Cloud Container Engine Autopilot

Service Overview

Issue 01

Date 2025-01-03

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

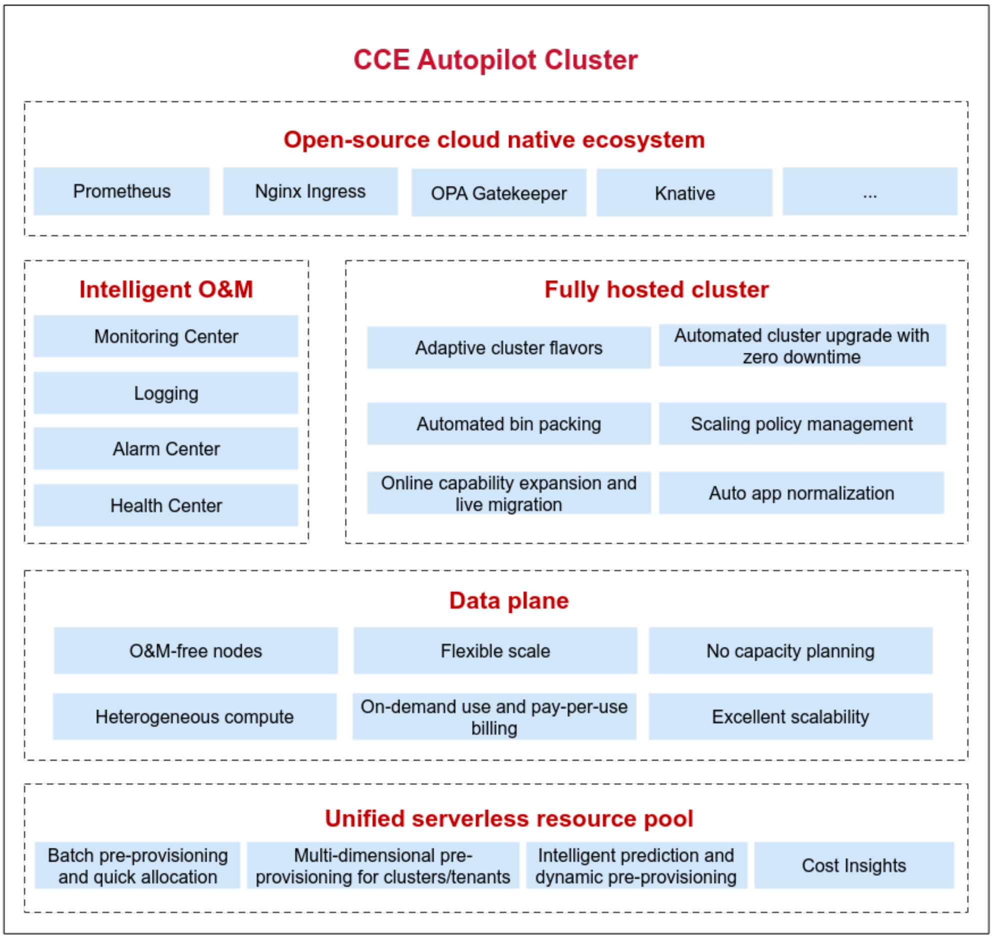
1 What Is a CCE Autopilot Cluster?

Introduction

CCE Autopilot allows you to create serverless clusters that offer optimized Kubernetes compatibility and free you from O&M. After a CCE Autopilot cluster is created, you can deploy applications without purchasing nodes or maintaining the deployment, management, and security of nodes. You only need to focus on the implementation of application service logic, which greatly reduces your O&M costs and improves the reliability and scalability of applications.

Product Architecture

Figure 1-1 Product architecture



Challenges with Traditional Serverful Container Clusters

Container technologies are driving the transformation of enterprise IT architecture in cloud computing due to their lightweight nature and high efficiency. However, traditional container services that rely on the serverful infrastructure are revealing the following issues, which severely hinder the pace of enterprise innovation:

