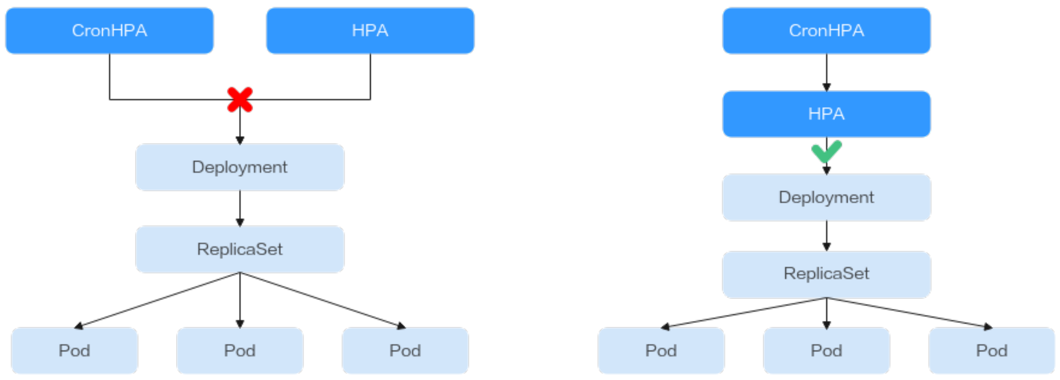
Prerequisites

The **CCE Advanced HPA** add-on has been installed in the cluster.

Using CronHPA to Adjust the HPA Scaling Scope

CronHPA can periodically scale out/in pods in HPA policies to satisfy complex services.

HPA and CronHPA associate scaling objects using the **scaleTargetRef** field. If a Deployment is the scaling object for both CronHPA and HPA, the two scaling policies are independent of each other. The operation performed later overwrites the operation performed earlier. As a result, the scaling effect does not meet the expectation.



When CronHPA and HPA are used together, CronHPA rules take effect based on the HPA policy. CronHPA uses HPA to perform operations on the Deployment. Understanding the following parameters can better understand the working rules of the CronHPA.

- **targetReplicas**: Number of pods set for CronHPA. When CronHPA takes effect, this parameter adjusts the maximum or minimum number of pods in HPA policies to adjust the number of Deployment pods.
- **minReplicas**: Minimum number of Deployment pods.
- **maxReplicas**: Maximum number of Deployment pods.
- **replicas**: Number of pods in a Deployment before the CronHPA policy takes effect.

When the CronHPA rule takes effect, the maximum or minimum number of pods are adjusted by comparing the number of **targetReplicas** with the actual number of pods and combining the minimum or maximum number of pods in the HPA policy.

Figure 8-2 CronHPA scaling scenarios

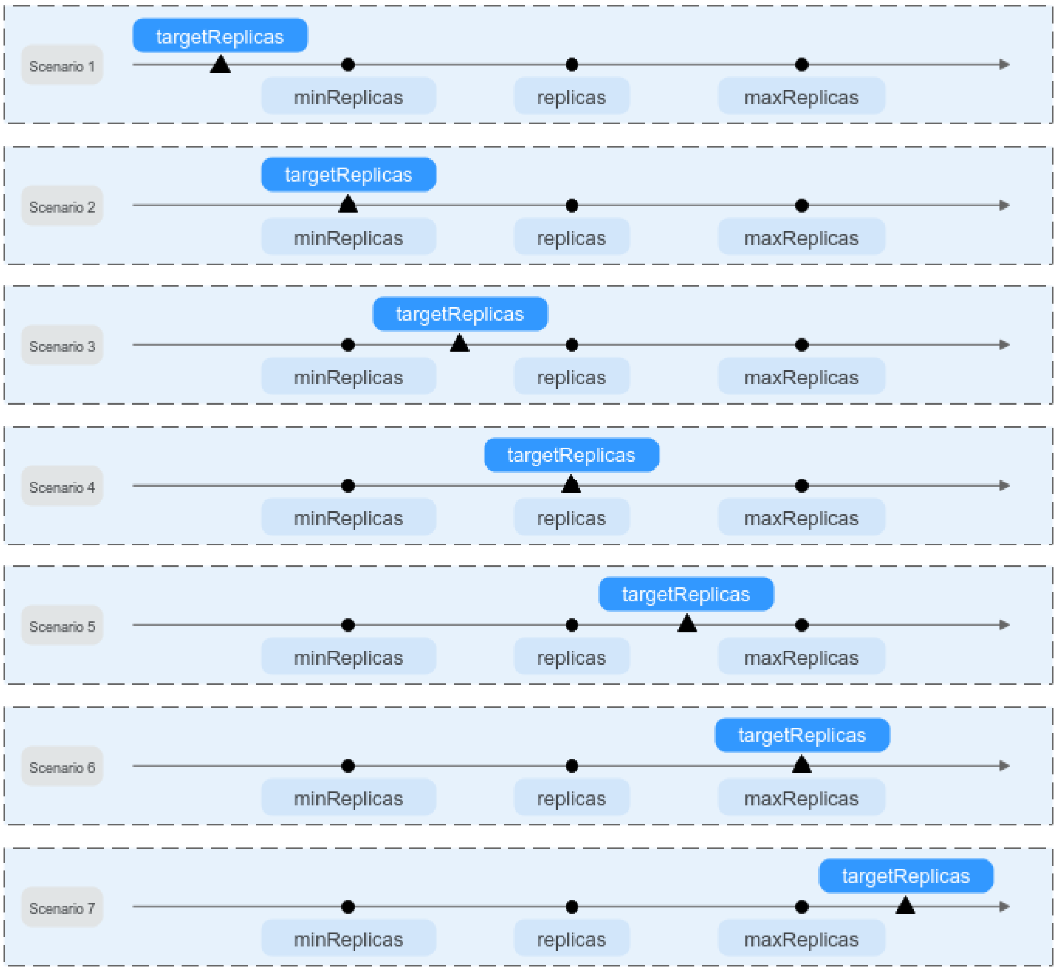


Figure 8-2 shows possible scaling scenarios. The following examples detail how CronHPA modifies the number of pods in HPAs.

Table 8-2 CronHPA scaling parameters

Scenario	Scenario Description	CronHPA (targetReplicas)	Deployment (replicas)	HPA (minReplicas / maxReplicas)	Result	Operation Description
1	**targetReplicas** < minReplicas ≤ replicas ≤ maxReplicas	4	5	5/10	HPA: 4/10 Deployments: 5	When the value of **targetReplicas** is smaller than that of **minReplicas**: ● Change the value of **minReplicas**. ● The value of **replicas** requires no change.
2	**targetReplicas** = minReplicas ≤ replicas ≤ maxReplicas	5	6	5/10	HPA: 5/10 Deployments: 6	When the value of **targetReplicas** is equal to that of **minReplicas**: ● The value of **minReplicas** requires no change. ● The value of **replicas** requires no change.
3	minReplicas < **targetReplicas** < replicas ≤ maxReplicas	4	5	1/10	HPA: 4/10 Deployments: 5	When the value of **targetReplicas** is greater than that of **minReplicas** and smaller than that of **replicates**: ● Change the value of **minReplicas**. ● The value of **replicas** requires no change.

Scenario	Scenario Description	CronHPA (targetReplicas)	Deployment (replicas)	HPA (minReplicas/maxReplicas)	Result	Operation Description
4	minReplicas < **targetReplicas** = replicas < maxReplicas	5	5	1/10	HPA: 5/10 Deployments: 5	When the value of **targetReplicas** is greater than that of **minReplicas** and equal to that of **replicates**: ● Change the value of **minReplicas**. ● The value of **replicas** requires no change.
5	minReplicas ≤ replicas < **targetReplicas** < maxReplicas	6	5	1/10	HPA: 6/10 Deployments: 6	When the value of **targetReplicas** is greater than that of **replicates** and less than that of **maxReplicas**: ● Change the value of **minReplicas**. ● Change the value of **replicas**.
6	minReplicas ≤ replicas < **targetReplicas** = maxReplicas	10	5	1/10	HPA: 10/10 Deployments: 10	When the value of **targetReplicas** is greater than that of **replicates** and equal to that of **maxReplicas**: ● Change the value of **minReplicas**. ● Change the value of **replicas**.

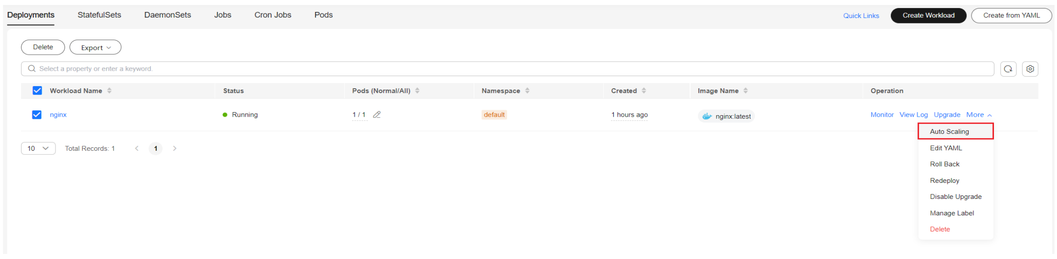
Scenario	Scenario Description	CronHPA (targetReplicas)	Deployment (replicas)	HPA (minReplicas/maxReplicas)	Result	Operation Description
7	minReplicas ≤ replicas ≤ maxReplicas < **targetReplicas**	11	5	5/10	HPA: 11/11 Deployments: 11	When the value of **targetReplicas** is greater than that of **maxReplicas**: ● Change the value of **minReplicas**. ● Change the value of **maxReplicas**. ● Change the value of **replicas**.

Using the CCE console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Choose **Workloads** in the navigation pane. Locate the target workload and choose **More** > **Auto Scaling** in the **Operation** column.

Figure 8-3 Scaling a workload



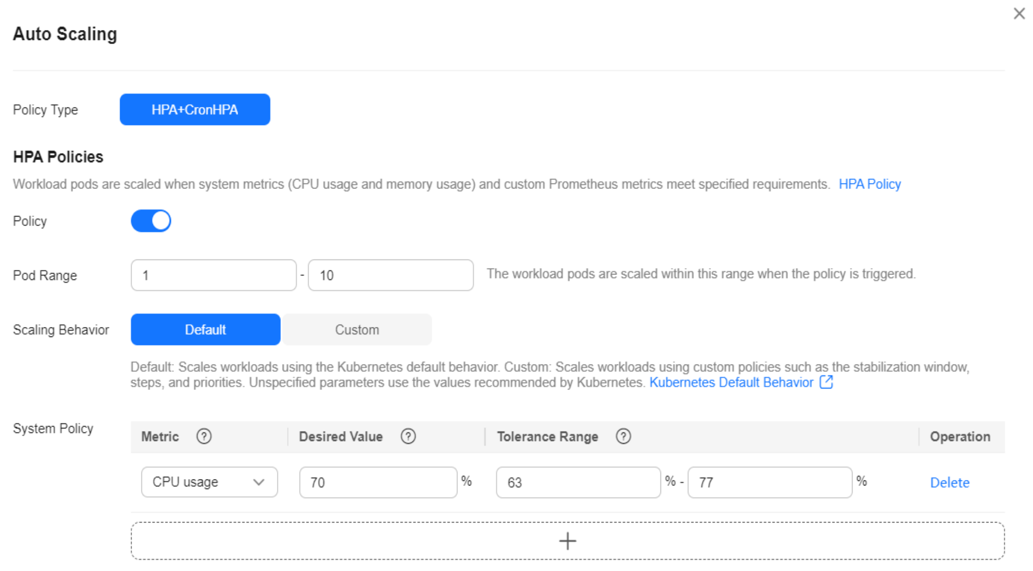
Workload Name	Status	Pods (Normal/All)	Namespace	Created	Image Name	Operation
nginx	Running	1 / 1	default	1 hours ago	nginx:latest	Monitor View Log Upgrade More

Step 3 Set **Policy Type** to **HPA+CronHPA** and enable HPA and CronHPA policies.

CronHPA periodically adjusts the maximum and minimum numbers of pods using the HPA policy.

Step 4 Configure the HPA policy. For details, see **HPA Policies**.

Figure 8-4 Enabling the HPA policy


Auto Scaling

Policy Type: HPA+CronHPA

HPA Policies

Workload pods are scaled when system metrics (CPU usage and memory usage) and custom Prometheus metrics meet specified requirements. HPA Policy

Policy: enabled

Pod Range: 1 - 10 The workload pods are scaled within this range when the policy is triggered.

Scaling Behavior: Default | Custom

Default: Scales workloads using the Kubernetes default behavior. Custom: Scales workloads using custom policies such as the stabilization window, steps, and priorities. Unspecified parameters use the values recommended by Kubernetes. Kubernetes Default Behavior

System Policy

Metric	Desired Value	Tolerance Range	Operation
CPU usage	70 %	63 % - 77 %	Delete

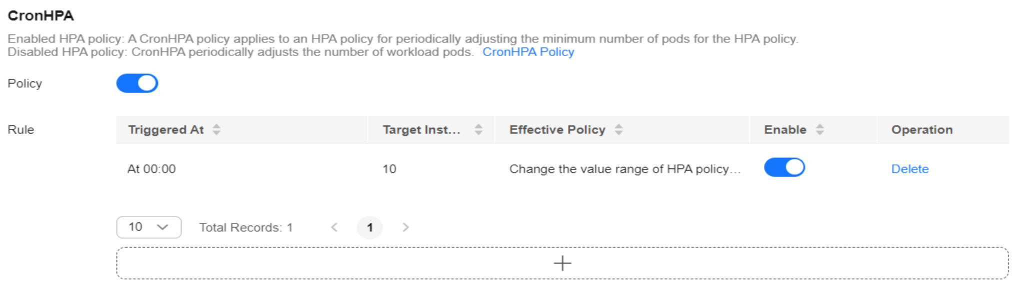
Table 8-3 HPA policy parameters

Parameter	Description
Pod Range	Minimum and maximum numbers of pods. When a policy is triggered, the pods are scaled within this range.
Scaling Behavior	● **Default**: Scale workloads using the Kubernetes default behavior. For details, see **Default Behavior**. ● **Custom**: Scale workloads using custom policies such as stabilization window, steps, and priorities. Unspecified parameters use the values recommended by Kubernetes. – **Disable scale-out/scale-in**: Select whether to disable scale-out or scale-in. – **Stabilization Window**: a period during which CCE continuously checks whether the metrics used for scaling keep fluctuating. CCE triggers scaling if the desired state is not maintained for the entire window. This window restricts the unwanted flapping of pod count due to metric changes. – **Step**: specifies the scaling step. You can set the number or percentage of pods to be scaled in or out within a specified period. If there are multiple policies, you can select the policy that maximizes or minimizes the number of pods.

Parameter	Description
System Policy	● **Metric**: You can select **CPU usage** or **Memory usage**. **NOTE** Usage = CPUs or memory used by pods/Requested CPUs or memory ● **Desired Value**: Enter the desired average resource usage. This parameter indicates the desired value of the selected metric. Number of pods to be scaled (rounded up) = (Current metric value/Desired value) x Number of current pods **NOTE** When calculating the number of pods to be added or reduced, the HPA policy uses the maximum number of pods in the last 5 minutes. ● **Tolerance Range**: Scaling is not triggered when the metric value is within the tolerance range. The desired value must be within the tolerance range. If the metric value is greater than the scale-in threshold and less than the scale-out threshold, no scaling is triggered. **This parameter is only supported in clusters v1.15 or later.**

Step 5 Click + in the CronHPA policy rule. In the dialog box displayed, configure scaling policy parameters.

Figure 8-5 Enabling the CronHPA policy


CronHPA
Enabled HPA policy: A CronHPA policy applies to an HPA policy for periodically adjusting the minimum number of pods for the HPA policy.
Disabled HPA policy: CronHPA periodically adjusts the number of workload pods. CronHPA Policy

Policy

Rule

Triggered At	Target Inst...	Effective Policy	Enable	Operation
At 00:00	10	Change the value range of HPA policy...		Delete

Table 8-4 CronHPA policy parameters

Parameter	Description
Target Instances	When the policy is triggered, CCE will adjust the number of HPA policy pods based on service requirements. For details, see **Table 8-2**.
Trigger Time	You can select a specific time every day, every week, every month, or every year. **NOTE** This time indicates the local time of where the node is deployed.
Enable	Enable or disable the policy rule.

Step 6 After configuring the preceding parameters, click **OK**. Then, the added policy rule is displayed in the rule list. Repeat the preceding steps to add multiple policy rules, but the triggering time of the policies must be different.

Step 7 Click **Create**.

----End

Using the kubectl command

When the CronHPA is compatible with the HPA policy, the **scaleTargetRef** field in CronHPA must be set to the HPA policy, and the **scaleTargetRef** field in the HPA policy must be set to Deployment. In this way, CronHPA adjusts the maximum and minimum numbers of pods in the HPA policy at a fixed time and the scheduled scaling is compatible with the auto scaling.

Step 1 Create an HPA policy for the Deployment.

```
apiVersion: autoscaling/v1
kind: HorizontalPodAutoscaler
metadata:
  name: hpa-test
  namespace: default
spec:
   maxReplicas: 10              #  Maximum number of pods
  minReplicas: 5               #  Minimum number of pods
  scaleTargetRef:              #  Associate a Deployment.
    apiVersion: apps/v1
    kind: Deployment
    name: nginx
  targetCPUUtilizationPercentage: 50
```

Step 2 Create a CronHPA policy and associate it with the HPA policy created in **Step 1**.

```
apiVersion: autoscaling.cce.io/v2alpha1
kind: CronHorizontalPodAutoscaler
metadata:
  name: ccetest
  namespace: default
spec:
   scaleTargetRef:                  # Associate an HPA policy.
      apiVersion: autoscaling/v1
      kind: HorizontalPodAutoscaler
      name: hpa-test
   rules:
   - ruleName: "scale-down"
     schedule: "15 * * * *"          # Running time and period of a job. For details, see Cron, for example, 0 * * *
* or @hourly.
     targetReplicas: 1               # Number of target pods
     disable: false
   - ruleName: "scale-up"
     schedule: "13 * * * *"
     targetReplicas: 11
     disable: false
```

Table 8-5 Key fields of CronHPA

Field	Description
apiVersion	API version. The value is fixed at **autoscaling.cce.io/v2alpha1**.
kind	API type. The value is fixed at **CronHorizontalPodAutoscaler**.

