

# **Tencent Kubernetes Engine**

## **TKE Edge Cluster Guide**

### **Product Documentation**



## Copyright Notice

©2013-2024 Tencent Cloud. All rights reserved.

Copyright in this document is exclusively owned by Tencent Cloud. You must not reproduce, modify, copy or distribute in any way, in whole or in part, the contents of this document without Tencent Cloud's the prior written consent.

## Trademark Notice



All trademarks associated with Tencent Cloud and its services are owned by Tencent Cloud Computing (Beijing) Company Limited and its affiliated companies. Trademarks of third parties referred to in this document are owned by their respective proprietors.

## Service Statement

This document is intended to provide users with general information about Tencent Cloud's products and services only and does not form part of Tencent Cloud's terms and conditions. Tencent Cloud's products or services are subject to change. Specific products and services and the standards applicable to them are exclusively provided for in Tencent Cloud's applicable terms and conditions.

# Contents

## TKE Edge Cluster Guide

### Overview

### Edge Cluster Management

#### Creating a Cluster

#### Enable Internet/Intranet Access

#### Connecting to a Cluster

### Edge Node Management

#### Add Node

##### Add ECM

##### Add CVM

##### Add 3rd Node using script

#### Node Management

#### Node Remote SSH

#### Independent ENI

### Edge and Multi-region Deployment

#### Node Pool Management

##### NodeUnit

##### NodeGroup

#### ServiceGroup Manual

#### Health Check

#### Service Access across Regions

#### Multi-Region Ingress

#### Kins Manual

### Edge Platform Capability

#### Monitor

#### Log

### O&M Management

#### Edgectl script manual

# TKE Edge Cluster Guide

## Overview

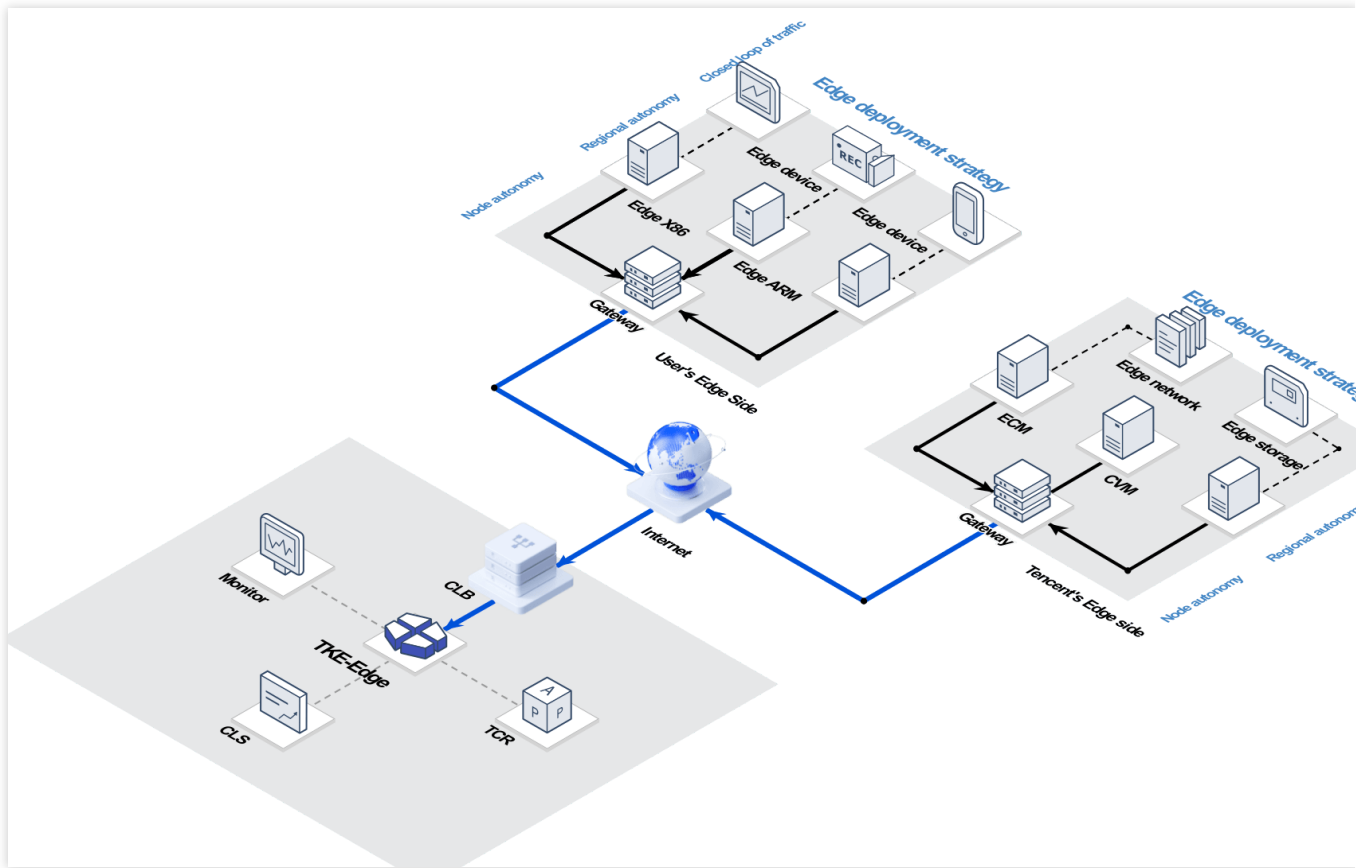
Last updated : 2023-06-01 11:22:54

## Introduction

Tencent Cloud Engine for Edge (TKE-Edge) is a container platform launched by Tencent Cloud that manages edge resources through a public cloud-hosted control plane, providing cloud-native capability support that integrates cloud and edge computing scenarios. TKE Edge Container is fully compatible with native Kubernetes, supports managing nodes in multiple data centers within the same cluster, allows one-click deployment of applications to all edge nodes, and features edge autonomy and distributed health check capabilities.

## Architecture





## Capabilities of TKE-Edge

### Native support

TKE-Edge is built upon the open-source SuperEdge project, closely following the community. It supports the latest Kubernetes versions and native Kubernetes cluster management methods. It is easy to use out of the box and requires only a minimal learning curve.

### High availability of the control plane

TKE-Edge is built upon the core Tencent Cloud TKE product, which hosts the Master control plane. Worker nodes located anywhere can connect to the control plane with a single click to start providing services, significantly improving the availability of the cluster.

### Cloud capabilities brought to the edge

TKE-Edge seamlessly reuses most of the supporting capabilities of TKE products, such as monitoring, logging, and app stores, quickly bringing standard public cloud capabilities to the edge side.

### **Heterogeneous resource support**

The connected edge nodes support various types of heterogeneous resources, including Tencent Cloud provided CVM/ECM machines, user-side owned IDC data center servers, PCs, IoT devices, etc. It also supports x86, ARM64, and ARM architecture, allowing for mixed deployment of applications on different types of resources.

### **Multi-region deployment capabilities**

The service supports the division of edge nodes into different node pools according to required features (such as region) and allows for batch application scheduling and deployment across multiple node pools. It also provides closed-loop traffic access capabilities within the node pool.

### **Secure and reliable**

TKE-Edge supports separated internal and external network certificates and minimized node permission control to avoid cluster access control permission leakage. Cloud-edge communication uses TLS encryption to prevent system management data leakage or tampering.

### **System disaster recovery**

In addition to the high availability of the control plane, the TKE-Edge provides reliable edge autonomy capabilities for weak network scenarios in the cloud and features distributed cluster health check capabilities to accurately determine pod migration timing.

### **Cloud-edge tunnel**

Leveraging Tencent's years of tunnel technology accumulation, even if edge devices do not have a public network address, the edge container service can support administrators directly logging into containers running on edge nodes from the cloud.

## Pricing

TKE-Edge's International site is temporarily free of charge.

Computing nodes are provided by users and therefore do not generate any additional charges.

## Use Cases

### **Edge computing**

TKE-Edge targets demands from customers in industries such as Industrial Internet, Edge IoT, and Edge AI. It supports customers in quickly creating cloud-based control planes, enabling unified management and control of various types of edge devices. Users can manage the application lifecycle in the cloud.

## Hybrid Cloud

TKE-Edge addresses the demands of large customers in hybrid cloud scenarios. It enables the rapid deployment of standard K8s clusters and cloud-standard services in user IDCs while seamlessly integrating with cloud-based SaaS services. This lowers deployment and operation costs and provides high availability support at the level of public cloud services.

## Additional Services

For more information on how to call TencentCloud APIs to access Tencent Cloud products and services, see [TencentCloud APIs](#).

# Edge Cluster Management

## Creating a Cluster

Last updated : 2023-07-14 16:30:12

This document describes how to create an edge cluster to use TKE Edge from the Tencent Cloud TKE console.

## Container Network Explanation

TKE Edge Container Service will utilize the node-side network to build an overlay network. Please ensure that the cluster network (control plane in VPC network) and container network (Pod CIDR/Service CIDR) do not conflict with the internal network of the edge nodes.

## Operation Steps

1. Log in to the [Tencent Cloud TKE console](#) and click **Edge Clusters** in the left sidebar.
2. On the Cluster Management page, click New.
3. Select Edge Cluster, then click 'Create'.

Select a cluster type

☐

**General cluster**

The default cluster type, which is fully compatible with standard features of the open-source Kubernetes clusters. The capabilities such as node management, cluster network and container scheduling are enhanced. Super nodes, native nodes, general CVM nodes and IDC nodes can be created and managed in a single cluster. You can configure the nodes based on your business needs to maximize the use of computing resources.

- It is applicable to all scenarios, and is fully compatible with standard capabilities of the open-source Kubernetes clusters.
- Supports super nodes, native nodes, registered nodes and CVMs.
- Provides resource visualization and optimization analysis to improve resource utilization.

Standard K8s clusterSupports super nodes

☐

**Serverless cluster**

You can quickly configure super nodes without cluster management fees to deploy massive security sandbox containers. You only need to pay for the actual running Pod resources. It reduces resource fees while ensuring stable business operations.

- It is applicable to resident applications with high stability and security as well as temporary computing tasks.
- Supports security sandbox containers. Strong isolation is applied to containers.
- It can start up 10K Pods in seconds, which are billed by actual usage.

Security sandbox containerSupports super nodesQuick start of containers

☒

**Edge cluster**

You can configure edge nodes to expand the Kubernetes cluster capacities to the edge region, and manage resources and application lifecycle in a cloud-native way. Also, you are provided with the multi-region edge autonomy, traffic loop and distributed health check capabilities.

- Edge Computing. You can manage computing resources at the edge (such as edge servers, edge IoT instances) in a cloud-native way, to solve the problems of weak network and node autonomy.
- Multi-region management. You can use a single cluster to manage resources in multiple regions and build a traffic loop within a region.

Edge ComputingEdge cloud nativeMulti-region managementTraffic loop

☐

**Registered cluster**

You can register the local Kubernetes clusters and those deployed in other cloud vendors' infrastructures to TKE for aligned management, and realize multi-cloud and multi-cluster management through Tencent Kubernetes Engine Distributed Cloud Center.

- Multi-cloud management. Access and manage all types of computing resources of the enterprise.
- High-availability disaster recovery. Multi-cluster applications, traffic and storage are managed collectively.
- Automatic release. It can be connected with the existing DevOps system to realize multi-cloud release.

Multi-cloud managementOpen EcosystemHigh-availability disaster recovery

Create

Cancel

4. On the Create Edge Cluster page, create an edge cluster according to the following information, as shown in the figure below:

The screenshot displays the Tencent Kubernetes Engine console for creating a new cluster. The form includes the following sections:

- Cluster name:** A text input field with a placeholder "Enter the cluster name (up to 50)".
- Kubernetes version:** A dropdown menu currently set to "1.20.6".
- Runtime components:** Two buttons, "docker" and "containerd", with a link "Learn more". A note states: "If you want to experience the latest K8s edge standalone cluster capabilities, choose Kubernetes version 1.22.5 and containerd runtime. The containerd is a more stable runtime component. It supports OCI standard and does not support docker API."
- Region:** Four buttons for "Guangzhou", "Beijing", "Shanghai", and "Singapore". A note below states: "Tencent Cloud resources in different regions cannot communicate via private network. The region cannot be changed after purchase. Please choose a region close to your end-users to minimize access latency and improve download speed."
- Billing method:** A button labeled "Pay-as-you-go".
- Cluster specification:** A row of buttons for "L5", "L20", "L50", "L100", "L200", "L500", and "L1000". A note below states: "Up to 5 nodes, 150 Pods, 32 ConfigMap and 150 CRDs are allowed under the current cluster specification. Please read Choosing Cluster Specification carefully before you make the choice. You can adjust the cluster specification manually, or enable Auto Cluster Upgrade to have it adjusted automatically."
- Auto Cluster Upgrade:** A checkbox that is checked.
- Cluster network:** A dropdown menu set to "Default-VPC" with a refresh icon and "CIDR: 172.22.0.0/16". A note below states: "If the current networks are not suitable, please go to the console to create a VPC."
- Pod CIDR:** A form with four input fields (192, 168, 0, 0) and a dropdown for "16". A note below states: "Please enter a valid private IP range. It should be a IP range and its subnet from below: 10.0.0.0/14, 172.16-31.0.0/16-24, 192.168.0.0/16-24, 9.0.0.0/14. It cannot duplicate with the IP range of the VPC or the IP range of Kubernetes clusters in the VPC. Please note that it cannot be changed after creation."
- Max Pods per node:** A dropdown menu set to "64". A note below states: "Up to 1024 nodes are allowed in the current container network configuration."
- Service CIDR:** A form with four input fields (192, 168, 0, 0) and a dropdown for "16". A note below states: "Please enter a valid private IP range. It should be a IP range and its subnet from below: 10.0.0.0/14, 172.16-31.0.0/16-24, 192.168.0.0/16-24, 9.0.0.0/14. It cannot duplicate with the IP range of the VPC or the IP range of Kubernetes clusters in the VPC. Please note that it cannot be changed after creation."
- Kube-proxy proxy mode:** Two buttons, "iptables" and "ipvs".
- Cluster description:** A large text area with a placeholder "Please enter cluster description".

At the bottom left, there is a link "Advanced settings".

**Cluster name:** indicates the name of the edge cluster to be created, with a maximum length of 60 characters.

**Kubernetes version:** Kubernetes version 1.16 is currently supported. This version will be updated when a newer Kubernetes version is published by the Kubernetes community.

**Runtime Components:** Supports Docker runtime and Containerd runtime. For more information, please refer to [How to Choose Between Containerd and Docker](#).

**Region:** select the region that is closest to your location to minimize access latency and improve the download speed.

**Billing Method:** The current Edge Cluster supports pay-as-you-go billing.

**Cluster Specifications:** You can choose different cluster specifications according to your needs. Different specifications have restrictions on the number of access nodes, as well as the number of pods, configmaps, and CRDs. For more details, please refer to the [Purchase Cluster Quota Limitations](#).

**Auto Upgrade Cluster Specifications:** If you enable this feature, when the cluster resources exceed the maximum limit of the current specification, the cluster specification will be automatically upgraded to a higher level and billed accordingly. Later on, you can reclaim resources, but the platform will not automatically downgrade. You can manually adjust the specifications to a more appropriate lower level. If this feature is not enabled, when the cluster resources exceed the maximum limit of the current specification, the newly added nodes will be in a "locked" state, not allowing application scheduling to that node. You will have to manually upgrade the cluster specification for the nodes to automatically return to normal.

**Cluster network:** assign a network for the cluster according to the internal network management of edge servers.

**Pod CIDR:** you need to assign a container network for the cluster according to the internal network management of edge servers. Therefore, plan the cluster size in advance to assign an IP range with sufficient IP addresses for the container network. **The pod CIDR block cannot overlap with IP ranges used by a VPC instance and existing Kubernetes clusters in the VPC instance. In addition, it cannot be modified once created.**

**Service CIDR:** you need to assign a service network for the cluster according to the internal network management of edge servers. Therefore, plan the cluster size in advance to assign an IP range with sufficient IP addresses for the service network. **The service CIDR block cannot overlap with IP ranges used by a VPC instance and existing Kubernetes clusters in the VPC instance. In addition, it cannot be modified once created.**

**Kube-proxy Mode:** You can choose between iptables mode and ipvs mode.

**Cluster description:** indicates information about the cluster, which is displayed on the **Cluster Information** page.

5. Click **Done** to finish creating the Master components of the cluster. You can check the progress of cluster creation on the "Edge Clusters" page.

## Node Management

### Note :

After the cluster is created, you need to enable the cluster's internal/external access in order to provide services to the outside world properly. For more details, please refer to [Enabling Internal and External Access for the Cluster](#).

# Enable Internet/Intranet Access

Last updated : 2023-06-01 11:28:46

## Enable Add Node Capability

If you need to enable the Add Node capability, you first need to open the cluster API Server service for external access, which includes opening internal network access (within the VPC) and opening external network access (to the Internet) two different types.

### Enable External Network Access

External network access means that the user's nodes are outside the Tencent Cloud public cloud system, but can access Tencent Cloud services through various types of Internet access. This situation applies to most edge nodes.

1. Log in to the [Tencent Kubernetes Engine](#), and select the Cluster option in the left-side navigation menu.
2. In the cluster list, click on the Edge Cluster ID to enter the cluster details page.
3. In the cluster "Basic Information", locate the "Cluster API Server Information" and enable external network access.



← Cluster(Singapore) / cls-3gmgtyoi(edge-demo)

|                            |  |  |
|----------------------------|--|--|
| Basic information          | Deployment type                          | Edge cluster   |
| Node management ▼          | Region                                   | Southeast Asia(Singapore)  |
| Namespace                  | Cluster network                          | <a href="#">vpc-c5ynz7i5</a>   |
| Service Group ▼            | Pod CIDR                                 | 10.33.0.0/16   |
| Workload ▼                 | Service CIDR block                       | 10.44.0.0/16   |
| Auto scaling               | Kube-proxy proxy mode                    | iptables   |
| Service ▼                  | Enable remote login                      | <input checked="" type="checkbox"/> Enabled  |
| Configuration management ▼ |  | For detailed directions on remote login, see <a href="#">Remote Logging in to an Edge Node</a> .   |
| Add-on management          | Cluster specification                    | L5<br>The application size does not exceed the recommended management size.<br>Up to 5 nodes, 150 Pods, 32 ConfigMap and 150 CRDs are allowed under the current cluster specification. Please read <a href="#">Choosing Cluster Specification</a> carefully before you make the  |
| Log                        |  | <input checked="" type="checkbox"/> Auto Cluster Upgrade<br>After the feature is enabled, it upgrades the cluster specification automatically when the load plane components reaches the threshold or the number of nodes reaches the upper limit. During the upgrade, the plane (master node) components are updated on a rolling basis, which may cause temporary downtime. It is recommended that you stop other operations (such as creating a workload) during the upgrade. |
| Event                      | Cluster Credential                       | <a href="#">View cluster credential</a> / <a href="#">Update Credential</a>  |
|                            | Time created                             | 2023-05-29 11:29:30  |
|                            | Description                              | N/A  |
|                            | Tencent Cloud tags                       | -  |
|                            | Kube-APIServer Custom Parameters         | N/A  |
|                            | Kube-ControllerManager Custom Parameters | N/A  |
|                            | Kube-Scheduler custom parameters         | N/A  |

After successfully enabling it, users can add ECM, CVM, and third-party nodes (via script) through the Internet link.

## Enable Internal Network Access

**Internal network access** is provided to facilitate users to add CVM machines under the same VPC to the cluster, and register them to the edge cluster via the VPC intranet, thereby saving public network traffic. Edge Container provides the capability to add nodes through the intranet, and you can use it according to the following steps.

1. Log in to the [Tencent Kubernetes Engine](#), and select the Cluster option in the left-side navigation menu.
2. In the cluster list, click on the Edge Cluster ID to enter the cluster details page.
3. In the cluster "Basic Information", locate the "Cluster API Server Information" and enable internal network access.

Workload

Auto scaling

Service

Configuration management

Add-on management

Log

Event

The application size does not exceed the i  
Up to 5 nodes, 150 Pods, 32 ConfigMap  
specification. Please read [Choosing Clust](#)

Auto Cluster Upgrade

After the feature is enabled, it upgrades th  
plane components reaches the threshold <  
the details of configuration modification or  
plane (master node) components are upda  
is recommended that you stop other operi

Cluster Credential

[View cluster credential](#) / [Update Credentia](#)

Time created

2023-05-29 11:29:30

Description

N/A

Tencent Cloud tags

-

Kube-APIServer Custom Parameters

N/A

Kube-ControllerManager Custom Parameters

N/A

Kube-Scheduler custom parameters

N/A

**Cluster APIServer information**

Internet access

Private network access

4. In the "Internal Network Access Settings", select a subnet under the VPC where the cluster is located. You need to select a specific subnet (the subnet here refers to the subnet where the subsequently added CVM is located).

**Private network access settings**

Subnet

subnet-73jspl1e(Default-...

CIDR block: 172.22.0.0/20; remaining available IPs in the subnet: 4093

Enable private network access. IPs will be assigned in the selected subnet.

Confirm

Cancel

5. After clicking Confirm, you can use the script to add CVM nodes under this VPC through the internal network.

#### Note

After enabling internal network access, you can only add CVM nodes within the VPC through the Add Nodes via Script method on the node page. You cannot use the capabilities to create ECM nodes and create CVM nodes.

## Next Steps

After enabling internal/external network access, you can add nodes to the cluster using the following methods:

[Create ECM node](#)

[Create CVM node](#)

[Add third-party nodes via script](#)

# Connecting to a Cluster

Last updated : 2023-06-01 11:22:54

## Scenario

This document describes how to connect a local client to an edge cluster through kubectl, which is the Kubernetes command-line tool.

## Prerequisites

The cURL software program has been installed.

Select the appropriate way to obtain kubectl based on the operating system:

### Note :

Replace `v1.18.4` in the command with the kubectl version required by your business.

### MacOS X

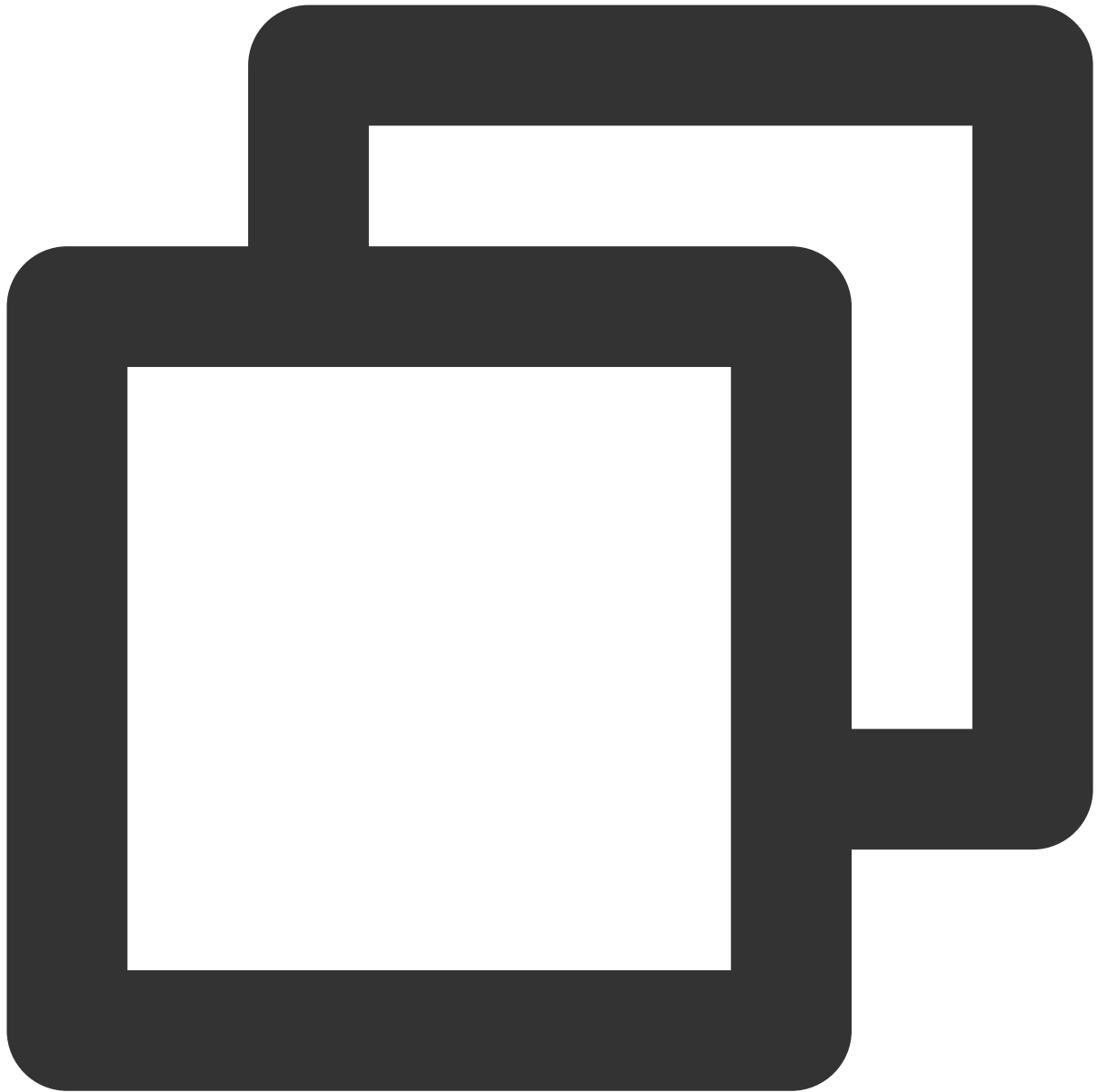
Run the following command to obtain kubectl:



```
curl -LO https://storage.googleapis.com/kubernetes-release/release/v1.18.4/bin/da
```

## Linux

Run the following command to obtain kubectl:



```
curl -LO https://storage.googleapis.com/kubernetes-release/release/v1.18.4/bin/li
```

## Windows

Run the following command to obtain kubectl:



```
curl -LO https://storage.googleapis.com/kubernetes-release/release/v1.18.4/bin/wi
```