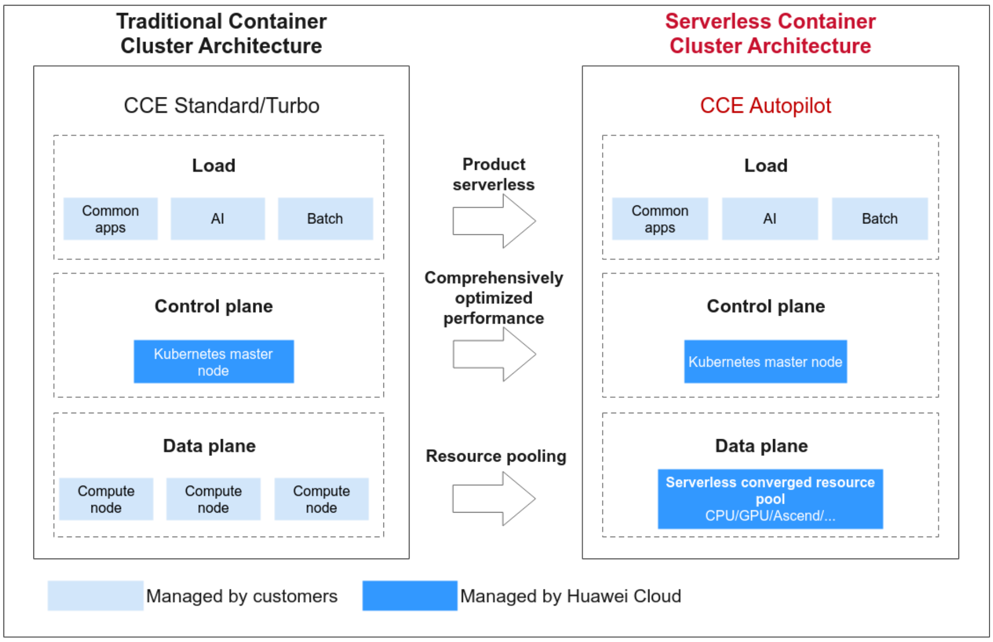
- O&M management: Enterprises need to manually manage server resource allocation and expansion, which involves complex capacity planning and resource scheduling, as well as continuous O&M such as node monitoring, troubleshooting, and system upgrades. This is expensive and requires a large workforce and many resources.
- Scalability: Enterprises need to create joint scaling policies for nodes and workloads. However, worker nodes must be scaled beforehand, which takes a few minutes and affects efficiency and response speed.
- Cost control: Enterprises need to allocate resources to nodes in advance. Unfortunately, those resources are often underutilized, or when there are heavy workloads, resources may be insufficient. This makes it hard to maximize cost effectiveness.

Benefits of Using the CCE Autopilot Cluster Architecture

Compared with CCE standard and Turbo clusters, CCE Autopilot clusters have the following advantages:

- Serverless evolution: Worker nodes are fully hosted on Huawei Cloud, so you do not have to maintain node deployment, manage nodes, or worry about security issues. Cluster flavors are adaptive.
- Resource pooling: A serverless converged resource pool is used to manage resources such as CPUs, memory, and GPUs, reducing resource fragments and enabling on-demand use of container resources.
- Comprehensive performance optimization: Resource pool resources are pre-provisioned to enable fast resource allocation, allowing for the scaling of more containers within seconds based on workload size.

Figure 1-2 Differences between CCE Standard/Turbo and CCE Autopilot



Comparison Between CCE Autopilot Clusters and Traditional Serverful Container Clusters

Category	Serverless Container Cluster	Traditional Serverful Container Cluster	
	CCE Autopilot	CCE Standard	CCE Turbo
Node management	Worker nodes are fully managed. CCE Autopilot takes care of node scaling and pre-binding.	You need to take care of the management and O&M of worker nodes.	You need to take care of the management and O&M of worker nodes.

Category	Serverless Container Cluster	Traditional Serverful Container Cluster	
	CCE Autopilot	CCE Standard	CCE Turbo
Node OSs	There are dedicated OSs that use containerd as the container engine.	You can select an OS and container engine.	You can select an OS and container engine.
Node specifications	Node specifications are adaptive to the workload scale.	You can select the node specifications as needed.	You can select the node specifications as needed.
Node upgrade and maintenance	Nodes are upgraded and recovered automatically.	Nodes need to be reset for upgrade.	Nodes need to be reset for upgrade.
Container network model	Cloud native 2.0 network	● VPC network ● Tunnel network	Cloud native 2.0 network
Network performance	The VPC network and container network are flattened into one for zero performance loss.	The container network is overlaid with the VPC network, causing performance loss.	The VPC network and container network are flattened into one for zero performance loss.
Network isolation	Pods can be associated with security groups for isolation.	● Tunnel network model: network policies for communications within a cluster ● VPC network model: isolation not supported	Pods can be associated with security groups for isolation.

2 Product Highlights

Intelligent, Reliable, Automated O&M

CCE Autopilot enhances your cluster experience with stable, secure, and intelligent features, such as automatic version upgrades, vulnerability fixing, and parameter tuning. CCE Autopilot is a serverless solution that simplifies capacity planning and node purchasing for fully hosted clusters. You do not need to manage or maintain the underlying resources, so there is much less O&M to deal with. You can focus on developing and deploying service logic.