Field	Description
metadata.name	Name of a CronHPA policy.
metadata.namespace	Namespace to which the CronHPA policy belongs.
spec.scaleTargetRef	Specifies the scaling object of CronHPA. The following fields can be configured: ● **apiVersion**: API version of the CronHPA scaling object. ● **kind**: API type of the CronHPA scaling object. ● **name**: Name of the CronHPA scaling object. CronHPA supports HPA policies or Deployments. For details, see **Using CronHPA to Adjust the HPA Scaling Scope** or **Using CronHPA to Directly Adjust the Number of Deployment Pods**.
spec.rules	CronHPA policy rule. Multiple rules can be added. The following fields can be configured for each rule: ● **ruleName**: CronHPA rule name, which must be unique. ● **schedule**: Running time and period of a job. For details, see **Cron**, for example, 0 * * * * or @hourly. **NOTE** This time indicates the local time of where the node is deployed. ● **targetReplicas**: indicates the number of pods to be scaled in or out. ● **disable**: The value can be **true** or **false**. **false** indicates that the rule takes effect, and **true** indicates that the rule does not take effect.

----End

Using CronHPA to Directly Adjust the Number of Deployment Pods

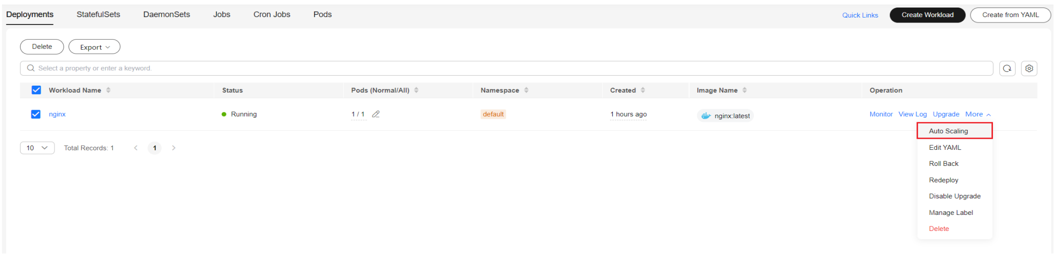
CronHPA adjusts the associated Deployments separately to periodically adjust the number of Deployment pods. The method is as follows:

Using the CCE console

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 Choose **Workloads** in the navigation pane. Locate the target workload and choose **More** > **Auto Scaling** in the **Operation** column.

Figure 8-6 Scaling a workload



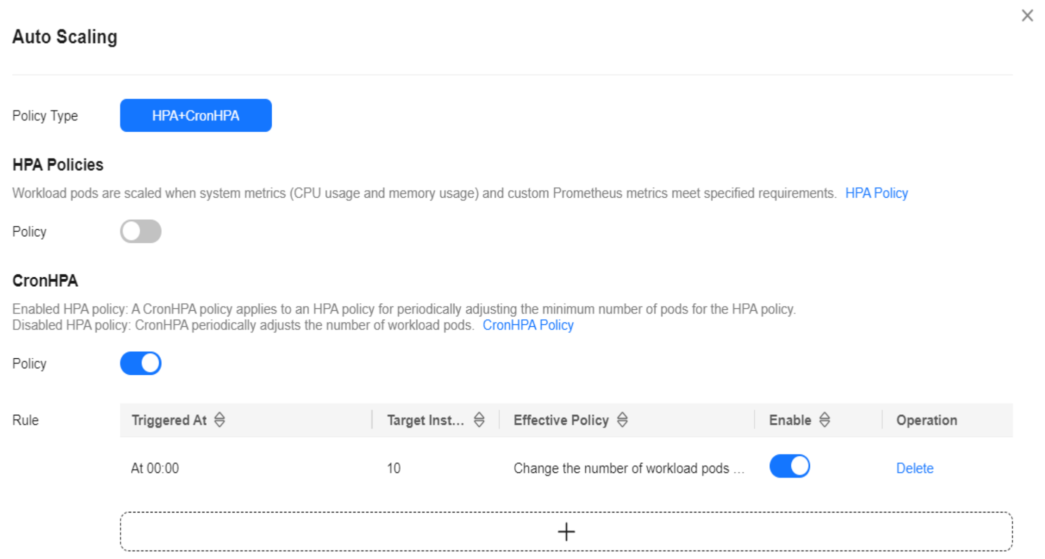
Workload Name	Status	Pods (Normal/All)	Namespace	Created	Image Name	Operation
nginx	Running	1 / 1	default	1 hours ago	nginx:latest	Monitor View Log Upgrade More

Step 3 Set **Policy Type** to **HPA+CronHPA**, disable HPA, and enable CronHPA.

CronHPA periodically adjusts the number of workload pods.

Step 4 Click + in the CronHPA policy rule. In the dialog box displayed, configure scaling policy parameters.

Figure 8-7 Using CronHPA to adjust the number of workload pods



Policy Type: HPA+CronHPA

Workload pods are scaled when system metrics (CPU usage and memory usage) and custom Prometheus metrics meet specified requirements.

Policy: disabled

Enabled HPA policy: A CronHPA policy applies to an HPA policy for periodically adjusting the minimum number of pods for the HPA policy.
Disabled HPA policy: CronHPA periodically adjusts the number of workload pods.

Policy: enabled

Rule

Triggered At	Target Inst...	Effective Policy	Enable	Operation
At 00:00	10	Change the number of workload pods ...	●	Delete

Table 8-6 CronHPA policy parameters

Parameter	Description
Target Instances	When a policy is triggered, the number of workload pods will be adjusted to the value of this parameter.
Trigger Time	You can select a specific time every day, every week, every month, or every year. **NOTE** This time indicates the local time of where the node is deployed.
Enable	Enable or disable the policy rule.

Step 5 After configuring the preceding parameters, click **OK**. Then, the added policy rule is displayed in the rule list. Repeat the preceding steps to add multiple policy rules, but the triggering time of the policies must be different.

Step 6 Click **Create**.

----End

Using kubectl commands

```
apiVersion: autoscaling.cce.io/v2alpha1
kind: CronHorizontalPodAutoscaler
metadata:
  name: ccetest
  namespace: default
spec:
   scaleTargetRef:            # Associate a Deployment.
      apiVersion: apps/v1
      kind: Deployment
      name: nginx
   rules:
   - ruleName: "scale-down"
     schedule: "08 * * * *"     # Running time and period of a job. For details, see Cron, for example, 0 * * * * or
@hourly.
     targetReplicas: 1
     disable: false
   - ruleName: "scale-up"
     schedule: "05 * * * *"
     targetReplicas: 3
     disable: false
```

8.1.4 Managing Workload Scaling Policies

Scenario

You can **view an HPA policy**, **edit an HPA policy on the console**, **edit an HPA policy using YAML**, **clone an HPA policy**, and **delete an HPA policy**.

Viewing an HPA Policy

You can view the rules, status, and events of an HPA policy and handle exceptions based on the error information displayed.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **HPA Policies** tab, locate the HPA policy you want to view and click ⌄ next to the policy.

Step 3 In the expanded area, choose **View Events** in the **Operation** column. If the policy malfunctions, locate and rectify the fault based on the error message displayed on the page.

NOTE

You can also view an HPA policy on the workload details page.

1. Log in to the CCE console and click the cluster name to access the cluster console.
2. In the navigation pane on the left, choose **Workloads**. Click the workload name to view its details.
3. On the workload details page, switch to the **Policies** tab to view the HPA policy. You can also view the policies you configured in **Workload Scaling**.

Table 8-7 Event types and names

Event Type	Event Name	Description
Normal	SuccessfulRescale	The scaling is performed successfully.
Abnormal	InvalidTargetRange	Invalid target range.
	InvalidSelector	Invalid selector.
	FailedGetObjectMetric	Objects fail to be obtained.
	FailedGetPodsMetric	Pods fail to be obtained.
	FailedGetResourceMetric	Resources fail to be obtained.
	FailedGetExternalMetric	External metrics fail to be obtained.
	InvalidMetricSourceType	Invalid metric source type.
	FailedConvertHPA	HPA conversion failed.
	FailedGetScale	The scale fails to be obtained.
	FailedComputeMetricsReplicas	Failed to calculate metric-defined replicas.
	FailedGetScaleWindow	Failed to obtain ScaleWindow.
	FailedRescale	Failed to scale the workload.

----End

Editing an HPA Policy on the Console

To edit an HPA policy on the console, take the following steps:

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **HPA Policies** tab, locate the HPA policy you want to edit and choose **More** > **Edit** in the **Operation** column.

Step 3 On the **Edit HPA Policy** page, modify the parameter settings based on **Table 8-1**.

Step 4 Click **OK**.

----End

Editing an HPA Policy Using YAML

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **HPA Policies** tab, locate the HPA policy you want to edit and click **Edit YAML** in the **Operation** column.

Step 3 In the **Edit YAML** dialog box displayed, edit or download the YAML file.

----End

Cloning an HPA Policy

Copy the configuration of an existing HPA policy to create a new one, and associate it with a different namespace or workload.

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **HPA Policies** tab, locate the HPA policy you want to edit and choose **More** > **Clone** in the **Operation** column.

Step 3 On the **Create HPA Policy** page, set basic policy information by referring to **Step 3**.

Step 4 Click **OK**.

----End

Deleting an HPA Policy

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 In the navigation pane on the left, choose **Policies**. On the **HPA Policies** tab, locate the HPA policy you want to edit and choose **More** > **Delete** in the **Operation** column.

Step 3 In the dialog box displayed, click **Yes**.

----End

9 Add-ons

9.1 CoreDNS

Introduction

CoreDNS is a DNS server that provides domain name resolution for Kubernetes clusters through a chain add-on.

CoreDNS is an open-source software and has been a part of CNCF. It provides a means for cloud services to discover each other in cloud-native deployments. Each of the plugins chained by CoreDNS provides a particular DNS function. You can integrate CoreDNS with only the plugins you need to make it fast, efficient, and flexible. When used in a Kubernetes cluster, CoreDNS can automatically discover services in the cluster and provide domain name resolution for these services. By working with DNS server, CoreDNS can resolve external domain names for workloads in a cluster.

This add-on is installed by default during cluster creation.

Kubernetes backs CoreDNS as the official default DNS for all clusters going forward.

CoreDNS official website: **https://coredns.io/**

Open-source community: **https://github.com/coredns/coredns**

NOTE

For more information about CoreDNS, see **DNS**.

Constraints

To run CoreDNS or upgrade CoreDNS in a cluster, ensure the number of available nodes in the cluster is greater than or equal to the number of CoreDNS instances and all CoreDNS instances are running. Otherwise, the add-on will malfunction or the upgrade will fail.

Editing the Add-on

This add-on is installed by default. To modify its settings, perform the following steps:

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **CoreDNS** on the right and click **Edit**.

Step 2 On the **Edit Add-on** page, modify the specifications.

Table 9-1 CoreDNS specifications

Parameter	Description
Pods	Number of pods for the add-on. High availability is not possible if there is only one pod. If an error occurs on the node where the pod runs, the add-on will fail.
Containers	CPU and memory quotas of the containers for running the add-on. Queries per second (QPS) of the add-on is positively correlated with the CPU consumption. If the number of nodes or containers in the cluster grows, the CoreDNS pods will bear heavier workloads.

Step 3 Modify the parameters.

Table 9-2 CoreDNS add-on parameters

Parameter	Description
Stub Domain	A domain name server for a custom domain name. The format is a key-value pair. The key is a domain name suffix, and the value is one or more DNS IP addresses, for example, **acme.local -- 1.2.3.4,6.7.8.9**. For details, see **Configuring the Stub Domain for CoreDNS**.

Parameter	Description

(Optional) Extended Parameter Settings	● **parameterSyncStrategy**: indicates whether to configure consistency check when the add-on is upgraded.
	– **ensureConsistent**: indicates that the configuration consistency check is enabled. If the delivered configuration differs from the current effective configuration, the current effective configuration will be replaced. However, if the delivered configuration is the same as the current effective configuration, the current effective configuration will be preserved. The **ensureConsistent** parameter ensures the configuration consistency. If a ConfigMap is modified manually, the add-on cannot be upgraded. In such cases, you will need to use the **force** or **inherit** policy to upgrade the add-on.
	– **force**: indicates that the configuration consistency check is ignored during an upgrade. In this case, you must ensure that the current effective configuration is the same as the original configuration. After the add-on is upgraded, restore the value of **parameterSyncStrategy** to **ensureConsistent** to enable the configuration consistency check again.
	– **inherit**: indicates that custom settings are automatically inherited during an upgrade. After the add-on is upgraded, restore the value of **parameterSyncStrategy** to **ensureConsistent** to enable the configuration consistency check again.
	● **stub_domains**: indicates the subdomains, which allow you to configure a domain name server for a custom domain name. A subdomain is in the format of a key-value pair, where the key is the suffix of a DNS domain name and the value is one or more DNS IP addresses.
	● **upstream_nameservers**: indicates the IP address of the upstream DNS server.
	● **servers**: indicates the name servers, which are available in CoreDNS v1.23.1 and later versions. You can customize nameservers. For details, see **dns-custom-nameservers**. **plugins** indicates the configuration of each component in CoreDNS. Retain the default settings typically to prevent CoreDNS from being unavailable due to configuration errors. Each plugin component contains **name**, **parameters** (optional), and **configBlock** (optional). The format of the generated Corefile is as follows:
	`$name  $parameters {`
	`$configBlock`
	`}`
	Table 9-3 describes common plugins. For details, see **Plugins**.
	Example:
	`{`
	`    "servers": [`
	`        {`

Parameter	Description
	```
	            "plugins": [
	                {
	                    "name": "bind",
	                    "parameters": "{$POD_IP}"
	                },
	                {
	                    "name": "cache",
	                    "parameters": 30
	                },
	                {
	                    "name": "errors"
	                },
	                {
	                    "name": "health",
	                    "parameters": "{$POD_IP}:8080"
	                },
	                                {
	                    "name": "ready",
	                    "{$POD_IP}:8081"
	                },
	                {
	                    "configBlock": "pods insecure\nfallthrough in-addr.arpa ip6.arpa",
	                    "name": "kubernetes",
	                    "parameters": "cluster.local in-addr.arpa ip6.arpa"
	                },
	                {
	                    "name": "loadbalance",
	                    "parameters": "round_robin"
	                },
	                {
	                    "name": "prometheus",
	                    "parameters": "{$POD_IP}:9153"
	                },
	                {
	                    "configBlock": "policy random",
	                    "name": "forward",
	                    "parameters": ". /etc/resolv.conf"
	                },
	                {
	                    "name": "reload"
	                }
	            ],
	            "port": 5353,
	            "zones": [
	                {
	                    "zone": "."
	                }
	            ]
	        }
	    ],
	    "stub_domains": {
	        "acme.local": [
	            "1.2.3.4",
	            "6.7.8.9"
	        ]
	    },
	    "upstream_nameservers": ["8.8.8.8", "8.8.4.4"]
	}
	```

Table 9-3 Default plugin configuration of the active CoreDNS zone

Plugin Name	Description
bind	Host IP address listened by CoreDNS. Retain the default value **{$POD_IP}**. For details, see **bind**.
cache	Enables DNS cache. For details, see **cache**.
errors	Errors are logged to stdout. For details, see **errors**.
health	Health check for CoreDNS. {$POD_IP}:8080 is listened to. Retain the default setting. Otherwise, the CoreDNS health check will fail and the add-on will restart repeatedly. For details, see **health**.
ready	Whether the backend server is ready to receive traffic. {$POD_IP}:8081 is listened to. If the backend server is not ready, CoreDNS will suspend DNS resolution until the backend server is ready. For details, see **ready**.
kubernetes	CoreDNS Kubernetes plugin, which provides the service parsing capability in a cluster. For details, see **kubernetes**.
loadbalance	Round-robin DNS load balancer that randomizes the order of A, AAAA, and MX records in an answer. For details, see **loadbalance**.
prometheus	API for obtaining CoreDNS metrics. {$POD_IP}:9153 is listened to in the default zone. Retain the default setting. Otherwise, Prometheus cannot collect CoreDNS metrics. For details, see **Prometheus**.
forward	Forwards any queries that are not within the cluster domain of Kubernetes to predefined resolvers (**/etc/resolv.conf**). For details, see **forward**.
reload	Automatically reloads modified Corefiles. After you modify a ConfigMap, wait for two minutes for the modification to take effect. For details, see **reload**.

log	Enables CoreDNS logging. For details, see **log**.
	Example:
	`{`
	`  "name": "log"`
	`}`

template	A quick response template, where **AAAA** indicates an IPv6 request. If **NXDOMAIN** is returned in an **rcode** response, no IPv6 resolution result is returned. For details, see **Template**.
	Example:
	`{`
	`  "configBlock": "rcode NXDOMAIN",`
	`  "name": "template",`
	`  "parameters": "ANY AAAA"`
	`}`

Step 4 Click **Install**.

----End

Components

Table 9-4 CoreDNS components

Component	Description	Resource Type
CoreDNS	DNS server for clusters	Deployment

How Does Domain Name Resolution Work in Kubernetes?

DNS policies can be configured for each pod. Kubernetes supports DNS policies **Default**, **ClusterFirst**, **ClusterFirstWithHostNet**, and **None**. For details, see **DNS for Services and Pods**. These policies are specified in the **dnsPolicy** field in the pod-specific.

- **Default**: Pods inherit the name resolution configuration from the node that the pods run on. The custom upstream DNS server and the stub domain cannot be used together with this policy.
- **ClusterFirst**: Any DNS query that does not match the configured cluster domain suffix, such as **www.kubernetes.io**, is forwarded to the upstream name server inherited from the node. Cluster administrators may have extra stub domains and upstream DNS servers configured.
- **ClusterFirstWithHostNet**: For pods running with hostNetwork, set its DNS policy **ClusterFirstWithHostNet**.
- **None**: It allows a pod to ignore DNS settings from the Kubernetes environment. All DNS settings should be provided using the **dnsPolicy** field in **dnsConfigPod**.

NOTE

- Clusters of Kubernetes v1.10 and later support **Default**, **ClusterFirst**, **ClusterFirstWithHostNet**, and **None**. Clusters earlier than Kubernetes v1.10 support only **Default**, **ClusterFirst**, and **ClusterFirstWithHostNet**.
- **Default** is not the default DNS policy. If **dnsPolicy** is not explicitly specified, **ClusterFirst** is used.