## Operation Steps

### Installing kubectl

1. Install kubectl as instructions above.
2. Run the following commands to grant permissions to use kubectl.



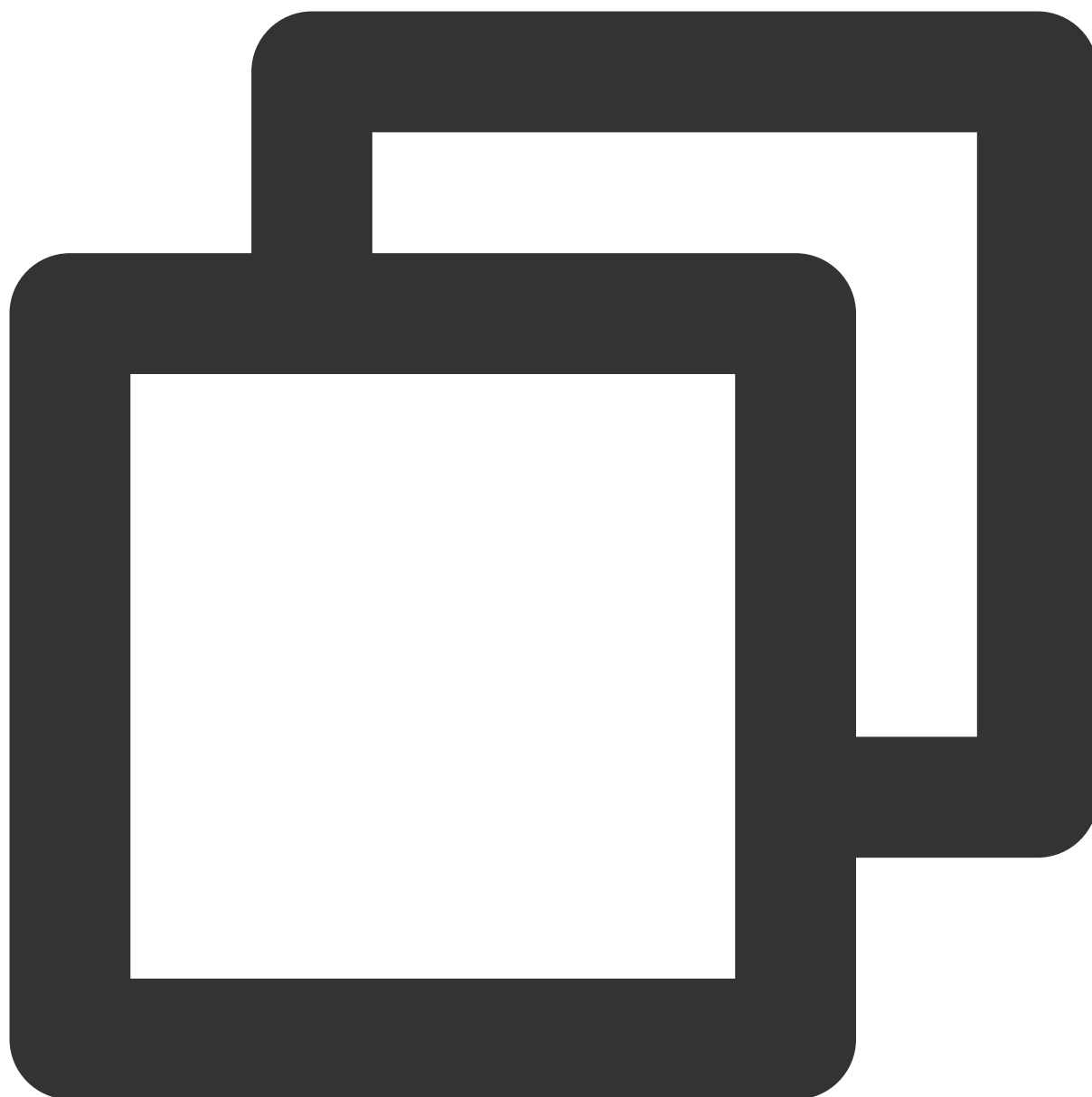
```
chmod +x ./kubectl
```





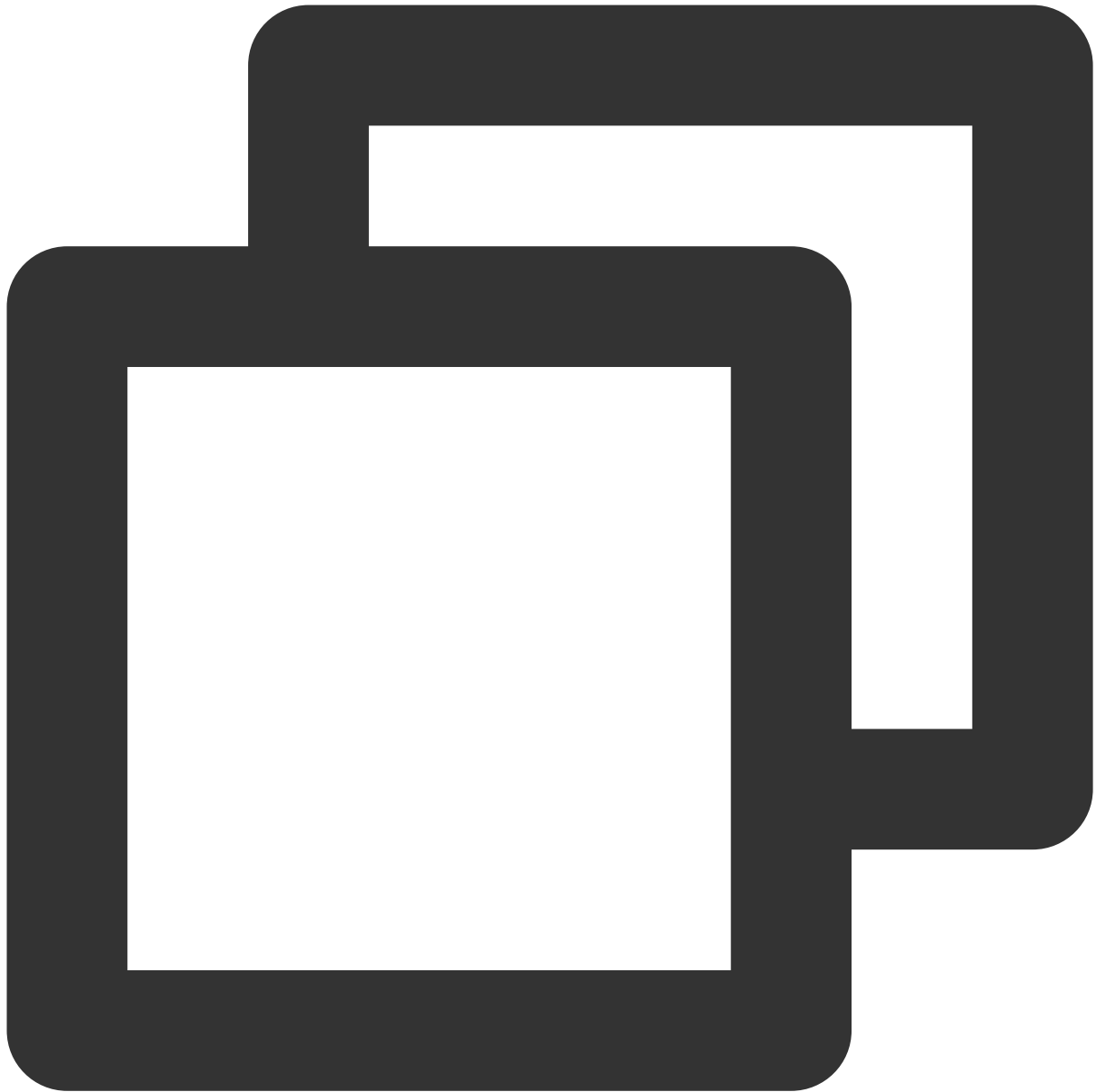
```
sudo mv ./kubectl /usr/local/bin/kubectl
```

3. Run the following command to check the installation result.



```
kubectl version
```

If the output is similar to the following version information, the installation was successful.



```
Client Version: version.Info{Major:"1", Minor:"5", GitVersion:"v1.5.2", GitCommit:"
```

## Obtaining cluster certificate information

1. Log in to the [Tencent Cloud TKE console](#) and click **Edge Clusters** in the left sidebar.
2. On the **Edge Clusters** page, click **View the cluster credential** for the cluster to be connected.

| Basic information                        |   |
|--|---|
| Cluster name                             | edge-demo   |
| Cluster ID                               | cls-3gmgtyoi  |
| Status                                   | Running...  |
| Runtime components                       | containerd  |
| K8s version                              | 1.22.5  |
| Deployment type                          | Edge cluster  |
| Region                                   | Southeast Asia(Singapore)   |
| Cluster network                          | <a href="#">vpc-c5ynz7i5</a>  |
| Pod CIDR                                 | 10.33.0.0/16  |
| Service CIDR block                       | 10.44.0.0/16  |
| Kube-proxy proxy mode                    | iptables  |
| Enable remote login                      | <input type="checkbox"/> Disabled   |
|  | For detailed directions on remote login, see <a href="#">Remote Logging in to an Edge Node</a> .  |
| Cluster specification                    | <p>L5 </p> <p>The application size does not exceed the recommended management size.</p> <p>Up to 5 nodes, 150 Pods, 32 ConfigMap and 150 CRDs are allowed under the current cluster specification. Please read <a href="#">Choosing Cluster Specification</a>  carefully before you make the choice.</p> <p><input checked="" type="checkbox"/> Auto Cluster Upgrade</p> <p>After the feature is enabled, it upgrades the cluster specification automatically when the load on the control plane components reaches the threshold or the number of nodes reaches the upper limit. You can view the details of configuration modification on the cluster details page. During the upgrade, the master plane (master node) components are updated on a rolling basis, which may cause temporary downtime. It is recommended that you stop other operations (such as creating a workload) during the period.</p> |
| Cluster Credential                       | <a href="#">View cluster credential / Update Credential</a>   |
| Time created                             | 2023-05-29 11:29:30   |
| Description                              | N/A   |
| Tencent Cloud tags                       | -   |
| Kube-APIServer Custom Parameters         | N/A   |
| Kube-ControllerManager Custom Parameters | N/A   |
| Kube-Scheduler custom parameters         | N/A   |

3. In the **Cluster Credential** window that appears, you can view, copy, and download the credential.

Note :

You can save the cluster access credential locally by clicking **Copy** or **Download** as needed.

4. In **Internet access**, click

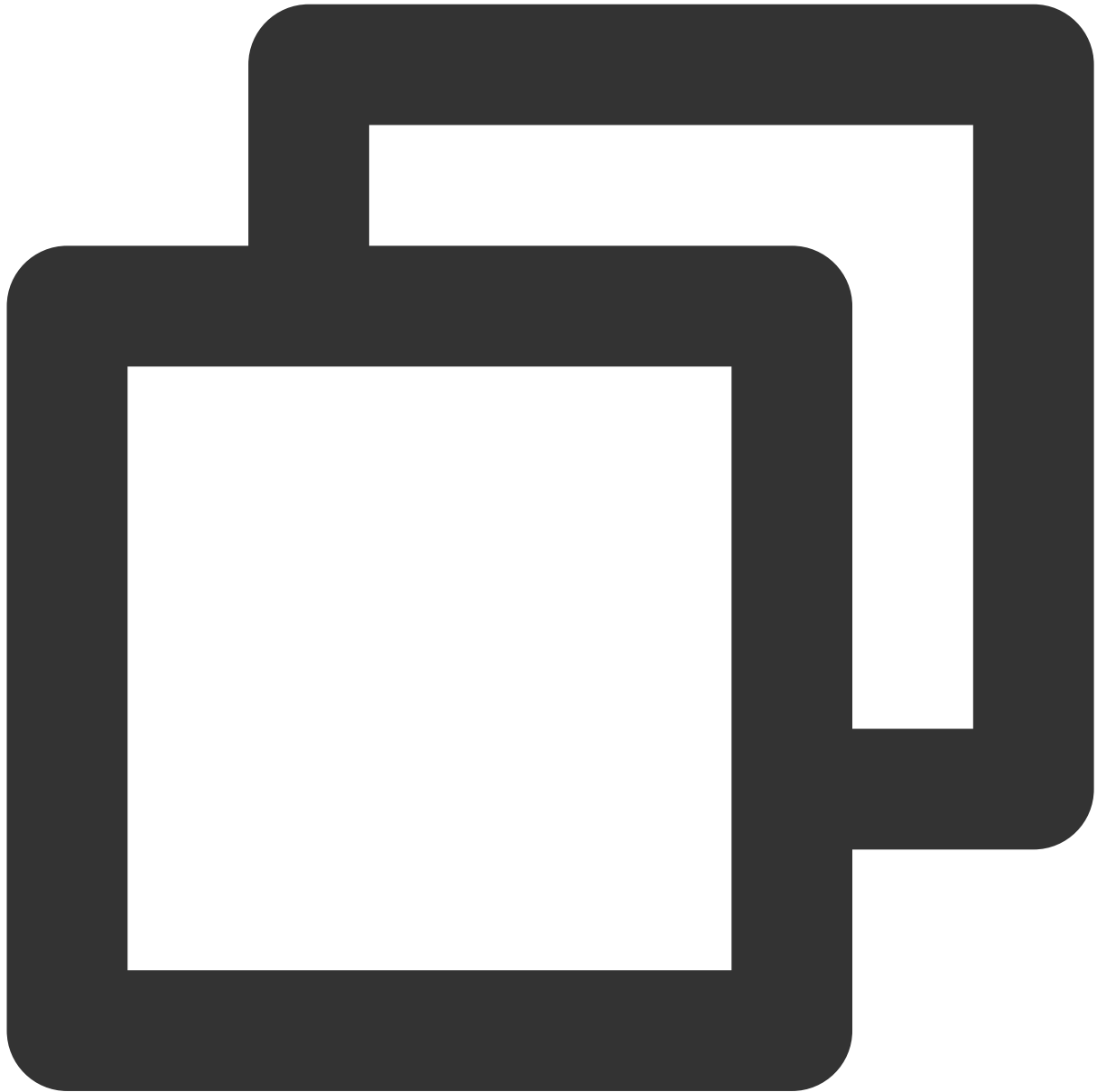


to enable internet access for the cluster. You can also see [Configuring kubectl autocomplete](#) for accessing with the cluster access credential.

## Using kubectl to manipulate a cluster through certificate information

### Request method

The kubectl command format is as follows:



```
--kubeconfig=<Local cluster access credential>
```

### Example

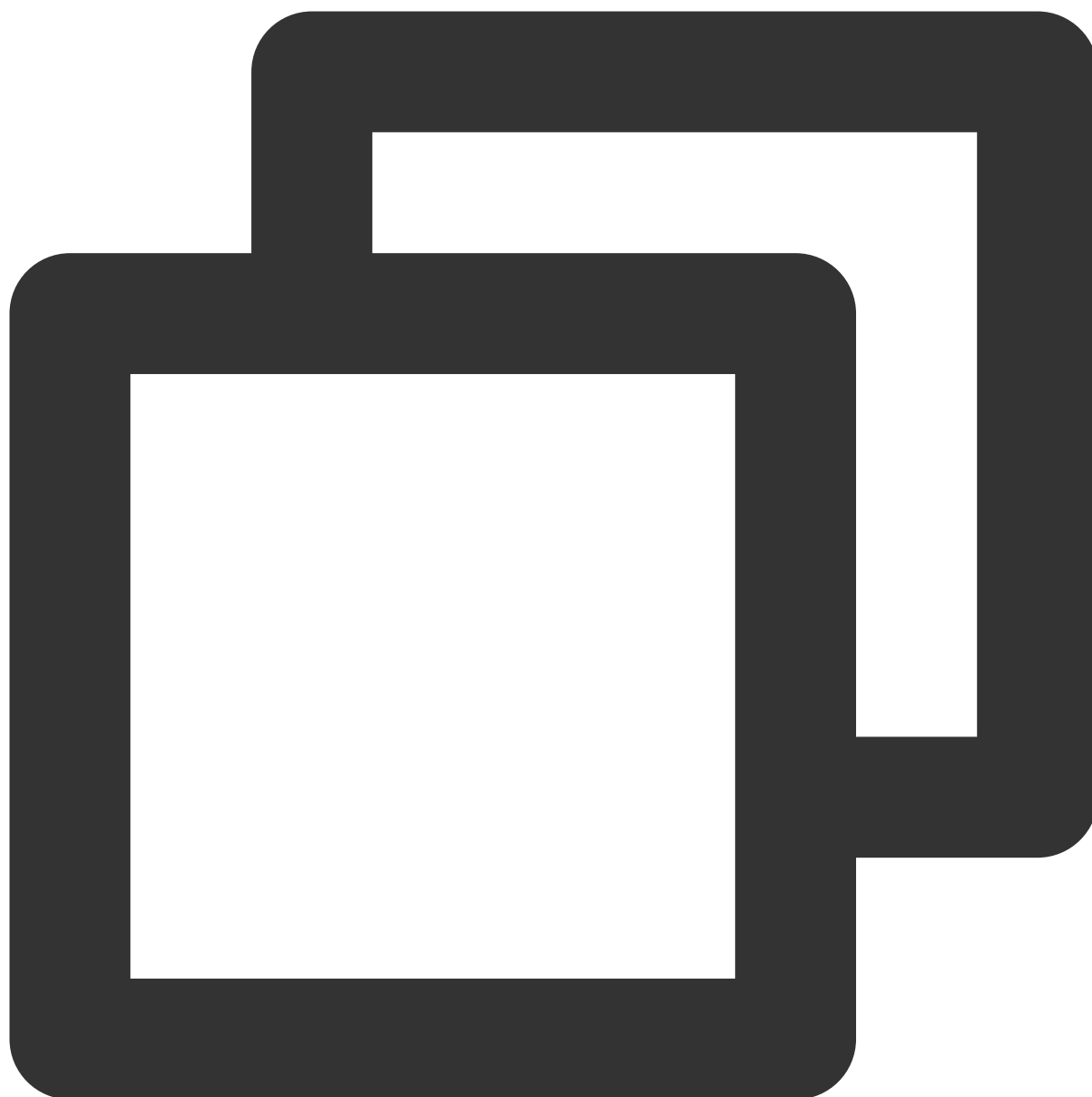
Run the following command to view existing namespaces under the cluster.



```
kubectl get namespace --kubeconfig=cls-8ipgf8u4.kubeconfig
```

The cluster credential used in this example is `cls-8ipgf8u4.kubeconfig` . In real-life cases, replace it with the actual credential.

If a message similar to the following is returned, the request was successful.



| NAME        | STATUS | AGE |
|-------------|--------|-----|
| default     | Active | 11d |
| kube-system | Active | 11d |

## Configuring kubectl autocomplete

You can configure kubectl autocomplete to improve usability by running the following command.



```
source <(kubectl completion bash)
```



# Edge Node Management

## Add Node

## Add ECM

Last updated : 2023-06-01 11:22:54

This article describes how to add an ECM node to the edge container cluster. ECM is short for Tencent Cloud Edge Computing Machine. Based on Tencent Cloud's edge nodes distributed in various regions, ECM provides cloud computing and networking services. The edge nodes are widely distributed, closer to users, and have the characteristics of low network latency, high availability, and low cost, which can seamlessly integrate with the edge container service.

## Prerequisites

Refer to Cluster [Enable Internet/Intranet Access](#) to enable **external network access** capabilities.

## Create an ECM node

1. Log in to the Container Service console and select the cluster in the left navigation bar.
2. In the cluster list, click the ID of the edge cluster to enter the cluster details page.
3. Select Node Management > Node on the left side of the page to enter the node list page, and click Create ECM Node.
4. On the "Create Node" page, configure relevant parameters according to actual needs, as shown in the following figure:

←

Create node

Edge Module

Please select the edge module

↻ Create Edge Module

Default instance configurations:

CPU cores: core | MEM: GB | System disk storage: GB | Data Disk Storage: GB

ECM Instance Image

Tencent Linux Release 2.2 (Final)

Default bandwidth cap

25

358

691

1024

25

Mbps

ECM Instance Name

Please enter the ECM instance name

Custom password

Enter the CVM password

Confirm password

Please enter the node password

Security group

Select Security Group

Security groups work as a virtual firewall for network access control of instances. You can go to the console to [Create Security Group](#)

Display Advanced configurations

☒ Free host security: Install components to enable Cloud Workload Protection Basic [Learn more](#)

☒ Free monitoring: enable analysis and alarm service for cloud products for free (component installation required) [Learn more](#)

Estimated cost:

Configuration fee:

0USD/day

(fee details)

Bandwidth fee:

please refer to the pricing of the [region of the node](#)

Confirm

Cancel

The main parameter information is as follows:

**Edge Module:** Select the edge module to determine the ECM edge node specifications.

**ECM Instance Image:** Tencent Cloud provides public images and custom images. The default is the same image as the module to which it belongs. Please select it according to actual needs.

**Default bandwidth cap:** Limits the bandwidth upper limit. If it exceeds this limit, packets will be lost by default. The default is 25Mbps, and the upper limit is 1024Mbps.

**ECM Instance Name:** Indicates the name of the instance to be created, user-defined.

**Set password and confirm password:** Define the password for logging in to the instance.

**Security group:** A security group is a virtual firewall used for network access control of instances. The default selected security group is the security group of the module to which it belongs, and users can modify the security group settings by themselves.

**Advanced settings:** You can modify the default settings. Please select according to actual needs:

**The maximum number of Pods:** determines the upper limit of the number of Pods allocated to each Node.

**Container directory:** Set container and image storage directories, and it is recommended to store them on data disks. For example, /var/lib/docker.

**Launch configuration:** shell script format.

**Province of the Node:** It is recommended to select the province closest to your customers, which can reduce access latency and increase access speed.

**Node region:** Please select according to actual needs.

**Network type:** Please select public network operators according to actual needs.

**Free host security:** Default selected, it helps users build a server security protection system to prevent data leakage.

**Free monitoring:** Default selected, it opens free cloud product monitoring, installs components to obtain host monitoring indicators and displays them in the form of monitoring icons, and supports setting custom alarm thresholds.

5. After clicking OK, wait for the corresponding ECM instance to be created successfully and its status to become Ready.

# Add CVM

Last updated : 2023-06-01 11:22:54

This article describes how to add Tencent Cloud Standard CVM nodes to an existing edge cluster. You can directly purchase Tencent Cloud CVMs in the required region and add them to the edge cluster to provide services. The difference between this and a standard TKE managed cluster is that you can purchase CVM nodes from different regions and VPCs and add them to the same edge cluster, rather than being limited to the VPC selected when creating the edge cluster.

## Prerequisites

Please refer to [Enable Internet/Intranet Access](#) to enable external network access capability.

## Creating a CVM Node

1. Log in to the Container Service console and select "Clusters" from the left navigation bar.
2. On the Cluster Management page, click on the edge cluster ID for which you want to create a CVM node to enter the details page for that cluster.
3. Select "**Node Management > Nodes**" from the left-hand side of the page to enter the Node List page. Click on "**Create CVM Node**", as shown in the figure below:

← Create node

Billing mode

Pay-as-you-go

Region

South China

East China

North China region

Southwest China

Hong Kong, Macau and Taiwan (China)

Northeast Asia

Guangzhou

Shanghai

Nanjing

Beijing

Chengdu

Chongqing

Hong Kong, China

Seoul

Tokyo

Southeast Asia

US West

Europe

South Asia

US East

South America

North America

Singapore

Bangkok

Jakarta

Silicon Valley

Frankfurt

Mumbai

Virginia

São Paulo

Toronto

Tencent Cloud resources in different regions cannot communicate via private network. The region cannot be changed after purchase. Please choose a region close to your end-users to minimize acc improve download speed.

Availability zone

Singapore Zone 1

Singapore Zone 2

Singapore Zone 3

Singapore Zone 4

Cluster network

Default-VPC

The VPC does not have valid subnets in the current availability zone. You can [create a new one now](#).

CIDR:172.22.0.0/16

If the current networks are not suitable, please go to the console to [create a VPC](#) or [create a subnet](#).

Image provider

Public image

Marketplace

Operating system

Please selectOperating system

Choosing an Image (TencentOS Server is recommended)

Model configuration

Select a model

Instance name

Auto-generated

Custom name

The CVM will be automatically named in the format of "tkeedge\_cluster id\_worker".

Login method

SSH key pair

Random password

Custom password

SSH key

willzgll | skey-12zpdvdj

[Instruction](#)

If existing keys are not suitable, you can [create a new one](#).

Security group

Select Security Group

Security groups work as a virtual firewall for network access control of instances. You can go to the console to [Create Security Group](#).

Amount

1

Create node

Cancel

4. On the "New Node" page, configure the relevant parameters based on your actual requirements. The parameters are similar to adding nodes in TKE, please refer to [Adding Nodes](#) for details.

# Add 3rd Node using script

Last updated : 2023-07-14 18:10:08

This article describes how to use the `edgectl` script to add user's third-party nodes to an edge cluster. Unlike CVMs and ECMs, these nodes belong to the user and may be located in the user's IDC room, office, or various types of edge regions. These machines can access Tencent Cloud services over the Internet. In this scenario, users can download the `edgectl` script from the cloud and execute it locally to add the node to the edge cluster.

## Prerequisites

Please refer to [Enable Internet/Intranet Access](#) to enable external network access or internal network access capability.

### Note :

After enabling internal network access, if you need to add internal network CVM machines, you can only do so using the script to add nodes for now.

## Prerequisites

Prepare the edge node according to the following conditions:

Node source: You can use existing servers from the [Tencent Cloud Console](#) or [Edge Computing Machine Console](#), servers from other platforms, or servers from self-built facilities.

Node processors: x86\_64, ARM64, ARM.

The following operating systems are supported (tested and adapted for use):

Ubuntu 20.04/18.04/16.04

CentOS 8.0/7.8/7.6/7.2

TencentOS Server 3.2/3.1/2.6/2.4

Tencent Linux Release 2.2 (Final)

Debian 10.2/9.0

SUSE Linux Enterprise Server 12 SP3

Please ensure that the node you want to add has installed `wget`, `systemctl`, and `iptables`.

The node's network needs to have the ability to access the internet actively.

### Note :

Due to different operating system environments deployed by different users, it is not ruled out that certain configurations may affect the successful addition of nodes. In this case, please check the deployment log of `edgectl` on

the node (/tmp/tkeedge-install.log) for detailed troubleshooting.

## Procedure

1. Log in to the [Tencent Cloud Container Service console](#) and select "Clusters" from the left navigation panel.
2. On the Cluster Management page, click on the Edge Cluster ID that needs to add an edge node to enter the details page of that cluster.
3. Select "**Node Management**" > "Nodes" on the left side of the page and enter the node list page. Click "**Script to Add Node**".
4. In the "Add Edge Node" window that pops up, configure according to the following steps to obtain the initialization script for the node.
  - 4.1 In the "Configuration" step, obtain the initial configuration and modify parameter values as follows:

### Add Edge Node ✕

1 Basic Information >

2 Custom script >

3 Installation Script

Container directory

Expiry time

24

hour ▼

Back

Next

**Container directory:** Set the container and image storage directory, it is recommended to store it in the data disk, for example `/var/lib/docker`.

**Expiration time:** Set the expiration time of the "installation script" command, the configurable range is from 1 minute to 24 hours, and the default configuration is 24 hours.

4.2 After clicking "Next", add custom scripts:



### Add Edge Node ✕

✓ Basic Information >

**2 Custom script** >

3 Installation Script

**pre-install script**

Please input pre-install script

If you want to access the cluster in VPC, please copy these commands to pre-install script:

```
sed -i "/cls-3gmgtyoi.ccs.tencent-cloud.com/d"
/etc/hosts
echo "172.22.0.33 cls-3gmgtyoi.ccs.tencent-cloud.com" >>
/etc/hosts
```

Copy

**post-install script**

Please input post-install script

Back Next

**pre-install script:** Here, you can set up pre-installation scripts for node access, which can be used to perform some cleanup or pre-configuration work under special conditions to enhance adaptability in different scenarios.

**post-install script:** Here, you can set up post-installation scripts for node access, which can be used to perform some subsequent cleanup or configuration work after the node is successfully added, enhancing adaptability in different scenarios.

**Note :**

In the above figure, because "Internal Network Access" is enabled, it will prompt for the execution of commands. If the user wants to add CVM nodes through the internal network, please copy and paste these instructions into the pre-install script. If adding nodes through the external network, these instructions can be ignored.

4.3 In the "Generate Installation Script" step, copy the script command to obtain the corresponding node initialization script.

### Add Edge Node

✓ Basic Information >

✓ Custom script >

3 Installation Script

Step 1: Copy the script

```
wget --header="x-cos-token:Gvi0KSUXkQYe8zbu" https://tke-edge-1253687700.cos.accelerate.myqcloud.com/user-pkgs%2Fcls-3gmgtyoi200021449168200022964241ap-singapore%2Fedgetl?sign=q-sign-algorithm%3Dsha1%26q-ak%3DAKIDUnpa4HDtIQHsDH4g15s0edLlvoQlpc8N%26q-sign-time%3D1685346978%3B1685433378%26q-key-time%3D1685346978%3B1685433378%26q-header-list%3Dx-cos-token%26q-url-param-list%3D%26q-signature%3Dcc1343b8047e1cbc5ca879e5d67a14a068049747 -O edgetl && chmod +x edgetl
```

Copy

Step 2: Execute the script

```
./edgetl install -n [nodename] -i [flannel eth]
```

Copy

[nodename] is the name of the registered node and need to be unique; [flannel eth] is the ethernet of the the node, used by flannel to complete pods communication of two different nodes

Back

Disable

5. Log in to the prepared server and switch to the root account to execute the copied command.

**Note :**

Executing this command will automatically download the edgetl edge node management tool.

The edgetl obtained in this step is valid for the time period set by the "Expiration Time" parameter. Please complete the node addition within this time period. After the expiration time, please execute the script to add the node operation again to obtain edgetl again.

6. To view the edgetl command, run the "Initialize Node" operation. For more operations with edgetl, please refer to the [Edgetl script manual](#)



```
./edgectl -h
```

Usage:

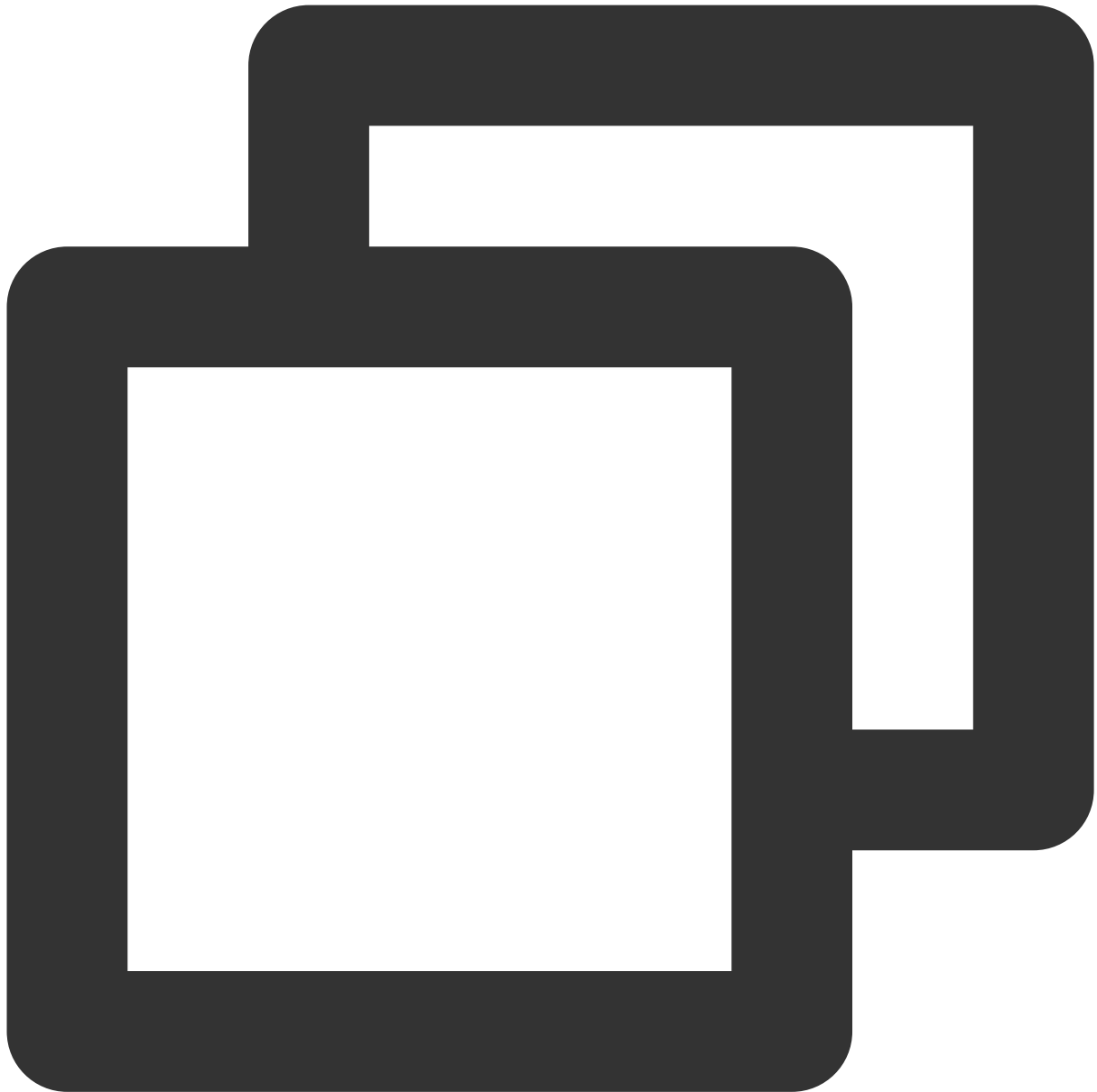
```
edgectl command [flags]
```

Available Commands:

|         |  |
|---------|--|
| check   | Check the edge node if to be add to clusters |
| install | Install components to edge node              |
| clear   | Clear edge node and recovery as usual        |

Flags:

```
-h, --help    Help for edgectl
```



```
./edgectl install -n Your-Node-Name -i Your-Interface-Name
```

**Note :**

Executing this command will check the host environment, install the required components for the edge node and automatically register the node to the edge cluster.