Excellent Scalability with Continuous Iteration

CCE Autopilot prioritizes performance by setting up a serverless resource foundation in collaboration with underlying services. It uses multi-level resource pool pre-provisioning technology to meet diverse heterogeneous resource requirements and continuously improves performance through iterations. There is no need for you to plan or reserve resources beforehand to handle traffic bursts, seasonal fluctuations, and long-term service growth. Containers can be automatically scaled in or out within seconds based on your workloads. This ensures continuity and the optimal performance for your services. With CCE Autopilot, you can quickly launch new applications or services in response to market changes.

Compatibility with Cloud Native Open-Source Ecosystem

CCE Autopilot leverages a serverless architecture and a cloud native open source ecosystem to offer serverless clusters that are compatible with the Kubernetes ecosystem. This allows you to flexibly expand functions as needed. CCE Autopilot can also keep up with all Kubernetes community versions, so you always get the latest technical updates and security patches in a timely manner. It makes it easier to keep current with cutting-edge technologies. CCE Autopilot is always integrating mainstream open source security, application management, scalability, CI/CD, and AI software, such as OPA Gatekeeper and Knative. Integrating all these tools gets you an out-of-the-box application framework for easier application management.

Flexible Specifications and Per-Second Billing

CCE Autopilot delivers a flexible, efficient, and cost-effective cloud service experience. It dynamically adjusts cluster specifications and removes traditional specification tier limitations. You can easily adjust specifications with as little as 0.25 vCPUs and configure any resource ratio needed. The pay-per-use billing of CCE Autopilot (measured in seconds) means you save money because you only pay for the resources you actually use.

Security Isolation and Automatic Warnings

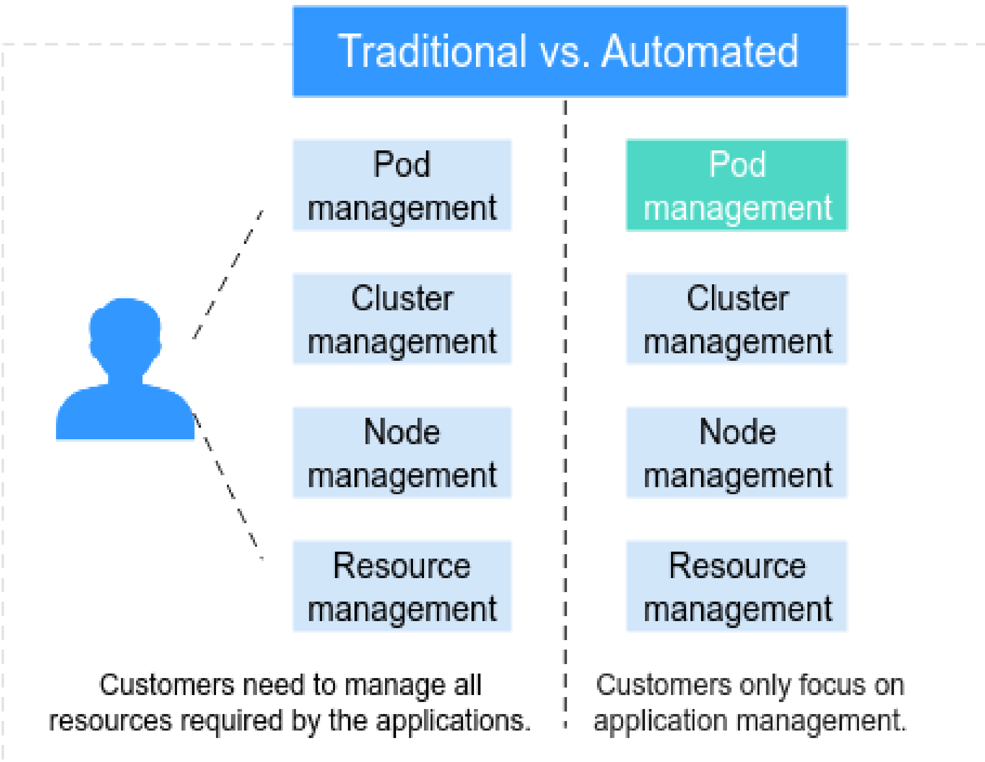
CCE Autopilot, built on the QingTian architecture, improves application security through virtual machine isolation. It provides a simplified and secure running environment using a dedicated container OS. With the underlying resource pool, CCE Autopilot supports quick fault isolation and rectification, ensuring continuous, secure, and stable application performance. Built-in automatic warnings can identify and prevent overloads on the control plane in a timely manner. The control plane components can be automatically scaled out to handle extra loads, ensuring service stability and reliability.

3 Applications Scenarios

O&M and Iterations of SaaS/Enterprise Platforms

CCE Autopilot is ideal for SaaS and enterprise platforms, especially for enterprises who have large resource pools that need to be iterated frequently. Traditionally, you need to handle your own O&M and upgrades, which drive labor costs through the roof. The automated O&M of CCE Autopilot puts an end to this issue. Internet finance enterprises have demanding compliance requirements. Traditionally, it is hard to develop compliance capabilities for OSs. CCE Autopilot simplifies node management and improves system security and compliance. This enables you to focus better on the innovation and development of core services.