Routing

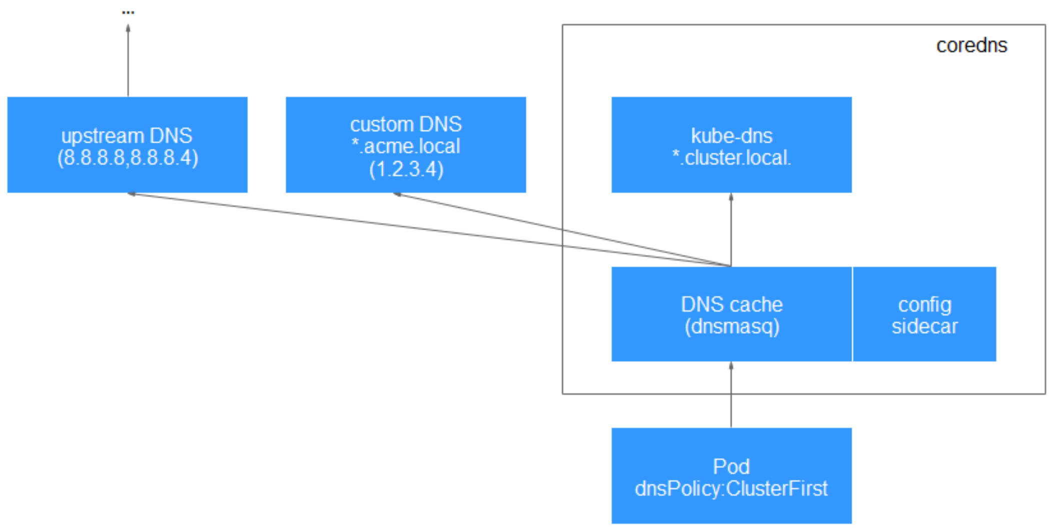
Without stub domain configurations: Any query that does not match the configured cluster domain suffix, such as **www.kubernetes.io**, is forwarded to the upstream DNS server inherited from the node.

With stub domain configurations: If stub domains and upstream DNS servers are configured, DNS queries are routed according to the following flow:

1. The query is first sent to the DNS caching layer in CoreDNS.
2. From the caching layer, the suffix of the request is examined and then the request is forwarded to the corresponding DNS:

- Names with the cluster suffix, for example, **.cluster.local**: The request is sent to CoreDNS.
- Names with the stub domain suffix, for example, **.acme.local**: The request is sent to the configured custom DNS resolver that listens, for example, on 1.2.3.4.
- Names that do not match the suffix (for example, **widget.com**): The request is forwarded to the upstream DNS.

Figure 9-1 Routing



9.2 CCE Container Storage (Everest)

Introduction

Everest is a cloud native container storage add-on that enables Kubernetes clusters to access cloud storage services through CSI.

NOTE

In clusters of v1.27.5-r0, v1.28.3-r0, and later versions, this add-on is automatically configured by the system and does not need to be manually installed or updated.

Editing the Add-on

This add-on is installed by default. To modify its settings, perform the following steps:

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **CCE Container Storage (Everest)** on the right and click **Edit**.

Step 2 On the **Edit Add-on** page, modify the specifications.

Table 9-5 Everest specifications

Parameter	Description
Pods	Number of pods for the add-on. High availability is not possible if there is only one pod. If an error occurs on the node where the pod runs, the add-on will fail.
Containers	CPU and memory quotas of the containers for running the add-on. The add-on contains the everest-csi-controller and everest-csi-driver components. For details, see **Components**.

Step 3 Modify the parameters.

Table 9-6 Everest parameters

Parameter	Description
csi_attacher_worker_threads	Number of worker nodes that can be concurrently processed by Everest for attaching EVS volumes. The default value is **60**.
csi_attacher_detach_worker_threads	Number of worker nodes that can be concurrently processed by Everest for detaching EVS volumes. The default value is **60**.
volume_attaching_flow_ctrl	Maximum number of EVS volumes that can be attached by the Everest add-on within 1 minute. The default value is **0**, indicating that the performance of attaching EVS volumes is determined by the underlying storage resources.
cluster_id	Cluster ID
default_vpc_id	ID of the VPC to which the cluster belongs
disable_auto_mount_secret	Whether the default AK/SK can be used when an object bucket or parallel file system is mounted. The default value is **false**.
enable_node_attacher	Whether to enable the attacher on the agent to process the **VolumeAttachment**.
flow_control	This field is left blank by default. You do not need to configure this parameter.
over_subscription	Overcommitment ratio of the local storage pool (**local_storage**). The default value is **80**. If the size of the local storage pool is 100 GiB, it can be overcommitted to 180 GiB.
project_id	ID of the project to which a cluster belongs

NOTE

The performance of attaching a large number of EVS volumes has been optimized. The following parameters can be configured:

- csi_attacher_worker_threads
- csi_attacher_detach_worker_threads
- volume_attaching_flow_ctrl

The preceding parameters are associated with each other and are constrained by the underlying storage resources in the region where the cluster is located. To attach a large number of volumes (more than 500 EVS volumes per minute), contact customer service and configure the parameters under their guidance to prevent the Everest add-on from running abnormally due to improper parameter settings.

Step 4 Click **OK**.

----End

Components

Table 9-7 Everest components

Component	Description	Resource Type
everest-csi-controller	Used to create, delete, snapshot, expand, attach, and detach storage volumes.	Deployment
everest-csi-driver	Used to mount and unmount PVs and resize file systems.	DaemonSet

9.3 Kubernetes Metrics Server

Kubernetes provides resource usage metrics, such as container CPU usage and memory usage, through the Metrics API. These metrics can be directly accessed by users (by running the **kubectl top** command) or used by a controller in the cluster (for example, HPA) to make decisions.

Kubernetes Metrics Server is an aggregator for monitoring data of core cluster resources. You can quickly install this add-on on the CCE console.

After this add-on is installed, you can create HPA policies. For details, see **HPA Policies**.

The official community project and documentation are available at **https://github.com/kubernetes-sigs/metrics-server**.

Editing the Add-on

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **Kubernetes Metrics Server** on the right and click **Edit**.

Step 2 On the **Edit Add-on** page, modify the specifications.

Table 9-8 Kubernetes Metrics Server specifications

Parameter	Description
Pods	Number of pods for the add-on.
Containers	CPU and memory quotas of the containers for running the add-on.

Step 3 Click **OK**.

----End

Components

Table 9-9 metrics-server components

Component	Description	Resource Type
metrics-server	Aggregator for the monitored data of cluster core resources, which is used to collect and aggregate resource usage metrics obtained through the Metrics API in the cluster	Deployment

9.4 Cloud Native Cluster Monitoring

Introduction

The Cluster Native Cluster Monitoring add-on (kube-prometheus-stack) uses Prometheus-operator and Prometheus to provide easy-to-use, end-to-end Kubernetes cluster monitoring.

This add-on allows the monitoring data to be interconnected with Monitoring Center so that you can view monitoring data and configure alarms in Monitoring Center.

Open-source community: **https://github.com/prometheus/prometheus**

Installing the Add-on

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **Cloud Native Cluster Monitoring** on the right and click **Install**.

Step 2 In the **Data Storage Configuration** area, enable **Report Monitoring Data to AOM** and select an AOM instance.

Reporting Monitoring Data to AOM: If this option is enabled, Prometheus data will be reported to AOM, and you need to select an AOM instance. Basic metrics are free, but custom metrics are billed based on the standard pricing of AOM. For details, see **AOM Pricing Details**. To interconnect with AOM, you must have certain permissions. Only Huawei Cloud accounts, HUAWEI IDs, and IAM users in the **admin** user group can perform this operation.

Step 3 Configure the add-on specifications as needed.

- **Prometheus HA**: The prometheus-lightweight, prometheus-operator, custom-metrics-apiserver, and kube-state-metrics components are deployed in multi-instance mode in the cluster.
- **Configuration List**: You can configure vCPU and memory quotas for each component as needed.

Step 4 Configure the add-on parameters as needed.

- **User-defined indicator HPA**: Application metrics are automatically collected in the form of service discovery. If this option is enabled, you need to add related configurations to the target application. For details, see **Creating an HPA Policy Using Custom Metrics**.
- **Collection Interval**: The value ranges from **10** to **60**.

Step 5 Click **Install**.

After the add-on is installed, you may need to perform the following operations:

If you want to use custom metrics to create an HPA policy, you need to aggregate the custom metrics collected by Prometheus to the API server. For details, see **Creating an HPA Policy Using Custom Metrics**.

----End

Components

All Kubernetes resources created during kube-prometheus-stack add-on installation are created in the namespace named **monitoring**.

Table 9-10 kube-prometheus-stack components

Component	Description	Resource Type
prometheusOperator (workload name: prometheus-operator)	Deploys and manages the Prometheus Server based on CustomResourceDefinitions (CRDs), and monitors and processes the events related to these CRDs. It is the control center of the entire system.	Deployment
prometheus-lightweight (workload name: prometheus-lightweight)	A Prometheus Server cluster deployed by the operator based on the Prometheus CRDs that can be regarded as StatefulSets.	StatefulSet

Component	Description	Resource Type
kubeStateMetrics (workload name: kube-state-metrics)	Converts the Prometheus metric data into a format that can be identified by Kubernetes APIs. **NOTE** If the components run in multiple pods, only metrics in one pod are collected.	Deployment
adapter (workload name: custom-metrics-apiserver)	Aggregates custom metrics to the native Kubernetes API Server. This is closely related to custom metric HPA. The adapter component needs to be installed only when custom metric HPA is enabled.	Deployment

Providing Resource Metrics Through the Metrics API

Resource metrics of containers and nodes, such as CPU and memory usage, can be obtained through the Kubernetes Metrics API. Resource metrics can be directly accessed, for example, by using the **kubectl top** command, or used by HPA or CustomedHPA policies for auto scaling.

The add-on can provide the Kubernetes Metrics API that is disabled by default. To enable the API, create the following APIService object:

```
apiVersion: apiregistration.k8s.io/v1
kind: APIService
metadata:
  labels:
    app: custom-metrics-apiserver
    release: cceaddon-prometheus
  name: v1beta1.metrics.k8s.io
spec:
  group: metrics.k8s.io
  groupPriorityMinimum: 100
  insecureSkipTLSVerify: true
  service:
    name: custom-metrics-apiserver
    namespace: monitoring
    port: 443
  version: v1beta1
  versionPriority: 100
```

You can save the object as a file, name it **metrics-apiservice.yaml**, and run the following command:

```
kubectl create -f metrics-apiservice.yaml
```

Run the **kubectl top pod -n monitoring** command. If information similar to the following is displayed, the Metrics API can be accessed:

```
# kubectl top pod -n monitoring
NAME                                          CPU(cores)   MEMORY(bytes)
......
custom-metrics-apiserver-d4f556ff9-l2j2m            38m          44Mi
......
```

NOTICE

To uninstall the add-on, run the following kubectl command and delete the APIService object. Otherwise, the metrics-server add-on cannot be installed due to residual APIService resources.

```
kubectl delete APIService v1beta1.metrics.k8s.io
```

Creating an HPA Policy Using Custom Metrics

Step 1 Log in to the CCE console and click the cluster name to access the cluster console.

Step 2 You can configure collection rules. The collected metrics can be used as monitoring metrics of HPA policies to control workload auto scaling.

- To create a ConfigMap for collection rules, perform the following steps:

  In the navigation pane on the left, choose **ConfigMaps and Secrets**. Then click **Create from YAML**.

  You can add multiple collection rules by adding multiple configurations under the **rules** field. For details, see **Metrics Discovery and Presentation Configuration**. The following is an example of a custom collection rule:

```
kind: ConfigMap
apiVersion: v1
metadata:
  name: user-adapter-config  # Configuration item name, which cannot be changed.
  namespace: monitoring # Namespace of the ConfigMap
data:
  config.yaml: |-
    rules: # Collection rule
    - seriesQuery: 'container_network_receive_bytes_total{namespace!="",pod!=""}' # Original metrics required for scale-out (metrics from kubelet)
      seriesFilters: []
      resources:
        overrides:  # Specifies pod and namespace resources.
          namespace:
            resource: namespace
          pod:
            resource: pod
      name:
        matches: container_(.*)_total  # Matches metrics that start with container_ and end with _total.
        as: "pod_${1}_per_second"  # Metric alias
      metricsQuery: sum(rate(<<.Series>>{<<.LabelMatchers>>}[30s])) by (<<.GroupBy>>) # Metric change rate of all containers for a workload within 30s
```

  NOTE

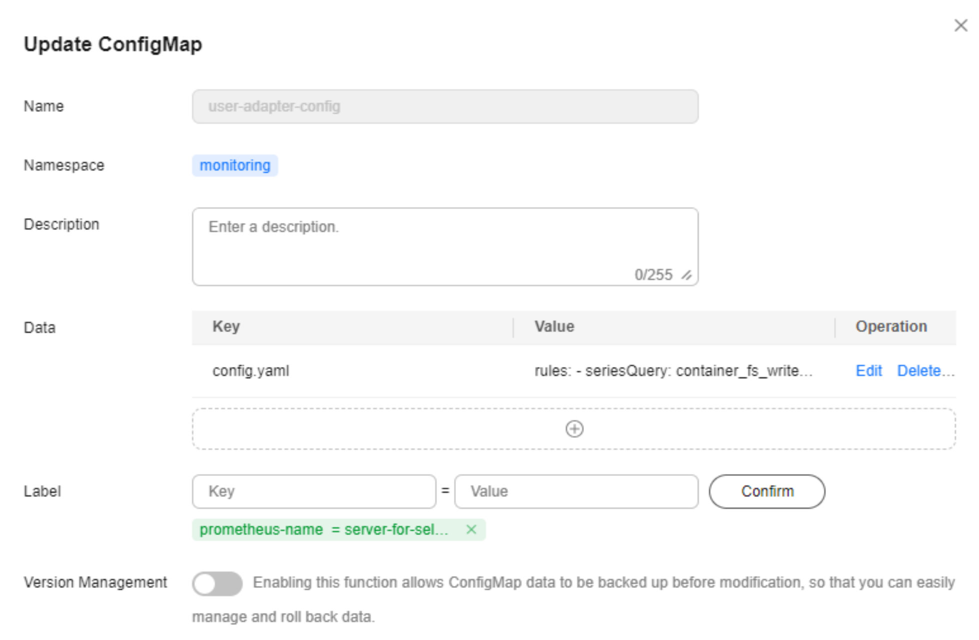
  In the preceding example, the aggregation time is 30s. If this time is set to a value less than 15s (a collection period), the metrics may be inaccurate.

- If a ConfigMap already exists, perform the following steps to modify the collection rule:

  In the navigation pane on the left, choose ConfigMaps and Secrets. Switch to the **monitoring** namespace, locate the target ConfigMap on the **ConfigMaps** tab, and click **Update**.

  On the **Update ConfigMap** page, locate the target data, click **Edit** to modify the collection rule.

Figure 9-2 Updating a ConfigMap



Update ConfigMap

Name	user-adapter-config
Namespace	monitoring
Description	Enter a description. 0/255

Data

Key	Value	Operation
config.yaml	rules: - seriesQuery: container_fs_write...	Edit Delete...

Label: Key = Value Confirm

prometheus-name = server-for-sel...

Version Management: Enabling this function allows ConfigMap data to be backed up before modification, so that you can easily manage and roll back data.

Step 3 In the navigation pane, choose **Add-ons**. Locate the Cloud Native Cluster Monitoring add-on, click **Edit**, and enable the custom metric HPA.

Click **OK**.

NOTE

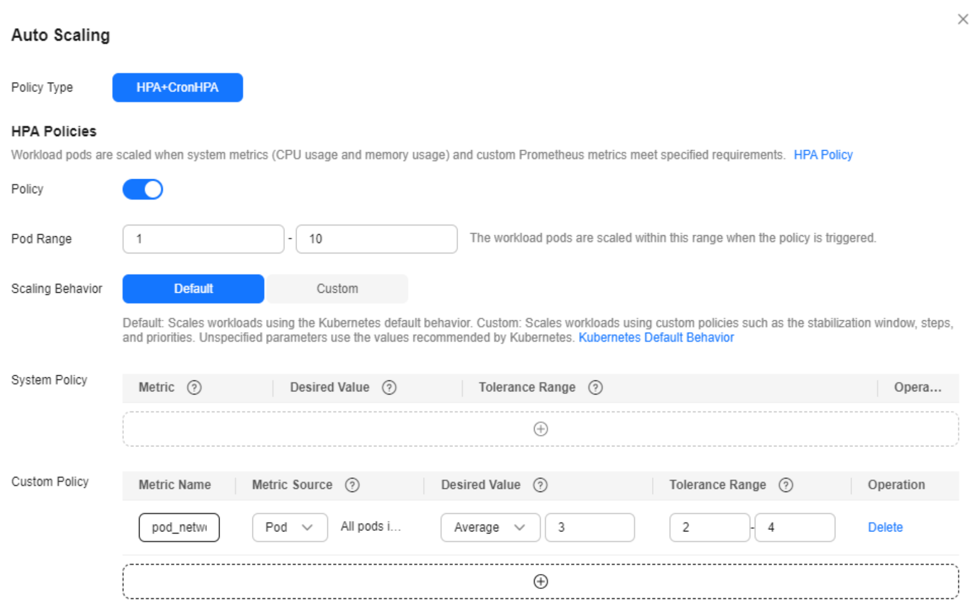
You need to create collection rules described in **Step 2** and then enable the custom metric HPA to trigger add-on redeployment for custom metric collection to take effect.

Step 4 In the navigation pane on the left, choose **Workloads**. Locate the workload for which you want to create an HPA policy and click **Auto Scaling**. In the **Custom Policy** area, you can create the auto scaling policy in **Step 2**.

NOTE

After a workload is created, create an HPA policy unless all pods for that workload are ready and the metrics of the first collection period are collected.

Figure 9-3 Creating an HPA policy


Auto Scaling

Policy Type: HPA+CronHPA

HPA Policies

Workload pods are scaled when system metrics (CPU usage and memory usage) and custom Prometheus metrics meet specified requirements. HPA Policy

Policy: enabled

Pod Range: 1 - 10 The workload pods are scaled within this range when the policy is triggered.

Scaling Behavior: Default | Custom

Default: Scales workloads using the Kubernetes default behavior. Custom: Scales workloads using custom policies such as the stabilization window, steps, and priorities. Unspecified parameters use the values recommended by Kubernetes. Kubernetes Default Behavior

System Policy

Metric	Desired Value	Tolerance Range	Opera...

Custom Policy

Metric Name	Metric Source		Desired Value		Tolerance Range		Operation
pod_netw	Pod	All pods i...	Average	3	2	4	Delete

----End

9.5 Cloud Native Log Collection

Introduction

The Cloud Native Log Collection add-on (log-agent) is developed based on Fluent Bit and OpenTelemetry for collecting logs and Kubernetes events. After the add-on is installed, the fluent-bit container component is automatically created for log collection. This add-on supports CRD-based log collection policies. It collects and forwards standard output logs, container file logs, and Kubernetes event logs in a cluster based on configured policies. It also reports all abnormal Kubernetes events and some normal Kubernetes events to AOM. For details about how to collect logs, see **Collecting Logs**.

NOTE

The fluent-bit container component is provided by Huawei Cloud for free.

Constraints

The constraints on using the log-agent add-on are as follows:

- A maximum of 50 log collection rules can be configured for each cluster.
- log-agent cannot collect .gz, .tar, or .zip log files.
- If the container runtime is containerd, stdout logs cannot be multi-line.
- In each cluster, up to 10,000 single-line logs can be collected per second, and up to 2,000 multi-line logs can be collected per second.