Due to network and installation speed factors, adding nodes may take several minutes.

To confirm the installation process information, view the log file at `/tmp/tke-edge-install.log`.

7. After confirming the command has been executed successfully, go to the "Node List" page and refresh to view the newly added nodes. You can also perform other operations on the node, such as eviction, removal, blocking, editing tags or unblocking.

## Related Operations

### Close cluster Internet and Private access

To disable both external and internal network access to the cluster, go to the "Basic Information" page of the cluster and turn off "Internet Network Access" and "Private Network Access".

The screenshot displays the 'Basic Information' page of a Tencent Kubernetes Engine cluster. The left sidebar contains navigation links: Basic information, Node management, Namespace, Service Group, Workload, Auto scaling, Service, Configuration management, Add-on management, Log, and Event. The main content area shows the following settings:

- Service CIDR block: 10.44.0.0/16
- Kube-proxy proxy mode: iptables
- Enable remote login: ☒ Enabled. For detailed directions on remote login, see [Remote Logging in to an Edge Node](#).
- Cluster specification: L5. The application size does not exceed the recommended management size. Up to 5 nodes, 150 Pods, 32 ConfigMap and 150 CRDs are allowed under the current cluster specification. Please read [Choosing Cluster Specification](#) carefully before you make the choice.
- Auto Cluster Upgrade: ☒ Auto Cluster Upgrade. After the feature is enabled, it upgrades the cluster specification automatically when the load on control plane components reaches the threshold or the number of nodes reaches the upper limit. You can check the details of configuration modification on the cluster details page. During the upgrade, the management plane (master node) components are updated on a rolling basis, which may cause temporary disruption. It is recommended that you stop other operations (such as creating a workload) during the period.
- Cluster Credential: [View cluster credential / Update Credential](#)
- Time created: 2023-05-29 11:29:30
- Description: N/A
- Tencent Cloud tags: -
- Kube-APIServer Custom Parameters: N/A
- Kube-ControllerManager Custom Parameters: N/A
- Kube-Scheduler custom parameters: N/A

A red box highlights the 'Cluster API Server information' section at the bottom, which contains two toggle switches:

- Internet access: ☒
- Private network access: ☒

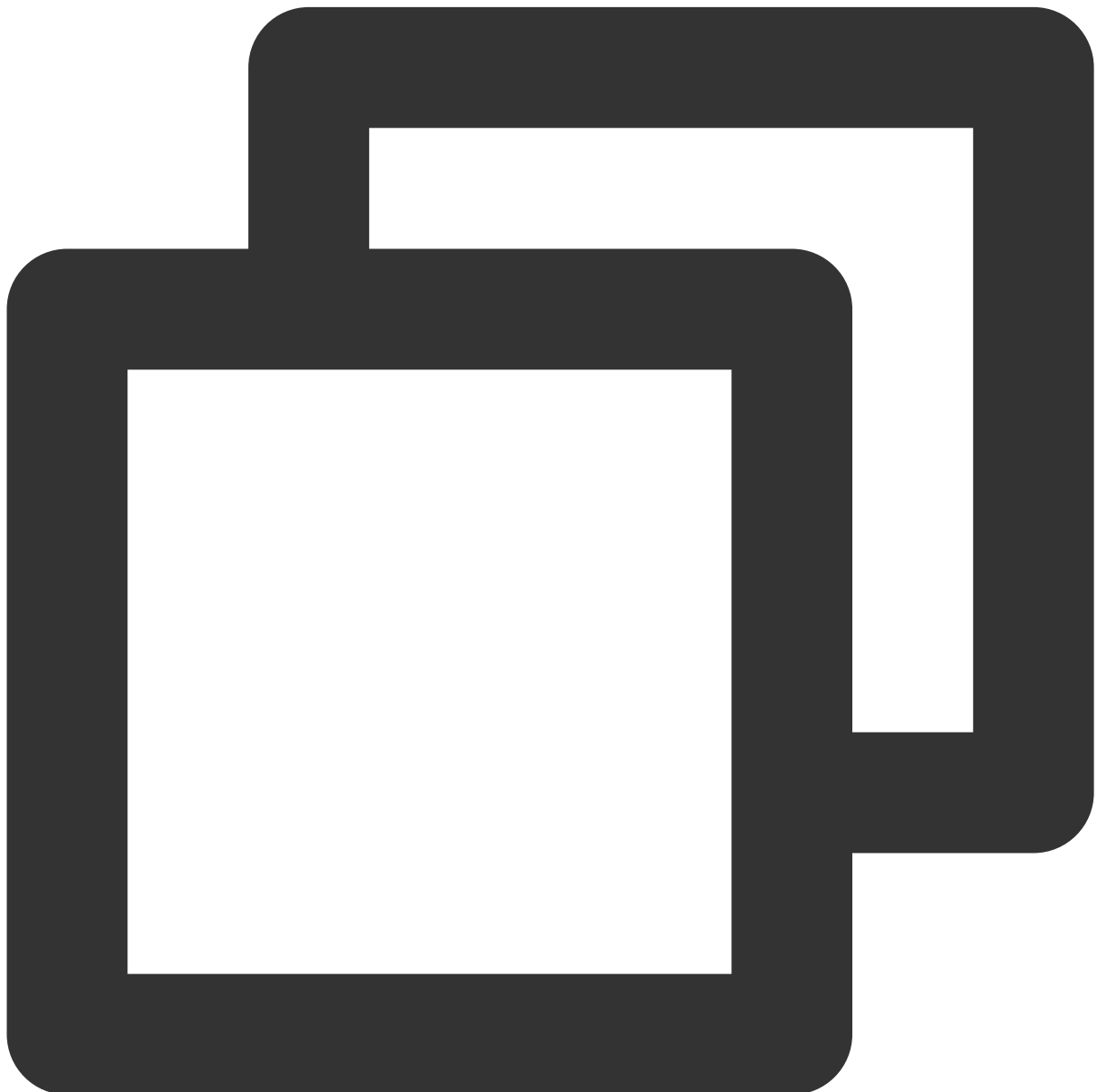
### Edge Node Authority

1. The edge nodes managed by the edge cluster K8S have the default permission setting of "system:node". For details on the "system:node" permission, please refer to [Using Node Authorization](#).

2. The kubeconfig file on the edge node will use the "system:node" permission by default. If you need to obtain the cluster certificate for operating the cluster, please refer to [Connecting a Cluster](#).

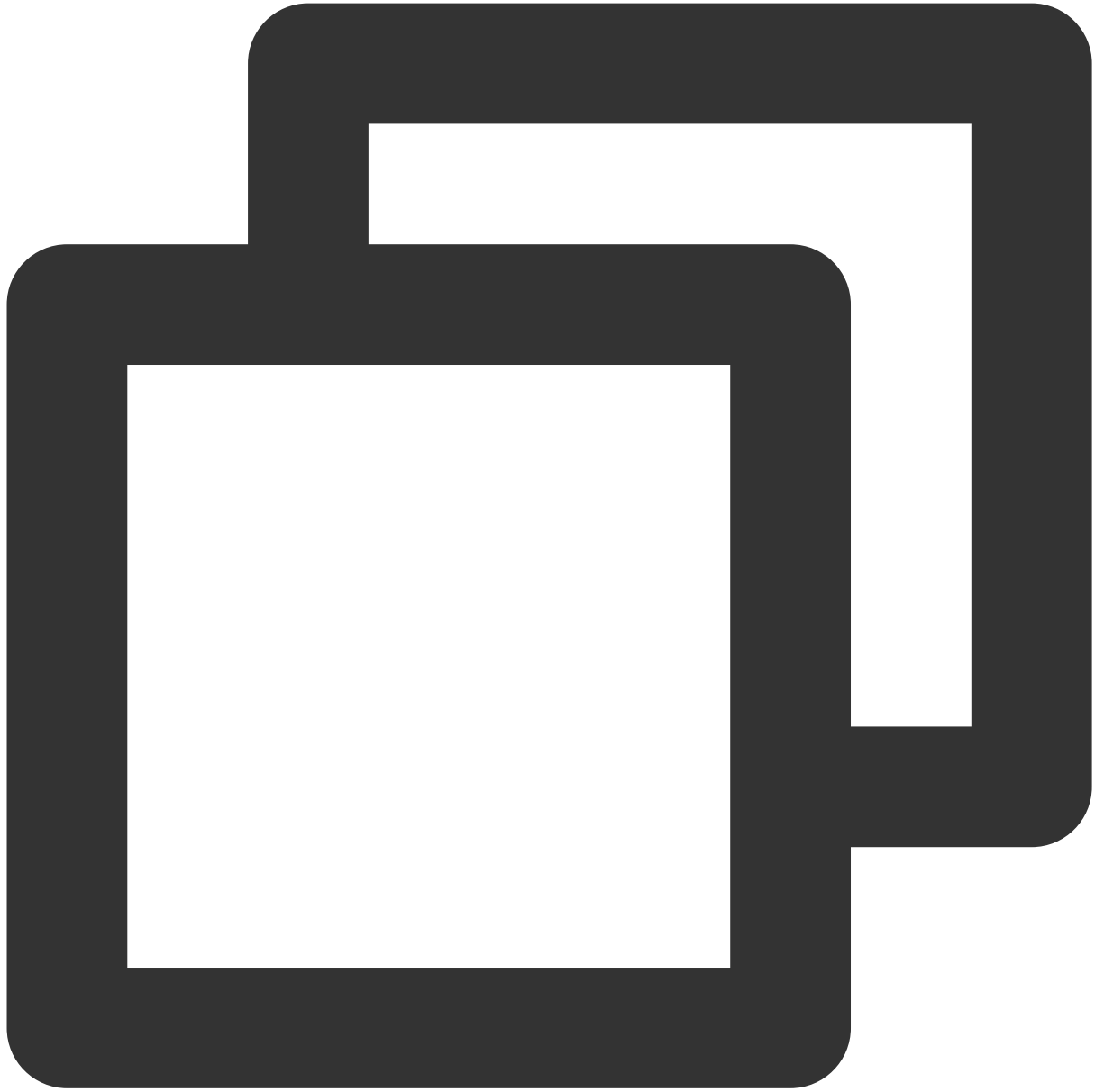
## Edge GPU Node

The edge cluster already has a nvidia-device-plugin DaemonSet submitted with image version nvidia-device-plugin:v0.9.0. If an edge node needs to deploy a DaemonSet, execute the following command to automatically deploy it to the corresponding node:



```
kubectl label nodes ${NODE_NAME} nvidia-device-enable=enable
```

To remove the taint added by the previous command, execute the following command:



```
kubectl label nodes ${NODE_NAME} nvidia-device-enable-
```

In the above commands, replace `\${NODE\_NAME}` with the name of the node where the operation will be performed.

### Versions and Operating Systems supported by Tencent Cloud GPU machines

The supported NVIDIA GPU series are:

[GN6](#)

[GN7](#)

[GN8](#)

[GN10X](#), [GN10Xp](#)

[GI3X](#)

**The following Linux systems are supported:**

CentOS 7.2 64-bit

CentOS 7.6 64-bit

CentOS 7.7 64-bit

CentOS 7.8 64-bit

CentOS 7.9 64-bit

Ubuntu Server 18.04.1 LTS 64-bit

TencentOS Server 2.4

Tencent Linux Release 2.4

If a corresponding system with an NVIDIA GPU version is used, the edge cluster will automatically install GPU drivers. Users of other systems can install the corresponding GPU drivers themselves. When adding an edge node via script, the edge cluster is not responsible for installing GPU drivers. The user should install the GPU drivers themselves.

#### **To add a CVM node with GPU support via script**

For edge nodes added with scripts, the edge cluster is not responsible for installing the GPU driver, which needs to be installed by the user.

#### **Note :**

The CVM system version with GPU is a subset of the [edge node system](#). If the system of the edge node does not support this version, adding a CVM node with GPU may fail. Therefore, please follow the requirements of the [edge node system](#) carefully.



# Node Management

Last updated : 2023-06-01 11:22:54

This article introduces how to manage the nodes that have already been added to the cluster.

## View node information

1. Log into the [Tencent Cloud Container Service console](#)
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Node management" > "Node"** on the left side of the page to enter the node list page.
4. Click on the node name you want to view, and enter the node details page. The detailed information of the node is displayed on this page:

**Pod Management:** View the Pod list and running status under this node.

**Events:** Display the Events information of this node within the Kubernetes cluster.

**Details:** Display the relevant host information and Kubernetes information of this node within the Kubernetes cluster.

**YAML:** Display the YAML information of this node's resource.

## Cordon Node

If you need to block a node and prevent subsequent applications from being scheduled to it, you can evict the node by adding **label:SchedulingDisabled** to it.

1. Log in to the TKE-Edge and select "Clusters" from the left navigation bar.
2. In the Cluster Management page, click on the Edge Cluster ID to enter the cluster details page.
3. Select **"Node Management" > "Nodes"** on the left side of the page to enter the node list page.
4. Choose the node you want to evict, and select **"Operation" > "Cordon"**. The node will enter the "Cordoned" state, and subsequent applications will not be scheduled to this node.
5. If you want to unblock the node, select **"Operation" > "Uncordon"**, and the node will return to the normal state.

## Delete Node

If a node is no longer needed within the cluster, you can delete it from the cluster.

1. Log in to the [Tencent Cloud Container Service console](#) and select "Clusters" from the left navigation bar.
2. In the Cluster Management page, click on the Edge Cluster ID to enter the cluster details page.
3. Select **"Node Management" > "Nodes"** on the left side of the page to enter the node list page.

4. Choose the node you'd like to delete and select **"Operation" > "Remove"**. The node will then be deleted from the Kubernetes cluster as a Node resource.

#### Caution

For nodes added using a script: This removal operation will only delete the Node from the Kubernetes cluster, equivalent to `kubectl delete node`, but it will not clean up the node data. Users are required to clean up the node data manually.

For nodes added using ECM/CVM: This removal operation will not only delete the Node from the Kubernetes cluster but will also release the applied ECM/CVM resources; users do not need to manually release these resources.

# Node Remote SSH

Last updated : 2023-06-01 11:36:31

## Scenario

This article describes how to remotely log in to edge nodes via SSH in the cloud. The edge nodes can be public cloud nodes or local intranet nodes.

[Enable node remote login](#)

[Close node remote login](#)

[Log in to the edge node remotely](#)

## Procedure

### Enable node remote login

1. Log into the [Tencent Cloud Container Service console](#)
2. On the Cluster Management page, click the cluster ID of the node you wish to remotely log into to enter the details page for that cluster.
3. Select **Node Management** > **Nodes** from the left-hand side of the page to enter the Nodes list page, and confirm whether the login hyperlink is displayed on the operation column. By default, the hyperlink is not displayed, as shown in the figure below:

Node list

Create ECM Node

Create CVM Node

Add nodes using scripts

Remove

Uncordon

Cordon

Search by node name

| <input type="checkbox"/> | Name | Availability Zone | Status  | Cluster Version    | IP             | PodCIDR      | CPU/MEM           | Time created      | Operation                                    |
|--------------------------|------|-------------------|---------|--------------------|----------------|--------------|-------------------|-------------------|--|
| <input type="checkbox"/> | gz-1 | -                 | Running | v1.22.5-tke.1.9... | 192.168.122.21 | 10.33.0.0/26 | 3.94-core/3.59... | 2023-05-29 11:... | <a href="#">Remove</a> <a href="#">Drain</a> |

Total items: 1

20 / page

1

/ 1

4. Select Basic Information from the left-hand side of the page to enter the Cluster **Basic Information** page, and click on the remote login switch to **enable the remote login** feature, as shown in the figure below:

The screenshot shows the 'Basic Information' page for a Kubernetes cluster. The left sidebar contains a navigation menu with options like 'Basic information', 'Node management', 'Namespace', 'Service Group', 'Workload', 'Auto scaling', 'Service', 'Configuration management', 'Add-on management', 'Log', and 'Event'. The main content area displays the following details:

| Basic information     |   |
|-----------------------|---|
| Cluster name          | edge-demo                                   |
| Cluster ID            | cls-3gmgtioi                                |
| Status                | Running...                                  |
| Runtime components    | containerd                                  |
| K8s version           | 1.22.5                                      |
| Deployment type       | Edge cluster                                |
| Region                | Southeast Asia(Singapore)                   |
| Cluster network       | <a href="#">vpc-c5ynz7i5</a>                |
| Pod CIDR              | 10.33.0.0/16                                |
| Service CIDR block    | 10.44.0.0/16                                |
| Kube-proxy proxy mode | iptables                                    |
| Enable remote login   | <input checked="" type="checkbox"/> Enabled |

For detailed directions on remote login, see [Remote Logging in to an Edge Node](#).

5. Return to **Node Management > Nodes** from the left-hand side of the page to access the Node List page again. Confirm that the **login** hyperlink is displayed, as shown in the figure below:

The screenshot shows the 'Node list' page. The left sidebar has 'Node management' expanded, with 'Node' selected. The main content area includes buttons for 'Create ECM Node', 'Create CVM Node', 'Add nodes using scripts', 'Remove', 'Uncordon', and 'Cordon'. Below these is a table of nodes:

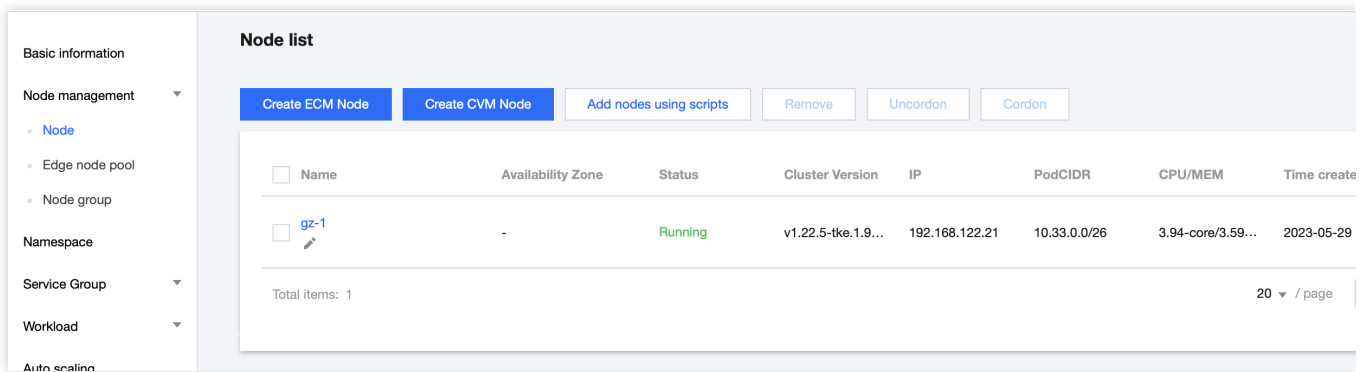
| Name | Availability Zone | Status  | Cluster Version    | IP             | PodCIDR      | CPU/MEM           | Time created      | Operation              |
|------|-------------------|---------|--------------------|----------------|--------------|-------------------|-------------------|------------------------|
| gz-1 | -                 | Running | v1.22.5-tke.1.9... | 192.168.122.21 | 10.33.0.0/26 | 3.94-core/3.59... | 2023-05-29 11:... | <a href="#">Log in</a> |

Total items: 1

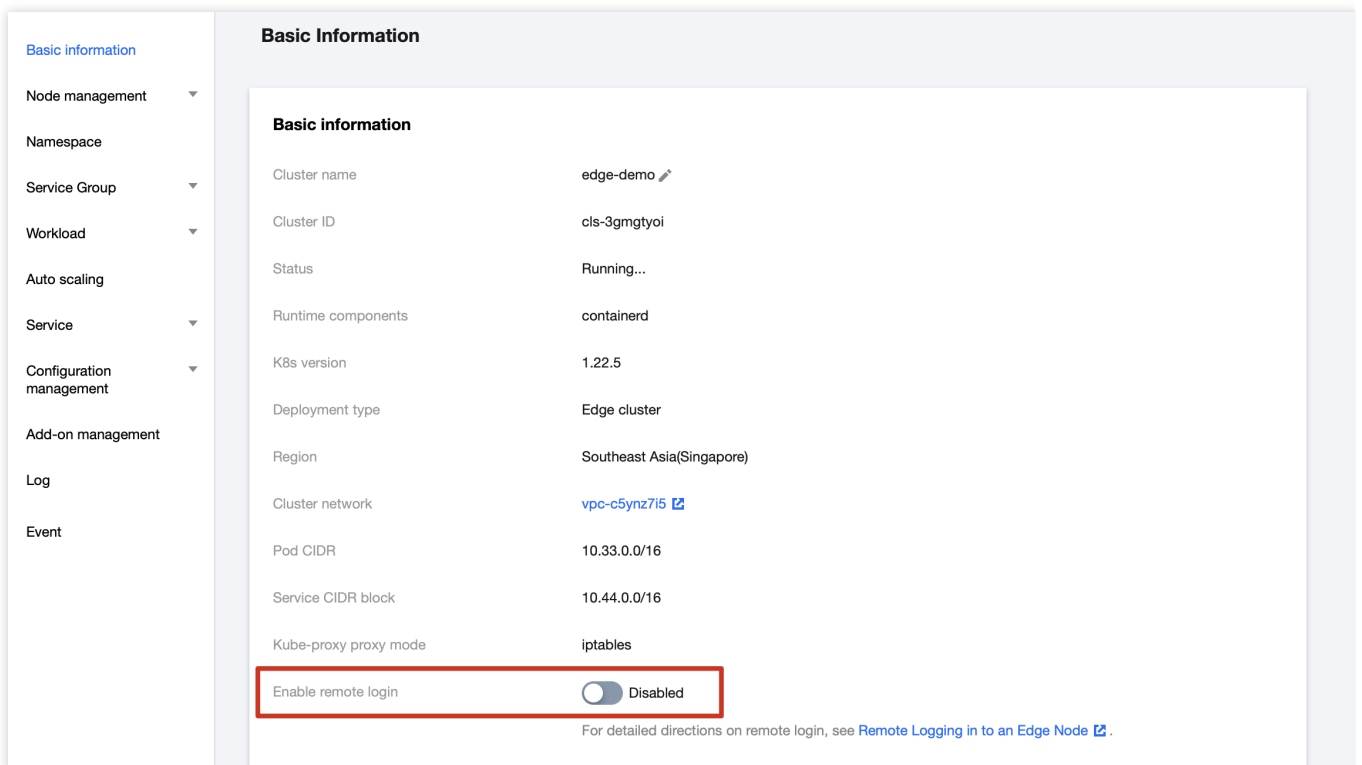
20 / page

## Close node remote login

1. Log into the [Tencent Cloud Container Service console](#).
2. On the Cluster Management page, click on the cluster ID of the node you wish to remotely log into to enter the details page for that cluster.
3. Select **Node Management > Nodes** from the left-hand side of the page to enter the Node List page, and confirm that the **login** hyperlink is displayed on the operation column. As shown in the figure below:



4. Select Basic Information from the left-hand side of the page to enter the Cluster **Basic Information** page, click on the remote login switch to **disable remote login** feature, as shown in the figure below:



5. Return to **Node Management > Nodes** from the left-hand side of the page to access the Node List page again, and confirm that the **login** hyperlink is hidden. As shown in the figure below:

**Node list**

Create ECM Node Create CVM Node Add nodes using scripts Remove Uncordon Cordon Search by node name

| <input type="checkbox"/> | Name | Availability Zone | Status  | Cluster Version    | IP             | PodCIDR      | CPU/MEM           | Time created      | Operation    |
|--------------------------|------|-------------------|---------|--------------------|----------------|--------------|-------------------|-------------------|--------------|
| <input type="checkbox"/> | gz-1 | -                 | Running | v1.22.5-tke.1.9... | 192.168.122.21 | 10.33.0.0/26 | 3.94-core/3.59... | 2023-05-29 11:... | Remove Drain |

Total items: 1 20 / page 1 / 1

## Log in to the edge node remotely

1. Log into the [Tencent Cloud Container Service console](#).
2. On the Cluster Management page, click on the cluster ID of the node you wish to remotely log into to enter the details page for that cluster.
3. Select **Node Management** > **Nodes** from the left-hand side of the page to enter the Node List page, and confirm that the **login** hyperlink is displayed on the operation column. As shown in the figure below:

**Node list**

Create ECM Node Create CVM Node Add nodes using scripts Remove Uncordon Cordon Search by node name

| <input type="checkbox"/> | Name | Availability Zone | Status  | Cluster Version    | IP             | PodCIDR      | CPU/MEM           | Time created      | Operation     |
|--------------------------|------|-------------------|---------|--------------------|----------------|--------------|-------------------|-------------------|---------------|
| <input type="checkbox"/> | gz-1 | -                 | Running | v1.22.5-tke.1.9... | 192.168.122.21 | 10.33.0.0/26 | 3.94-core/3.59... | 2023-05-29 11:... | Log in Remove |

Total items: 1 20 / page 1 / 1

4. On the "Node List" page, select the desired node to log in and click on **"Log In"**. You will be redirected to the **remote node login console** page, as shown in the figure below:

**Standard Login|Linux Instance**

Password **SSH Key**

IP: 172.16.17.136

Port: 22

Username: root

Key:   
 [Select Key](#)

[Login](#)

If the SSH key login is successful, you will be directed to the console page, as shown in the figure below:

```
Warning: Permanently added '192.168.122.21' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.4 LTS (GNU/Linux 5.4.0-136-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

System information as of Mon 29 May 2023 11:03:48 PM CST

System load:  0.01               Processes:            208
Usage of /:   41.6% of 48.17GB   Users logged in:     1
Memory usage: 22%               IPv4 address for ens3: 192.168.122.21
Swap usage:   0%                IPv4 address for ens3: 192.168.122.200

149 updates can be applied immediately.
83 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable

Last login: Mon May 29 23:02:46 2023 from 192.168.122.19
dodia@gz-1:~$
```

Type in a command, such as 'kubectl get node', in the console and press enter to execute. View the results of the command, as shown in the figure below:

```
root@gz-1:~# kubectl get nodes
NAME    STATUS    ROLES    AGE    VERSION
gz-1    Ready    <none>    3h28m  v1.22.5-tke.1.9+5f76a1720f9839
```

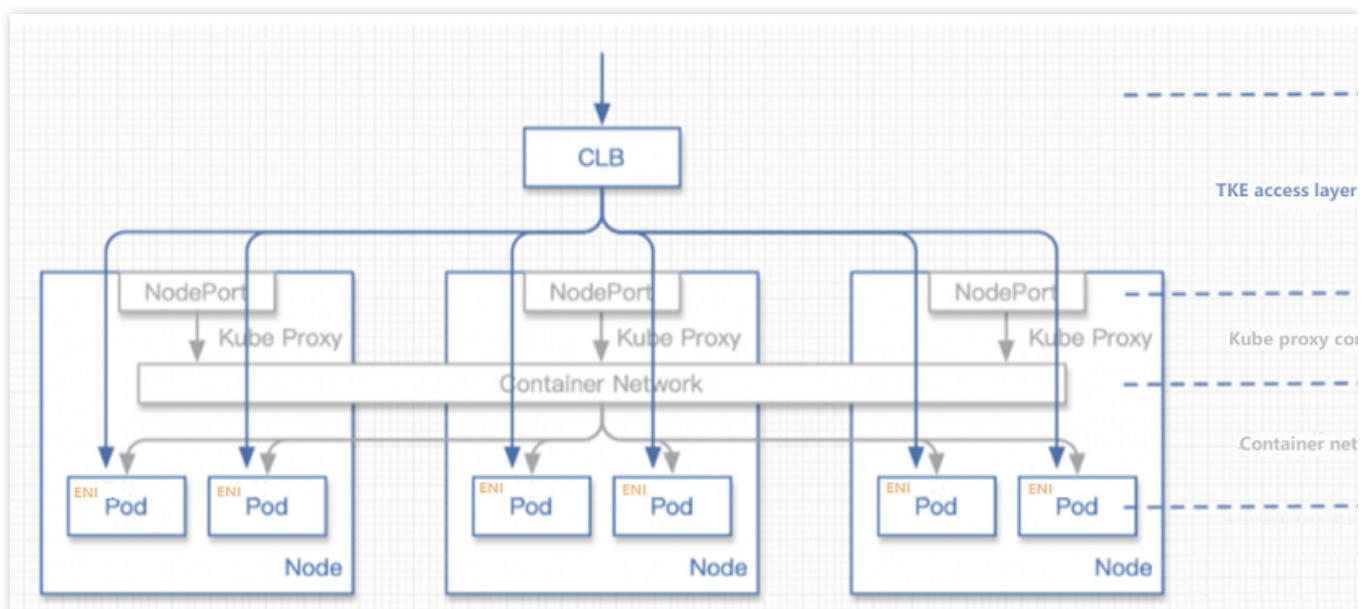
# Independent ENI

Last updated : 2023-06-01 11:39:14

## Overview

If you add Tencent Cloud CVMs to the edge cluster as edge nodes, the platform allows you to enable ENIs and bind Pods on CVM nodes with independent ENIs to implement a high-availability network scheme.