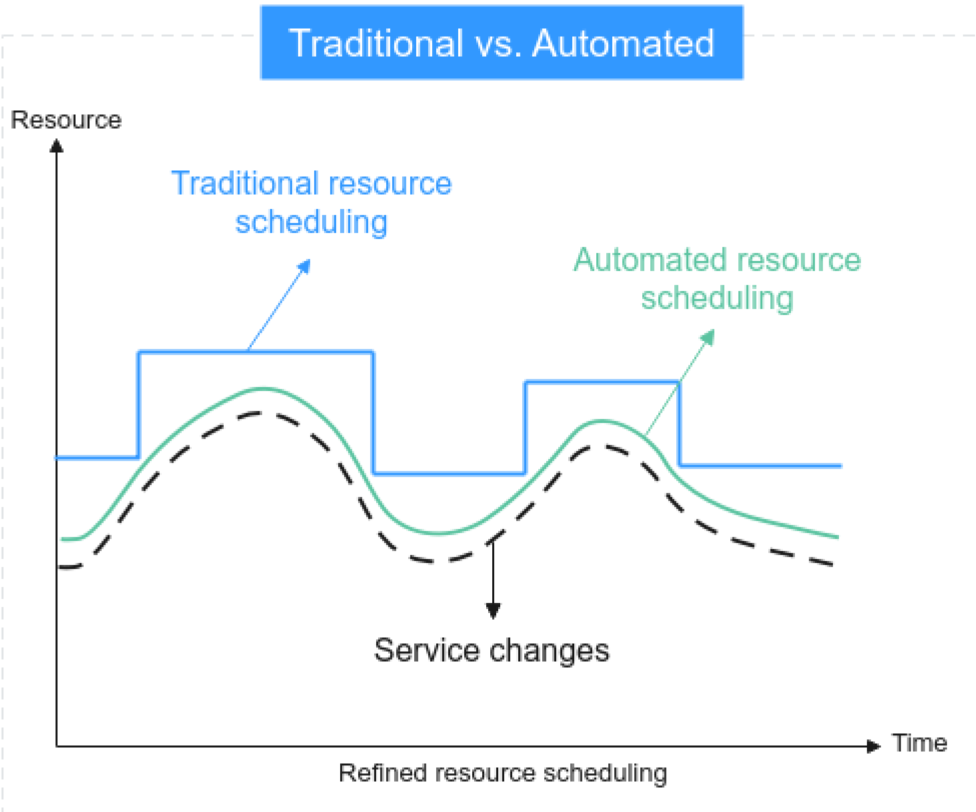
Figure 3-1 Resource management comparison between the traditional mode and the automated mode



Efficient, Auto Scaling of Services

For Internet entertainment, social networking, and ride-hailing, CCE Autopilot provides auto scaling to dynamically adjust resource configurations based on service characteristics and predicted traffic during traffic bursts. Unlike traditional scheduled scaling, CCE Autopilot scaling ensures efficient matching of resources to service requirements. This type of auto scaling allows you to optimize your enterprise's cost structure and reduce resource waste while still maintaining service continuity.

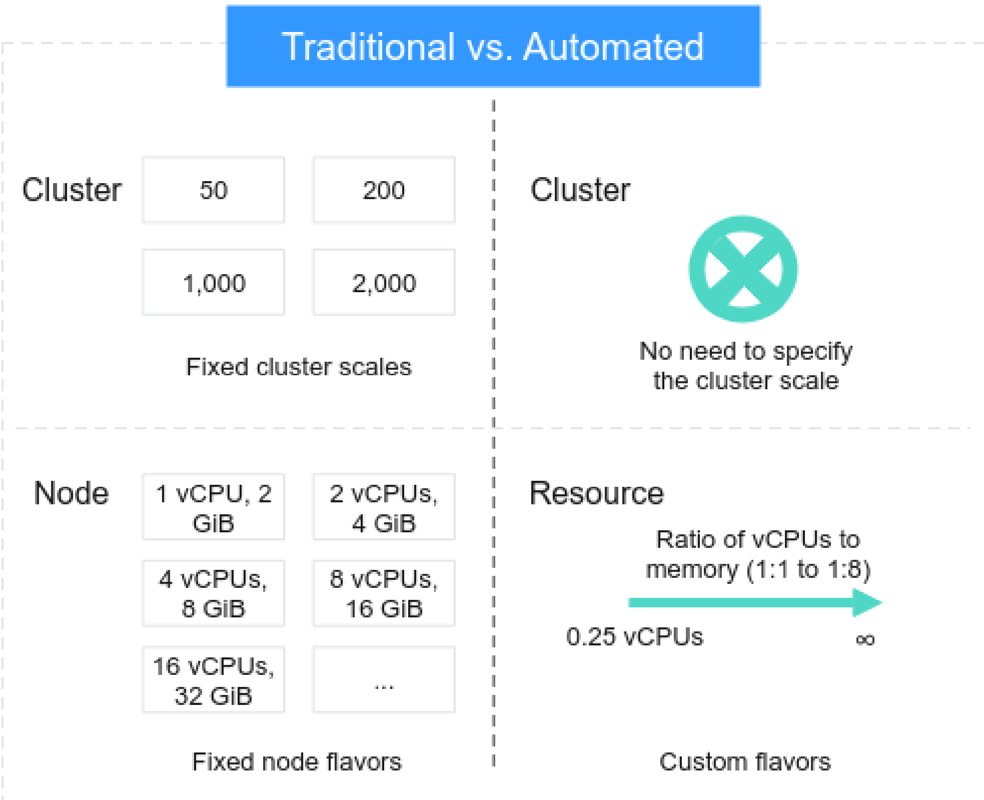
Figure 3-2 Resource scheduling comparison between the traditional mode and the automated mode