Permissions

The fluent-bit component reads and collects the standard output logs and container file logs based on the collection configuration.

The following permissions are required for running the fluent-bit component:

- CAP_DAC_OVERRIDE: ignores the discretionary access control (DAC) restrictions on files.
- CAP_FOWNER: ignores the restrictions that the file owner ID must match the process user ID.
- DAC_READ_SEARCH: ignores the DAC restrictions on file reading and catalog research.
- SYS_PTRACE: allows all processes to be traced.

Installing the Add-on

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **Cloud Native Log Collection** on the right and click **Install**.

Step 2 On the **Install Add-on** page, configure the specifications.

Table 9-11 Add-on specifications

Parameter	Description
Pods	Number of pods that will be created to match the selected add-on specifications. If you select **Custom**, you can adjust the number of pods as required.
Containers	The add-on contains the following container components, whose specifications can be adjusted as required: ● **log-operator**: parses and updates log rules. ● **otel-collector**: forwards logs collected by fluent-bit to LTS.

Step 3 Click **Install**.

----End

Components

Table 9-12 log-agent components

Component	Description	Resource Type
fluent-bit	A lightweight log collector and forwarder for collecting logs.	Pod

Component	Description	Resource Type
log-operator	Used to generate internal configuration files	Deployment
otel-collector	Used to collect logs from applications and services and report the logs to LTS	Deployment

9.6 NGINX Ingress Controller

Introduction

Kubernetes uses kube-proxy to expose Services and provide load balancing. The implementation is at the transport layer. When it comes to Internet applications, where a bucket-load of information is generated, forwarding needs to be more fine-grained, precisely and flexibly controlled by policies and load balancers to deliver higher performance.

This is where Ingresses enter. Ingresses provide application-layer forwarding functions, such as virtual hosts, load balancing, SSL proxy, and HTTP routing, for Services that can be directly accessed outside a cluster.

Kubernetes has officially released the Nginx-based Ingress controller. CCE NGINX Ingress Controller uses community templates and images. This add-on generates Nginx configuration and stores the configuration using ConfigMaps. The configuration will be written to Nginx pods through the Kubernetes API. In this way, the Nginx configuration is modified and updated. For details, see **How Nginx Ingress Works**.

You can visit the **open-source community** for more information.

NOTE

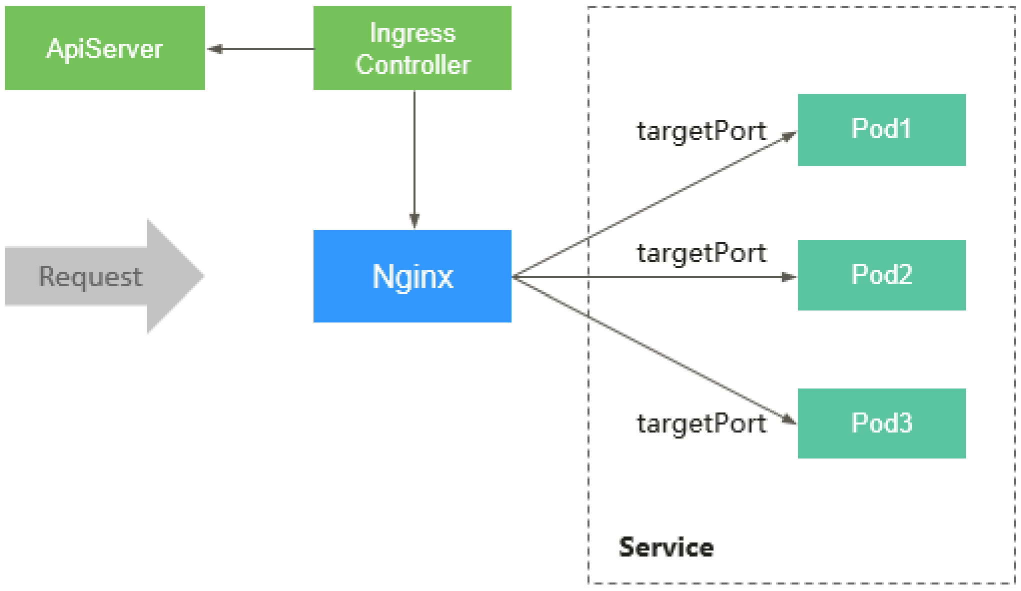
- When installing NGINX Ingress Controller, you can specify Nginx parameters. These parameters take effect globally and are contained in the **nginx.conf** file. You can search for the parameters in **ConfigMaps**. If the parameters are not included in ConfigMaps, the configurations will not take effect.
- After the add-on is installed, you can interconnect the **Ingress** you create on the CCE console with Nginx and set Nginx Ingress functions using **Annotations**. For details about the supported annotation fields, see **Annotations**.
- Do not manually modify or delete the load balancer and listener that is automatically created by CCE. If the load balancer or listener is deleted, the workload will be abnormal. If you have modified or deleted them by mistake, uninstall NGINX Ingress Controller and re-install it.

How Nginx Ingress Works

An Nginx Ingress consists of the Ingress object, Ingress Controller, and Nginx. Ingress Controller assembles Ingress into the Nginx configuration file (**nginx.conf**) and reloads Nginx to make the configuration changes apply. When Ingress Controller detects that a pod in a Service changes, it dynamically changes the upstream server group configuration of Nginx. In this case, the Nginx process does not need to be reloaded. **Figure 9-4** shows how an Nginx Ingress works.

- An Ingress is a group of access rules that forward requests to specified Services based on domain names or URLs. Ingresses are stored in etcd and can be added, deleted, modified, and queried through APIs.
- Ingress Controller monitors the changes of resource objects such as Ingresses, Services, endpoints, secrets (mainly TLS certificates and keys), nodes, and ConfigMaps in real time and automatically performs operations on Nginx.
- Nginx implements load balancing and access control at the application layer.

Figure 9-4 How Nginx Ingress works



Constraints

- Dedicated load balancers with private IP addresses bound and used for network load balancing (load balancing over TCP or UDP) should be selected.
- During the upgrade of NGINX Ingress Controller, 10s is reserved for deleting NGINX Ingress Controller deployed on the backend servers associated with a load balancer.
- The timeout duration for the graceful exit of NGINX Ingress Controller is 300s. If the timeout duration is longer than 300s during the upgrade of the NGINX Ingress Controller, persistent connections will be disconnected, and connectivity will be interrupted for a short period of time.

Prerequisites

Before creating a workload, you must have an available cluster. If no cluster is available, create one by performing the operations in **Buying a CCE Autopilot Cluster**.

Installing the Add-on

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **NGINX Ingress Controller** on the right and click **Install**.

Step 2 On the **Install Add-on** page, configure the specifications.

Table 9-13 NGINX Ingress Controller specifications

Parameter	Description
Pods	You can adjust the number of add-on instances as required.
Containers	You can adjust the container specifications of an add-on instance as required.

Step 3 Configure the add-on parameters.

- **Load Balancer**: Select a dedicated load balancer. If no load balancer is available, create one. The load balancer has at least two listeners, and ports 80 and 443 are not occupied by listeners.
- **Admission Check**: If this option is enabled, an admission check will be performed on the configuration of Nginx Ingresses. If the configuration files the admission check, requests to the Ingresses will be intercepted. For details about verification, see **Access Control**.

  NOTE

  – Admission checks slow down the responses to Ingress requests.
  – Admission checks are only available for the add-on v2.4.1 or later.
- **Nginx Parameters**: Configuring the **nginx.conf** file will affect all managed Ingresses. You can search for related parameters through **ConfigMaps**. If the parameters you configured are not included in the options listed in ConfigMaps, the parameters will not take effect.

  For example, you can use the **keep-alive-requests** parameter to describe how to set the maximum number of requests for keeping active connections to 100.

  ```
  {
     "keep-alive-requests": "100"
  }
  ```
- **Default 404 Service**: By default, the 404 service provided by the add-on is used. To customize the 404 service, enter the namespace/service name. If the service does not exist, the add-on installation will fail.

Step 4 Click **Install**.

----End

Components

Table 9-14 NGINX Ingress Controller components

Component	Description	Resource Type
cceaddon-nginx-ingress-controller	NGINX Ingress Controller that provides flexible routing and forwarding for clusters.	Deployment
cceaddon-nginx-ingress-default-backend	Default backend of the Nginx Ingress. The message "default backend - 404" is returned.	Deployment
cceaddon-nginx-ingress-admission-create	After webhooks are enabled, you can create a certificate for webhook verification.	Job
cceaddon-nginx-ingress-admission-patch	After webhooks are enabled, you can update the created certificate to the webhook configuration.	Job

9.7 CCE Advanced HPA

CCE Advanced HPA (cce-hpa-controller) is an in-house add-on, which can be used to flexibly scale in or out Deployments based on metrics such as CPU usage and memory usage.

Main Functions

- Scaling can be performed based on the percentage of the current number of pods.
- The minimum scaling step can be set.
- Different scaling operations can be performed based on the actual metric values.

Constraints

To use CCE Advanced HPA, install an add-on that provides metrics APIs. Select one of the following add-ons based on your cluster version and actual requirements.

- **Kubernetes Metrics Server**: provides basic resource usage metrics, such as container CPU and memory usage. It is supported by all cluster versions.

- **Creating an HPA Policy Using Custom Metrics**: Custom metrics need to be aggregated to the Kubernetes API server for auto scaling.

Installing the Add-on

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **Add-ons**. Locate **CCE Advanced HPA** on the right and click **Install**.

Step 2 On the **Install Add-on** page, configure the specifications.

Table 9-15 CCE Advanced HPA specifications

Parameter	Description
Pods	Number of pods for the add-on. High availability is not possible if there is only one pod. If an error occurs on the node where the pod runs, the add-on will fail.
Containers	CPU and memory quotas of the container allowed for the selected add-on specifications. You can adjust the container specifications of an add-on instance as required.

Step 3 Click **Install**.

----End

Components

Table 9-16 cce-hpa-controller components

Component	Description	Resource Type
customedhpa-controller	CCE auto scaling component, which scales in or out Deployments based on metrics such as CPU usage and memory usage	Deployment

10 Helm Chart

10.1 API Resource Restrictions on a Template

Resource	Restriction Item	Description	Recommended Alternative Solution
namespaces	-	Supported	For security purposes, CCE Autopilot does not allow you to deployment workloads in the system namespace (such as **kube-system**). Also, you cannot create, modify, delete, or execute any resources.
nodes	-	Supported	You can query nodes but cannot create, delete, and modify nodes.
persistentvolumeclaims	-	Supported	-
persistentvolumes	-	Supported	-
pods	hostPath	Mounting a file on the local host to a pod is not allowed.	Use emptyDir or cloud storage.
	HostNetwork	Mapping the host port to a pod is not allowed.	Use load balancing (**type=LoadBalancer**).
	HostPID	Sharing the host's PID namespace to pods is not allowed.	Users are unaware of the node. There is no need to use the restriction item.

Resource	Restriction Item	Description	Recommended Alternative Solution
	HostIPC	Container processes are not allowed to communicate with processes on the host.	Users are unaware of the node. There is no need to use the restriction item.
	NodeName	Scheduling pods to specific nodes is not allowed.	Users are unaware of the node. There is no need to use the restriction item.
	Privileged containers	Not supported	-
	Linux capabilities	**SETPCAP**, **MKNOD**, **AUDIT_WRITE**, **CHOWN**, **DAC_OVERRIDE**, **FOWNER**, **FSETID**, **KILL**, **SETGID**, **SETUID**, **NET_BIND_SERVICE**, **SYS_CHROOT**, **SETFCAP**, and **SYS_PTRACE** are supported. You can also enable **NET_RAW**, **SYS_PTRACE**, and **NET_ADMIN** by setting **SecurityContext**.	Use allowed values.
	Node affinity and anti-affinity	Pods cannot be scheduled to specified nodes or nodes with certain labels, or a batch of pods cannot be scheduled to nodes with certain labels. The node affinity or the **nodeSelector** field does not take effect in CCE Autopilot clusters.	● You do not need to specify a node for scheduling, but you can specify a pod to an AZ. ● A batch of pods can be scheduled to multiple AZs.
	Pod affinity and anti-affinity	Ineffective	You do not need to set this parameter.
	allowPrivilegeEscalation (whether privilege escalation is allowed)	Not supported	Keep the default settings.

Resource	Restriction Item	Description	Recommended Alternative Solution
	RuntimeClassName	This parameter does not need to be configured. When RuntimeClassName is specified by an application (except pods), the value is automatically changed to **runc** supported by the system.	You do not need to set this parameter.
	Time zone synchronization (the **/etc/localtime** file on the host)	Not supported	Keep the default settings.
serviceaccounts	-	System configurations cannot be modified, and system-defined roles cannot be bound.	Keep the default settings.
services	-	Services of the NodePort type are not allowed, and only dedicated load balancer can be used for Services.	Use load balancing (**type=LoadBalancer**).
daemonsets	apps	DaemonSets are not allowed.	Deploy multiple images in a pod using sidecars.
deployments	apps	Supported. The restricted fields are the same as those in **pods**.	Use allowed values.
replicasets	apps	Supported. The restricted fields are the same as those in **pods**.	Use allowed values.
statefulsets	apps	Supported. The restricted fields are the same as those in **pods**.	Use allowed values.
cronjobs	batch	Supported. The restricted fields are the same as those in **pods**.	Use allowed values.
jobs	batch	Supported. The restricted fields are the same as those in **pods**.	Use allowed values.

Resource	Restriction Item	Description	Recommended Alternative Solution
clusterrolebindings	rbac.authorization.k8s.io	Supported. The system group, system user, and cce-service group cannot be bound.	Use allowed values.
rolebindings	rbac.authorization.k8s.io	Supported. The system group, system user, and cce-service group cannot be bound.	Use allowed values.
storageclasses	storage.k8s.io	OBS and EVS storage classes cannot be created. Other functions are supported.	Use allowed values.

10.2 Deploying an Application from a Chart

On the CCE console, you can upload a Helm chart package, deploy it, and manage the deployed pods.

Constraints

- The number of charts that can be uploaded by a single user is limited. The value displayed on the console of each region is the allowed quantity.
- CCE uses Helm v3.8.2 and allows uploading Helm v3 chart packages.
- A chart with multiple versions consumes the same amount of portion of chart quota.
- Users with chart operation permissions can perform multiple operations on clusters. Therefore, exercise caution when assigning users the chart lifecycle management permissions, including uploading charts and creating, deleting, and updating chart releases.

Chart Specifications

The Redis workload is used as an example to illustrate the chart specifications.

- **Naming Requirement**

  A chart package is named in the format of **{name}-{version}**.tgz, where **{version}** indicates the version number in the format of *Major version number.Minor version number.Revision number*, for example, **redis-0.4.2.tgz**.

NOTE

The chart name {name} can contain a maximum of 64 characters.

The version number must comply with the **semantic versioning** rules.

- The main and minor version numbers are mandatory, and the revision number is optional.
- The major and minor version numbers and revision number must be integers, greater than or equal to 0, and less than or equal to 99.

- **Directory Structure**

  The directory structure of a chart is as follows:

```
redis/
  templates/
  values.yaml
  README.md
  Chart.yaml
  .helmignore
```

As listed in **Table 10-1**, the parameters marked with * are mandatory.

Table 10-1 Parameters in the directory structure of a chart

Parameter	Description
* templates	Stores all templates.
* values.yaml	Describes configuration parameters required by templates. **NOTICE** Make sure that the image address set in the **values.yaml** file is the same as the image address in the container image repository. Otherwise, an exception occurs when you create a workload, and the system displays a message indicating that the image fails to be pulled. To obtain the image address, perform the following operations: Log in to the CCE console. In the navigation pane on the left, choose **Image Repository** to access the SWR console. Choose **My Images** > **Private Images** and click the name of the uploaded image. On the **Image Tags** tab, obtain the image address from the pull command. You can click ⧉ to copy the command in the **Image Pull Command** column.
README.md	A markdown file, including: ● The workload or services provided by the chart. ● Prerequisites for running the chart. ● Configurations in the **values.yaml** file. ● Information about chart installation and configuration.
* Chart.yaml	Basic information about the chart. Note: The API version of Helm v3 is switched from v1 to v2.
.helmignore	Files or data that does not need to read templates during workload installation.

Uploading a Chart

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **App Templates**. Then click **Upload Chart** in the upper right corner.

Step 2 Click **Select File**, select the chart to be uploaded, and click **Upload**.

----End

Creating a Release

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **App Templates**.

Step 2 On the **My Charts** tab, click **Install** of the target chart.

Step 3 Set workload installation parameters by referring to **Table 10-2**.

Table 10-2 Installation parameters

Parameter	Description
Instance	Unique name of the chart release.
Namespace	Namespace to which the workload will be deployed.
Select Version	Version of a chart.
Configuration File	You can import and replace the **values.yaml** file or directly edit the chart parameters online. **NOTE** An imported **values.yaml** file must comply with YAML specifications, that is, KEY:VALUE format. The fields in the file are not restricted. The key value of the imported values.yaml must be the same as that of the selected chart package. Otherwise, the values.yaml does not take effect. That is, the key cannot be changed. 1. Click **Select File**. 2. Select the corresponding **values.yaml** file and click **Open**.

Step 4 Click **Install**.

On the **Releases** tab, you can view the installation status of the release.

----End

Upgrading a Chart-based Workload

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **App Templates**. Then click the **Releases** tab.

Step 2 Click **Upgrade** in the row where the desired workload resides and set the parameters for the workload.

Step 3 Select a chart version for **Chart Version**.

Step 4 Follow the prompts to modify the chart parameters. Click **Upgrade** and then click **Submit**.

Step 5 Click **Back to Release List**. If the chart status changes to **Upgrade successful**, the workload is successfully upgraded.

----End

Rolling Back a Chart-based Workload

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **App Templates**. Then click the **Releases** tab.

Step 2 Click **More** > **Roll Back** for the workload to be rolled back, select the workload version, and click **Roll back to this version**.

In the workload list, if the status is **Rollback successful**, the workload is rolled back successfully.

----End

Uninstalling a Chart-based Workload

Step 1 Log in to the CCE console and click the cluster name to access the cluster console. In the navigation pane on the left, choose **App Templates**. Then click the **Releases** tab.

Step 2 Click **More** > **Uninstall** next to the release to be uninstalled, and click **Yes**. Exercise caution when performing this operation because releases cannot be restored after being uninstalled.

----End

11 Permissions

11.1 Permission Overview

CCE permissions management allows you to assign permissions to IAM users and user groups under your tenant accounts. CCE combines the advantages of Identity and Access Management (IAM) and Kubernetes Role-based Access Control (RBAC) authorization to provide a variety of authorization methods, including IAM fine-grained authorization, IAM token authorization, cluster-scoped authorization, and namespace-wide authorization.