The network architecture is as follows:



You can also use ENIs in the VPC where the CVMs reside to open Pods to external users and then bind different ENIs to the CLB to enable high-performance network forwarding.

[Enabling the independent ENI](#)

[Disabling the independent ENI](#)

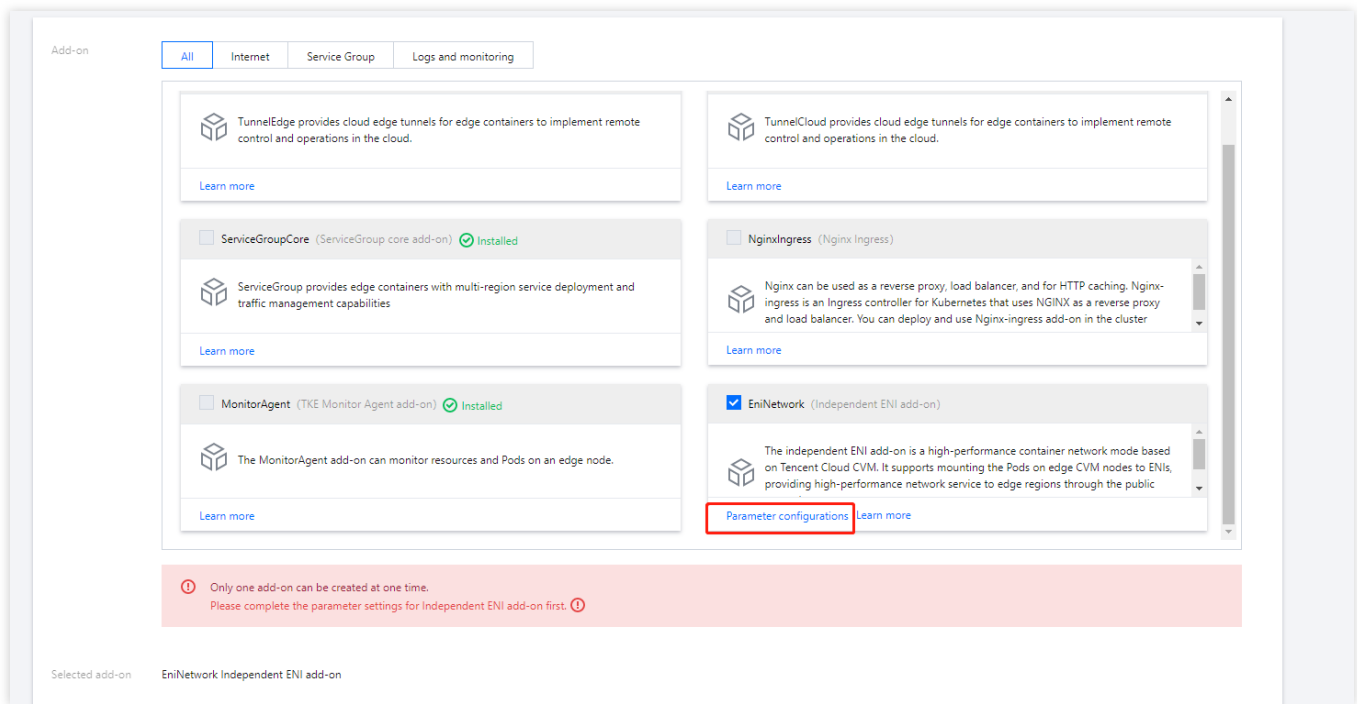
## Directions

### Enabling the independent ENI

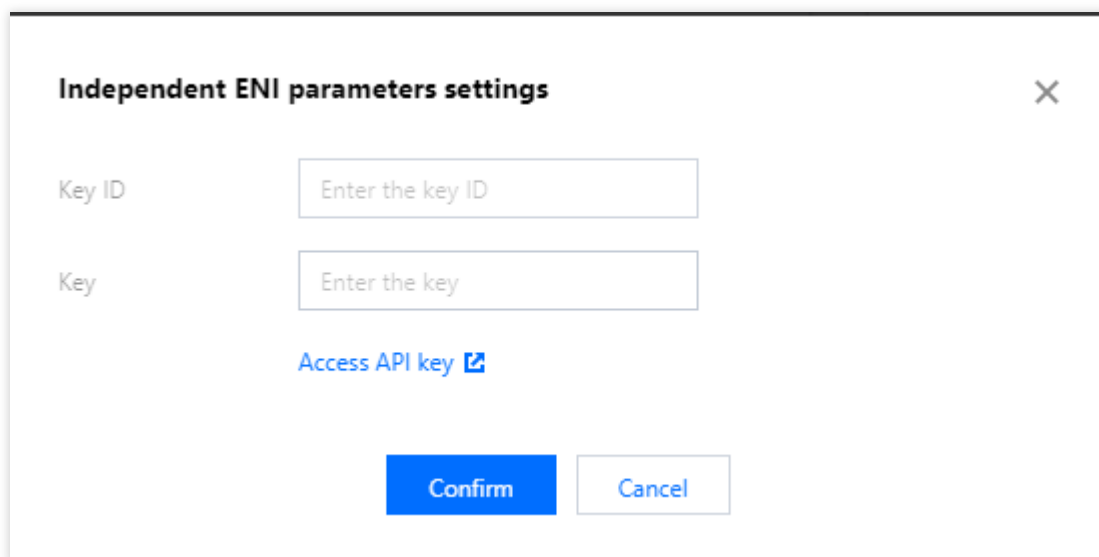
1. Log in to the [TKE console](#).



2. On the **Cluster Management** page, click the ID of the target cluster to enter its details page.
3. Select **Add-On Management** on the left of the page and click **Create** on the add-on list page.
4. On the **Create Add-On** page, select **EniNetwork (independent ENI add-on)** and click **Parameter Configurations** as shown below:



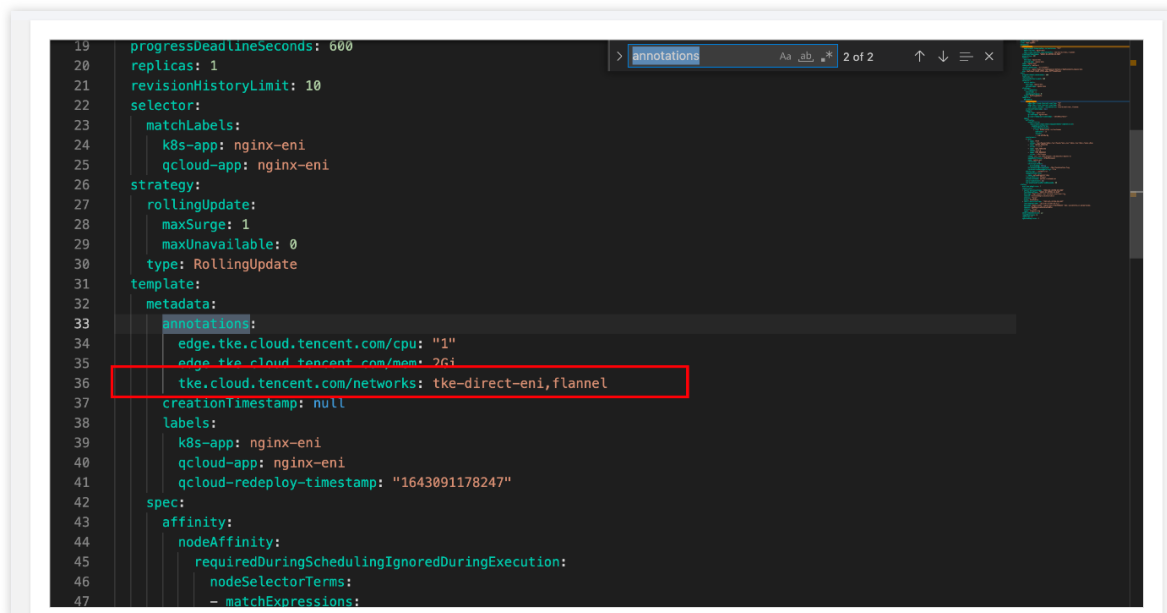
- 4.1 In the **Independent ENI parameter settings** pop-up window, click **Access API key** to enter the key information page.



- 4.2 Create a key and copy the `SecretId` and `SecretKey` in [API Key Management](#).
- 4.3 In the **Independent ENI parameter settings** pop-up window, enter `SecretId` and `SecretKey` and click **Confirm**.

5. On the **Create Add-On** page, click **Done** to enable the ENI.
6. Select **Workload > Deployment** on the left of the page to enter the Deployment list page. If there is already a Deployment in the list, skip this step; otherwise, create one as instructed in [Deployment Management](#).
7. Select **Node Management > Node** on the left of the page to enter the node list page. If there is already a CVM node in the list, skip this step; otherwise, [create one](#).
8. Configure the ENI in the Pod of the target edge cluster.

The configuration is as follows:

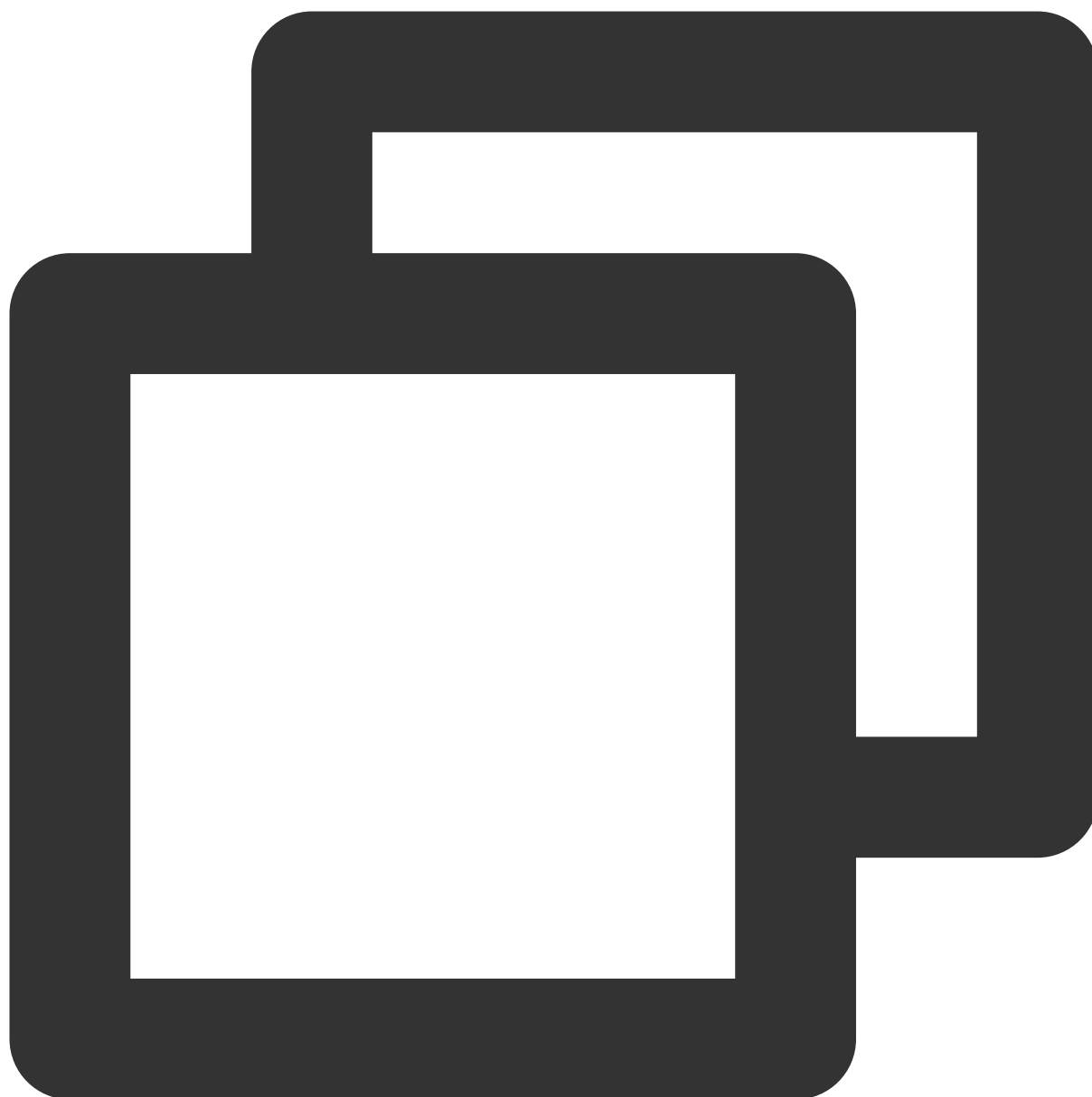


```
19 progressDeadlineSeconds: 600
20 replicas: 1
21 revisionHistoryLimit: 10
22 selector:
23   matchLabels:
24     k8s-app: nginx-eni
25     qcloud-app: nginx-eni
26 strategy:
27   rollingUpdate:
28     maxSurge: 1
29     maxUnavailable: 0
30   type: RollingUpdate
31 template:
32   metadata:
33     annotations:
34       edge.tke.cloud.tencent.com/cpu: "1"
35       edge.tke.cloud.tencent.com/mem: "2Gi"
36       tke.cloud.tencent.com/networks: tke-direct-eni, flannel
37   creationTimestamp: null
38   labels:
39     k8s-app: nginx-eni
40     qcloud-app: nginx-eni
41     qcloud-redeploy-timestamp: "1643091178247"
42 spec:
43   affinity:
44     nodeAffinity:
45       requiredDuringSchedulingIgnoredDuringExecution:
46         nodeSelectorTerms:
47           - matchExpressions:
```

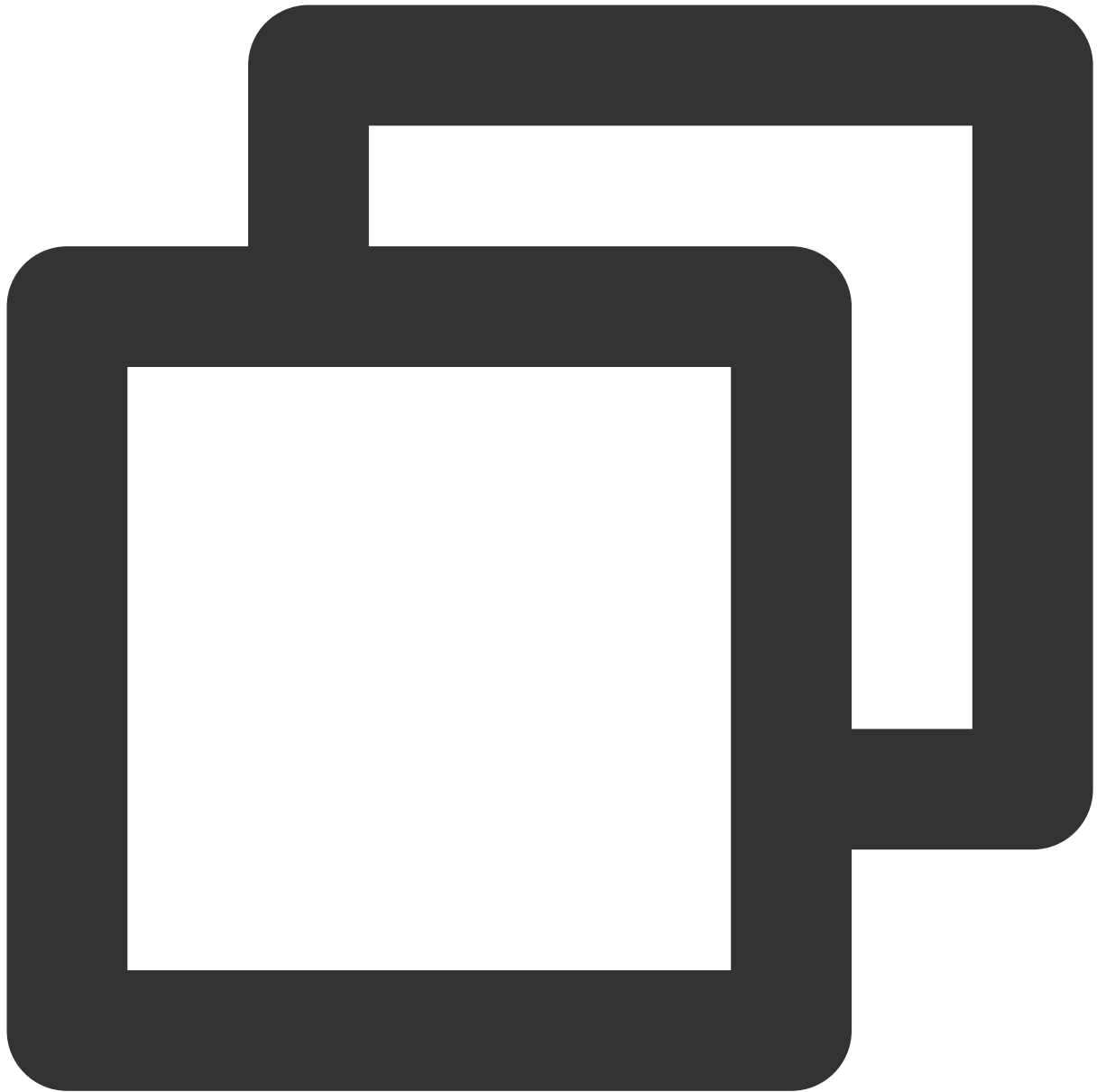
The ENI capability of edge clusters is supported only by Tencent Cloud CVM node resources. Therefore, when deploying an application, you need to use the `nodeAffinity` capability to schedule the Pod mounted with the ENI to the real CVM edge node (you can enter multiple CVM node IDs).

```
34   edge.tke.cloud.tencent.com/mem: 2Gi
35   creationTimestamp: null
36   labels:
37     k8s-app: nginx-eni
38     qcloud-app: nginx-eni
39     qcloud-redeploy-timestamp: "1643091178247"
40   spec:
41     affinity:
42       nodeAffinity:
43         requiredDuringSchedulingIgnoredDuringExecution:
44           nodeSelectorTerms:
45             - matchExpressions:
46               - key: kubernetes.io/hostname
47                 operator: In
48                 values:
49                   - cvm-2cxqi4ow #访问目标的cvm节点ID
50   containers:
51   - env:
52     - name: PATH
53       value: /usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin
54     - name: NGINX_VERSION
55       value: 1.21.5
56     - name: NJS_VERSION
57       value: 0.7.1
58     - name: PKG_RELEASE
59       value: 1~bullseye
60     image: ccr.ccs.tencentyun.com/wdjtest/nginx:v1
61     imagePullPolicy: IfNotPresent
```

Below is the actual code:



```
template:
  metadata:
    annotations:
      tke.cloud.tencent.com/networks: tke-direct-eni,flannel
```



```
spec:
  affinity:
    nodeAffinity:
      requiredDuringSchedulingIgnoredDuringExecution:
        nodeSelectorTerms:
          - matchExpressions:
              - key: kubernetes.io/hostname
                operator: In
                values:
                  - cvm-2cxgi4ow # CVM node ID of the access target
```

## Disabling the independent ENI

1. Log in to the [TKE console](#).
2. On the cluster management page, click the target cluster ID to enter its details page.
3. Select **Add-On Management** on the left of the page and click **Delete** on the right of the target add-on on the add-on list page.
4. In the **Delete Resource** pop-up window, click **Confirm**.

# Edge and Multi-region Deployment

## Node Pool Management

### NodeUnit

Last updated : 2023-06-01 11:22:54

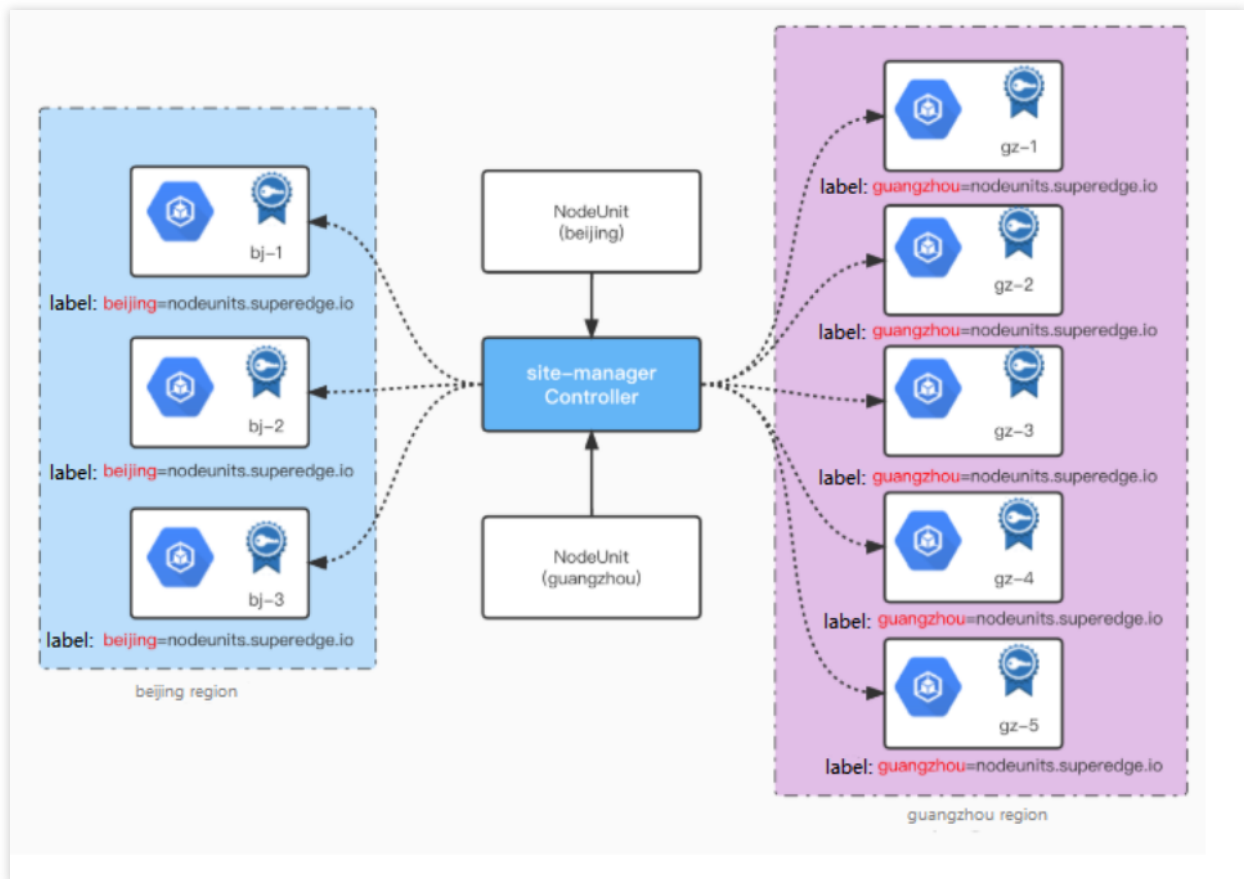
## Operation Scenario

This document describes how to manage edge node pools of an edge container in the TKE console.

This update reconstructs the UI interaction between the earlier versions of NodeGroup and NodeUnit. Clusters created after March 29, 2022 will use the new interaction logic, and clusters on earlier versions will not be affected. In this update, the edge node pool corresponds to the design of NodeUnit in SuperEdge, and node pool category corresponds to the design of NodeGroup in SuperEdge.

## Overview

In edge scenarios, nodes are assigned more attributes. For example, different nodes can be placed in different network environments, architectures, and cloud services. In some specific scenarios, users can divide nodes with the same characteristics into different groups. For example, if three nodes are located in the Beijing region and the other five nodes are located in the Guangzhou region, we can divide them into two node pools, Beijing and Guangzhou, according to their geographical attributes. Different node pools can schedule and deploy different applications. We name a node pool 'NodeUnit' resource, as a CRD resource implemented for Kubernetes. The platform interacts with users through this CRD and manages nodes in groups by labeling nodes through Kubernetes, as shown in the figure below:



## Directions

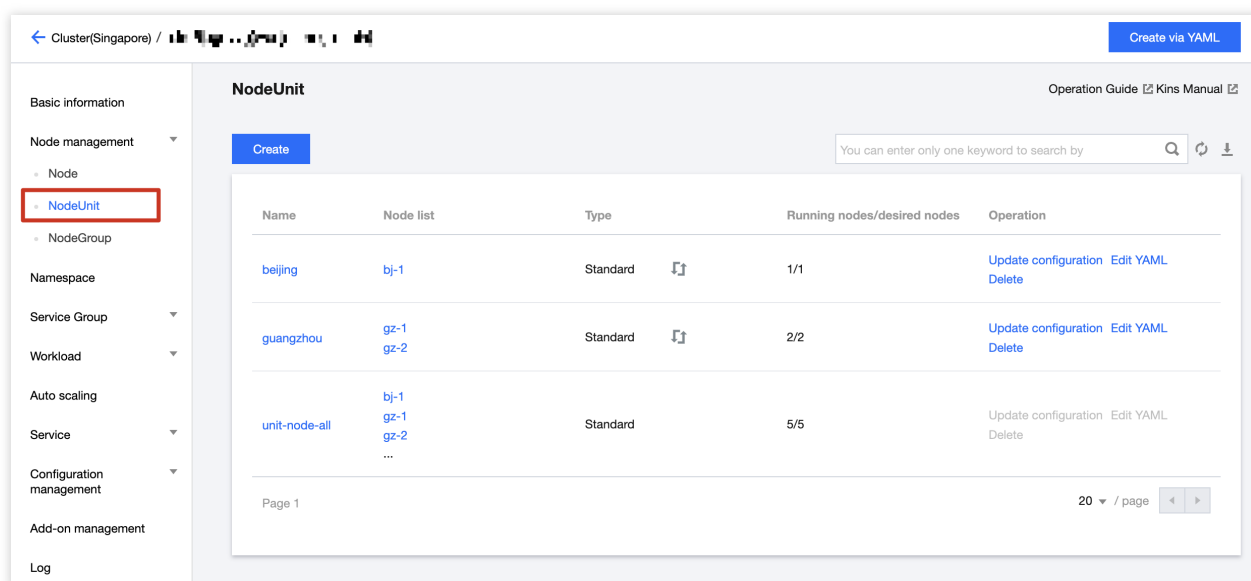
Note :

As operations on an edge node pool will affect the labels of the nodes, perform such operations with caution.

### Creating an edge node pool

1. Log in to the [TKE console](#).
2. On the **Cluster Management** page, click the ID of the target cluster to go to its details page.
3. Select **Node management > NodeUnit** in the left sidebar to go to the NodeUnit page.





Each cluster has a default node pool ( `unit-node-all` ), which contains all edge nodes added to the cluster.

4. Click **Create**. In the **Create edge node pool** pop-up window, enter the node pool name and add available nodes as needed.

### Create node pool

Node pool name

The name can contain up to 60 characters. It supports lower-case letters, digits and "-", and must start with a letter and end with a digit or letter.

Node list

Available nodes of the cluster:1/1 loaded0 items selected

Separate multiple keywords with "|"

| <input type="checkbox"/> | Node ID/name | Status  |
|--------------------------|--------------|---------|
| <input type="checkbox"/> | gz-1         | Running |

| Node ID/name | Status |
|--------------|--------|
|--------------|--------|

Press and hold Shift key to select more

Done

Cancel

5. Click **Done**. You can view added nodes in the edge node pool list.

## Managing edge node pools

1. Log in to the [TKE console](#).
2. On the **Cluster Management** page, click the ID of the target cluster to go to its details page.
3. Select **Node management** > **NodeUnit** in the left sidebar to go to the NodeUnit list page.
4. Click **Update configuration** on the right of an edge node pool. On the **Update NodeUnit** page, you can add and delete existing nodes.

## Deleting an edge node pool

1. Log in to the [TKE console](#).
2. On the **Cluster Management** page, click the ID of the target cluster to go to its details page.
3. Select **Node management** > **NodeUnit** in the left sidebar to go to the **NodeUnit list** page.
4. Click **Delete** on the right of an NodeUnit to delete the NodeUnit.