Cost Optimization Configuration

CCE Autopilot offers enterprises with cost optimization requirements a range of flexible resource configuration options. It meets requirements for affordable learning and adaptable resource configurations, and also supports automatic service scaling to adapt to rapid service growth. With CCE Autopilot, start-ups can enjoy high-performance and reliable services, even when their resource needs are minimal. As their services grow, resources can be scaled to ensure cost-effectiveness and service continuity.

Figure 3-3 Resource configuration comparison between the traditional mode and the automated mode



4 Billing

The price of a CCE Autopilot cluster consists of the following parts: cluster management, pods, and cloud resources (VPC endpoints and other cloud resources).

NOTE

The billed items marked with asterisks (*) are mandatory.

Table 4-1 Price of a CCE Autopilot cluster

Billed Item	Description	Billing Mode	Formula
*Cluster management	The expenses for managing the cluster **NOTE** If a cluster is frozen, workloads in the cluster will be in the pending state and will not be rescheduled until the cluster is unfrozen.	Pay-per-use	**Unit price of the cluster specification x Required duration** For details about the unit prices of cluster specifications, see **Unit Prices in Pay-per-Use Billing**.
*Pods	Pods are billed by specification. **NOTICE** If a specification is not supported, it will be automatically upgraded to a higher one. For example, if all containers in a pod require 2 vCPUs and 3 GiB of memory, the specification is automatically upgraded to 2 vCPUs and 4 GiB of memory. **Specification Description** lists the specifications supported by CCE Autopilot.	Pay-per-use	**Unit price of the pod specification x Required duration** For details about the unit prices of pod specifications, see **Unit Prices in Pay-per-Use Billing**.

Billed Item	Description	Billing Mode	Formula
*VPC endpoints	CCE Autopilot clusters connect to other cloud services such as SWR through VPC endpoints, which are billed separately based on the number of VPC endpoints you use.	Pay-per-use	**Unit price of the VPC endpoint x Required duration** **NOTE** • If a VPC endpoint connects to a VPC endpoint service other than DNS or OBS, you will be billed for how long you use this VPC endpoint. • If a VPC endpoint connects to DNS or OBS, you will not be billed for this VPC endpoint. See the pricing on the VPC Endpoint console.
Other cloud resources	Resources of cloud services used by a cluster such as Elastic Load Balance (ELB) are billed based on their pricing rules, no matter whether these resources are automatically created or manually added during cluster creation and use. Although cloud resources can be created on the CCE console, their billed items and bills are independent of those of CCE clusters.	Billing mode of each cloud service	For details, see **Price Calculator**.

Unit Prices in Pay-per-Use Billing

NOTICE

By default, 30 GiB of ephemeral storage is allocated to each pod for free (with an IOPS upper limit of 2,500 and a burst limit of 16,000). Any storage space that exceeds the 30 GiB quota will be billed separately.