CCE allows you to manage permissions on clusters and related resources at a finer granularity, for example, to control the access of employees in different departments to cloud resources.

This section describes the CCE permissions management mechanism and related concepts. If your account has met your service requirements, you can skip the configurations in this chapter.

NOTE

CCE policies apply to all CCE products, including CCE standard, CCE Turbo, and CCE Autopilot clusters. Any permission that is specific for CCE Autopilot clusters will be described in special notes.

CCE Permissions Management

CCE permissions are described as follows:

- **Cluster-level permissions**: Cluster-level permissions management evolves out of the system policy authorization feature of IAM. IAM users in the same user group have the same permissions. A user group is simply a group of users. By granting cluster permissions to specific user groups, you can enable those users to perform various operations on clusters, including creating or deleting clusters, nodes, node pools, charts, and add-ons. In the meantime, you can restrict other user groups to only view clusters.

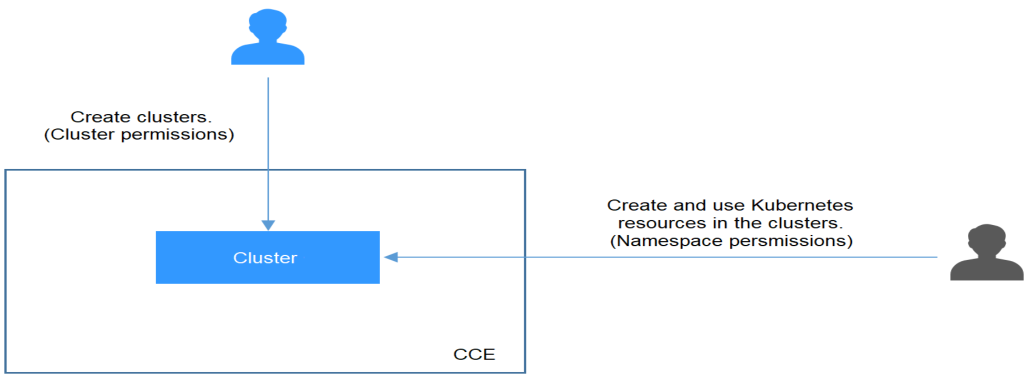
  Cluster-level permissions involve CCE non-Kubernetes APIs and support fine-grained IAM policies and enterprise project management capabilities.
- **Namespace-level permissions**: You can regulate users' or user groups' access to Kubernetes resources in a single namespace based on their Kubernetes

RBAC roles. CCE has also been enhanced based on open-source capabilities. It supports RBAC authorization based on IAM user or user group, and RBAC authentication on access to APIs using IAM tokens.

Namespace-level permissions involve CCE and Kubernetes APIs and are enhanced based on the Kubernetes RBAC capabilities. Namespace-level permissions can be granted to IAM users or user groups for authentication and authorization, but are independent of fine-grained IAM policies.

In general, you configure CCE permissions in two scenarios. The first is creating and managing clusters and related resources, such as nodes. The second is creating and using Kubernetes resources in the cluster, such as workloads and Services.

Figure 11-1 Illustration on CCE permissions



These permissions allow you to manage resource users at a finer granularity.

Cluster Permissions (IAM-based) and Namespace Permissions (Kubernetes RBAC-based)

Users with different cluster permissions (assigned using IAM) have different namespace permissions (assigned using Kubernetes RBAC). **Table 11-1** lists the namespace permissions of different users.

Table 11-1 Differences in namespace permissions

User	CCE Standard, CCE Turbo, and CCE Autopilot (v1.13 or Later)
User with the Tenant Administrator permissions (for example, an account)	All namespace permissions
IAM user with the CCE Administrator role	All namespace permissions
IAM user with the CCE FullAccess or CCE ReadOnlyAccess role	Kubernetes RBAC authorization
IAM user with the Tenant Guest role	Kubernetes RBAC authorization

kubectl Permissions

You can use **kubectl** to access Kubernetes resources in a cluster.

When you access a cluster using kubectl, CCE uses **kubeconfig.json** generated on the cluster for authentication. This file contains user information, based on which CCE determines which Kubernetes resources can be accessed by kubectl. The permissions recorded in a **kubeconfig.json** file vary from user to user. The permissions that a user has are listed in **Table 11-1**.

Federated Users

IAM provides the identity provider function to implement federated identity authentication based on Security Assertion Markup Language (SAML) or OpenID Connect. This allows users in your management system to access the cloud platform through single sign-on (SSO).

Users who log in through federated identity authentication are called federated users. Federated users are equivalent to IAM users.

Pay attention to the following for federated users to use CCE:

- When a user creates a CCE cluster, the cluster-admin permission is granted to the user by default. The user ID of a federated user changes upon each login and logout. Therefore, the user is displayed as deleted on the **Permissions** page of the CCE console. Do not manually delete the permission, otherwise, the authentication fails. You are advised to grant the cluster-admin permission to a user group on CCE and add federated users to the user group.
- Federated users cannot create permanent access keys (AKs/SKs). In scenarios where AKs/SKs are required (for example, when creating OBS-related PVs/PVCs), only you or an IAM user can create the AK/SK and share them with the federated users. An access key contains the permissions granted to a user, so it is recommended that the federated user request an IAM user in the same group to create an access key.

Supported Actions in IAM

CCE provides system-defined policies that can be directly used in IAM. You can also create custom policies to supplement system-defined policies for more refined access control. Operations supported by policies are specific to APIs. The following are common concepts related to policies:

- Permissions: statements in a policy that allow or deny certain operations
- APIs: REST APIs that can be called by a user who has been granted specific permissions.
- Actions: specific operations that are allowed or denied.
- Dependencies: actions which a specific action depends on. When allowing an action for a user, you also need to allow any existing action dependencies for that user.
- IAM projects/Enterprise projects: the authorization scope of a custom policy. A custom policy can be applied to IAM projects or enterprise projects or both. Policies that contain actions for both IAM and enterprise projects can be used and applied for both IAM and Enterprise Management. Policies that contain actions only for IAM projects can be used and applied to IAM only. For details

about the differences between IAM and enterprise management, see **What Are the Differences Between IAM and Enterprise Management?**

NOTE

The check mark (√) and cross symbol (x) indicate that an action takes effect or does not take effect for the corresponding type of projects.

CCE supports the following actions in custom policies.

Table 11-2 Cluster management actions

Permission	API	Action	IAM Project	Enterprise Project
Obtaining clusters in a project	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters	cce:cluster:list	√	√
Obtaining a cluster	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id} **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}	cce:cluster:get	√	√
Creating a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters	cce:cluster:create	√	√
Updating a cluster	**[CCE standard and CCE Turbo clusters]** PUT /api/v3/projects/{project_id}/clusters/{cluster_id} **[CCE Autopilot clusters]** PUT /autopilot/v3/projects/{project_id}/clusters/{cluster_id}	cce:cluster:update	√	√

Permission	API	Action	IAM Project	Enterprise Project

Deleting a cluster	**[CCE standard and CCE Turbo clusters]**
	DELETE /api/v3/projects/{project_id}/clusters/{cluster_id}
	[CCE Autopilot clusters]
	DELETE /autopilot/v3/projects/{project_id}/clusters/{cluster_id}
		cce:cluster:delete	√	√

Upgrading a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade	cce:cluster:upgrade	√	√
Waking up a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/awake **[CCE Autopilot clusters]** N/A	cce:cluster:start	√	√
Hibernating a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/hibernate **[CCE Autopilot clusters]** N/A	cce:cluster:stop	√	√
Changing the specifications of a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v2/projects/{project_id}/clusters/:clusterid/resize **[CCE Autopilot clusters]** N/A	cce:cluster:resize	√	√

Permission	API	Action	IAM Project	Enterprise Project
Binding/ Unbinding the public API server address of a cluster	**[CCE standard and CCE Turbo clusters]** PUT /api/v3/projects/{project_id}/clusters/{cluster_id}/mastereip **[CCE Autopilot clusters]** PUT /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/mastereip	cce:cluster:update	√	√
Obtaining the certificate of a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/clustercert **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/clustercert	cce:cluster:get	√	√

Table 11-3 Node

Permissions	API	Action	IAM Project	Enterprise Project
Obtaining all nodes in a cluster	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/nodes **[CCE Autopilot clusters]** N/A	cce:node:list	√	√
Obtaining a node	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/nodes/{node_id} **[CCE Autopilot clusters]** N/A	cce:node:get	√	√

Permissions	API	Action	IAM Project	Enterprise Project
Creating a node	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/nodes **[CCE Autopilot clusters]** N/A	cce:node:create	√	√ **NOTE** If you use enterprise project authorization to create a node, you need to add the global permission of **evs:quota:get**.
Updating a node	**[CCE standard and CCE Turbo clusters]** PUT /api/v3/projects/{project_id}/clusters/{cluster_id}/nodes/{node_id} **[CCE Autopilot clusters]** N/A	cce:node:update	√	√
Deleting a node	**[CCE standard and CCE Turbo clusters]** DELETE /api/v3/projects/{project_id}/clusters/{cluster_id}/nodes/{node_id} **[CCE Autopilot clusters]** N/A	cce:node:delete	√	√

Table 11-4 Job

Permissions	API	Action	IAM Project	Enterprise Project
Obtaining information about a job	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/jobs/{job_id} **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/jobs/{job_id}	cce:job:get	√	√

Permissions	API	Action	IAM Project	Enterprise Project
Listing all jobs	**[CCE standard and CCE Turbo clusters]** GET /api/v2/projects/{project_id}/jobs **[CCE Autopilot clusters]** GET /autopilot/v2/projects/{project_id}/jobs	cce:job:list	√	√
Deleting one or all jobs	**[CCE standard and CCE Turbo clusters]** DELETE /api/v2/projects/{project_id}/jobs DELETE /api/v2/projects/{project_id}/jobs/{job_id} **[CCE Autopilot clusters]** DELETE /autopilot/v2/projects/{project_id}/jobs DELETE /autopilot/v2/projects/{project_id}/jobs/{job_id}	cce:job:delete	√	√

Table 11-5 Nodepool

Permission	API	Action	IAM Project	Enterprise Project
Obtaining all node pools in a cluster	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/nodepools **[CCE Autopilot clusters]** N/A	cce:nodepool:list	√	√
Obtaining a node pool	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/nodepools/{nodepool_id} **[CCE Autopilot clusters]** N/A	cce:nodepool:get	√	√

Permission	API	Action	IAM Project	Enterprise Project
Creating a node pool	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/nodepools **[CCE Autopilot clusters]** N/A	cce:nodepool:create	√	√
Updating a node pool	**[CCE standard and CCE Turbo clusters]** PUT /api/v3/projects/{project_id}/clusters/{cluster_id}/nodepools/{nodepool_id} **[CCE Autopilot clusters]** N/A	cce:nodepool:update	√	√
Deleting a node pool	**[CCE standard and CCE Turbo clusters]** DELETE /api/v3/projects/{project_id}/clusters/{cluster_id}/nodepools/{nodepool_id} **[CCE Autopilot clusters]** N/A	cce:nodepool:delete	√	√

Table 11-6 Chart

Permission	API	Action	IAM Project	Enterprise Project
Updating a chart	**[CCE standard and CCE Turbo clusters]** PUT /v2/charts/{id} **[CCE Autopilot clusters]** POST /autopilot/v2/charts	cce:chart:update	√	×
Uploading a chart	**[CCE standard and CCE Turbo clusters]** POST /v2/charts **[CCE Autopilot clusters]** POST /autopilot/v2/charts	cce:chart:upload	√	×

Permission	API	Action	IAM Project	Enterprise Project
Downloading a chart	**[CCE standard and CCE Turbo clusters]** GET /v2/charts/{id}/archive **[CCE Autopilot clusters]** GET /autopilot/v2/charts/{id}/archive	cce:chart:get	√	×
Listing all charts	**[CCE standard and CCE Turbo clusters]** GET /v2/charts **[CCE Autopilot clusters]** GET /autopilot/v2/charts	cce:chart:list	√	×
Obtaining information about a chart	**[CCE standard and CCE Turbo clusters]** GET /v2/charts/{id} **[CCE Autopilot clusters]** GET /autopilot/v2/charts/{id}	cce:chart:get	√	×
Obtaining a chart	**[CCE standard and CCE Turbo clusters]** GET /v2/charts/{id}/values **[CCE Autopilot clusters]** GET /autopilot/v2/charts/{id}/values	cce:chart:get	√	×
Deleting a chart	**[CCE standard and CCE Turbo clusters]** DELETE /v2/charts/{id} **[CCE Autopilot clusters]** DELETE /autopilot/v2/charts/{id}	cce:chart:delete	√	×
Obtaining the quota of a chart	**[CCE standard and CCE Turbo clusters]** GET /v2/charts/{project_id}/quotas **[CCE Autopilot clusters]** GET /autopilot/v2/charts/{project_id}/quotas	cce:chart:list	√	×

Table 11-7 Release

Permission	API	Action	IAM Project	Enterprise Project
Updating a release	**[CCE standard and CCE Turbo clusters]** PUT /cce/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name} PUT /v2/releases/{name} (deprecated) **[CCE Autopilot clusters]** PUT /autopilot/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name}	cce:release:update	√	√
Listing all releases	**[CCE standard and CCE Turbo clusters]** GET /cce/cam/v3/clusters/{cluster_id}/releases GET /v2/releases (deprecated) **[CCE Autopilot clusters]** GET /autopilot/cam/v3/clusters/{cluster_id}/releases	cce:release:list	√	√
Creating a release	**[CCE standard and CCE Turbo clusters]** POST /cce/cam/v3/clusters/{cluster_id}/releases POST /v2/releases (deprecated) **[CCE Autopilot clusters]** POST /autopilot/cam/v3/clusters/{cluster_id}/releases	cce:release:create	√	√

Permission	API	Action	IAM Project	Enterprise Project
Obtaining information about a release	**[CCE standard and CCE Turbo clusters]** GET /cce/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name} GET /v2/releases/{name} (deprecated) **[CCE Autopilot clusters]** GET /autopilot/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name}	cce:release:get	√	√
Querying historical records of a release	**[CCE standard and CCE Turbo clusters]** GET /cce/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name}/history GET /v2/releases/{name}/history (deprecated) **[CCE Autopilot clusters]** GET /autopilot/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name}	cce:release:get	√	√
Deleting a release	**[CCE standard and CCE Turbo clusters]** DELETE /cce/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name} DELETE /v2/releases/{name} (deprecated) **[CCE Autopilot clusters]** DELETE /autopilot/cam/v3/clusters/{cluster_id}/namespace/{namespace}/releases/{name}	cce:release:delete	√	√

Table 11-8 Storage

Permission	API	Action	IAM Project	Enterprise Project
Creating a PVC (to be discarded)	**[CCE standard and CCE Turbo clusters]** POST /api/v1/namespaces/{namespace}/cloudpersistent-volumeclaims **[CCE Autopilot clusters]** N/A	cce:storage:create	√	√
Deleting a PVC (to be discarded)	**[CCE standard and CCE Turbo clusters]** DELETE /api/v1/namespaces/{namespace}/cloudpersistent-volumeclaims/{name} **[CCE Autopilot clusters]** N/A	cce:storage:delete	√	√
Listing all volumes	**[CCE standard and CCE Turbo clusters]** GET /storage/api/v1/namespaces/{namespace}/listvolumes **[CCE Autopilot clusters]** N/A	cce:storage:list	√	√

Table 11-9 Addon

Permission	API	Action	IAM Project	Enterprise Project
Listing all add-on templates	**[CCE standard and CCE Turbo clusters]** GET /api/v3/addontemplate **[CCE Autopilot clusters]** GET /autopilot/v3/addontemplates	cce:addonTemplate:get	√	x
Creating an add-on instance	**[CCE standard and CCE Turbo clusters]** POST /api/v3/addons **[CCE Autopilot clusters]** POST /autopilot/v3/addons	cce:addonInstance:create	√	√

Permission	API	Action	IAM Project	Enterprise Project
Obtaining an add-on instance	**[CCE standard and CCE Turbo clusters]** GET /api/v3/addons/{id}?cluster_id={cluster_id} **[CCE Autopilot clusters]** GET /autopilot/v3/addons/{id}	cce:addonInstance:get	√	√
Listing all add-on instances	**[CCE standard and CCE Turbo clusters]** GET /api/v3/addons?cluster_id={cluster_id} **[CCE Autopilot clusters]** GET /autopilot/v3/addons?cluster_id={cluster_id}	cce:addonInstance:list	√	√
Deleting an add-on instance	**[CCE standard and CCE Turbo clusters]** DELETE /api/v3/addons/{id} **[CCE Autopilot clusters]** DELETE /autopilot/v3/addons/{id}	cce:addonInstance:delete	√	√
Updating an add-on instance	**[CCE standard and CCE Turbo clusters]** PUT /api/v3/addons/{id} **[CCE Autopilot clusters]** PUT /autopilot/v3/addons/{id}	cce:addonInstance:update	√	√

Table 11-10 Quota

Permission	API	Action	IAM Project	Enterprise Project
Obtaining quota details	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/quotas **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/quotas	cce:quota:get	√	√

Table 11-11 Label

Permission	API	Action	IAM Project	Enterprise Project
Adding resource tags to a cluster in batches	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/tags/create **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/tags/create	cce:tag:operate	√	√
Deleting resource tags from a cluster in batches	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/tags/delete **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/tags/delete	cce:tag:operate	√	√

Table 11-12 Upgrade

Permission	API	Action	IAM Project	Enterprise Project
Obtaining the details about a cluster upgrade task	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/tasks/{task_id} **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/tasks/{task_id}	cce:cluster:get	√	√

Permission	API	Action	IAM Project	Enterprise Project
Obtaining a list of cluster upgrade task details	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/tasks **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/tasks	cce:cluster:get	√	√
Retrying a cluster upgrade task	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/retry **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgrade/retry	cce:cluster:upgrade	√	√
Performing a pre-upgrade check for a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck	cce:cluster:upgrade	√	√

Permission	API	Action	IAM Project	Enterprise Project
Obtaining details about a pre-upgrade check task of a cluster	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck/tasks/{task_id} **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck/tasks/{task_id}	cce:cluster:get	√	√
Obtaining a list of pre-upgrade check tasks of a cluster	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck/tasks **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/precheck/tasks	cce:cluster:get	√	√
Performing a post-upgrade check for a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/postcheck **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/postcheck	cce:cluster:upgrade	√	√

Permission	API	Action	IAM Project	Enterprise Project
Backing up a cluster	**[CCE standard and CCE Turbo clusters]** POST /api/v3.1/projects/{project_id}/clusters/{cluster_id}/operation/snapshot **[CCE Autopilot clusters]** POST /autopilot/v3.1/projects/{project_id}/clusters/{cluster_id}/operation/snapshot	cce:cluster:upgrade	√	√
Obtaining a list of cluster backup task details	**[CCE standard and CCE Turbo clusters]** GET /api/v3.1/projects/{project_id}/clusters/{cluster_id}/operation/snapshot/tasks **[CCE Autopilot clusters]** GET /autopilot/v3.1/projects/{project_id}/clusters/{cluster_id}/operation/snapshot/tasks	cce:cluster:get	√	√
Obtaining details about a cluster upgrade task	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/upgradeinfo **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/upgradeinfo	cce:cluster:get	√	√
Enabling the cluster upgrade booting task	**[CCE standard and CCE Turbo clusters]** POST /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows **[CCE Autopilot clusters]** POST /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows	cce:cluster:upgrade	√	√

Permission	API	Action	IAM Project	Enterprise Project
Obtaining a list of historical cluster upgrade booting tasks	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows	cce:cluster:get	√	√
Obtaining details about a specified cluster upgrade booting task	**[CCE standard and CCE Turbo clusters]** GET /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows/{upgrade_workflow_id} **[CCE Autopilot clusters]** GET /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows/{upgrade_workflow_id}	cce:cluster:get	√	√
Updating the status of a cluster upgrade booting task	**[CCE standard and CCE Turbo clusters]** PATCH /api/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows/{upgrade_workflow_id} **[CCE Autopilot clusters]** PATCH /autopilot/v3/projects/{project_id}/clusters/{cluster_id}/operation/upgradeworkflows/{upgrade_workflow_id}	cce:cluster:upgrade	√	√

11.2 Cluster Permissions (IAM Authorization)

CCE cluster-level permissions are assigned based on IAM system-defined policies and custom policies. You can use user groups to assign permissions to IAM users.