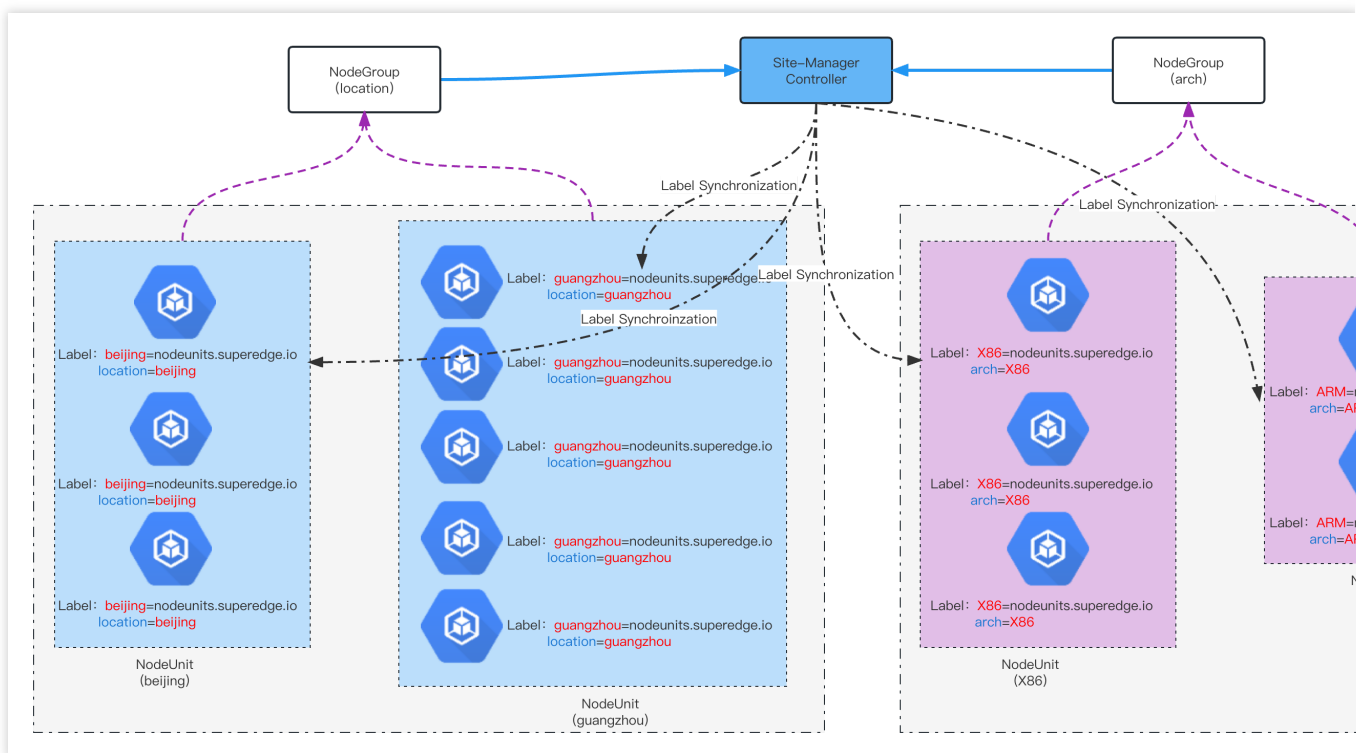
# NodeGroup

Last updated : 2023-06-01 11:44:35

This article introduces how to manage multiple edge node pools(NodeUnit) in the TKE-Edge through the console UI for group management.

## Overview

Node pool classification(NodeGroup) allows further grouping of different node pools(NodeUnit). By adding mutually exclusive node pools to the same group, applications can be deployed in this group, enabling one-click batch deployment capabilities across multiple node pools(NodeUnit). The architecture is as following:



For example, you can classify the nodes in the cluster according to two types of scenarios based on your requirements. One scenario is by location, and the other scenario is by architecture. Each scenario represents a NodeGroup Custom Resource (CR) instance - location and arch. Then, you can add different NodeUnits to the corresponding NodeGroup categories.

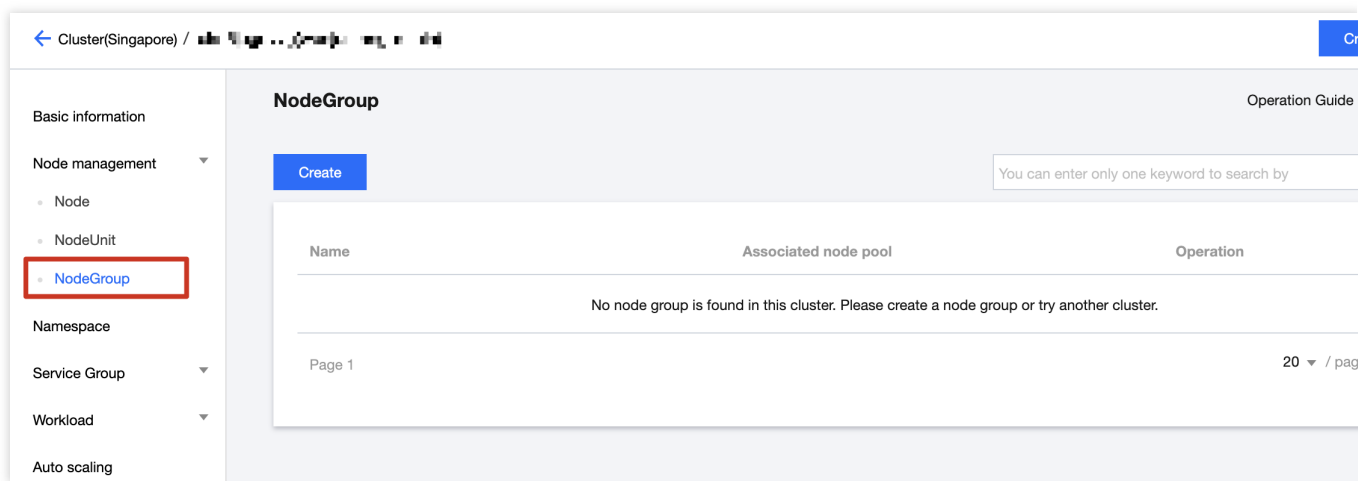
After creating different NodeGroups based on this structure, you can use the TKE-Edge's DeploymentGrid/StatefulsetGrid/ServiceGrid capabilities for multi-region deployment. For instance, you can create a DeploymentGrid for Nginx instances, bind it to the location-based NodeGroup, and both beijing and guangzhou

NodeUnits will deploy completely identical Nginx Deployment instances, providing access services across both regions simultaneously.

## Operation Steps

### Create NodeGroup

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Node Management" > "Node group"**, enter the NodeGroup list page, as follow:



4. Click **"Create"**, select the NodeUnit name on the "Create node pool" page, and associate the NodeUnit as required.

### Create node group

Node group name

location

The name can contain up to 60 characters. It supports lower-case letters, digits and "-", and must start with a letter and end with a digit or letter.

Associated node pool

Available nodes of the cluster:3/3 loaded2 items selected

Separate multiple keywords with "|"

Node ID/name

☒ beijing

☒ guangzhou

☐ unit-node-all

Node ID/name

beijing

guangzhou

Press and hold Shift key to select more

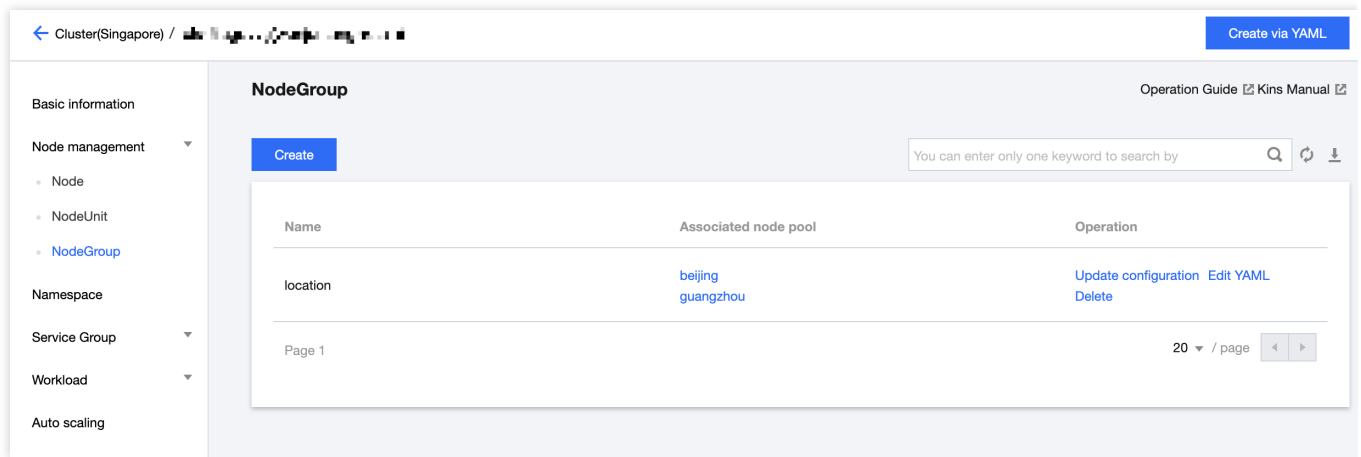
Done

Cancel

5. Click **"Done"**, and this "location" NodeGroup will be shown in the list page.

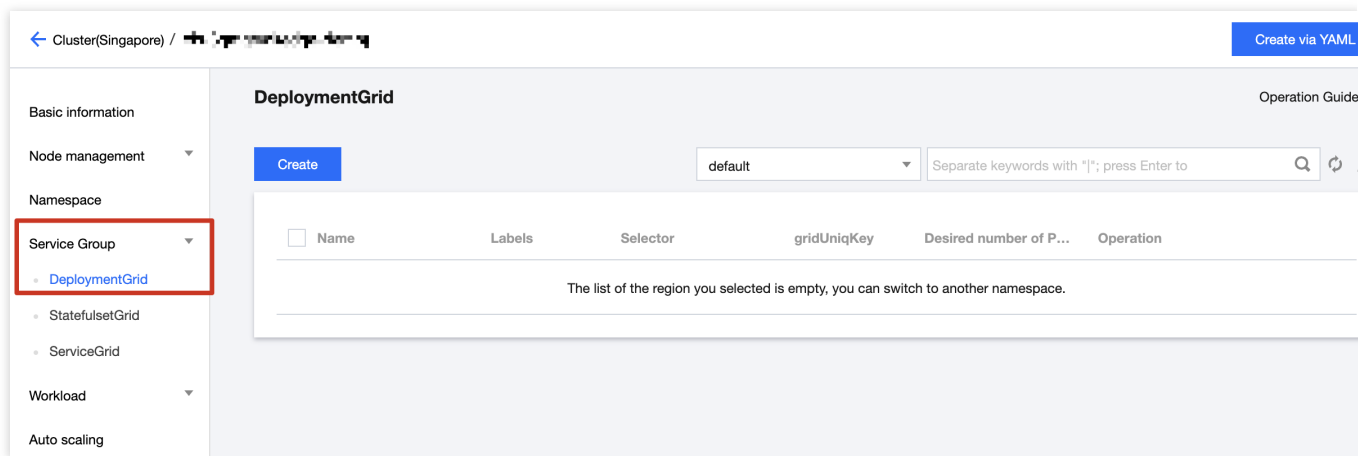
## Manage NodeGroup

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Node Management"** > **"Node group"**, enter the NodeGroup list page.
4. Select the "NodeGroup" you want to update, click **"Update configuration"** on the right side and you can modify the NodeGroup configuration as you want.



## How to use NodeGroup

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Service Group" > "DeploymentGrid"**, step into the list page.



4. Click **"Create"**, and input the parameters on the "Create DeploymentGrid" page as needed, select the NodeGroup as you want, for example "location". The deployment will be automatically created on all corresponding NodeUnit:

[←](#) Create DeploymentGrid

Workload name   
Up to 40 chars. It supports only lower-case chars, numbers and hyphens. It must start with a lower-case letter and end with a number or lower-case letter.

NodeGroup

Description

Label  =  [×](#)  
Add variable  
The label name can only contain letters, numbers and symbols ("-", "\_", ".", "/"), and must start and end with a number or letter.  
The label value can only contain letters, numbers and symbols ("-", "\_", ".", "/"), and must start and end with a number or letter.

Namespace

Volume (optional) [Add volume](#)  
Used for container storage. It supports temp directory, NFS, config file, PVC, and should be mounted to the specified directory of the container. [Instruction](#)

Containers in the Pod [✓](#) [×](#)

Name

[←](#) Cluster(Singapore) /        [Create via YAML](#)

Basic information

Node management [▼](#)

Namespace


Service Group [▼](#)• [DeploymentGrid](#)

• StatefulsetGrid

• ServiceGrid

## DeploymentGrid

[Operation Guide](#)[Create](#)

| <input type="checkbox"/> Name  | Labels           | Selector                | gridUniqKey | Desired number of P... | Operation  |
|--|------------------|-------------------------|-------------|------------------------|--|
| <input type="checkbox"/> <a href="#">nginx</a>  | k8s-app:nginx... | k8s-app:nginx, qclou... | location    | 1                      | <a href="#">Edit YAML</a> <a href="#">Delete</a> |





# ServiceGroup Manual

Last updated : 2023-07-14 18:37:26

## Features

In edge computing scenarios, it is common to manage multiple edge sites within the same cluster, with each edge site containing one or more computing nodes.

It is also desirable to run a set of interconnected services within each site, with the services within each site providing a complete set of features for users.

Due to network constraints, interconnected services may not want or be able to access each other across sites.

Considering these three characteristics of edge computing, TKE-Edge has specifically designed a custom resource logic called **ServiceGroup** to address application distribution and service governance issues encountered in multi-regional scenarios.

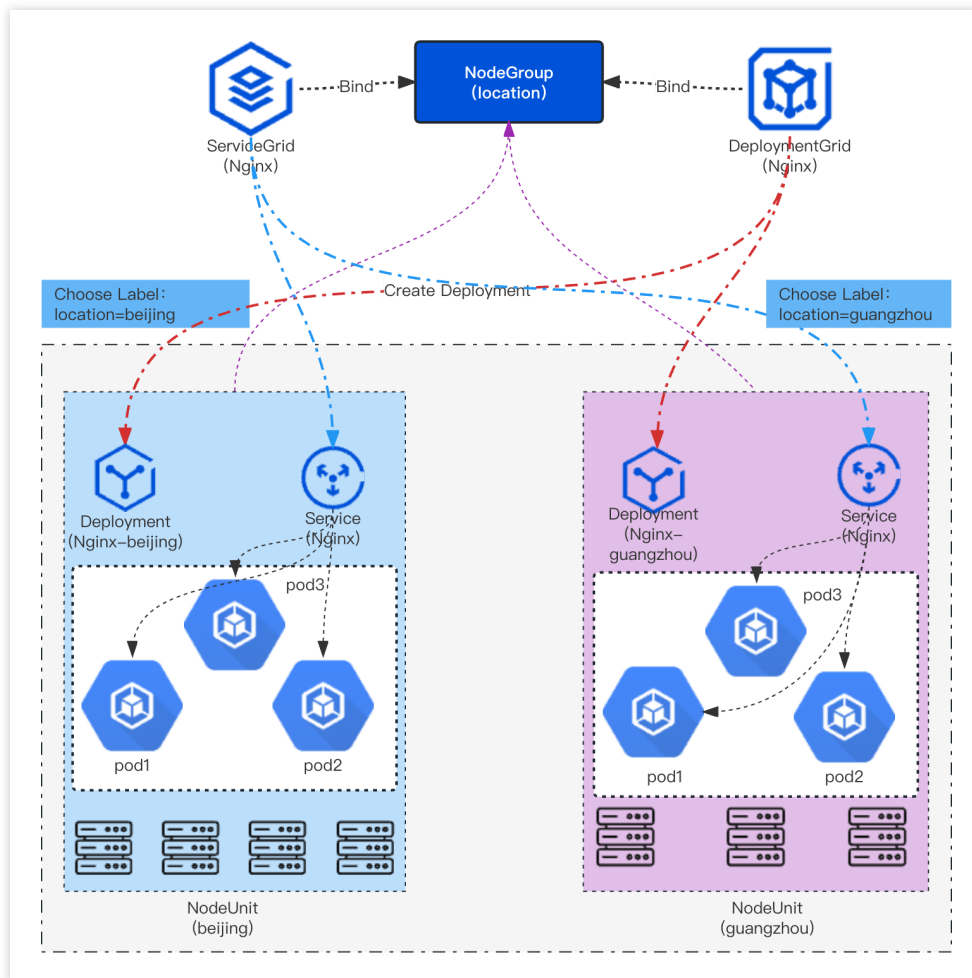
## Scenarios

**ServiceGroup** allows for conveniently deploying a set of services within different data centers or regions that belong to the same cluster. It ensures that requests between services can be completed within the same data center or region, avoiding cross-regional access.

Native Kubernetes cannot directly control the specific node locations where Deployment Pods are created. This must be done indirectly by orchestrating node affinity planning. When the number of edge sites and the number of services to be deployed are too high, the management and deployment aspects can become extremely complex or even only theoretically feasible. Furthermore, in order to limit service interactions within a certain range, the business side needs to create dedicated Services for each Deployment, resulting in a huge and error-prone workload that may lead to online business anomalies.

ServiceGroup is designed specifically for this scenario. Users only need to use the two TKE-Edge self-developed Kubernetes resources provided by ServiceGroup, **DeploymentGrid** and **ServiceGrid**, to conveniently deploy services to these node groups and control service traffic. This approach also ensures the number of services and disaster recovery in each region.

## Architecture



## Basic Concept

ServiceGroup needs to be used in conjunction with NodeUnit and NodeGroup. The concepts are detailed as follows:

### NodeUnit

NodeUnit typically consists of one or more computing resource instances located within the same edge site. It is essential to ensure that nodes within a NodeUnit have an internal network connection.

For more detail information, please refer to: [NodeUnit](#)

### NodeGroup

Multiple NodeUnits: NodeGroup includes one or more NodeUnits.

Deploying services on every NodeUnit: It ensures that the services in the ServiceGroup are deployed on every NodeUnit within the group.

Automatic deployment to new NodeUnits: When a NodeUnit is added to the cluster, the services within the ServiceGroup are automatically deployed to the newly added NodeUnit.

For more detail information, please refer to: [NodeGroup](#)

## ServiceGroup

ServiceGroup is not a definition of a concrete resource, but rather a collection of Kubernetes custom resources. It is an abstract resource, and multiple ServiceGroups can be created within a cluster.

A ServiceGroup contains one or more business services and is applicable to the following scenarios:

Business services need to be packaged for deployment.

Business services need to run in each NodeUnit and ensure a specific number of pods.

Business services need to control the interactions between services within the same NodeUnit, and cannot forward traffic to other NodeUnits.

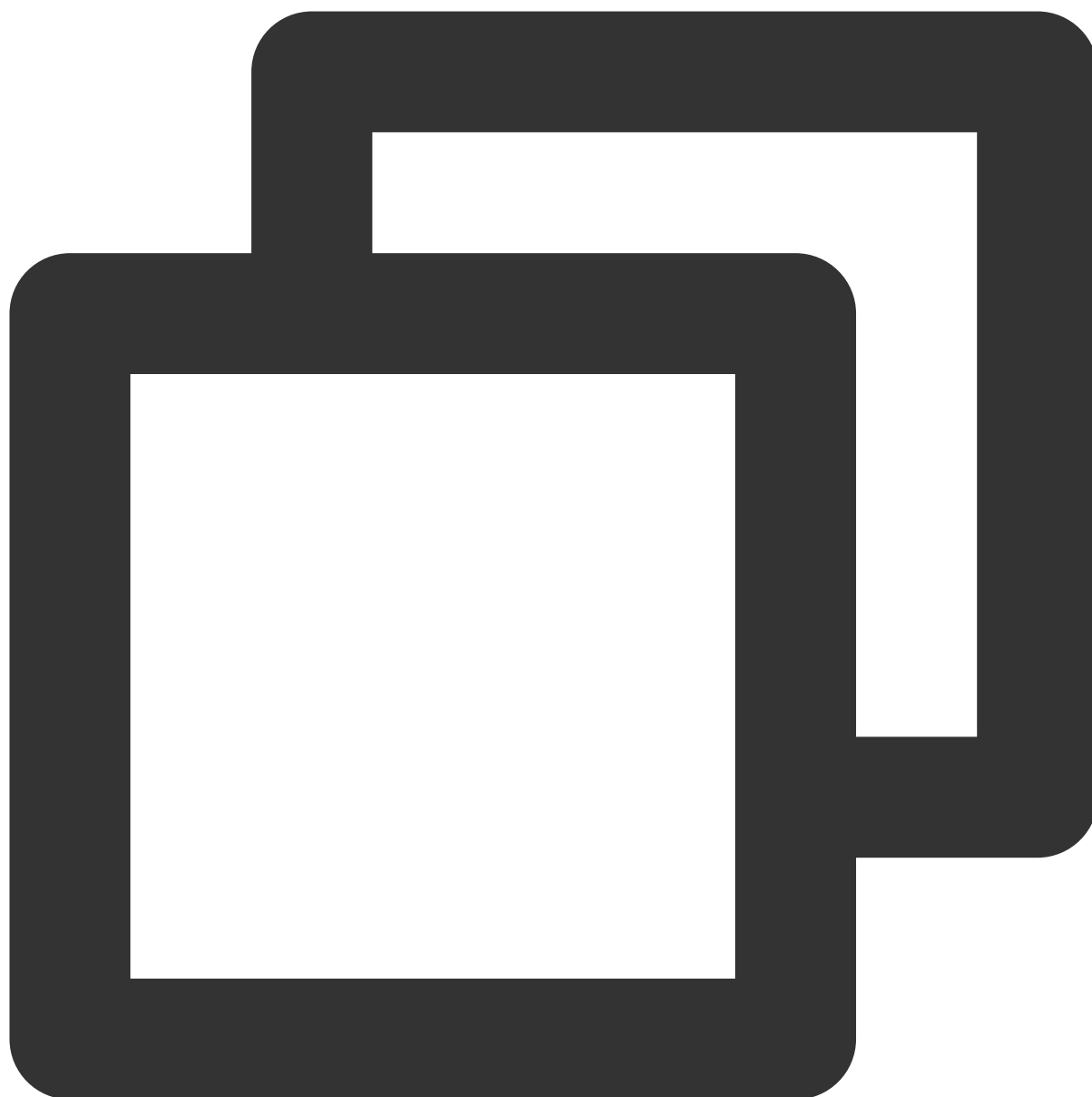
The resource types involved in ServiceGroup include the following three categories:

DeploymentGrid

StatefulSetGrid

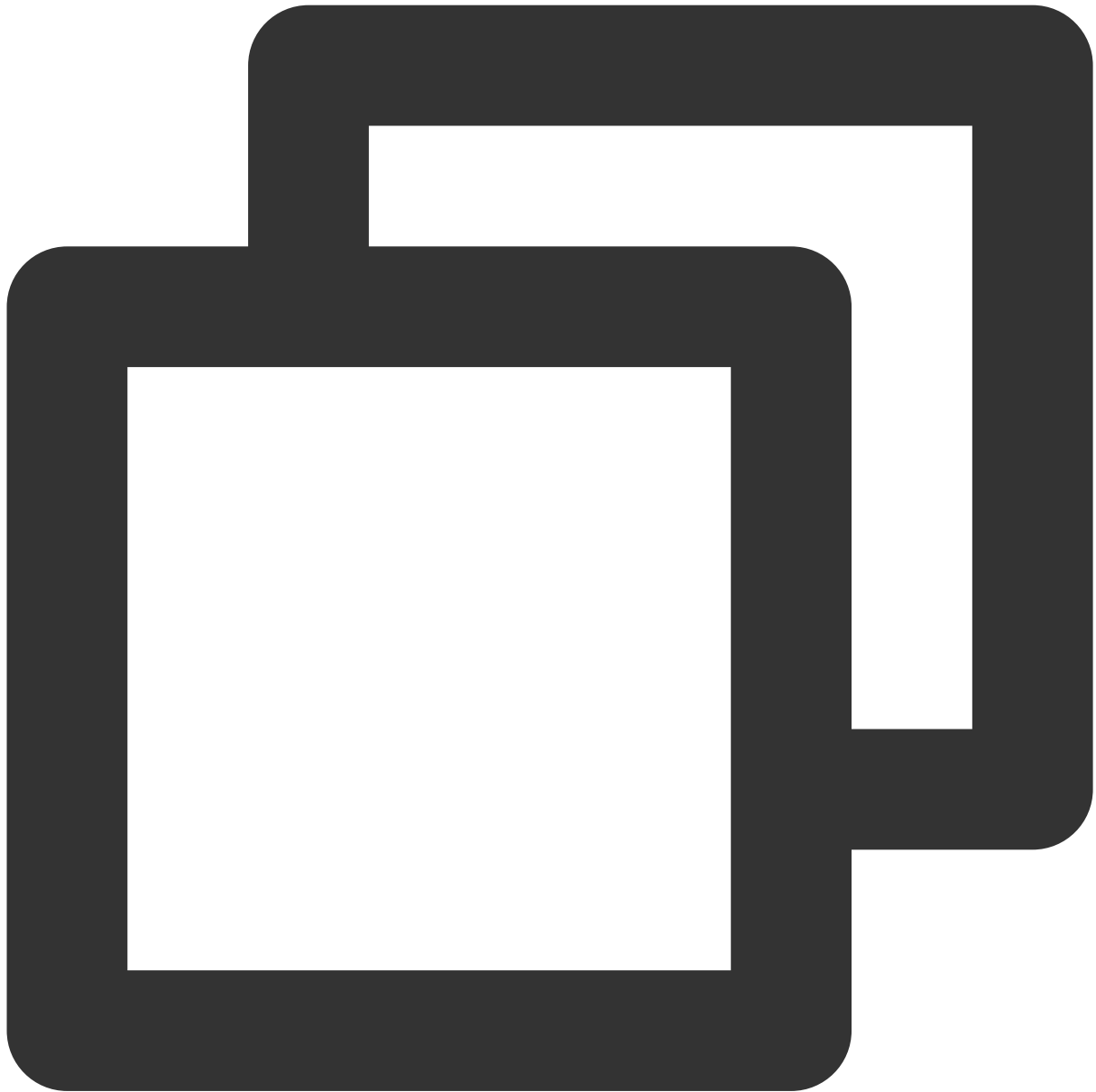
ServiceGrid

The format of DeploymentGrid is similar to Deployment. The `<deployment-template>` field corresponds to the original `template` field of the deployment. What is special about DeploymentGrid is the `gridUniqKey` field, which specifies the key value of the label for node grouping:



```
apiVersion: superedge.io/v1
kind: DeploymentGrid
metadata:
  name:
  namespace:
spec:
  gridUniqKey: <NodeLabel Key>
  <deployment-template>
```

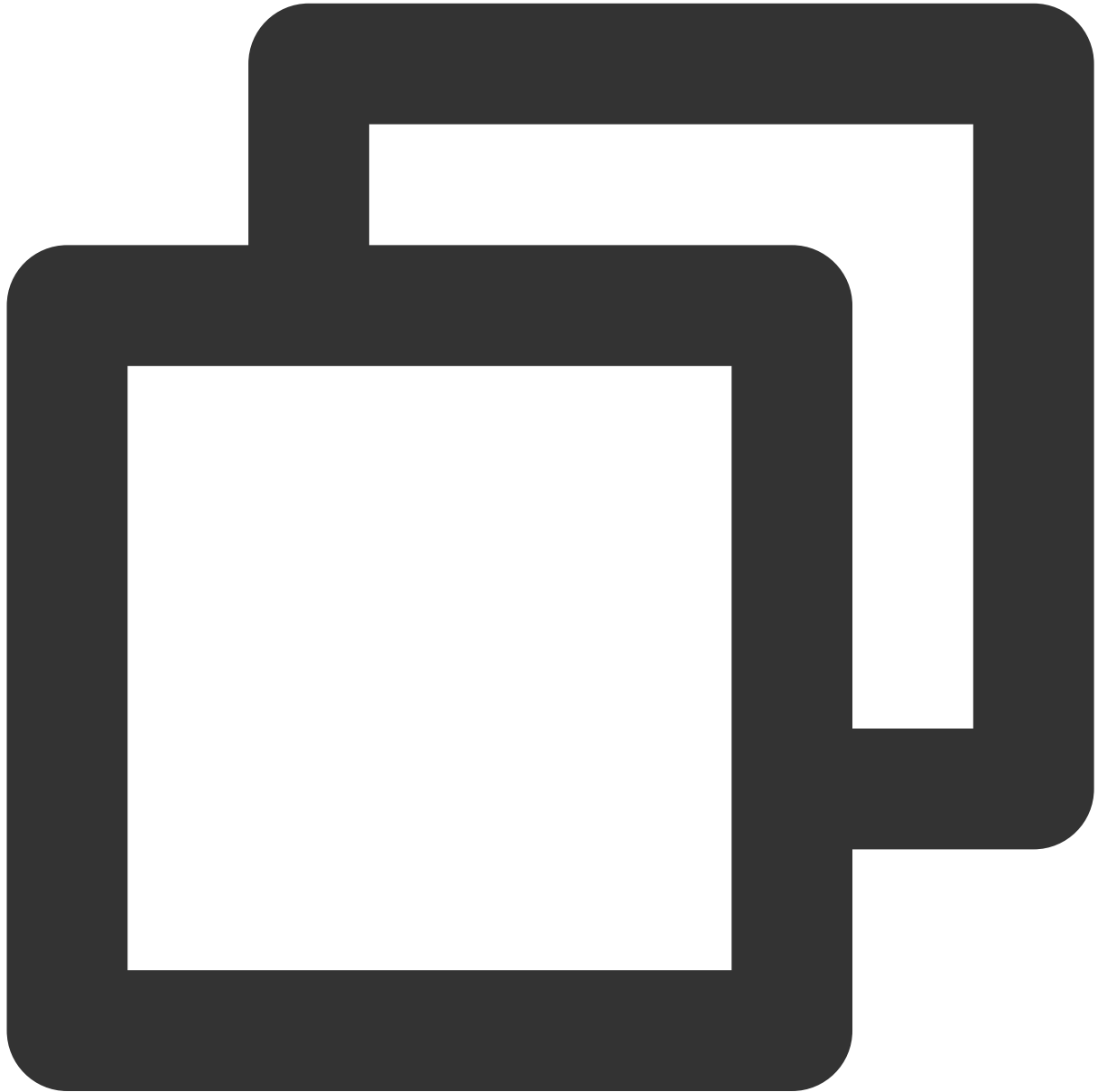
The format of StatefulSetGrid is similar to StatefulSet. The `<statefulset-template>` field corresponds to the original `template` field of the statefulset. What is special about StatefulSetGrid is the `gridUniqKey` field, which specifies the key value of the label for node grouping:



```
apiVersion: superedge.io/v1
kind: StatefulSetGrid
metadata:
  name:
  namespace:
spec:
  gridUniqKey: <NodeLabel Key>
```

```
<statefulset-template>
```