Table 4-2 Unit prices in pay-per-use billing

Region	Cluster Management	Pod
AP-Singapore	$0.1 USD/hour	• vCPU: $0.045 USD/hour per vCPU • Memory: $0.005 USD/hour per GiB • Storage: $0.00028 USD/hour per GiB
AP-Bangkok	$0.1 USD/hour	• vCPU: $0.043 USD/hour per vCPU • Memory: $0.005 USD/hour per GiB • Storage: $0.00027 USD/hour per GiB
AP-Jakarta	$0.1 USD/hour	• vCPU: $0.045 USD/hour per vCPU • Memory: $0.005 USD/hour per GiB • Storage: $0.000294 USD/hour per GiB
AF-Johannesburg	$0.1 USD/hour	• vCPU: $0.049 USD/hour per vCPU • Memory: $0.005 USD/hour per GiB • Storage: $0.0003204 USD/hour per GiB
CN-Hong Kong	$0.1 USD/hour	• vCPU: $0.05 USD/hour per vCPU • Memory: $0.006 USD/hour per GiB • Storage: $0.000294 USD/hour per GiB
CN Southwest-Guiyang1	$0.1 USD/hour	• vCPU: $0.025 USD/hour per vCPU • Memory: $0.003 USD/hour per GiB • Storage: $0.00022 USD/hour per GiB
CN South-Guangzhou	$0.1 USD/hour	• vCPU: $0.028 USD/hour per vCPU • Memory: $0.003 USD/hour per GiB • Storage: $0.00022 USD/hour per GiB
CN East-Shanghai1	$0.1 USD/hour	• vCPU: $0.028 USD/hour per vCPU • Memory: $0.003 USD/hour per GiB • Storage: $0.00022 USD/hour per GiB
CN North-Beijing4	$0.1 USD/hour	• vCPU: $0.028 USD/hour per vCPU • Memory: $0.003 USD/hour per GiB • Storage: $0.00022 USD/hour per GiB

Specification Description

CCE Autopilot automatically upgrades the specifications that are not supported to higher ones to ensure that the pods always have the required resources.

Table 4-3 Combinations of vCPUs and memory supported by CCE Autopilot

vCPU	Memory (GiB)
0.25 vCPUs	0.5, 1, and 2
0.5 vCPUs	1, 2, 3, and 4
1 vCPU	2 to 8 (increment: 1 GiB)
2 vCPUs	4 to 16 (increment: 1 GiB)
4 vCPUs	8 to 32 (increment: 1 GiB)
8 vCPUs	8 to 64 (increment: 4 GiB)
16 vCPUs	16 to 128 (increment: 8 GiB)
32 vCPUs	32, 64, 128, and 256
48 vCPUs	96, 192, and 384
64 vCPUs	128, 256, and 512

5 Permissions

CCE Autopilot permissions management allows you to assign permissions to IAM users and user groups under your tenant accounts. It combines the advantages of IAM and RBAC to provide a variety of authorization methods, including IAM fine-grained/token authorization and cluster-/namespace-scoped authorization.

CCE Autopilot permissions are 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, charts, and add-ons. In the meantime, you can restrict other user groups to only view clusters.

  Cluster-level permissions involve non-Kubernetes native APIs and support fine-grained IAM policies and enterprise project management.

- **Namespace-level permissions**: You can regulate users' or user groups' access to **Kubernetes resources**, such as workloads, jobs, and Services, in a single namespace based on their Kubernetes RBAC roles. CCE Autopilot has also been enhanced based on open-source capabilities. It supports RBAC authorization based on IAM users or user groups, and RBAC authentication on access to APIs using IAM tokens.

  Namespace-level permissions involve CCE 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. For details, see **Using RBAC Authorization**.

CAUTION

- **Cluster-level permissions** are configured only for cluster-related resources (such as clusters and charts). You must also configure **namespace permissions** to operate Kubernetes resources (such as workloads, jobs, and Services).
- After you create a cluster, CCE Autopilot will automatically grant you the cluster-admin permission. This gives you complete control over all resources in all namespaces within the cluster.
- When viewing CCE resources on the console, the resources displayed depend on the namespace permissions. If no namespace permissions are granted, the console will not show you the resources.

Cluster-level Permissions (Assigned by Using IAM System Policies)

New IAM users do not have any permissions assigned by default. You need to first add them to one or more groups and then attach policies or roles to these groups. The users then inherit permissions from the groups and can perform specified operations on cloud services based on the permissions they have been assigned.

CCE is a project-level service deployed for specific regions. When you set **Scope** to **Region-specific projects** and select the specified projects in the specified regions, the users only have permissions for CCE in the selected projects. If you select **All projects**, the users have permissions for CCE in all region-specific projects. When accessing CCE, the users need to switch to the authorized region.

You can grant permissions by using roles and policies.

- Roles: A coarse-grained authorization strategy that defines permissions by job responsibility. Only a limited number of service-level roles are available for authorization. Cloud services often depend on each other. When you grant permissions using roles, you also need to attach any existing role dependencies. Roles are not ideal for fine-grained authorization and least privilege access.
- Policies: A fine-grained authorization strategy that defines permissions required to perform operations on specific cloud resources under certain conditions. This type of authorization is more flexible and is ideal for least privilege access. For example, you can grant users only permission to manage a certain type of clusters and nodes. A majority of fine-grained policies contain permissions for specific APIs, and permissions are defined using API actions. For the API actions supported by CCE, see **Permissions and Supported Actions**.