CAUTION

- Cluster permissions are granted to users for operating cluster-related resources only (such as clusters and nodes). To operate Kubernetes resources like workloads and Services, you must be granted **namespace permissions** as well.
- When viewing a cluster on the CCE console, the information displayed depends on the namespace permissions. If you have no namespace permissions, you cannot view the resources in the cluster.

Prerequisites

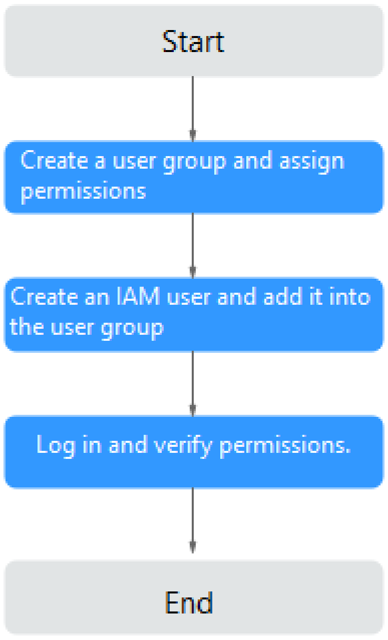
- Before granting permissions to user groups, you need to get familiar with the system policies listed in **Permissions** for CCE. To grant permissions for other services, you need to learn about all **system-defined permissions** supported by IAM.
- A user with the Security Administrator role (for example, your account) has all IAM permissions except role switching. Only these users can view user groups and their permissions on the **Permissions** page on the CCE console.

Configuration Description

On the CCE console, when you choose **Permissions** > **Cluster-Level Permissions** to create a user group, you will be directed to the IAM console to complete the process. After the user group is created and its permissions are configured, you can view the information on the **Cluster-Level Permissions** tab. This section describes the operations in IAM.

Process Flow

Figure 11-2 Process of granting CCE permissions using identity policy-based authorization


Start
Create a user group and assign permissions
Create an IAM user and add it into the user group
Log in and verify permissions.
End

1. **Create a user group and assign permissions**.

   Create a user group on the IAM console, and assign CCE permissions, for example, the **CCE ReadOnlyAccess** policy to the group.

   NOTE

   CCE is deployed by region. On the IAM console, select **Region-specific projects** when assigning CCE permissions.

2. **Create an IAM user and add it to the user group**.

   Create a user on the IAM console and add it to the user group created in **1**.

   NOTICE

   IAM users need programmatic and management console access to use CCE.

3. **Log in** and verify permissions.

   In the authorized region, perform the following operations:

   – Choose **Service List** > **Cloud Container Engine**. Then click **Buy Cluster** on the CCE console. If a message appears indicating that you have insufficient permissions to perform the operation, the **CCE ReadOnlyAccess** policy is in effect.

   – Choose another service from **Service List**. If a message appears indicating that you have insufficient permissions to access the service, the **CCE ReadOnlyAccess** policy is in effect.

System-defined Roles

Roles are a type of coarse-grained authorization mechanism that defines service-level permissions based on user responsibilities. Only a limited number of service-level roles are available for authorization. However, roles are not an ideal choice for fine-grained authorization and secure access control.

The preset system role for CCE in IAM is **CCE Administrator**. When assigning this role to a user group, you must also select other roles and policies on which this role depends, such as **Tenant Guest**, **Server Administrator**, **ELB Administrator**, **OBS Administrator**, **SFS Administrator**, **SWR Admin**, and **APM FullAccess**. For more information about dependencies, see **System Permissions**.

System Policy

The system policies preset for CCE in IAM are **CCE FullAccess** and **CCE ReadOnlyAccess**.

- **CCE FullAccess**: common operation permissions on CCE cluster resources, excluding the namespace-level permissions for the clusters (with Kubernetes RBAC enabled) and the privileged administrator operations, such as agency configuration and cluster certificate generation
- **CCE ReadOnlyAccess**: permissions to view CCE cluster resources, excluding the namespace-level permissions of the clusters (with Kubernetes RBAC enabled)

NOTE

When purchasing a CCE standard or CCE Turbo cluster or a node that is billed on a yearly/monthly basis, add **custom policies** and configure payment permissions such as **bss:*:*** for the Billing Center.

In **Table 11-13** and **Table 11-14**, "√" indicates that the action is supported for the cluster type, and "x" indicates that the action is not supported for the cluster type.

Table 11-13 Permissions granted by CCE FullAccess

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
cce:*:*	cce:cluster:create	√	√	Creating a cluster
	cce:cluster:delete	√	√	Deleting a cluster
	cce:cluster:update	√	√	Updating a cluster, for example, updating cluster node scheduling parameters and providing RBAC support to clusters
	cce:cluster:upgrade	√	√	Upgrading a cluster
	cce:cluster:start	√	×	Waking up a cluster
	cce:cluster:stop	√	×	Hibernating a cluster
	cce:cluster:list	√	√	Listing all clusters
	cce:cluster:get	√	√	Obtaining cluster details
	cce:node:create	√	×	Adding a node
	cce:node:delete	√	×	Deleting one or more nodes
	cce:node:update	√	×	Updating a node, for example, updating the node name
	cce:node:get	√	×	Obtaining node details
	cce:node:list	√	×	Listing all nodes

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
	cce:nodepool:create	√	×	Creating a node pool
	cce:nodepool:delete	√	×	Deleting a node pool
	cce:nodepool:update	√	×	Updating a node pool
	cce:nodepool:get	√	×	Obtaining a node pool
	cce:nodepool:list	√	×	Listing all node pools in a cluster
	cce:release:create	√	√	Creating a release
	cce:release:delete	√	√	Deleting a release
	cce:release:update	√	√	Updating a release
	cce:job:list	√	√	Listing all cluster jobs
	cce:job:delete	√	√	Deleting one or more cluster jobs
	cce:job:get	√	√	Obtaining job details
	cce:storage:create	√	√	Creating a storage volume
	cce:storage:delete	√	√	Deleting a storage volume
	cce:storage:list	√	√	Listing all storage volumes
	cce:addonInstance:create	√	√	Creating an add-on instance
	cce:addonInstance:delete	√	√	Deleting an add-on instance
	cce:addonInstance:update	√	√	Updating an add-on instance

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
	cce:addonInstance:get	√	√	Obtaining an add-on instance
	cce:addonTemplate:get	√	√	Obtaining an add-on template
	cce:addonInstance:list	√	√	Listing all add-on instances
	cce:addonTemplate:list	√	√	Listing all add-on templates
	cce:chart:list	√	√	Listing all charts
	cce:chart:delete	√	√	Deleting a chart
	cce:chart:update	√	√	Updating a chart
	cce:chart:upload	√	√	Uploading a chart
	cce:chart:get	√	√	Obtaining chart details
	cce:release:get	√	√	Obtaining release details
	cce:release:list	√	√	Listing all releases
	cce:userAuthorization:get	√	√	Obtaining CCE user authorization
	cce:userAuthorization:create	√	√	Creating CCE user authorization
nat:*:*	-	×	√	Performing all operations on NAT Gateway resources
nat:*:get	-	√	√	Viewing NAT Gateway resource details
nat:*:list	-	√	√	Listing all NAT Gateway resources
vpcep:*:*	-	×	√	Performing all operations on VPC Endpoint resources
ecs:*:*	-	√	√	Performing all operations on ECSs

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
evs:*:*	-	√	√	Performing all operations on EVS disks EVS disks can be attached to cloud servers and expanded to a higher capacity whenever needed.
vpc:*:*	-	√	√	Performing all operations on a VPC, including the shared load balancers in the VPC. A cluster must run in a VPC. When creating a namespace, you need to create or associate a VPC for the namespace so that all containers in the namespace will run in the VPC.
bms:*:get*	-	√	√	Viewing BMS resource details
bms:*:list*	-	√	√	Listing all BMS resources
ims:*:get*	-	√	√	Viewing IMS resource details
ims:*:list*	-	√	√	Listing all IMS resources
elb:*:get	-	√	√	Viewing ELB resource details
elb:*:list	-	√	√	Listing all ELB resources
sfs:*:get*	-	√	√	Viewing SFS resource details
sfs:shares:ShareAction	-	√	√	Sharing SFS resources for scaling
sfsturbo:*:get*	-	√	√	Viewing SFS Turbo resource details

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
sfsturbo:shares:ShareAction	-	√	√	Sharing SFS Turbo resources for scaling
tms:resourceTags:list	-	√	√	Listing TMS resources
kps:domainKeypairs:list	-	√	√	Listing DEW SSH keys
kps:domainKeypairs:get	-	√	√	Viewing DEW SSH keys
kms:cmk:get	-	√	√	Viewing DEW keys
kms:cmk:list	-	√	√	Listing DEW keys
aom:*:get	-	√	√	Viewing AOM resource details
aom:*:list	-	√	√	Listing AOM resources
aom:autoScalingRule:*	-	√	√	Performing all operations on AOM auto scaling rules
apm:icmgr:*	-	√	√	Performing operations on the ICAgent in APM
lts:*:*	-	√	√	Performing all operations on LTS
smn:*:*	-	√	√	Performing all operations on SMN

Table 11-14 Permissions granted by CCE ReadOnlyAccess

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
cce:*:get	cce:cluster:get	√	√	Obtaining cluster details
	cce:node:get	√	×	Obtaining node details
	cce:job:get	√	√	Obtaining job details
	cce:addonInstance:get	√	√	Obtaining an add-on instance
	cce:addonTemplate:get	√	√	Obtaining an add-on template
	cce:chart:get	√	√	Obtaining chart details
	cce:nodepool:get	√	×	Obtaining a node pool
	cce:release:get	√	√	Obtaining release details
	cce:userAuthorization:get	√	√	Obtaining CCE user authorization
cce:*:list	cce:cluster:list	√	√	Listing all clusters
	cce:node:list	√	×	Listing all nodes
	cce:job:list	√	√	Listing all cluster jobs
	cce:addonInstance:list	√	√	Listing all add-on instances
	cce:addonTemplate:list	√	√	Listing all add-on templates
	cce:chart:list	√	√	Listing all charts
	cce:nodepool:list	√	×	Listing all node pools in a cluster
	cce:release:list	√	√	Listing all releases

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
	cce:storage:list	√	√	Listing all volumes
cce:kubernetes:*	-	√	√	Performing operations on all Kubernetes resources. For details, see **Namespace Permissions**.
nat:*:get	-	√	√	Viewing NAT Gateway resource details
nat:*:list	-	√	√	Listing all NAT Gateway resources
vpcep:*:get	-	×	√	Viewing VPC Endpoint resource details
vpcep:*:list	-	×	√	Listing all VPC Endpoint resources
ecs:*:get	-	√	√	Viewing details of all resources on an ECS
ecs:*:list	-	√	√	Listing all resources on an ECS
bms:*:get*	-	√	√	Viewing BMS resource details
bms:*:list	-	√	√	Listing all BMS resources
ims:*:get*	-	√	√	Viewing IMS resource details
ims:*:list*	-	√	√	Listing all IMS resources
evs:*:get	-	√	√	Viewing details about all EVS disk resources. EVS disks can be attached to cloud servers and expanded to a higher capacity whenever needed.
evs:*:list	-	√	√	Listing all EVS resources
evs:*:count	-	√	√	-

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
vpc:*:get	-	√	√	Viewing details about all VPC resources A cluster must run in a VPC. When creating a namespace, you need to create or associate a VPC for the namespace so that all containers in the namespace will run in the VPC.
vpc:*:list	-	√	√	Listing all VPC resources
elb:*:get	-	√	√	Viewing ELB resource details
elb:*:list	-	√	√	Listing all ELB resources
sfs:*:get*	-	√	√	Viewing SFS resource details
sfs:shares:ShareAction	-	√	√	Sharing SFS resources for scaling
sfsturbo:*:get*	-	√	√	Viewing SFS Turbo resource details
sfsturbo:shares:ShareAction	-	√	√	Sharing SFS Turbo resources for scaling
tms:resourceTags:list	-	√	√	Listing all TMS resources
kps:domainKeypairs:list	-	√	√	Listing DEW SSH keys
kps:domainKeypairs:get	-	√	√	Viewing DEW SSH keys
kms:cmk:get	-	√	√	Viewing DEW keys
kms:cmk:list	-	√	√	Listing DEW keys

Action	Specific Action	CCE Standard or CCE Turbo Clusters	CCE Autopilot Clusters	Description
aom:*:get	-	√	√	Viewing details about all AOM resources
aom:*:list	-	√	√	Listing all AOM resources
aom:autoScalingRule:*	-	√	√	Performing all operations on AOM auto scaling rules
lts:*:get	-	√	√	Viewing details about all LTS resources
lts:*:list	-	√	√	Listing all LTS resources
smn:*:get	-	√	√	Viewing details about all SMN resources
smn:*:list	-	√	√	Listing SMN resources.

Custom Policies

Custom policies can be created as a supplement to the system-defined policies of CCE. For details about actions supported in custom policies, see **Permissions Policies and Supported Actions**.

To create a custom policy, choose either visual editor or JSON.

- Visual editor: Select cloud services, actions, resources, and request conditions. This does not require knowledge of policy syntax.
- JSON: Edit JSON policies from scratch or based on an existing policy.

For details, see **Creating a Custom Policy**. The following lists examples of common CCE custom policies.

Examples

- Example 1: Grant permission to create a cluster named **test**.
```
{
    "Version": "1.1",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "cce:cluster:create"
            ]
        }
    ]
}
```
- Example 2: Grant permission to deny node deletion.

A policy with only "Deny" permissions must be used together with other policies. If the permissions granted to an IAM user contain both "Allow" and "Deny", the "Deny" permissions take precedence over the "Allow" permissions.

Assume that you want to grant the permissions of the **CCEFullAccess** to a user but want to prevent them from deleting nodes (**cce:node:delete**). You can create a custom policy for denying node deletion, and attach this policy together with the **CCEFullAccess** policy to the user. As an explicit deny in any policy overrides any allows, the user can perform all operations on nodes excepting deleting them. The following is an example of a deny policy:

```
{
    "Version": "1.1",
    "Statement": [
        {
            "Effect": "Deny",
            "Action": [
                "cce:node:delete"
            ]
        }
    ]
}
```

- Example 3: Create a custom policy containing multiple actions.

  A custom policy can contain the actions of multiple services that are of the global or project-level type. The following is an example policy containing actions of multiple services:

```
{
    "Version": "1.1",
    "Statement": [
        {
            "Action": [
                "ecs:cloudServers:resize",
                "ecs:cloudServers:delete",
                "ecs:cloudServers:delete",
                "ims:images:list",
                "ims:serverImages:create"
            ],
            "Effect": "Allow"
        }
    ]
}
```

CCE Cluster Permissions and Enterprise Projects

CCE supports resource management and permission allocation by cluster and enterprise project.

Note that:

- IAM projects are based on physical isolation of resources, whereas enterprise projects provide global logical groups of resources, which better meet the actual requirements of enterprises. In addition, IAM policies can be managed based on enterprise projects. Therefore, use enterprise projects for permissions management. For details, see **Creating an Enterprise Project**.
- When there are both IAM projects and enterprise projects, IAM preferentially matches the IAM project policies.
- When creating a cluster or node using purchased cloud resources, ensure that IAM users have been granted the required permissions in the enterprise project to use these resources. Otherwise, the cluster or node may fail to be created.

- If a resource does not support enterprise projects, the permissions granted to the resource will not take effect.

Table 11-15 Enterprise project resource pool

Support for Enterprise Projects	Resources	Description
Supported	cluster	Clusters
	node	Nodes
	nodepool	Node pools
	job	Jobs
	tag	Cluster labels
	addonInstance	Add-on instances
	release	Helm releases
	storage	Storage
Not supported	quota	Cluster quotas
	chart	Charts
	addonTemplate	Add-on templates

CCE Cluster Permissions and IAM RBAC

CCE is compatible with IAM system roles for permissions management. Use fine-grained policies provided by IAM to simplify permissions management.

CCE supports the following roles:

- Basic IAM roles:
  – **te_admin (Tenant Administrator)**: Users with this role can call all APIs of all services except IAM.
  – **readonly (Tenant Guest)**: Users with this role can call APIs with the read-only permissions of all services except IAM.
- Custom CCE administrator role: **CCE Administrator**

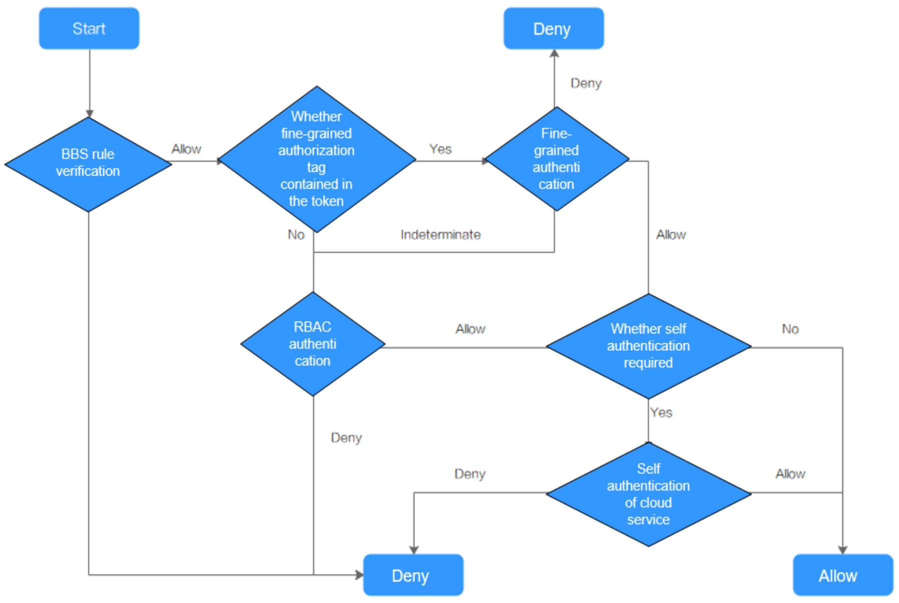
NOTE

If a user has the **Tenant Administrator** or **CCE Administrator** system-defined role, the user has the cluster-admin permissions in Kubernetes RBAC and the permissions cannot be removed after the cluster is created.