The format of ServiceGrid is similar to Service. The `<service-template>` field corresponds to the original `template` field of the service. What is special about ServiceGrid is the `gridUniqKey` field, which specifies the key value of the label for node grouping:



```
apiVersion: superedge.io/v1
kind: ServiceGrid
metadata:
  name:
  namespace:
```

```
spec:
  gridUniqKey: <NodeLabel Key>
  <service-template>
```

## Operation Steps

Taking the deployment of Nginx service at the edge as an example, if we want to deploy a complete set of Nginx services separately within multiple node pools, we need to perform the following operations:

### Create NodeUnit and NodeGroup

Let's take an edge cluster as an example and add the nodes within the cluster to the edge node pool (NodeUnit) and node pool classification (NodeGroup). This cluster consists of 4 edge nodes located in 2 regions, Beijing and Guangzhou. The node names are bj-1, bj-2, gz-1, gz-2, as shown in the picture below:

| Name | Availability | Status  | Cluster     | IP          | PodCIDR    | CPU/M...    | Time cr... | Operation         |
|------|--------------|---------|-------------|-------------|------------|-------------|------------|-------------------|
| gz-1 | -            | Running | v1.22.5-... | 172.16.1... | 10.33.0... | 1.94-cor... | 2023-05... | Remove Drain More |
| gz-2 | -            | Running | v1.22.5-... | 172.16.1... | 10.33.0... | 1.94-cor... | 2023-05... | Remove Drain More |
| bj-1 | -            | Running | v1.22.5-... | 172.16.1... | 10.33.1... | 1.94-cor... | 2023-05... | Remove Drain More |
| bj-2 | -            | Running | v1.22.5-... | 172.16.1... | 10.33.1... | 1.94-cor... | 2023-05... | Remove Drain More |

Total items: 4

20 / page 1 / 1 page

Create 2 NodeUnit: beijing and guangzhou, and add corresponding nodes to the NodeUnit:

Cluster(Singapore) / [Cluster Name]

Basic information

Node management

- Node
- NodeUnit**
- NodeGroup

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

Add-on management

### NodeUnit

Operation Guide

Create

You can enter only one keyword to search by

| Name          | Node list                   | Type     | Running nodes/desired nodes | Operation   |
|---------------|-----------------------------|----------|-----------------------------|---|
| beijing       | bj-1<br>bj-2                | Standard | 2/2                         | <a href="#">Update configuration</a> <a href="#">Edit Y</a><br><a href="#">Delete</a> |
| guangzhou     | gz-1<br>gz-2                | Standard | 2/2                         | <a href="#">Update configuration</a> <a href="#">Edit Y</a><br><a href="#">Delete</a> |
| unit-node-all | bj-1<br>bj-2<br>gz-1<br>... | Standard | 4/4                         | <a href="#">Update configuration</a> <a href="#">Edit Y</a><br><a href="#">Delete</a> |

Page 1

20 / pag

Create a NodeGroup (edge node pool classification) named location, and divide the two NodeUnit, beijing and guangzhou, into this classification, as shown in the following image:

Cluster(Singapore) / [Cluster Name]

Basic information

Node management

- Node
- NodeUnit
- NodeGroup**

Namespace

Service Group

Workload

Auto scaling

Service

### NodeGroup

Operation Guide

Create

You can enter only one keyword to search by

| Name     | Associated node pool | Operation   |
|----------|----------------------|---|
| location | beijing<br>guangzhou | <a href="#">Update configuration</a> <a href="#">Edit Y</a><br><a href="#">Delete</a> |

Page 1

20 / page

After performing the above operations, each node will be labeled accordingly. The labels for node gz-2 are shown in the image below:



```
Name:          gz-2
Roles:         <none>
Labels:        beta.kubernetes.io/arch=amd64
               beta.kubernetes.io/os=linux
               guangzhou=nodeunits.superedge.io
               kubernetes.io/arch=amd64
               kubernetes.io/hostname=gz-2
               kubernetes.io/os=linux
               location=guangzhou
               unit-node-all=nodeunits.superedge.io
```

### Description :

The key of the label is the name of the NodeGroup, and the value is the name of the NodeUnit. Nodes with the same value belong to the same NodeUnit.

If there are multiple NodeGroups within the same cluster, create different NodeGroup names as unique identifiers. When deploying ServiceGroup-related resources, they will be bound to the designated NodeGroup for deployment through the unique identifier of the NodeGroup name.

## Stateless ServiceGroup

### Create DeploymentGrid

1. Select **"ServiceGroup"** > **"DeploymentGrid"**
2. Click "Create", create a DeploymentGrid called "nginx"

**Create DeploymentGrid**

Workload name:   
Up to 40 chars. It supports only lower-case chars, numbers and hyphens. It must start with a lower-case letter and end with a number or lower-case letter.

NodeGroup: location ▼

Description:

Label:  =  ✕  
[Add variable](#)  
The label name can only contain letters, numbers and symbols ("-", ".", "\_", "/",), and must start and end with a number or letter.  
The label value can only contain letters, numbers and symbols ("-", ".", "\_", "/",), and must start and end with a number or letter.

Namespace:  ▼

Volume (optional): [Add volume](#)  
Used for container storage. It supports temp directory, NFS, config file, PVC, and should be mounted to the specified directory of the container. [Instruction](#)

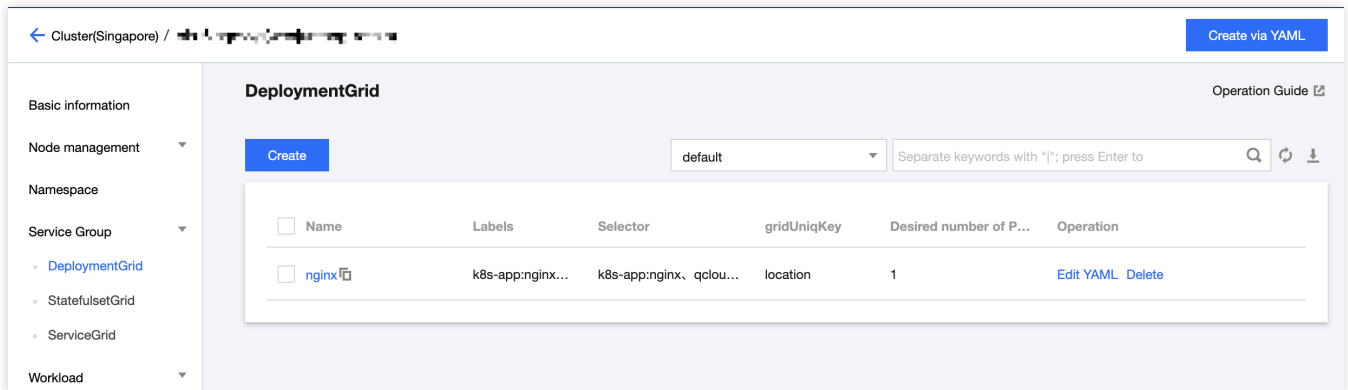
Containers in the Pod:

**NodeGroup:** Here, choose the NodeGroup group where you want to deploy the Nginx service in bulk. By selecting `location`, it means that the corresponding Deployment will be deployed under the two NodeUnits, `beijing`

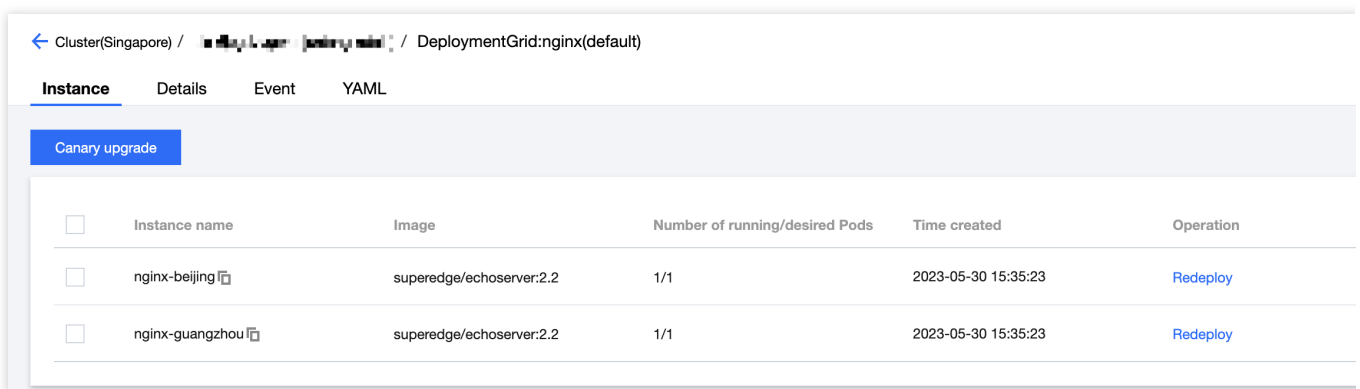
and `guangzhou` .

**Other parameters:** The remaining parameters are identical to the TKE application deployment method, so no further details are provided here.

3. Click "Create DeploymentGrid" and wait for finish.



4. Click on the "nginx" link to enter the detail page, where you can view the specifics of the created Deployment, as shown in the image below:



## Introduction :

The platform created a standard Kubernetes Deployment under each NodeUnit contained in the NodeGroup, named as DeploymentGrid-NodeUnit. According to the example, the names here are `nginx-beijing` and `nginx-guangzhou` respectively.

## Create ServiceGrid

1. Select "ServiceGroup" > "ServiceGrid", click "Create", as follow

### Basic information

Service name

nginx

The name can contain up to 59 characters. It supports lower-case letters, digits and "-", and must start with a letter and end with a digit or letter.

NodeGroup

location

Description

Up to 1000 characters

Namespace

default

### Access settings (Service)

Service access

☒ ClusterIP ☐ NodePort [Learn more](#)

☐ Headless Service (Headless Service can only be selected upon creation. The access method cannot be changed after the creation.)

Port mapping

| Protocol | Target port                     | Port                         |
|----------|---------------------------------|------------------------------|
| TCP      | Port listened by application in | Should be the same as the ta |

[Add port mapping](#)

[Advanced settings](#)

### Workload binding (Optional)

Selectors

k8s-app = nginx

qcloud-app = nginx

[Add](#) | [Select Workload](#)

**NodeGroup:** Choose the required NodeGroup, which should be the same NodeGroup as the one selected for DeploymentGrid previously.

**Access Settings:** Standard Kubernetes Service configuration information. Choose the required port; in this case, the nginx service is set to port 80 by default.

**Workload Binding:** Choose the Pods required by the Service using a Selector. You can manually enter and add them, or you can select 'Reference Workload' to add the previously deployed DeploymentGrid, in this case > nginx.

2. Click "**Create ServiceGrid**". After successful creation, the event details page is displayed.

3. You can view the specific created Service information under **Service > Service**, as shown in the following image:

Cluster(Singapore) / [breadcrumb]

Basic information

Node management ▾

Namespace

Service Group ▾

Workload ▾

Auto scaling

Service ▾

- Service
- Ingress

Configuration management ▾

### Service

Operation Guide

Create

default ▾ You can enter only one keyword to search by

| Name                       | Labels                               | Type ▾    | Selector                          | Access entry point ⓘ        | Time created     | Operation  |
|----------------------------|--------------------------------------|-----------|-----------------------------------|-----------------------------|------------------|--|
| <a href="#">kubernetes</a> | component:ap...<br>provider:kuber... | ClusterIP | -                                 | - 10.44.0.1 (Service IP)    | 2023-05-30 12... | <a href="#">Update configuration</a> <a href="#">Edit</a> <a href="#">Delete</a> |
| <a href="#">nginx-svc</a>  | superedge.io/...<br>superedge.io/... | ClusterIP | k8s-app:nginx<br>qcloud-app:ng... | - 10.44.159.74 (Service IP) | 2023-05-30 15... | <a href="#">Update configuration</a> <a href="#">Edit</a> <a href="#">Delete</a> |

Page 1

20 / 2

4. By using the **nginx-svc** Service, the following objectives can be achieved:

When accessing this Service from a Pod in the Beijing region, only the Pods within the Beijing region will be accessed by the backend.

When accessing this Service from a Pod in the Guangzhou region, only the Pods within the Guangzhou region will be accessed by the backend.

Access to this Service in each region will be restricted within the local NodeUnit scope.

# Health Check

Last updated : 2023-06-01 11:22:54

In edge scenarios, weak networks can trigger the Kubernetes eviction mechanism, causing unexpected pod eviction behavior. In edge computing scenarios, the network environment between edge nodes and the cloud is highly complex, the network quality cannot be guaranteed, and issues such as API Server and node connection interruptions can occur easily. Using native Kubernetes without modification may lead to frequent node status anomalies, activating the Kubernetes eviction mechanism, resulting in pod eviction, missing endpoints, and eventually causing service interruptions and fluctuations.

To address this issue, [TKE-Edge](#) introduces a distributed node status determination mechanism for the first time. This mechanism better identifies eviction timing, ensuring the system runs smoothly under weak network conditions, and avoiding service interruptions and fluctuations.

## Pain Points of the Requirements

### Native Kubernetes Approach

Weak cloud-edge networks impact communication between the kubelet running on edge nodes and the cloud API Server. If the cloud API Server cannot receive heartbeat signals from the kubelet or renew leases, it cannot accurately obtain the running status of the node and the pods on the node. If this lasts longer than the set threshold, the API Server will consider the node unavailable and take the following actions:

The disconnected node's status is set to NotReady or Unknown, and taints NoSchedule and NoExecute are applied.

Pods on the disconnected node are evicted and rebuilt on other nodes.

Pods on the disconnected node are removed from the Service's Endpoint list.

## Solutions

### Design Principles

In edge computing scenarios, relying solely on the connection between the edge node and the API Server to determine whether a node is normal is not reasonable. To make the system more robust, it's necessary to introduce additional judgment mechanisms.

Compared to the cloud and edge nodes, the network between edge nodes is more stable. Thus, a more stable infrastructure can be leveraged to improve accuracy. TKE-Edge pioneered the edge health distributed node status determination mechanism. In addition to considering the connection between nodes and the API Server, it introduces edge nodes as evaluation factors to provide a more comprehensive status assessment of the nodes. Extensive tests and practical demonstrations have proven that this mechanism improves the accuracy of node status determination

under weak cloud-edge network conditions, ensuring stable operation of services. The main principles of this mechanism are as follows:

Each node periodically probes the health status of other nodes.

All nodes within the cluster regularly vote to determine the status of each node.

Both cloud and edge nodes collectively determine the node's status.

Firstly, nodes internally probe and vote for the specific status of a node, and the unanimous judgment of the majority ensures the accurate status. Secondly, although the network status between nodes is typically better than the cloud-edge network, we should consider the complexity of the edge nodes' network environment, which is not 100% reliable. Therefore, it is not possible to fully trust the network between nodes, and the node status cannot be decided solely by the nodes themselves. A joint decision is more reliable. With this in mind, the following design is proposed:

| Status Determination            | Cloud Determination Normal | Cloud Determination Abnormal   |
|---------------------------------|----------------------------|--|
| Intranet Determination Normal   | Normal                     | K8s shows <b>NotReady</b> , but the behavior is different from the standard K8s process: No longer scheduling new Pods to the node, but not evicting Pods and handling services.             |
| Intranet Determination Abnormal | Normal                     | K8s shows <b>NotReady</b> , with behavior consistent with the standard K8s process: evicting existing Pods; removing them from the Endpoint list; no longer scheduling new Pods to the node. |

## Prerequisite

This feature requires opening port 51005 on the node so that the nodes can perform distributed smart health detection amongst themselves.

## Operation Steps

### Caution :

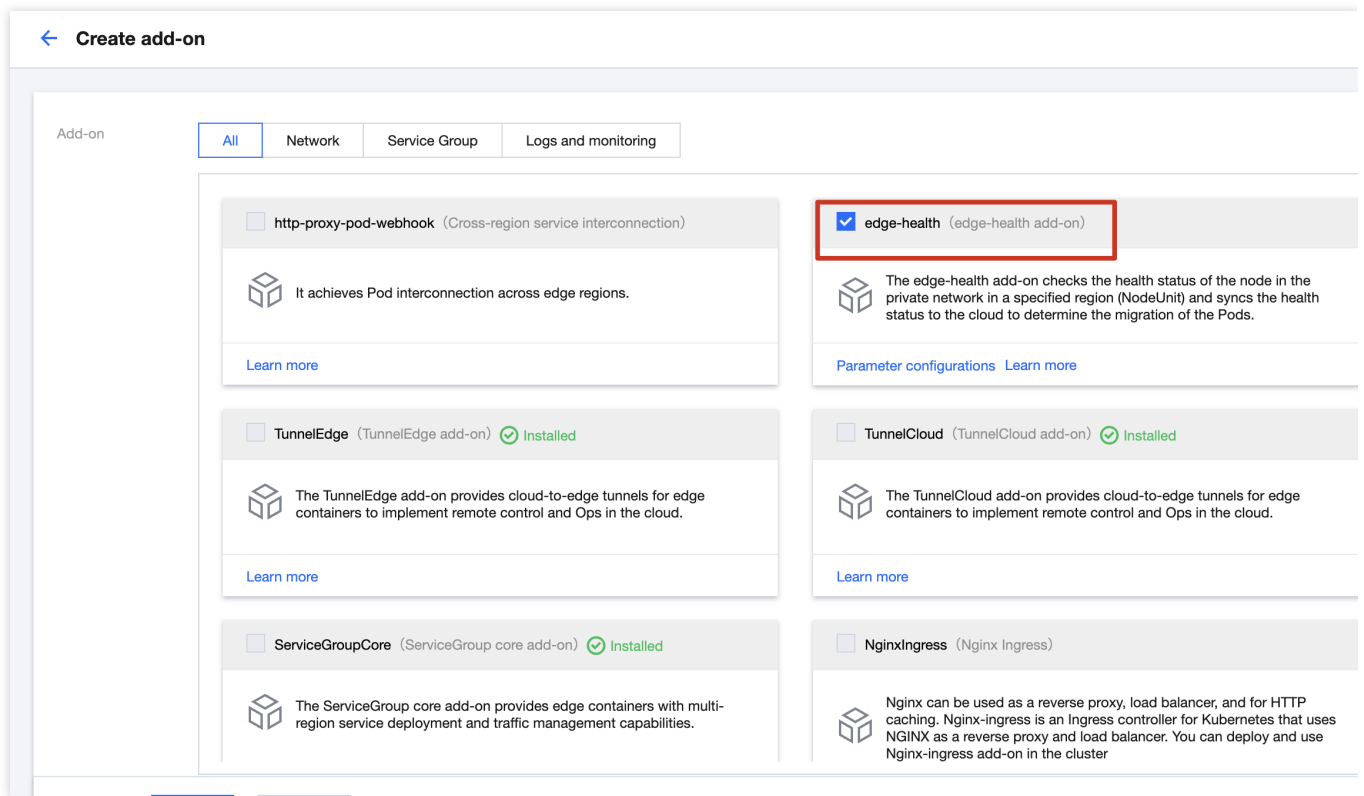
The edge health check and multi-region check features require some deployment and configuration time and will not take effect immediately.

### Enable Global Health-Check

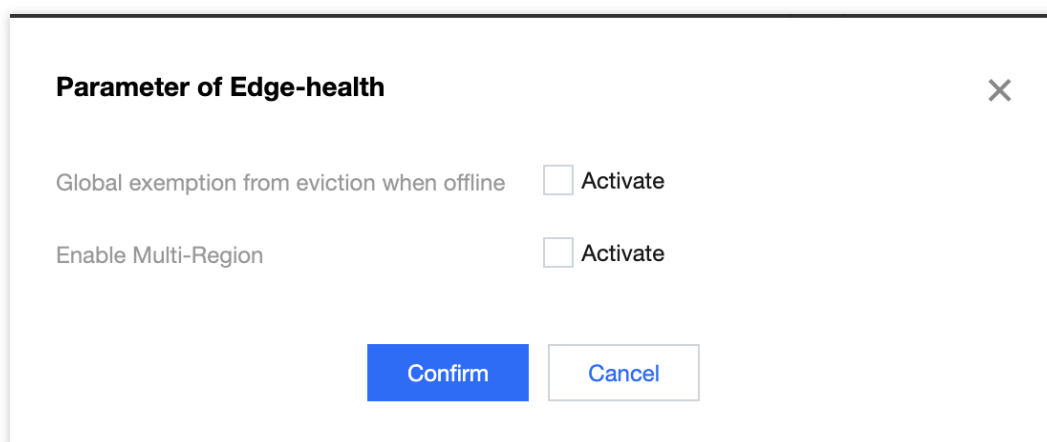
The edge health check feature is turned off by default. Please follow the steps below to enable it manually:

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.

3. Select **"Add-on management"** and enter the add-on list page.
4. Click **"Create"**, enter the "Create add-on" page.
5. Select **"edge-health"** and click **"Done"**, as follow:



On "Parameter configurations" pages, you can enable "Global exemption from eviction when offline" and "Multi-Region" capabilities, as follow:



6. Click "Done" and "edge-health" add-on will be deployed on the edge cluster.

## Global exemption from eviction when offline

After enabling this feature, all eviction flags, specifically the **NoExecute** taints, will be intercepted through the webhook mechanism. This prevents Pods from being evicted and rescheduled due to frequent node disconnections and reconnections. This feature is mutually exclusive with Multi-Region checks. When the eviction prevention feature is enabled, the actual health status of edge nodes will not be detected.

## Multi-Region:

The concept of multi-region here actually refers to NodeUnit in edge-health, which distinguishes the node's regional attributes through NodeUnit. For details, please refer to the [NodeUnit](#). Enabling Edge Health capability deploys edge-health health monitoring Pods on all edge nodes. If the multi-region capability is not enabled at this time, all nodes' statuses will be detected by default, and there is no multi-region concept. All nodes can be considered as nodes within the same region for mutual detection.

If you want nodes within different regions to independently check each other's health status, you need to enable the **Multi-Region** capability. At this point, you'll modify an **edge-health-config** ConfigMap, notifying K8s to handle multi-region health check capability. Once enabled, nodes will be separated into different regions based on the **<nodeunit-name>: nodeunits.superedge.io** label on the node. For example, the label **beijing:nodeunits.superedge.io** indicates that the node belongs to NodeUnit Beijing. Nodes with the same label are considered part of the same NodeUnit. When enabling the multi-region functionality, nodes within the same region will probe and vote on each other's state.

### Note :

If you enable the Multi-Region capability and the node does not belong to any NodeUnit, the node will not perform health checks.

If the Multi-Region capability is not enabled, all nodes within a cluster will check each other, even if the nodes belong to a NodeUnit.

## Enable Health-Check on specified NodeUnit

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Node management" > "NodeUnit"**, and you'll see there is an additional switch called "Enable Edge Health Check", as follow:



← Cluster(Singapore) / **Cluster Name**

Basic information

Node management ▾

- Node
- NodeUnit**
- NodeGroup

Namespace

Service Group ▾

Workload ▾

Auto scaling

Service ▾

Configuration management ▾

Add-on management

Log

Event

### NodeUnit

Create

| Name          | Node list                   | Type     |   | Running nodes/desir... | Enable                              |
|---------------|-----------------------------|----------|---|------------------------|-------------------------------------|
| beijing       | bj-1                        | Standard | ↕ | 1/1                    | <input type="checkbox"/>            |
| guangzhou     | gz-1<br>gz-2<br>gz-3<br>... | Standard | ↕ | 4/4                    | <input checked="" type="checkbox"/> |
| unit-node-all | bj-1<br>gz-1<br>gz-2<br>... | Standard |   | 5/5                    | <input type="checkbox"/>            |

Page 1

4. You can choose the NodeUnit you want and enable the "Edge Health Check" switch. At that time, you can check on the edge-health of each node that all nodes under the NodeUnit will probe each other.

[Cluster\(Singapore\)](#) / [NodeUnit](#)

Basic information

Node management

Node

**NodeUnit**

NodeGroup

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

Add-on management

Log

Event

### NodeUnit

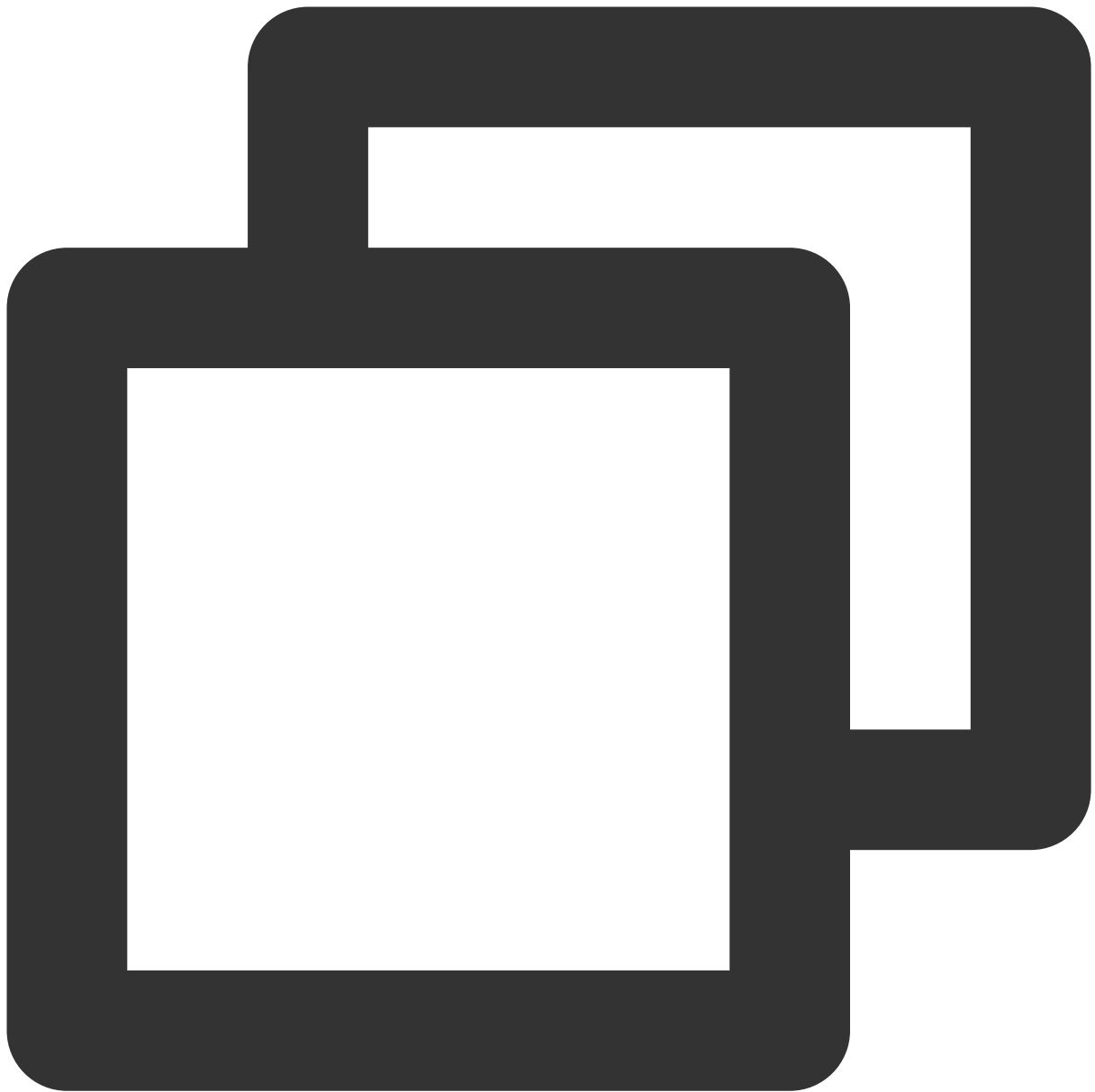
Create

| Name          | Node list                   | Type     | Running nodes/desir... | Enabl                               |
|---------------|-----------------------------|----------|------------------------|-------------------------------------|
| beijing       | bj-1                        | Standard | 1/1                    | <input type="checkbox"/>            |
| guangzhou     | gz-1<br>gz-2<br>gz-3<br>... | Standard | 4/4                    | <input checked="" type="checkbox"/> |
| unit-node-all | bj-1<br>gz-1<br>gz-2<br>... | Standard | 5/5                    | <input type="checkbox"/>            |

Page 1

5. At this point, the multi-regional health status capability is enabled. We can simulate the test using the gz-3 node, as follows:

You can use iptables rules to simulate disconnection from the apiserver:



```
iptables -I INPUT -s xxx.xxx.xxx.xxx -j DROP # This would be the IP address of the
```

After some time, you can observe that the `gz-3` node is in a `NotReady` state, but the Taint `node.kubernetes.io/unreachable:NoExecute` is not added. Therefore, the Pods on this node will not be evicted, and no new Pods will be scheduled on this node. This behavior demonstrates the proper handling of node unavailability by the cluster and its impact on workload management.