Table 5-1 lists all the system-defined permissions for CCE.

Table 5-1 System-defined permissions for CCE

Role/ Policy Name	Description	Type	Dependencies
CCE Administrator	Read and write permissions for CCE clusters and all resources (including workloads, jobs, and Services) in the clusters.	System-defined role	Users granted permissions of this policy must also be granted permissions of the following policies: **Global service project**: OBS Buckets Viewer and OBS Administrator **Region-specific projects**: Tenant Guest, Server Administrator, ELB Administrator, SFS Administrator, SWR Admin, and APM FullAccess **NOTE** To grant cluster namespace permissions to other users or user groups, an IAM user must have read-only permission.
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	Policy	None
CCE ReadOnly Access	Permissions to view CCE cluster resources, excluding the namespace-level permissions of the clusters (with Kubernetes RBAC enabled)	Policy	None

Table 5-2 Common operations supported by system-defined permissions

Operation	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
Creating a cluster	x	√	√
Deleting a cluster	x	√	√
Updating a cluster, for example, updating cluster node scheduling parameters and providing RBAC support to clusters	x	√	√
Upgrading a cluster	x	√	√
Listing all clusters	√	√	√
Obtaining cluster details	√	√	√
Listing all cluster jobs	√	√	√
Deleting one or more cluster jobs	x	√	√
Obtaining job details	√	√	√
Creating a storage volume	x	√	√
Deleting a storage volume	x	√	√
Performing operations on all Kubernetes resources	√ (Kubernetes RBAC authorization required)	√ (Kubernetes RBAC authorization required)	√
Viewing all resources in Monitoring Center	√	√	√
Performing operations on all resources in Monitoring Center	x	√	√
Performing all operations on NAT Gateway resources	x	√	√
Viewing NAT Gateway resource details	√	√	√
Listing all NAT Gateway resources	√	√	√

Operation	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
Performing all operations on VPC Endpoint resources	x	√	√
Viewing VPC Endpoint resource details	√	√	√
Listing all VPC Endpoint resources	√	√	√
Performing all operations on EVS disks EVS disks can be attached to workloads and scaled to a higher capacity whenever needed.	x	√	√
Performing all operations on VPCs A cluster must run in a VPC. When creating a namespace, create or associate a VPC for the namespace so that all containers in the namespace will run in the VPC.	x	√	√
Viewing details about all EVS disk resources. EVS disks can be attached to workloads and scaled to a higher capacity whenever needed.	√	√	√
Listing all EVS resources	√	√	√
Viewing details about all VPC resources A cluster must run in a VPC. When creating a namespace, create or associate a VPC for the namespace so that all containers in the namespace will run in the VPC.	√	√	√
Listing all VPC resources	√	√	√

Operation	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
Viewing details about all ELB resources	√	√	√
Listing all ELB resources	√	√	√
Viewing details about all SFS resources	√	√	√
Listing all SFS resources	√	√	√
Viewing details about all AOM resources	√	√	√
Listing all AOM resources	√	√	√
Performing all operations on AOM auto scaling rules	√	√	√

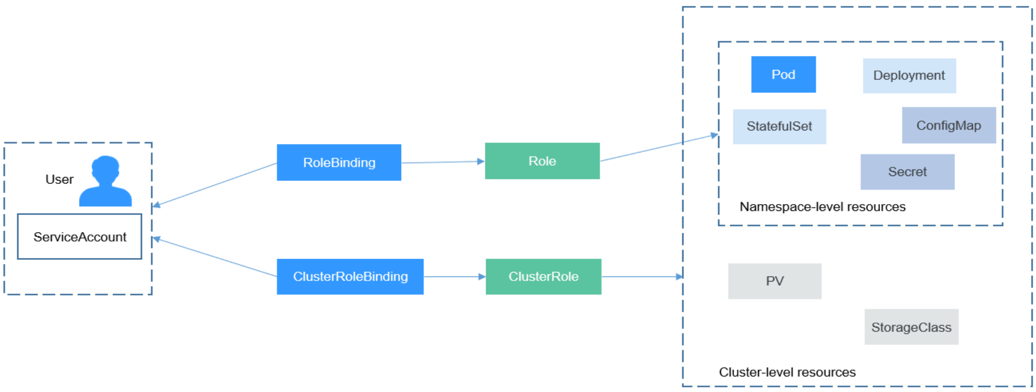
Namespace-level Permissions (Assigned by Using Kubernetes RBAC)

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 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. See the following figure.

Figure 5-1 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.

In addition to the preceding typical ClusterRoles, you can define Role and RoleBinding to grant permissions to add, delete, modify, and obtain global resources (such as PVs and CustomResourceDefinitions) and different resources (such as pods, Deployments, and Services) within specific namespaces. This allows for more precise permission control.