If the user is the cluster creator, the **cluster-admin** permissions in Kubernetes RBAC are granted to the user by default. The permissions can be manually removed after the cluster is created.

- Method 1: Choose **Permissions Management** > **Namespace-Level Permissions** > **Delete** in the same role as **cluster-creator** on the CCE console.
- Method 2: Delete **ClusterRoleBinding: cluster-creator** through the API or kubectl.

When RBAC and IAM policies co-exist, the backend authentication logic for open APIs or console operations on CCE is as follows.


Start
Deny
Deny
BBS rule verification
Allow
Whether fine-grained authorization tag contained in the token
Yes
Fine-grained authentication
No
Indeterminate
Allow
RBAC authentication
Allow
Whether self authentication required
No
Yes
Deny
Deny
Self authentication of cloud service
Allow
Deny
Allow

11.3 Namespace Permissions (Kubernetes RBAC-based)

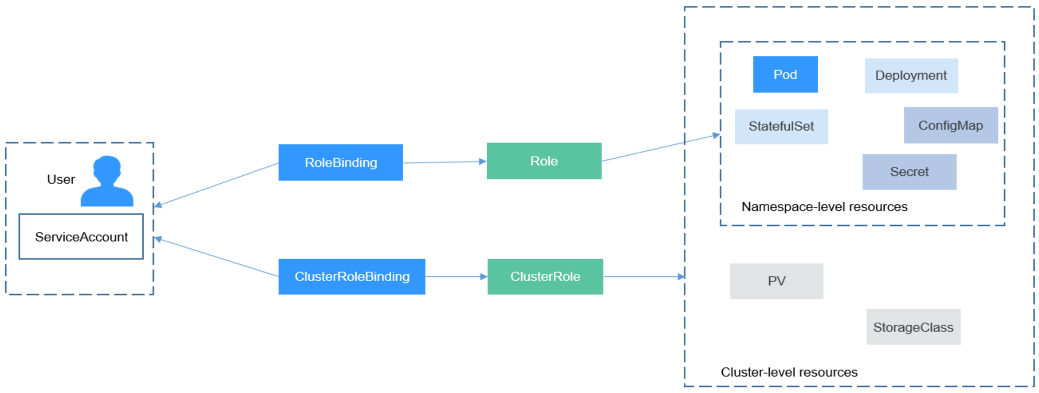
Namespace Permissions (Kubernetes RBAC-based)

You can regulate users' or user groups' access to Kubernetes resources in a single namespace based on their Kubernetes RBAC roles. The RBAC API declares four kinds of Kubernetes objects: Role, ClusterRole, RoleBinding, and ClusterRoleBinding, which are described as follows:

- Role: defines a set of rules for accessing Kubernetes resources in a namespace.
- RoleBinding: defines the relationship between users and roles.
- ClusterRole: defines a set of rules for accessing Kubernetes resources in a cluster (including all namespaces).
- ClusterRoleBinding: defines the relationship between users and cluster roles.

Role and ClusterRole specify actions that can be performed on specific resources. RoleBinding and ClusterRoleBinding bind roles to specific users, user groups, or ServiceAccounts.

Figure 11-3 Role binding



On the CCE console, you can assign permissions to a user or user group to access resources in one or multiple namespaces. By default, the CCE console provides the following ClusterRoles:

- view (read-only): read-only permission on most resources in all or selected namespaces.
- edit (development): read and write permissions on most resources in all or selected namespaces. If this ClusterRole is configured for all namespaces, its capability is the same as the O&M permission.
- admin (O&M): read and write permissions on most resources in all namespaces, and read-only permission on nodes, storage volumes, namespaces, and quota management.
- cluster-admin (administrator): read and write permissions on all resources in all namespaces.
- drainage-editor: node drainage operation permission.
- drainage-viewer: read-only permission for node drainage. Users with this permission can only view the node drainage status but cannot drain a node.
- geip-editor-role: read and write permissions on global EIPs in a cluster.
- icagent-clusterRole: reporting Kubernetes event logs to LTS.

In addition to the preceding typical ClusterRoles, you can define Role and RoleBinding to grant the permissions to add, delete, modify, and obtain global resources (such as nodes, PVs, and CustomResourceDefinitions) and different resources (such as pods, Deployments, and Services) in namespaces for refined permission control.

Cluster Permissions (IAM-based) and Namespace Permissions (Kubernetes RBAC-based)

Users with different cluster permissions (assigned using IAM) have different namespace permissions (assigned using Kubernetes RBAC). **Table 11-16** lists the namespace permissions of different users.

Table 11-16 Differences in namespace permissions

User	Clusters of v1.13 and Later
User with the Tenant Administrator permissions (for example, an account)	All namespace permissions
IAM user with the CCE Administrator role	All namespace permissions
IAM user with the CCE FullAccess or CCE ReadOnlyAccess role	Requires Kubernetes RBAC authorization.
IAM user with the Tenant Guest role	Requires Kubernetes RBAC authorization.

Precautions

- After you create a cluster, CCE automatically assigns the cluster-admin permission to you, which means you have full control on all resources in all namespaces in the cluster. The ID of a federated user changes upon each login and logout. Therefore, the user with the permissions is displayed as deleted. In this case, do not delete the permissions. Otherwise, the authentication fails. You are advised to grant the cluster-admin permission to a user group on CCE and add federated users to the user group.
- A user with the Security Administrator role has all IAM permissions except role switching. For example, an account in the admin user group has this role by default. Only these users can assign permissions on the **Permissions** page on the CCE console.

Configuring Namespace Permissions (on the Console)

You can regulate users' or user groups' access to Kubernetes resources in a single namespace based on their Kubernetes RBAC roles.

Step 1 Log in to the CCE console. In the navigation pane, choose **Permissions**.

Step 2 Select a cluster for which you want to add permissions from the drop-down list on the right.

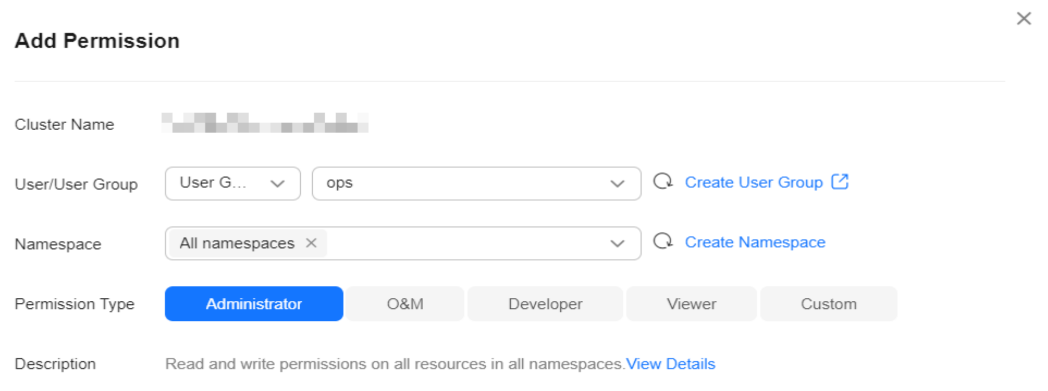
Step 3 In the upper right corner, click **Add Permission**.

Step 4 Confirm the cluster name and select the namespace to assign permissions for. For example, select **All namespaces**, the target user or user group, and select the permissions.

NOTE

If you do not have IAM permissions, you cannot select users or user groups when configuring permissions for other users or user groups. In this case, you can enter a user ID or user group ID.

Figure 11-4 Configuring namespace permissions


Add Permission

Cluster Name

User/User Group: User G... | ops

Namespace: All namespaces

Permission Type: Administrator | O&M | Developer | Viewer | Custom

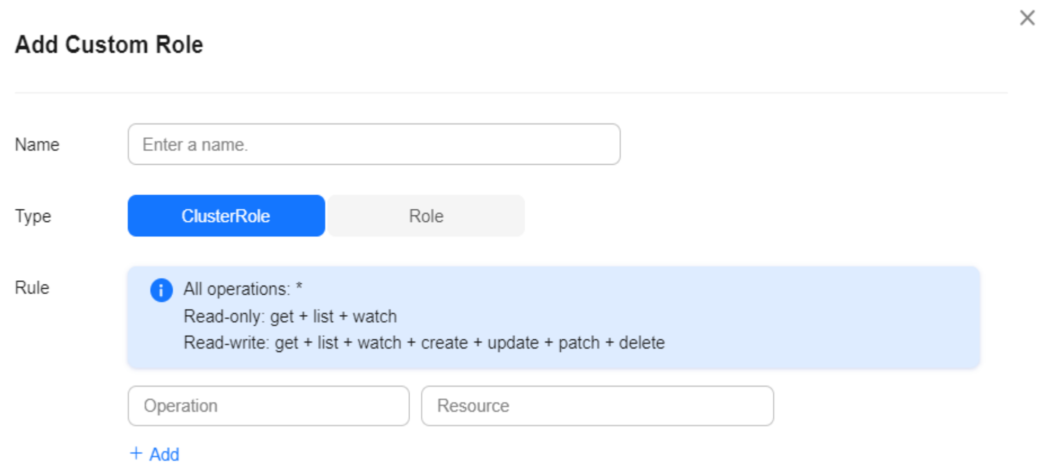
Description: Read and write permissions on all resources in all namespaces. View Details

Permissions can be customized as required. After selecting **Custom** for **Permission Type**, click **Add Custom Role** on the right of the **Custom** parameter. In the dialog box displayed, enter a name and select a rule. After the custom rule is created, you can select a value from the **Custom** drop-down list box.

Custom permissions are classified into ClusterRole and Role. Each ClusterRole or Role contains a group of rules that represent related permissions. For details, see **Using RBAC Authorization**.

- **ClusterRole**: a cluster-level resource that can be used to configure cluster access permissions.
- **Role**: used to configure access permissions in a namespace. When creating a Role, specify the namespace to which the Role belongs.

Figure 11-5 Custom


Add Custom Role

Name: Enter a name.

Type: ClusterRole | Role

Rule:
All operations: *
Read-only: get + list + watch
Read-write: get + list + watch + create + update + patch + delete

Operation | Resource

Step 5 Click **OK**.

----End

Using kubectl to Configure Namespace Permissions

NOTE

When you access a cluster using kubectl, CCE uses **kubeconfig.json** generated on the cluster for authentication. This file contains user information, based on which CCE determines which Kubernetes resources can be accessed by kubectl. The permissions recorded in a **kubeconfig.json** file vary from user to user. The permissions that a user has are listed in **Cluster Permissions (IAM-based) and Namespace Permissions (Kubernetes RBAC-based)**.

In addition to cluster-admin, admin, edit, and view, you can define Roles and RoleBindings to configure the permissions to add, delete, modify, and obtain resources, such as pods, Deployments, and Services, in the namespace.

The procedure for creating a Role is very simple. To be specific, specify a namespace and then define rules. The rules in the following example are to allow GET and LIST operations on pods in the default namespace.

```
kind: Role
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  namespace: default                          # Namespace
  name: role-example
rules:
- apiGroups: [""]
  resources: ["pods"]                         # The pod can be accessed.
  verbs: ["get", "list"]                      # The GET and LIST operations can be performed.
```

- **apiGroups** indicates the API group to which the resource belongs.
- **resources** indicates the resources that can be operated. Pods, Deployments, ConfigMaps, and other Kubernetes resources are supported.
- **verbs** indicates the operations that can be performed. **get** indicates querying a specific object, and **list** indicates listing all objects of a certain type. You can also select other value options such as **create**, **update**, and **delete**.

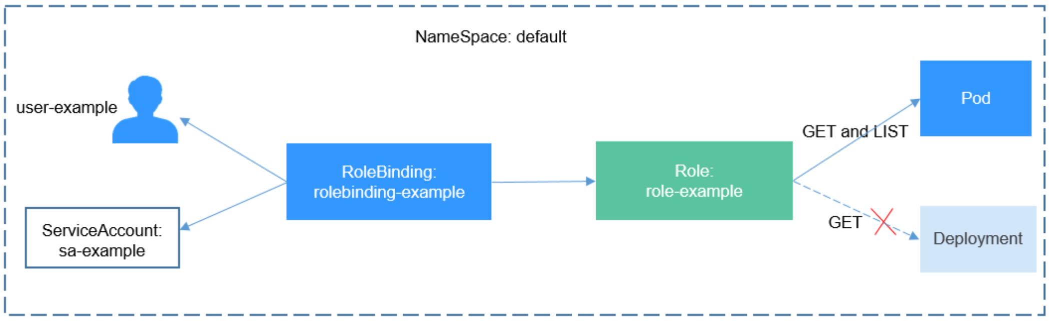
For details, see **Using RBAC Authorization**.

After creating a Role, you can bind the Role to a specific user, which is called RoleBinding. The following shows an example:

```
kind: RoleBinding
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  name: RoleBinding-example
  namespace: default
  annotations:
    CCE.com/IAM: 'true'
roleRef:
  kind: Role
  name: role-example
  apiGroup: rbac.authorization.k8s.io
subjects:
- kind: User
  name: 0c97ac3cb280f4d91fa7c0096739e1f8 # User ID of user-example
  apiGroup: rbac.authorization.k8s.io
```

The **subjects** section binds a Role with an IAM user so that the IAM user can obtain the permissions defined in the Role, like in the following figure.

Figure 11-6 Binding a role to a user



You can also specify a user group in the **subjects** section. In this case, all users in the user group obtain the permissions defined in the Role.

```
subjects:
- kind: Group
  name: 0c96fad22880f32a3f84c009862af6f7    # User group ID
  apiGroup: rbac.authorization.k8s.io
```

Use the IAM user user-example to connect to the cluster and obtain the pod information. The following is an example of the returned pod information.

```
# kubectl get pod
NAME                                    READY   STATUS    RESTARTS   AGE
deployment-389584-2-6f6bd4c574-2n9rk   1/1     Running   0          4d7h
deployment-389584-2-6f6bd4c574-7s5qw   1/1     Running   0          4d7h
deployment-3895841-746b97b455-86g77    1/1     Running   0          4d7h
deployment-3895841-746b97b455-twvpn    1/1     Running   0          4d7h
nginx-658dff48ff-7rkph                 1/1     Running   0          4d9h
nginx-658dff48ff-njdhj                 1/1     Running   0          4d9h
# kubectl get pod nginx-658dff48ff-7rkph
NAME                     READY   STATUS    RESTARTS   AGE
nginx-658dff48ff-7rkph   1/1     Running   0          4d9h
```

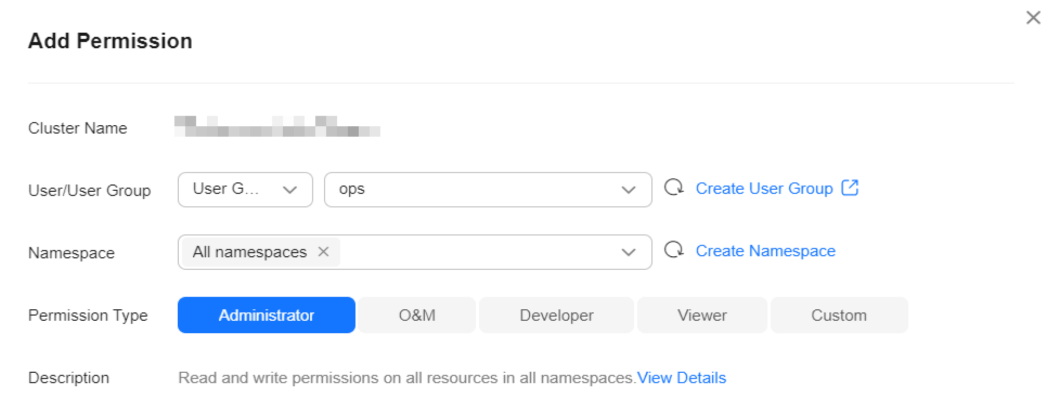
Try querying Deployments and Services in the namespace. The output shows that **user-example** does not have the required permissions. Try querying the pods in namespace kube-system. The output shows that **user-example** does not have the required permissions. This indicates that the IAM user **user-example** has only the GET and LIST Pod permissions in the **default** namespace, which is the same as expected.

```
# kubectl get deploy
Error from server (Forbidden): deployments.apps is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot list resource "deployments" in API group "apps" in the namespace "default"
# kubectl get svc
Error from server (Forbidden): services is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot list resource "services" in API group "" in the namespace "default"
# kubectl get pod --namespace=kube-system
Error from server (Forbidden): pods is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot list resource "pods" in API group "" in the namespace "kube-system"
```

Example: Assigning Cluster Administrator Permissions (cluster-admin)

You can use the cluster-admin role to assign all permissions on a cluster. This role contains the permissions for all cluster resources.