Every 2.0s: kubectl get nodes

| NAME | STATUS   | ROLES  | AGE   | VERSION                        |
|------|----------|--------|-------|--------------------------------|
| bj-1 | Ready    | <none> | 3h59m | v1.22.5-tke.1.9+5f76a1720f9839 |
| gz-1 | Ready    | <none> | 7h57m | v1.22.5-tke.1.9+5f76a1720f9839 |
| gz-2 | Ready    | <none> | 7h57m | v1.22.5-tke.1.9+5f76a1720f9839 |
| gz-3 | NotReady | <none> | 4h3m  | v1.22.5-tke.1.9+5f76a1720f9839 |
| gz-4 | Ready    | <none> | 4h3m  | v1.22.5-tke.1.9+5f76a1720f9839 |

```

Name:          gz-3
Roles:         <none>
Labels:        beta.kubernetes.io/arch=amd64
               beta.kubernetes.io/os=linux
               guangzhou=nodeunits.superedge.io
               kubernetes.io/arch=amd64
               kubernetes.io/hostname=gz-3
               kubernetes.io/os=linux
               location=guangzhou
               unit-node-all=nodeunits.superedge.io
Annotations:   flannel.alpha.coreos.com/backend-data: {"VtepMAC":"2a:44:f6:f2:ce"}
               flannel.alpha.coreos.com/backend-type: vxlan
               flannel.alpha.coreos.com/kube-subnet-manager: true
               flannel.alpha.coreos.com/public-ip: 172.16.16.106
               node.alpha.kubernetes.io/ttl: 0
               superedge.io/node-quota: false
               volumes.kubernetes.io/controller-managed-attach-detach: true
CreationTimestamp: Tue, 30 May 2023 16:04:42 +0800
Taints:        node.kubernetes.io/unreachable:NoSchedule
Unschedulable: false
  
```

To simulate a scenario where both cloud and internal network checks fail, shut down the `gz-3` node and observe its status. In this situation, the node state reflects the standard Kubernetes node failure state, and the Pods on this node will be evicted and rescheduled to other available nodes.

```
Name:          gz-3
Roles:         <none>
Labels:        beta.kubernetes.io/arch=amd64
               beta.kubernetes.io/os=linux
               guangzhou=nodeunits.superedge.io
               kubernetes.io/arch=amd64
               kubernetes.io/hostname=gz-3
               kubernetes.io/os=linux
               location=guangzhou
               unit-node-all=nodeunits.superedge.io
Annotations:   flannel.alpha.coreos.com/backend-data: {"VtepMAC":"2a:44:f6:f2:ce:c0"}
               flannel.alpha.coreos.com/backend-type: vxlan
               flannel.alpha.coreos.com/kube-subnet-manager: true
               flannel.alpha.coreos.com/public-ip: 172.16.16.106
               node.alpha.kubernetes.io/ttl: 0
               nodeunhealth: yes
               superedge.io/node-quota: false
               volumes.kubernetes.io/controller-managed-attach-detach: true
CreationTimestamp: Tue, 30 May 2023 16:04:42 +0800
Taints:        node.kubernetes.io/unreachable:NoExecute
               node.kubernetes.io/unreachable:NoSchedule
Unschedulable: false
```

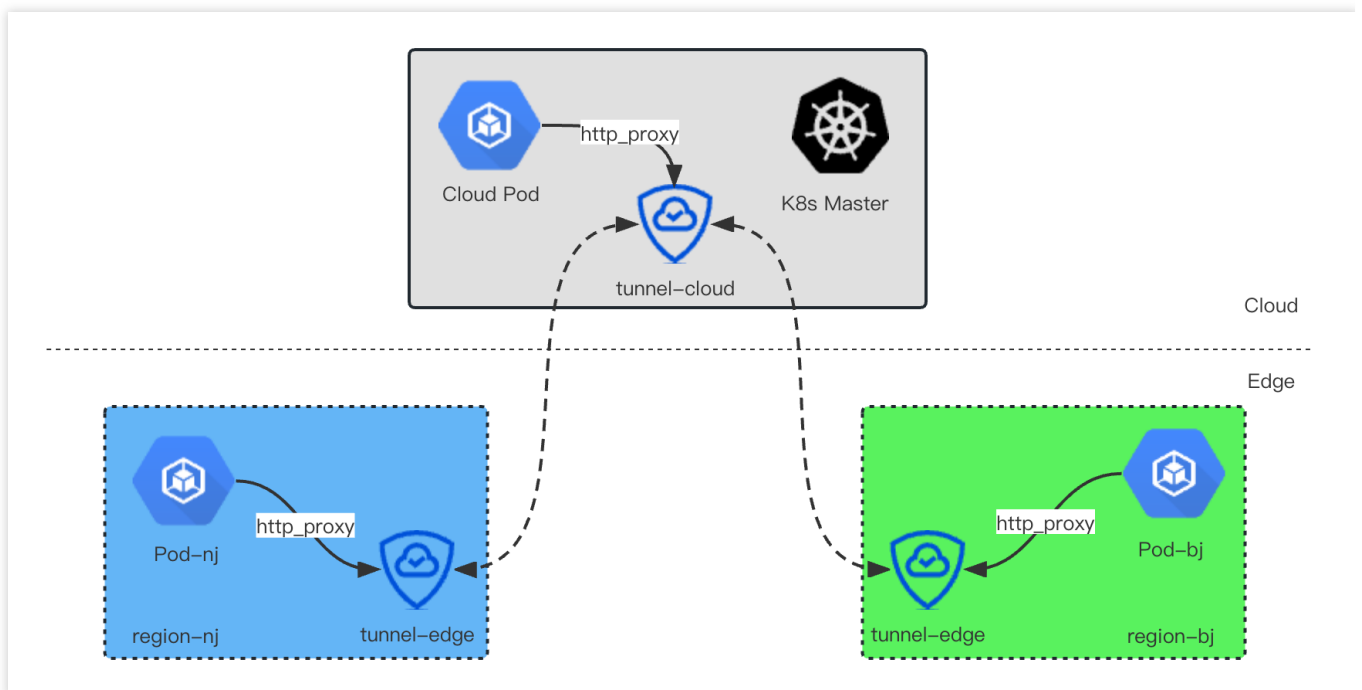
# Service Access across Regions

Last updated : 2023-06-01 11:22:54

## Operation Scenarios

This article introduces how to enable Pods/Svc from different edge regions to access each other through the cloud-edge Tunnel (currently only supporting Layer 7 http/https protocols). For example, a Pod in the Nanjing region can successfully access an Nginx service in the Beijing region, and vice versa.

## Architecture



As illustrated in the diagram above, the edge node tunnel-edge will establish a bidirectional tunnel with the cloud tunnel-cloud. The tunnel-edge will open the local 8080 port as the http/https proxy port. If an edge Pod needs to access other regions' Pod IPs or services through a proxy, you can specify the http\_proxy or https\_proxy environment variable within the Pod. This will forward the http/https traffic through the tunnel to the cloud or other edge nodes in different regions.

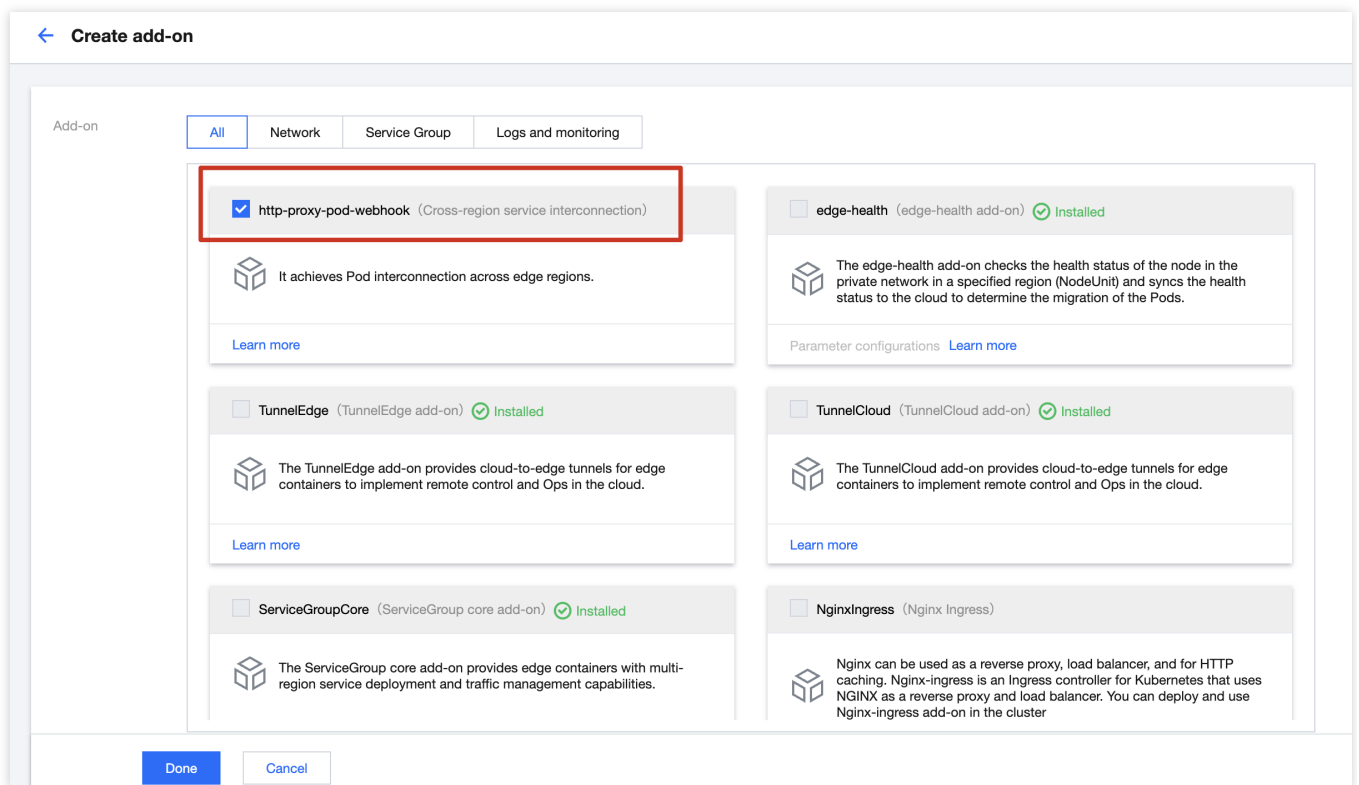
### Note :

At the product level, the platform uses the webhook method to conveniently specify proxies for you. You can set the label http-proxy=enable in the workload to enable proxy mode. When this is done, the corresponding Pods of the



workload will automatically inject the relevant environment variables, such as `http_proxy=169.254.20.11:8080`

## Operation Steps

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Select **"Add-on management"** and enter the add-on list page.
4. Click **"Create"**, enter the "Create add-on" page.
5. Select **"http-proxy-pod-webhook"** and click **"Done"**, as follow:



6. To verify the across-region access feature, create deployment in different regions as follows:

Cluster (Singapore) /   / Create Deployment


---


Name   
Up to 63 characters, including lowercase letters, numbers, and hyphens ("-"). It must begin with a lowercase letter, and end with a number or lowercase letter.

Description

Namespace

Labels  =  ×

[Add](#)  
The key name cannot exceed 63 chars. It supports letters, numbers, "/" and "-". "/" cannot be placed at the beginning. A prefix is supported. [Learn more](#)   
The label key value can only include letters, numbers and separators ("-", "\_", "."). It must start and end with letters and numbers.

Volume (optional) [Add volume](#)  
Used for container storage. It supports temp directory, NFS, config file, and should be mounted to the specified directory of the container. [Instruction](#) 

Containers in the Pod  × [+ Add container](#)

Name   
Up to 63 characters. It supports lower case letters, numbers, and hyphen ("-") and cannot start or end with "-".

Add "http-proxy=enable" label to the deployment, then the webhook of the step 5 will inject http\_proxy in the Pod environment, as follow:



```
[root@bj-1 ~]# crictl exec -it d9ba3b2aa2d7a /bin/sh
/ # env
KUBERNETES_PORT=tcp://10.44.0.1:443
KUBERNETES_SERVICE_PORT=443
HOSTNAME=echo-bj-55c4bc47bf-8sq5p
SHLVL=1
HOME=/root
ECHO_BJ_SERVICE_HOST=10.44.224.177
ECHO_BJ_PORT_8080_TCP_ADDR=10.44.224.177
ECHO_BJ_PORT_8080_TCP_PORT=8080
ECHO_BJ_PORT_8080_TCP_PROTO=tcp
http_proxy=http://169.254.20.11:8080
TERM=xterm
ECHO_BJ_SERVICE_PORT=8080
ECHO_BJ_PORT=tcp://10.44.224.177:8080
KUBERNETES_PORT_443_TCP_ADDR=10.44.0.1
NGINX_VERSION=1.15.3
ECHO_BJ_SERVICE_PORT_8080_8080_TCP_3T8IH1V68HM=8080
PATH=/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin
KUBERNETES_PORT_443_TCP_PORT=443
ECHO_BJ_PORT_8080_TCP=tcp://10.44.224.177:8080
KUBERNETES_PORT_443_TCP_PROTO=tcp
KUBERNETES_PORT_443_TCP=tcp://10.44.0.1:443
KUBERNETES_SERVICE_PORT_HTTPS=443
PWD=/
KUBERNETES_SERVICE_HOST=10.44.0.1
```

After the 2 deployments deployed, you'll check the status of these deployments:

Cluster(Singapore) / **Deployment** Create via YAM

Basic information Node management Namespaces Service Group Workload **Deployment** StatefulSet DaemonSet Job CronJob Auto scaling Service

**Deployment** Create Monitor default You can enter only one keyword to search by Q ↺

| <input type="checkbox"/> Name     | Labels   | Selector   | Number of runni... | Request/Limits                             | Operation   |
|-----------------------------------|--|--|--------------------|--|---|
| <input type="checkbox"/> echo-bj  | http-proxy:enable<br>k8s-app:echo-bj<br>qcloud-app:echo-bj   | http-proxy:enable<br>k8s-app:echo-bj<br>qcloud-app:echo-bj   | 1/1                | CPU: 0.25 / 0.5 core<br>MEM: 256 / 1024 Mi | <a href="#">Update Pod quantity</a><br><a href="#">Update Pod configuration</a> <span>More ▾</span> |
| <input type="checkbox"/> nginx-gz | http-proxy:enable<br>k8s-app:nginx-gz<br>qcloud-app:nginx-gz | http-proxy:enable<br>k8s-app:nginx-gz<br>qcloud-app:nginx... | 1/1                | CPU: 0.25 / 0.5 core<br>MEM: 256 / 1024 Mi | <a href="#">Update Pod quantity</a><br><a href="#">Update Pod configuration</a> <span>More ▾</span> |

Page 1 20 ▾ / page ◀ ▶

### Caution :

Each deployment must use the scheduler to assign the pod to the specific node. For example, echo-bj is assigned to "bj-1" and nginx-gz is assigned to "gz-2"

7. Login to "bj-1" node add access the pods of "gz-2", as follow:

| NAME                     | READY | STATUS  | RESTARTS | AGE   | IP         | NODE | NOMINATED NODE | REASON |
|--------------------------|-------|---------|----------|-------|------------|------|----------------|--------|
| echo-bj-55c4bc47bf-8sq5p | 1/1   | Running | 0        | 9m18s | 10.33.2.3  | bj-1 | <none>         | <none> |
| nginx-gz-d94bf4fc7-c492d | 1/1   | Running | 0        | 7m49s | 10.33.0.72 | gz-2 | <none>         | <none> |

When accessing the pod of gz-2, you'll see the request is redirect to **http\_proxy "169.254.20.11:8080"**

```
[root@bj-1 ~]# crictl exec -it d9ba3b2aa2d7a /bin/sh
/ # wget -O- http://10.33.0.72:8080
Connecting to 169.254.20.11:8080 (169.254.20.11:8080)

Hostname: nginx-gz-d94bf4fc7-c492d

Pod Information:
  -no pod information available-

Server values:
  server_version=nginx: 1.12.2 - lua: 10010

Request Information:
  client_address=10.33.0.65
  method=GET
  real path=/
  query=
  request_version=1.1
  request_scheme=http
  request_uri=http://10.33.0.72:8080/

Request Headers:
  connection=close
  host=10.33.0.72:8080
  user-agent=Wget

Request Body:
  -no body in request-

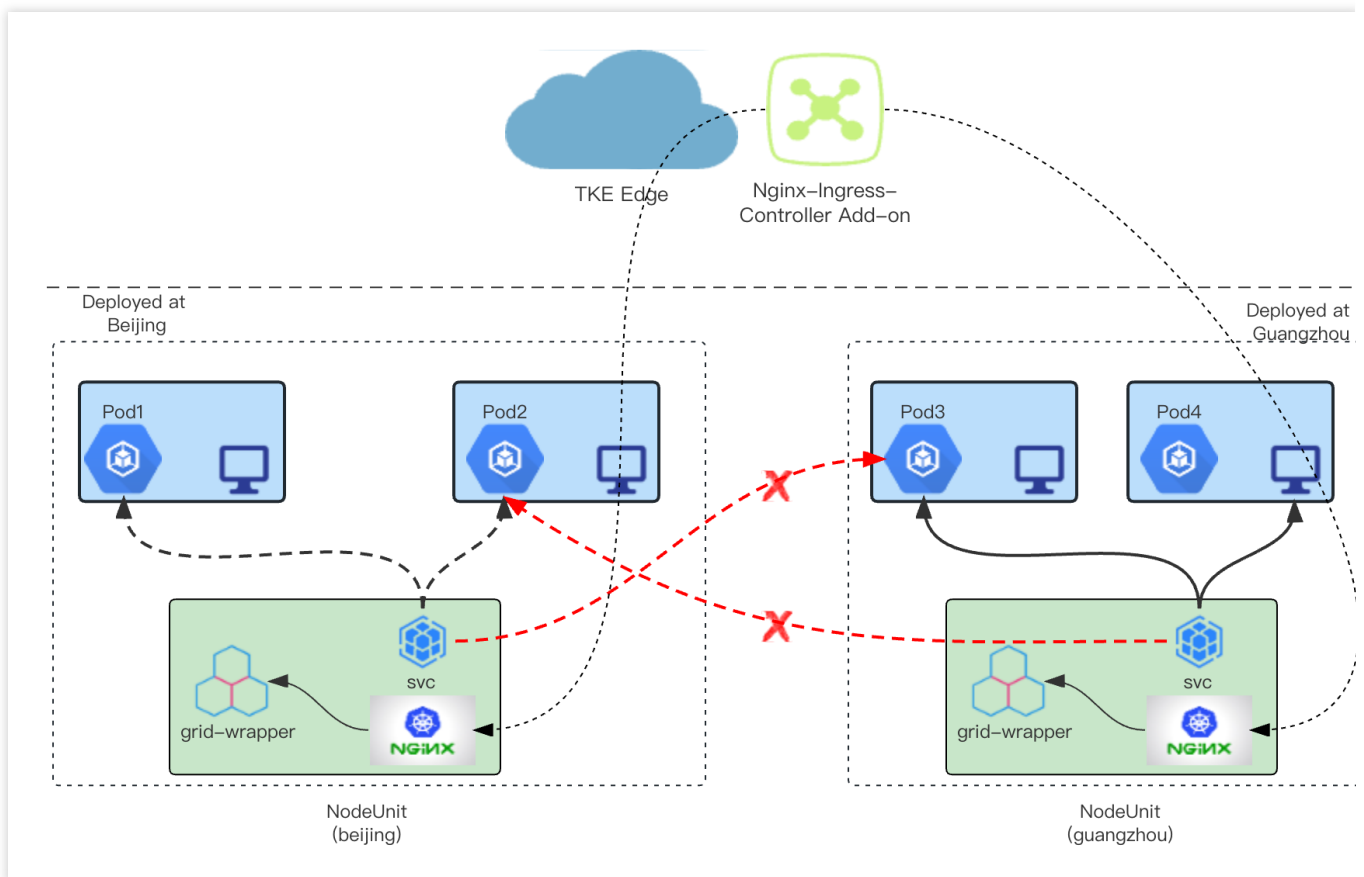
-
/ #
```

# Multi-Region Ingress

Last updated : 2023-06-01 11:22:54

## Scenarios

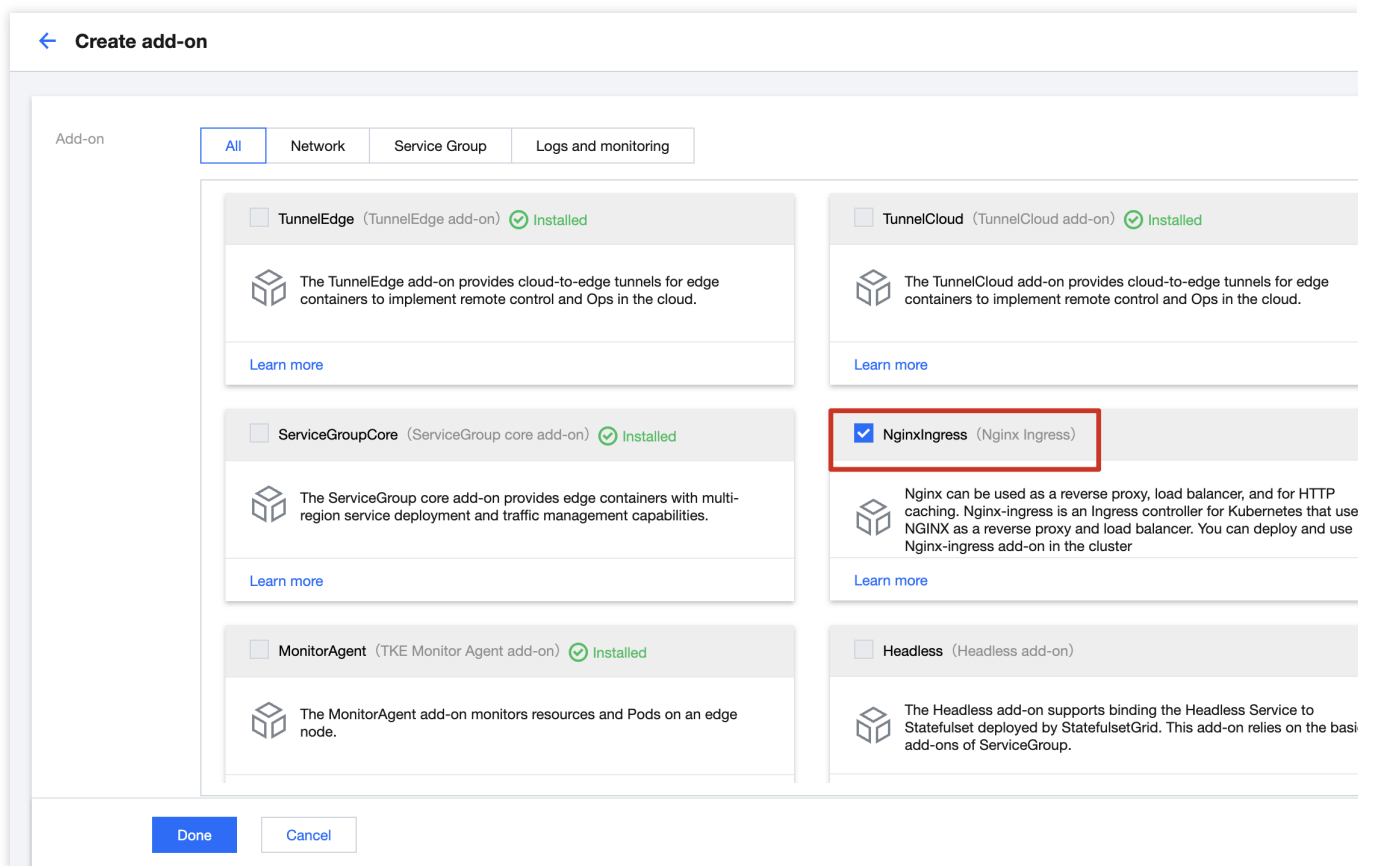
In the case of distinguishing multiple regions in edge nodes, each region has an independent network architecture and needs to provide Ingress service capabilities externally in the local region. Based on the concept of NodeUnit, TKE-Edge supports the ability to create **Nginx-Ingress-Controller** in different regions in the product. At the same time, through the application-grid-wrapper component, the Ingress-Controller's access to pods can be restricted within the local region (NodeUnit range). The specific architecture is shown in the following figure:



## Operation Steps

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.

3. Select **"Add-on management"** and enter the add-on list page.
4. Click **"Create"**, enter the "Create add-on" page.
5. Select **"NginxIngress"** and click **"Done"**, as follow:



6. Wait for the add-on is ready, as follow:

Cluster(Singapore) / [Cluster Icon]

Basic information

Node management

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

[Add-on management](#)

Log

Event

### Add-on management

Create

Separate keywords with "|"; press Enter to

| ID/name                      | Status     | Type            | Version | Time created           | Operation                                  |
|------------------------------|------------|-----------------|---------|------------------------|--|
| ingressnginx<br>ingressnginx | Successful | Enhanced add-on | 1.1.0   | 2023-05-29<br>17:58:56 | <a href="#">Update Nginx configuration</a> |
| edge-health<br>edge-health   | Successful | Enhanced add-on | 1.0.0   | 2023-05-29<br>16:32:34 | <a href="#">Delete</a>                     |

7. Click **"ingressnginx"** and enter the detailed page. As follow:

Cluster(Singapore) / [Cluster Icon] / NginxIngress:ingressnginx

**Nginx Ingress instance** Add-on details Nginx configuration

*You can deploy multiple Nginx Ingress instances in the cluster. When creating an Ingress object, you can specify the Nginx Ingress instance through the Ingress Class.*

Add Nginx Ingress instance

### Nginx Ingress CRD

| Name        | IngressClass | Time created | Operation |
|-------------|--------------|--------------|-----------|
| No data yet |              |              |           |

8. Click **"Add Nginx Ingress instance"**, and create Nginx-Ingress-Controller at specified NodeUnit, for example in bj NodeUnit:

### nginx-ingress-controller Parameter Settings

**NginxIngress name** 
  
The name can contain only lower-case letters, digits, hyphens ("-") and backslash ("\"), and must start with a lower-case letter, and end with a digit or lower-case letter.

**Namespace** 
  
Nginx Controller monitors and processes all Ingress resources under the specified namespace.

**Deploy modes** 
  
It is recommended to specify a separate node pool as DaemonSet to deploy Nginx-Ingress. When the node pool is scaled out, the Nginx-Ingress is scaled out as well.

**Node pool**

**Nginx configuration**

**CPU limit**

| request | limit | request | limit | unit |
|---------|-------|---------|-------|------|
| 0.25    | 256   | 1024    | MIB   |      |

**limits**

☐ All ☒ bj-1

**NginxIngress name:** Specifies the name of the deployed Ingress-Ingress-Controller instance, which creates the corresponding daemonset.

**Namespace:** Currently set to listen to Ingress resources in all namespaces by default.

**Deployment modes:** Currently, the default deployment mode is using DaemonSet. The user must manually choose one or more nodes under the NodeUnit to provide the Ingress Controller service.

**Node Pool:** Select the NodeUnit where you want to deploy, and then choose the nodes on which you want to deploy the Nginx-Ingress-Controller service.

#### Caution :

Make sure that the selected nodes have ports 80 and 443 available, as they should not be occupied. Otherwise, the Nginx-Ingress-Controller may fail to start.

9. After clicking "**Confirm**", you can view the created Ingress-Controller instance, as shown in the following image:

[← Cluster\(Singapore\)](#) / [...](#) / [NginxIngress:ingressnginx](#)

**Nginx Ingress instance**   Add-on details   Nginx configuration

**i** You can deploy multiple Nginx Ingress instances in the cluster. When creating an Ingress object, you can specify the Nginx Ingress instance through

[Add Nginx Ingress instance](#)

**Nginx Ingress CRD**

| Name                                | IngressClass | Time created           |
|-------------------------------------|--------------|------------------------|
| ingress-bj-nginx-ingress-controller | beijing      | 2023-05-30<br>17:58:23 |

10. By checking the component details, you can confirm the current deployment status of the controller. Once the number of running Pods reaches the expected count, it indicates that the deployment has succeeded, as shown in the following image:

Nginx Ingress instance
Add-on details
Nginx configuration

### Select Nginx Ingress instance

ingress-bj-nginx-ingress-...