Helpful Links

- **Identity and Access Management Service Overview**
- **Granting Cluster Permissions to an IAM User**
- **Permissions and Supported Actions**

6 Notes and Constraints

Use Restrictions

- Worker nodes in CCE Autopilot clusters are fully hosted. For this reason, certain features like hostPath and hostNetwork, which rely on node features, are not supported.

Unavailable Feature	Description	Recommended Alternative Solution
DaemonSet	Deploys pods on each node.	Deploy multiple images in a pod using sidecars.
Setting hostPath in a pod	Mounts local files of a node to a container.	Use emptyDir or cloud storage of any type.
Setting hostNetwork in a pod	Maps a node port to a container port.	Use Services of the LoadBalancer type.
NodePort Service	Makes a node port open to access containers.	Use Services of the LoadBalancer type.

- When CCE Autopilot is used, each pod in the workload has 30-GiB of free disk space (with an IOPS upper limit of 2,500 and a burst limit of 16,000) by default. In addition to the resources reserved for the OS and CCE platform, the total available space for images, containers, and temporary storage is about 20 GiB. If the container image is too large (generally greater than 5 GiB) or a large number of files are created in the container root directory or temporary storage, you are advised to increase the temporary storage capacity of the pod. For details, see **Increasing the Ephemeral Storage Space of a Pod**.
- By default, no more than 500 pods can be created in a CCE Autopilot cluster. The add-on instances may occupy some pod quotas. You need to plan the pod quota appropriately. To increase the quota, **submit a service ticket**.
- If a CCE Autopilot cluster is used, workloads that use Arm images are not supported.

Cloud Product Quota Limits

The table below shows the maximum number of resources that can be created per account in a region.

Category	Item	Quota
CCE cluster	Total number of clusters	50
VPC	VPCs	5
	Subnets	100
	Security groups	100
	Security group rules	5000
	Routes per route table	100
	Routes per VPC	100
	VPC peering connections	50
	Network ACLs	200
ELB	Load balancers	50
	Load balancer listeners	100
	Load balancer certificates	120
	Load balancer forwarding policies	500
	Load balancer backend server groups	500
	Load balancer backend servers	500
VPCEP	Endpoints	50
DNS	Private zones	50
	DNS record sets	500

NOTE

If your current quota is insufficient, submit a **service ticket** to request an increase.

7 CCE Autopilot and Other Services

Table 7-1 describes how CCE Autopilot collaborates with other services.

Table 7-1 Collaboration between CCE Autopilot and other services

Service	Relationship
VPC	CCE Autopilot clusters run in VPCs. Pods created in a cluster will be assigned private IP addresses from the CIDR block of the VPC where the cluster is running. For more information about VPC, see **Virtual Private Cloud**.
ELB	CCE Autopilot can work with ELB to improve service capabilities and fault tolerance by associating load balancers with workloads. You can access a container workload from an external network through a load balancer. For more information, see **Elastic Load Balance**.
SWR	An image repository is used to store and manage Docker images. You can create workloads using images in SWR. For more information about SWR, see **SoftWare Repository for Container**.
EVS	EVS disks can be attached to workloads and scaled to a higher capacity whenever needed. Up to 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. For more information about EVS disks, see **Elastic Volume Service**.

Service	Relationship
OBS	OBS is a scalable service that provides secure, reliable, and cost-effective cloud storage for massive amounts of data. With OBS, you can create, modify, and delete buckets, as well as uploading, downloading, and deleting objects. CCE Autopilot allows you to create an OBS volume and mount it to a path inside a container. For more information about OBS, see **Object Storage Service**.
SFS	SFS is a fully managed, shared file storage service that supports the Network File System protocol. SFS file systems can scale up to petabytes, ensuring optimal performance for data-intensive and bandwidth-intensive applications. You can use SFS file systems as persistent storage for containers and mount the file systems to containers when creating a workload. For more information about SFS, see **Scalable File Service**.
SFS Turbo	SFS Turbo provides fully managed shared file storage that can be automatically scaled to 320 TB. It provides as little as sub-millisecond latency and up to tens of millions of IOPS and hundreds GB/s of bandwidth. Its high availability and durability can provide strong support for applications dealing with massive small files and requiring low latency and high IOPS, such as AI training, autonomous driving, and rendering. In addition, SFS Turbo supports NFS and SMB protocols and is compatible with mainstream operating systems. The GUI is easy to use, reducing operation costs. 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. For more information about SFS Turbo, see **What Is SFS Turbo?**
CTS	CTS records operations on your cloud resources, allowing you to obtain, audit, and backtrack resource operation requests initiated from the public cloud management console or open APIs as well as responses to these requests. For more information about CTS, see **Cloud Trace Service (CTS)**.