Figure 11-7 Assigning cluster administrator permissions (cluster-admin)


Add Permission

Cluster Name

User/User Group: User G... ops

Namespace: All namespaces

Permission Type: Administrator

Description: Read and write permissions on all resources in all namespaces.View Details

In the following example kubectl output, a ClusterRoleBinding has been created and the cluster-admin role is bound to the user group **cce-role-group**.

```
# kubectl get clusterrolebinding
NAME                                                    ROLE                          AGE
clusterrole_cluster-admin_group0c96fad22880f32a3f84c009862af6f7   ClusterRole/cluster-admin      61s

# kubectl get clusterrolebinding clusterrole_cluster-admin_group0c96fad22880f32a3f84c009862af6f7 -oyaml
apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRoleBinding
metadata:
  annotations:
    CCE.com/IAM: "true"
  creationTimestamp: "2021-06-23T09:15:22Z"
  name: clusterrole_cluster-admin_group0c96fad22880f32a3f84c009862af6f7
  resourceVersion: "36659058"
  selfLink: /apis/rbac.authorization.k8s.io/v1/clusterrolebindings/clusterrole_cluster-
admin_group0c96fad22880f32a3f84c009862af6f7
  uid: d6cd43e9-b4ca-4b56-bc52-e36346fc1320
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: cluster-admin
subjects:
- apiGroup: rbac.authorization.k8s.io
  kind: Group
  name: 0c96fad22880f32a3f84c009862af6f7
```

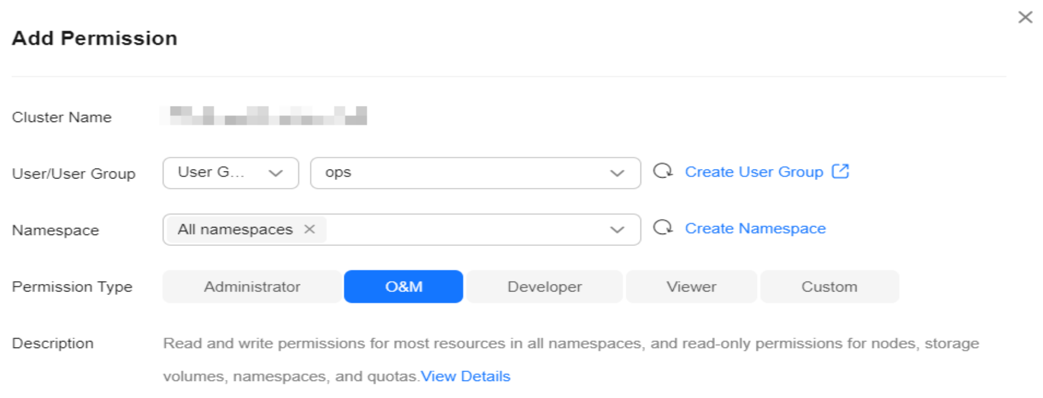
Connect to the cluster as an authorized user. If the PVs and StorageClasses can be queried, the permission configuration takes effect.

```
# kubectl get pv
No resources found
# kubectl get sc
NAME                PROVISIONER                   RECLAIMPOLICY   VOLUMEBINDINGMODE
ALLOWVOLUMEEXPANSION   AGE
csi-disk            everest-csi-provisioner        Delete          Immediate              true                75d
csi-disk-topology   everest-csi-provisioner        Delete          WaitForFirstConsumer   true                75d
csi-nas             everest-csi-provisioner        Delete          Immediate              true                75d
csi-obs             everest-csi-provisioner        Delete          Immediate              false               75d
csi-sfsturbo        everest-csi-provisioner        Delete          Immediate              true                75d
```

Example: Assigning Namespace O&M Permissions (admin)

The admin role has the read and write permissions on most namespace resources. You can grant the admin permission on all namespaces to a user or user group.

Figure 11-8 Assigning O&M permissions on all namespaces (admin)



Add Permission

Cluster Name [redacted]

User/User Group: User G... | ops — Create User Group

Namespace: All namespaces — Create Namespace

Permission Type: Administrator | **O&M** | Developer | Viewer | Custom

Description: Read and write permissions for most resources in all namespaces, and read-only permissions for nodes, storage volumes, namespaces, and quotas. View Details

In the following example kubectl output, a RoleBinding has been created and the admin role is bound to the user group **cce-role-group**.

```
# kubectl get rolebinding
NAME                                                      ROLE                  AGE
clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7   ClusterRole/admin    18s
# kubectl get rolebinding clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7 -oyaml
apiVersion: rbac.authorization.k8s.io/v1
kind: RoleBinding
metadata:
  annotations:
    CCE.com/IAM: "true"
  creationTimestamp: "2021-06-24T01:30:08Z"
  name: clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7
  resourceVersion: "36963685"
  selfLink: /apis/rbac.authorization.k8s.io/v1/namespaces/default/rolebindings/
clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7
  uid: 6c6f46a6-8584-47da-83f5-9eef1f7b75d6
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: admin
subjects:
- apiGroup: rbac.authorization.k8s.io
  kind: Group
  name: 0c96fad22880f32a3f84c009862af6f7
```

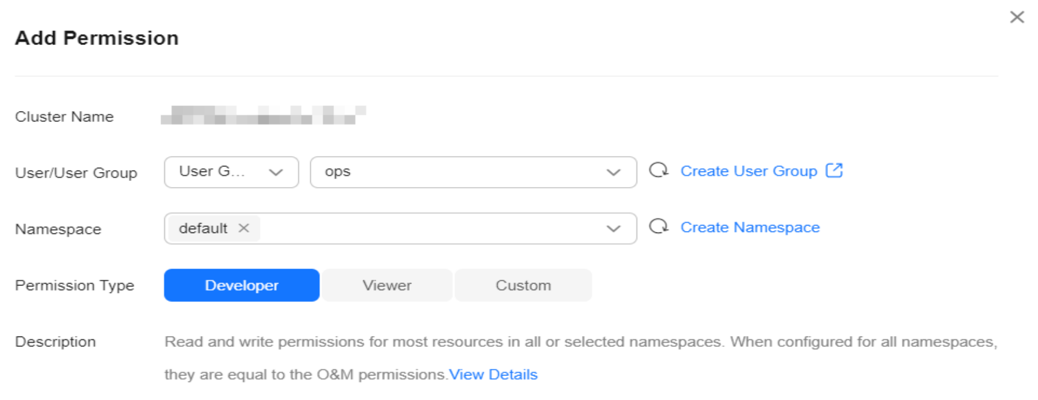
Connect to the cluster as an authorized user. If the PVs and StorageClasses can be queried but a namespace cannot be created, the permission configuration takes effect.

```
# kubectl get pv
No resources found
# kubectl get sc
NAME                PROVISIONER                   RECLAIMPOLICY   VOLUMEBINDINGMODE
ALLOWVOLUMEEXPANSION    AGE
csi-disk            everest-csi-provisioner       Delete          Immediate              true                  75d
csi-disk-topology   everest-csi-provisioner       Delete          WaitForFirstConsumer   true                  75d
csi-nas             everest-csi-provisioner       Delete          Immediate              true                  75d
csi-obs             everest-csi-provisioner       Delete          Immediate              false                 75d
csi-sfsturbo        everest-csi-provisioner       Delete          Immediate              true                  75d
# kubectl apply -f namespaces.yaml
Error from server (Forbidden): namespaces is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot
create resource "namespaces" in API group "" at the cluster scope
```

Example: Assigning Namespace Developer Permissions (edit)

The edit role has the read and write permissions on most namespace resources. You can grant the edit permission on all namespaces to a user or user group.

Figure 11-9 Assigning developer permissions on the default namespace (edit)


Add Permission

Cluster Name

User/User Group: User G... | ops — Create User Group

Namespace: default — Create Namespace

Permission Type: Developer | Viewer | Custom

Description: Read and write permissions for most resources in all or selected namespaces. When configured for all namespaces, they are equal to the O&M permissions.View Details

In the following example kubectl output, a RoleBinding has been created, the edit role is bound to the user group **cce-role-group**, and the target namespace is the default namespace.

```
# kubectl get rolebinding
NAME                                                   ROLE                 AGE
clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7   ClusterRole/admin    18s
# kubectl get rolebinding clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7 -oyaml
apiVersion: rbac.authorization.k8s.io/v1
kind: RoleBinding
metadata:
  annotations:
    CCE.com/IAM: "true"
  creationTimestamp: "2021-06-24T01:30:08Z"
  name: clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7
  namespace: default
  resourceVersion: "36963685"
  selfLink: /apis/rbac.authorization.k8s.io/v1/namespaces/default/rolebindings/
clusterrole_admin_group0c96fad22880f32a3f84c009862af6f7
  uid: 6c6f46a6-8584-47da-83f5-9eef1f7b75d6
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: edit
subjects:
- apiGroup: rbac.authorization.k8s.io
  kind: Group
  name: 0c96fad22880f32a3f84c009862af6f7
```

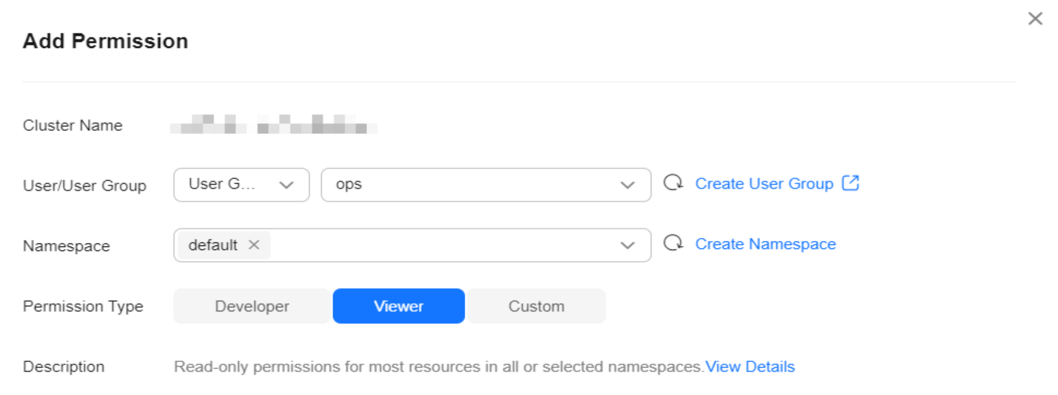
Connect to the cluster as an authorized user. In this example, you can create and obtain resources in the default namespace, but cannot query resources in the kube-system namespace or cluster resources.

```
# kubectl get pod
NAME                     READY   STATUS    RESTARTS   AGE
test-568d96f4f8-brdrp   1/1     Running   0          33m
test-568d96f4f8-cgjqp   1/1     Running   0          33m
# kubectl get pod -nkube-system
Error from server (Forbidden): pods is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot list
resource "pods" in API group "" in the namespace "kube-system"
# kubectl get pv
Error from server (Forbidden): persistentvolumes is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8"
cannot list resource "persistentvolumes" in API group "" at the cluster scope
```

Example: Assigning Read-Only Namespace Permissions (view)

The view role has the read-only permissions on a namespace. You can assign permissions to users to view one or multiple namespaces.

Figure 11-10 Assigning read-only namespace permissions (view)



Add Permission

Cluster Name

User/User Group: User G... ops

Namespace: default

Permission Type: Viewer

Description: Read-only permissions for most resources in all or selected namespaces.View Details

In the following example kubectl output, a RoleBinding has been created, the view role is bound to the user group **cce-role-group**, and the target namespace is the default namespace.

```
# kubectl get rolebinding
NAME                                                ROLE               AGE
clusterrole_view_group0c96fad22880f32a3f84c009862af6f7   ClusterRole/view   7s

# kubectl get rolebinding clusterrole_view_group0c96fad22880f32a3f84c009862af6f7 -oyaml
apiVersion: rbac.authorization.k8s.io/v1
kind: RoleBinding
metadata:
  annotations:
    CCE.com/IAM: "true"
  creationTimestamp: "2021-06-24T01:36:53Z"
  name: clusterrole_view_group0c96fad22880f32a3f84c009862af6f7
  namespace: default
  resourceVersion: "36965800"
  selfLink: /apis/rbac.authorization.k8s.io/v1/namespaces/default/rolebindings/
clusterrole_view_group0c96fad22880f32a3f84c009862af6f7
  uid: b86e2507-e735-494c-be55-c41a0c4ef0dd
roleRef:
  apiGroup: rbac.authorization.k8s.io
  kind: ClusterRole
  name: view
subjects:
- apiGroup: rbac.authorization.k8s.io
  kind: Group
  name: 0c96fad22880f32a3f84c009862af6f7
```

Connect to the cluster as an authorized user. In this example, you can query resources in the default namespace but cannot create resources.

```
# kubectl get pod
NAME                    READY   STATUS    RESTARTS   AGE
test-568d96f4f8-brdrp   1/1     Running   0          40m
test-568d96f4f8-cgjqp   1/1     Running   0          40m
# kubectl run -i --tty --image tutum/dnsutils dnsutils --restart=Never --rm /bin/sh
Error from server (Forbidden): pods is forbidden: User "0c97ac3cb280f4d91fa7c0096739e1f8" cannot create
resource "pods" in API group "" in the namespace "default"
```

Example: Assigning Permissions for a Specific Kubernetes Resource Object

You can assign permissions on a specific Kubernetes resource object, such as pod, Deployment, and Service. For details, see **Using kubectl to Configure Namespace Permissions**.

12 Settings

12.1 Dashboard

The **Dashboard** tab provides basic cluster configuration information, including **Cluster Information**, **Cluster Settings**, and **Installed Add-ons**.

Navigation Path

Step 1 Log in to the CCE console and click the name of the target cluster to access the cluster console.

Step 2 In the navigation pane, choose **Cluster** > **Settings**. Then click the **Dashboard** tab.

----End

Cluster Information

It includes:

- **ID**: uniquely identifies a cluster resource. It is automatically generated after a cluster is created and can be used in scenarios such as API calling.
- **Current Name**: After a cluster is created, you can click ✎ to change its name.
- **Original Name**: specifies a cluster's original name after its current name is changed. The current name of a cluster must be unique.
- **Status**: specifies the status of the cluster. For details, see **Cluster Lifecycle**.
- **Type**: This is a CCE Autopilot cluster.
- **Created On**: specifies the time when a cluster was created. You will be billed based on the creation time of clusters.

Cluster Settings

After a cluster is created, you can view the following information:

- **Cluster Version | Patch Version**: specifies the Kubernetes version and CCE patch version of the cluster.

- **Network Model**: Only cloud native network 2.0 is supported.
- **Billing Mode**: Only pay-per-use is supported.
- **Enterprise Project**: specifies the enterprise project to which a cluster belongs. For more details, see **Enterprise Management**.
- **Operation Protection**: If operation protection is enabled, you need to confirm the operation through the virtual MFA device, SMS messages, or emails. You can go to the IAM console, choose **Security Settings** > **Critical Operations**, and enable operation protection.
- **Resource Tag**: You can add resource tags to classify resources.
- **Cluster Description**: specifies the description that you entered for a cluster. A maximum of 200 characters are allowed.

Installed Add-ons

This module allows you to view the add-ons installed in the cluster. If there are add-ons that can be upgraded in the cluster, click **Go to Upgrade** to view the add-ons on the **Add-ons** page. In the upper right corner, select **All** to view more details on the **Add-ons** page.

12.2 Cluster Access

Access Mode

- **kubectl**: You need to download and configure the kubectl and kubeconfig configuration files first, and then use kubectl to access a Kubernetes cluster. For details, see **Connecting to a Cluster Using kubectl**.
- **EIP**: You can bind an EIP to the cluster API server. Then, the cluster API server can access the public network.

NOTE

- Binding an EIP to a cluster make the cluster prone to attacks from the public network. You are advised to change the inbound security group rule for port 5443 on the master nodes. For details, see **How Can I Configure a Security Group Rule in a Cluster?**
- This operation will restart kube-apiserver and update the cluster access certificate (kubeconfig). Do not perform any operations on the cluster during this time.

Authentication

CCE Autopilot allows you to download the X509 certificate, which contains the **client.key**, **client.crt**, and **ca.crt** files. Keep your certificate secure.

For details about how to use a certificate to access clusters, see **Connecting to a Cluster Using an X.509 Certificate**.

12.3 Network

On the Network tab, you can configure cluster and container CIDR blocks.

Cluster Network

Table 12-1 Parameters

Parameter	Description
VPC	VPC where the cluster resides. A VPC provides an isolated virtual network for cloud servers, cloud containers, and cloud databases. You can configure and manage the network as required. You can configure a CIDR block for the VPC, create subnets, configure security groups, and even assign EIPs and allocate bandwidth in your network, enabling secure and easy access to your service system.
VPC CIDR Block	IP address range of the VPC for the cluster.
Network Model	Only cloud native network 2.0 is supported.

Service Settings

Service CIDR Block: IP address range of the ClusterIP service for the cluster. The Service CIDR block cannot overlap with the container subnet CIDR block and can be used only in the cluster.

Container Network

If you need to set a different container subnet or security group for a namespace or workload, you can create a custom container network configuration and associate the namespace or workload. For details, see **Associating a Subnet and Security Group with a Namespace or Workload Using a Container Network Configuration**.

- Associating containers with a subnet: Container IP addresses are allocated from this subnet. Only the containers in the same namespace or for running the same workload can communicate with each other.
- Associating containers with a security group: You can configure security group rules for containers in the same namespace or for running the same workload to allow or deny traffic to and from the containers.

NOTE

Container subnets can be deleted for clusters v1.27.8-r0, v1.28.6-r0, or later.

12.4 Monitoring

CCE Autopilot monitors applications and resources and collects metrics and events to analyze application health. The **Monitoring** tab allows you to modify parameters for better O&M.

You will need to enable Monitoring Center by referring to **Enabling Cluster Monitoring**.

Monitoring Configuration

Collection Configuration

Default Collection Period: specifies the metric collection period of the Cloud Native Cluster Monitoring add-on. The default value is 15 seconds.

Interconnection with AOM

AOM uses Prometheus for container monitoring. To enable monitoring, you need to select an AOM instance first. After you select the option for reporting the monitoring data to AOM, metrics will be reported to the AOM instance you select. Basic container metrics are free, and other metrics are billed on a pay-per-use basis. For details about free metrics, see **Basic Container Metrics**.

Log Configuration

Collection Configuration

CCE Autopilot cluster logs, including container logs and Kubernetes events, are collected. You can quickly query and analyze the collected logs.

NOTE

A default log group named **k8s-logs-***{Cluster ID}* will be created for CCE Autopilot cluster logs reported to LTS, and you will be billed for log storage on LTS. For details about LTS pricing, see **Price Calculator**.

Table 12-2 Cluster logs and events

Item	Log Type	LTS Log Group Name	Status	Reference
Container logs	Container standard output	stdout-{clusterId}	To enable log collection, you need to install the Cloud Native Log Collection add-on by referring to **Cloud Native Log Collection**.	**Collecting Logs**
Kubernetes events	Kubernetes events	event-{clusterId}	To enable event collection, you need to install the Cloud Native Log Collection add-on by referring to **Cloud Native Log Collection**.	**Collecting Kubernetes Events**

Kubernetes Events Reported to AOM

Once the Cluster Native Logging add-on (**Cloud Native Log Collection**) is installed in a cluster, Kubernetes events are reported to LTS by default, while this feature can be used to report Kubernetes events to AOM.

- Abnormal events: This option is enabled by default. All abnormal events are reported to AOM. You can configure a blocklist to add events that do not need to be reported to the blocklist. You can query event names in **CCE Autopilot Cluster Events**.
- Normal events: If this option is enabled, normal events will be reported to AOM. CCE Autopilot is pre-configured to report some normal events. If you need custom events reporting, configure a trustlist and add the events that need to be reported to the trustlist. You can query event names in **CCE Autopilot Cluster Events**.