### DaemonSet

| Name                                | superegde.io/nginx-ingress:true | Selector  |
|-------------------------------------|---------------------------------|---|
| ingress-bj-nginx-ingress-controller | superegde.io/nginx-ingress:true | k8s-app:ingress-bj-nginx-ingress-controller, qc |

### ConfigMap

| Name                                    | Namespace   | Labels |
|---|-------------|--------|
| ingress-bj-nginx-ingress-controller     | kube-system | -      |
| ingress-bj-nginx-ingress-controller-tcp | kube-system | -      |
| ingress-bj-nginx-ingress-controller-udp | kube-system | -      |

## Deploy Service and Access it Using Ingress:

1. Here we take the Nginx service as an example, creating a Deployment where the Pods will be deployed to the beijing and guangzhou regions, as shown in the image below:

```

nginx-beijing-76cdbbc899-k66c8    1/1    Running    0    2m11s    10.33.2.4
nginx-guangzhou-75f4c49d47-tzkkg  1/1    Running    0    2m11s    10.33.0.6

```

Then use the ServiceGrid to provide internal access to the deployed service:



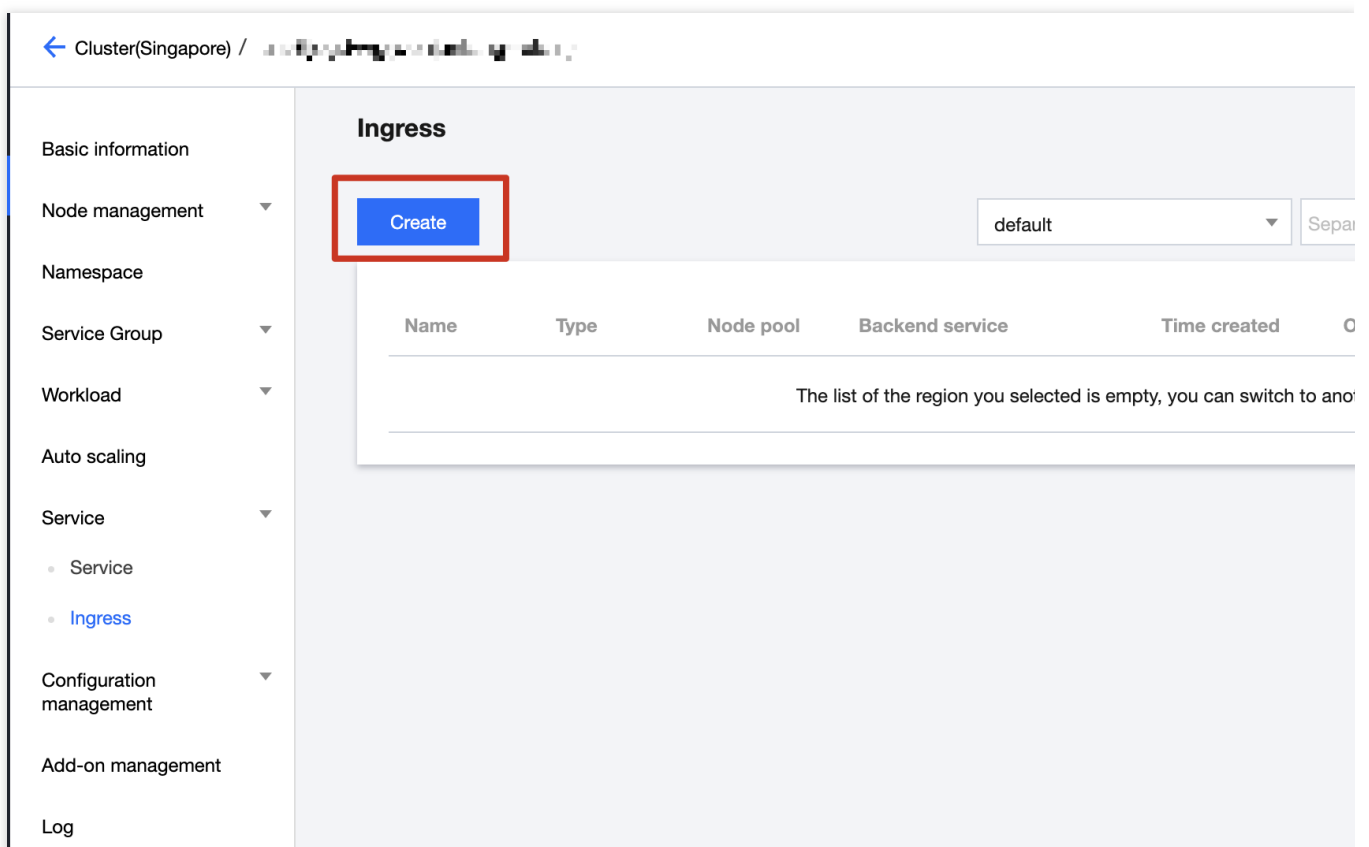
```

Name:          nginx-svc
Namespace:     default
Labels:        superedge.io/grid-selector=nginx
               superedge.io/grid-uniq-key=location
Annotations:   topologyKeys: ["location"]
Selector:      k8s-app=nginx,qcloud-app=nginx
Type:          ClusterIP
IP Family Policy: SingleStack
IP Families:   IPv4
IP:            10.44.6.125
IPs:           10.44.6.125
Port:          tcp-8080-8080 8080/TCP
TargetPort:    8080/TCP
Endpoints:     10.33.0.6:8080,10.33.2.4:8080
Session Affinity: None
Events:        <none>

```

In this scenario, we expect that when accessing Ingress from the Beijing region, only the Pods in the Beijing region will be accessed, for example, `nginx-beijing-76cdbbc899-k66c8`, and it won't access the Pods in the Guangzhou region.

2. To create an Ingress, navigate to the cluster details page **"Services" > "Ingress"**, and click **"Create"**. As shown in the image below:



3. Enter the required Ingress information, as shown in the following image:

[←](#) **Create Ingress**

Ingress name

echo-ingress

Up to 63 characters, including lowercase letters, numbers, and hyphens ("-"). It must begin with a lowercase letter, and end

Description

Up to 1000 characters

Ingress type

Nginx Ingress Controller

Namespace

default

Node pool

beijing

Forwarding configuration

| Protocol | Domain ⓘ    | Path  | Backend service ⓘ | Port |
|----------|-------------|-------|-------------------|------|
| HTTP ▾ ▾ | test.k8s.io | /echo | nginx-svc ▾       | 80   |

[Add forwarding rule](#)

**Ingress name:** Enter the name for the Ingress you want to create.

**Description:** Enter a descriptive remark.

**Ingress type:** Currently, only the Nginx Ingress Controller type is supported by default.

**Namespace:** Specify the namespace to create this Ingress.

**Node Pool:** Choose the Ingress-Controller to be bound to this Ingress, which refers to the Nginx-Ingress-Controller instance deployed earlier.

**Forwarding Configuration:** Input the specific configuration for the service here. For example, entering a custom domain name test.k8s.io, path set to echo, and backend service pointing to the previously created nginx-svc.

4. Access the Ingress service within the corresponding region. For example, in the Beijing region, access http://test.k8s.io/echo, as shown in the image below:

```
[root@bj-1 ~]# curl http://test.k8s.io/echo

Hostname: nginx-beijing-76cdbbc899-k66c8
Pod Information:
  -no pod information available-

Server values:
  server_version=nginx: 1.12.2 - lua: 10010

Request Information:
  client_address=10.33.2.1
  method=GET
  real_path=/
  query=
  request_version=1.1
  request_scheme=http
  request_uri=http://test.k8s.io:8080/

Request Headers:
  accept=/*/*
  host=test.k8s.io
  user-agent=curl/7.61.1
  x-forwarded-for=192.168.20.234
  x-forwarded-host=test.k8s.io
  x-forwarded-port=80
  x-forwarded-proto=http
  x-forwarded-scheme=http
  x-real-ip=192.168.20.234
  x-request-id=7fb10970f188c46323106db51c31d5d6
  x-scheme=http

Request Body:
  -no body in request-
```

You can perform multiple access tests and find that all accesses will be restricted to the Pods within the Beijing region, and will not access the Pods in the Guangzhou region. This demonstrates the effectiveness of the Ingress and ServiceGroup configuration in managing traffic and ensuring that interactions remain within the specified NodeUnit and region.

# Kins Manual

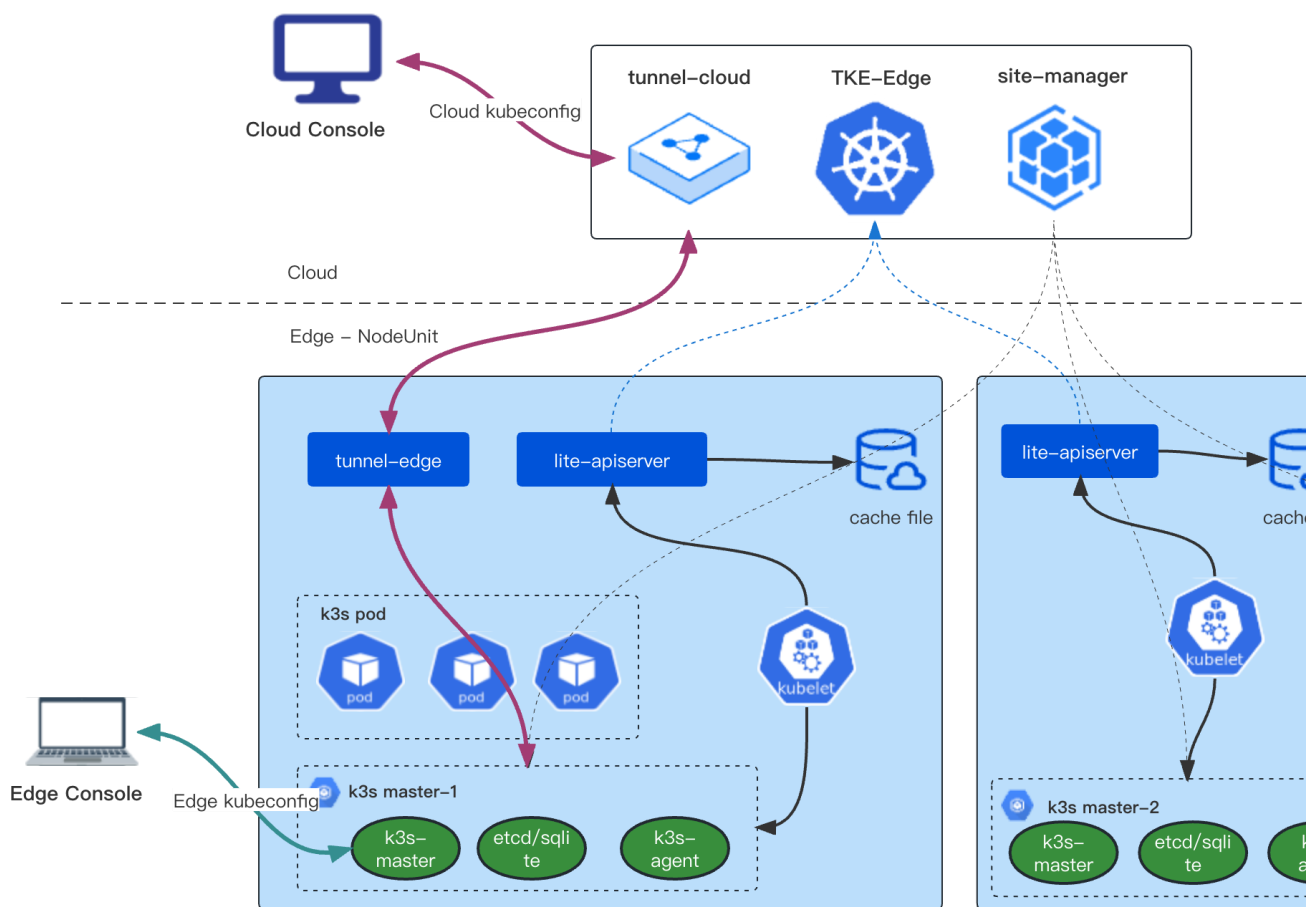
Last updated : 2023-06-01 11:22:54

## Scenarios

This article introduces a new feature of TKE-Edge - Kins, with the main functions described as follows: This capability can turn a standard NodeUnit edge node pool into an independent K3s cluster with a single click. This edge node pool can then be disconnected from the cloud control plane and used for long-term offline operation. During the offline process, the edge node pool can be independently operated and maintained as a standard K3s cluster. If there is a need for maintenance or upgrades in the future, this node pool can be reconnected to the cloud, enabling remote maintenance operations such as cloud-based synchronization upgrades. This feature comprehensively improves the previous capability deficiency of NodeUnit, which could not operate autonomously after disconnecting from the network.

## Architecture

The basic architecture of Kins is as follow:



For example, the three edge nodes can be divided into a NodeUnit. Through edge node pool-related operations, the K3s master and agent components can be deployed on these three nodes to form an independent single-master or 3-master K3s cluster. This edge K3s cluster can be accessed from the cloud via a tunnel, and it can also be accessed directly on the edge-side nodes.

## Operation Steps

### Caution :

The Kins capability is currently only supported for **Kubernetes version 1.22**, and the runtime must be **containerd**.

## Create edge K3s cluster

1. Log in to the [TKE console](#).
2. On the cluster management page, click the cluster ID to enter the cluster details page.
3. Add nodes into the cluster.
4. Click **"Node management" > "Edge node pool"**, and create standard NodeUnit called "demo".

Cluster(Singapore) / [Cluster Name]

Basic information

Node management

- Node
- Edge node pool
- Node group

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

### NodeUnit

Create

| Name          | Node list                   | Type     | Running nodes/desir... | Enable                              |
|---------------|-----------------------------|----------|------------------------|-------------------------------------|
| demo          | gz-1<br>gz-2<br>gz-3        | Standard | 3/3                    | <input checked="" type="checkbox"/> |
| unit-node-all | bj-1<br>gz-1<br>gz-2<br>... | Standard | 4/4                    | <input type="checkbox"/>            |

Page 1

5. Select "demo" and click "**Promote to Standalone NodeUnit(K3s cluster)**" to promote K3s cluster, as follow:

Cluster(Singapore) / [Cluster Name]

Basic information

Node management

- Node
- Edge node pool
- Node group

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

Add-on management

### NodeUnit

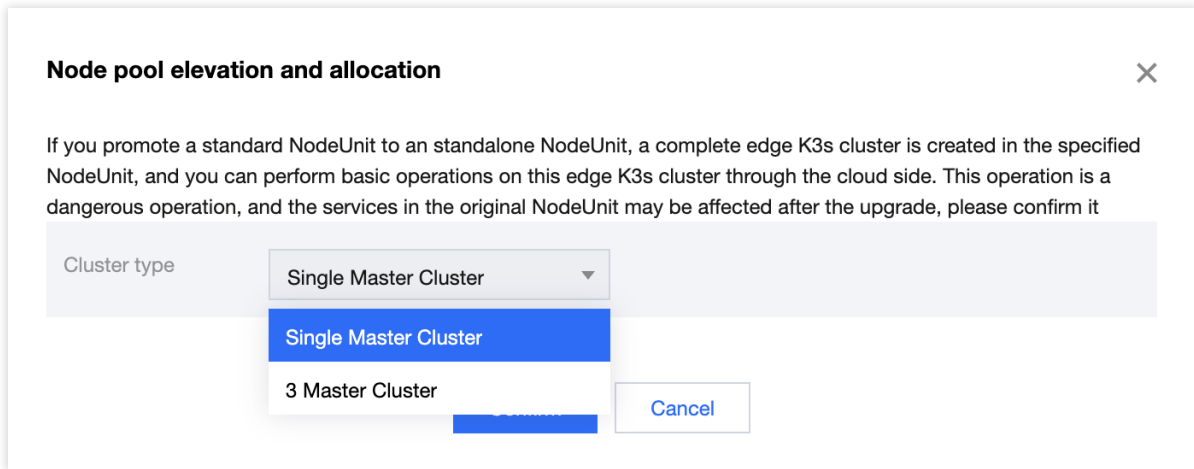
Create

| Name          | Node list                   | Type     | Running nodes/desir... | Enable                              |
|---------------|-----------------------------|----------|------------------------|-------------------------------------|
| demo          | gz-1<br>gz-2<br>gz-3        | Standard | 3/3                    | <input checked="" type="checkbox"/> |
| unit-node-all | bj-1<br>gz-1<br>gz-2<br>... | Standard | 4/4                    | <input type="checkbox"/>            |

Page 1

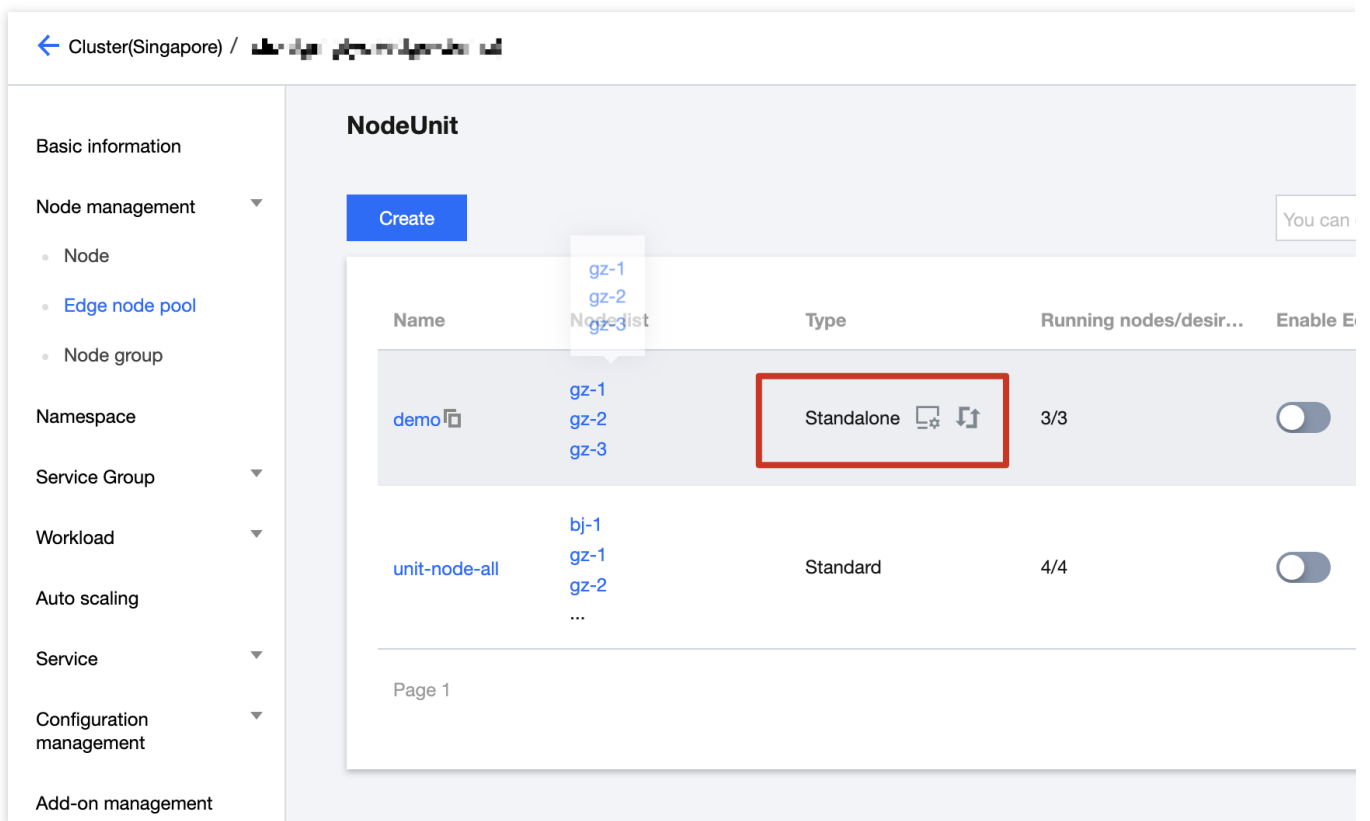
Promote to Standalone NodeUnit(K3s cluster)

6. At the "Node pool elevation and allocation" page, you can select **"Cluster type"** and Click **"Done"**



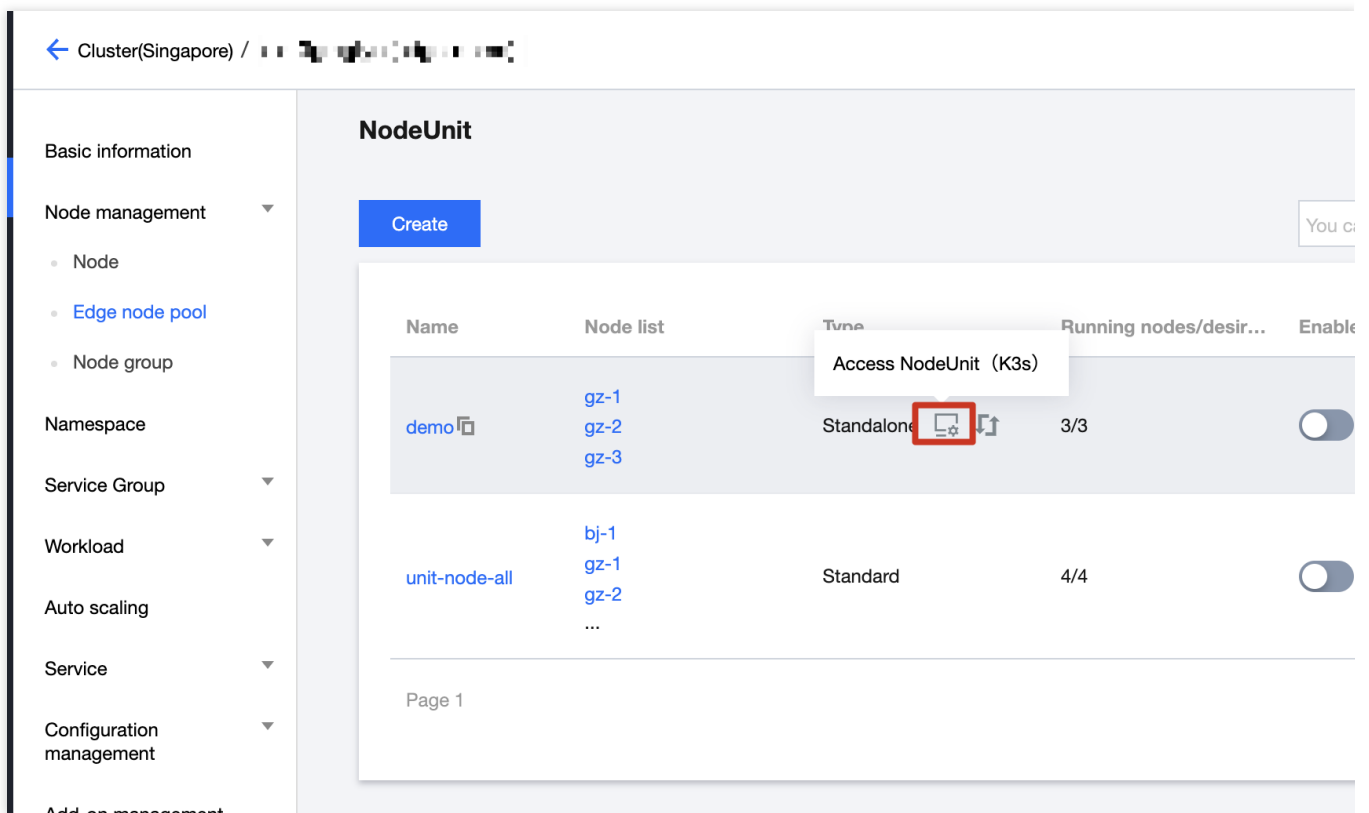
If there are only 1-2 nodes in the NodeUnit, you can only use a "single Master cluster" by default. If the NodeUnit contains 3 or more nodes, you can choose to use a "single Master cluster" or a "3 Master highly available cluster". You can make a choice based on your own needs.

7. After a moment, you'll see the type of "demo" NodeUnit will change to "Standalone":



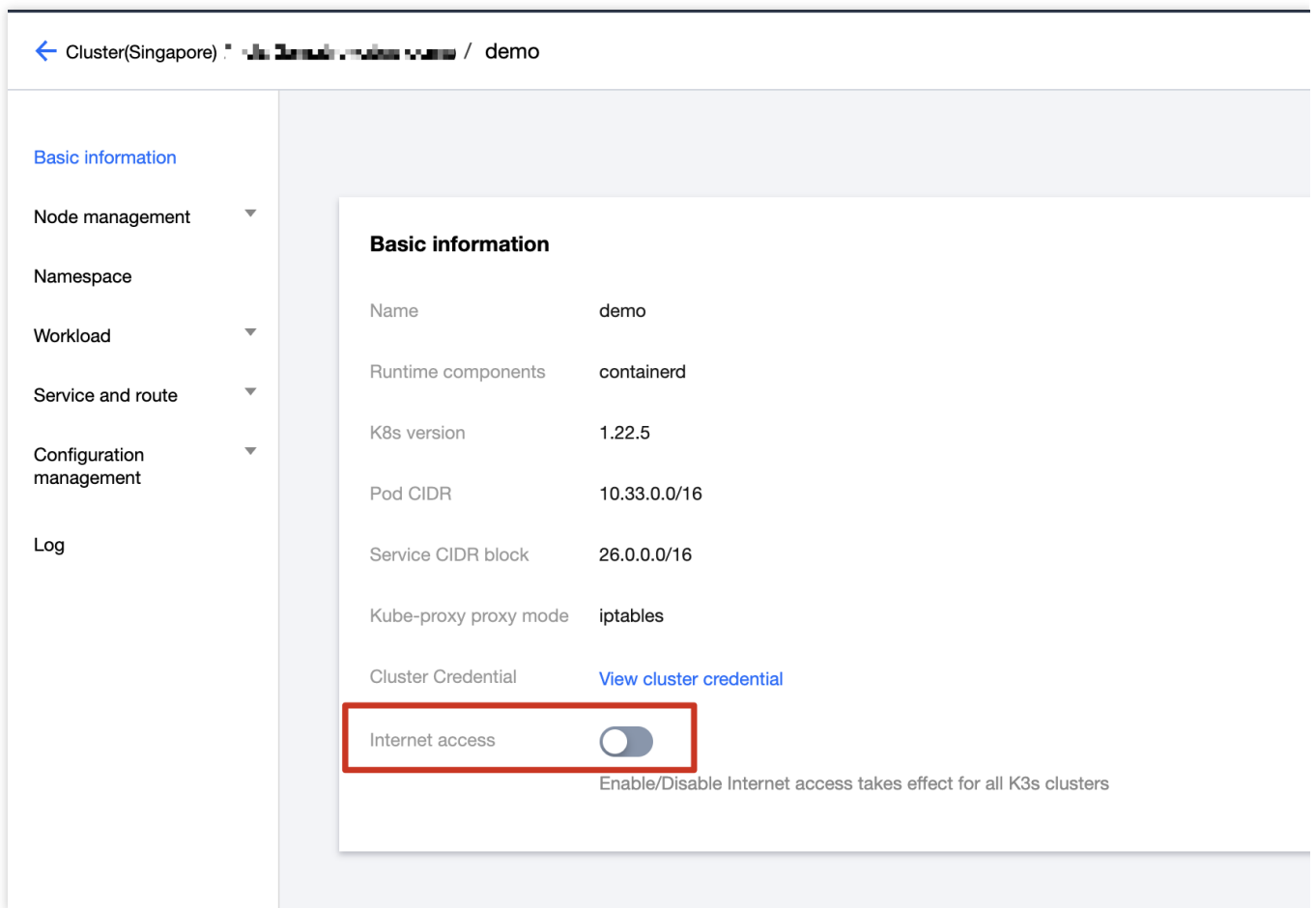
## Access the edge K3s cluster

1. At the NodeUnit page list, click "**Access NodeUnit(K3s)**", and you will be redirected to the console UI of the edge K3s cluster, as shown in the following figure:

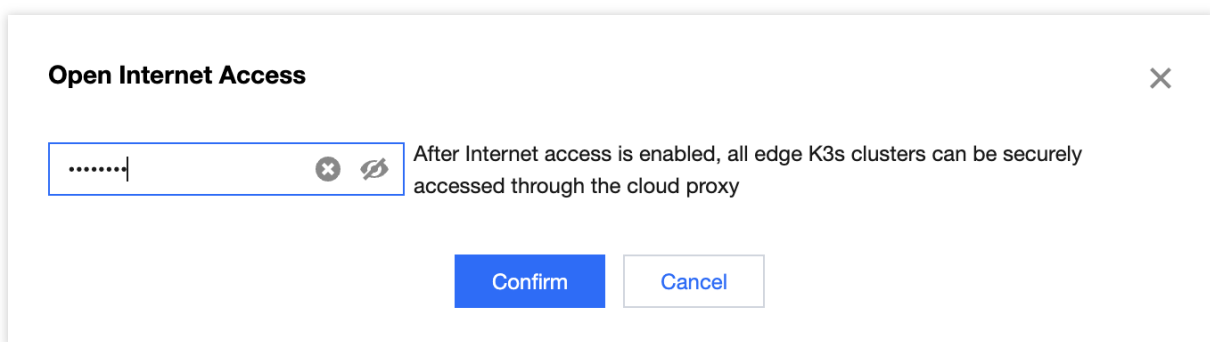


2. In the basic information of the edge K3s cluster, click "**Internet access**" to enable the external network access capability of the edge K3s cluster. Once enabled, you can use the `kubectl` command on any client in the public cloud, access the edge K3s cluster through TKE-Edge's external network proxy capability, as shown in the following figure:















At the "Open Internet Access" page, please enter the password required for proxy access, as shown in the following figure:



### Caution :

When enabling Internet network access, the system will provide http/https proxy access through the cloud CLB capability. Therefore, users are required to set a proxy password for authentication to ensure the security of the proxy.

3. You can log in to the edge K3s cluster using either the cloud kubeconfig or the edge-side kubeconfig. Click to view the cluster credentials, as shown in the following figure:

← Cluster(Singapore) /           / demo

Basic information

Node management ▼

Namespace



Workload ▼

Service and route ▼

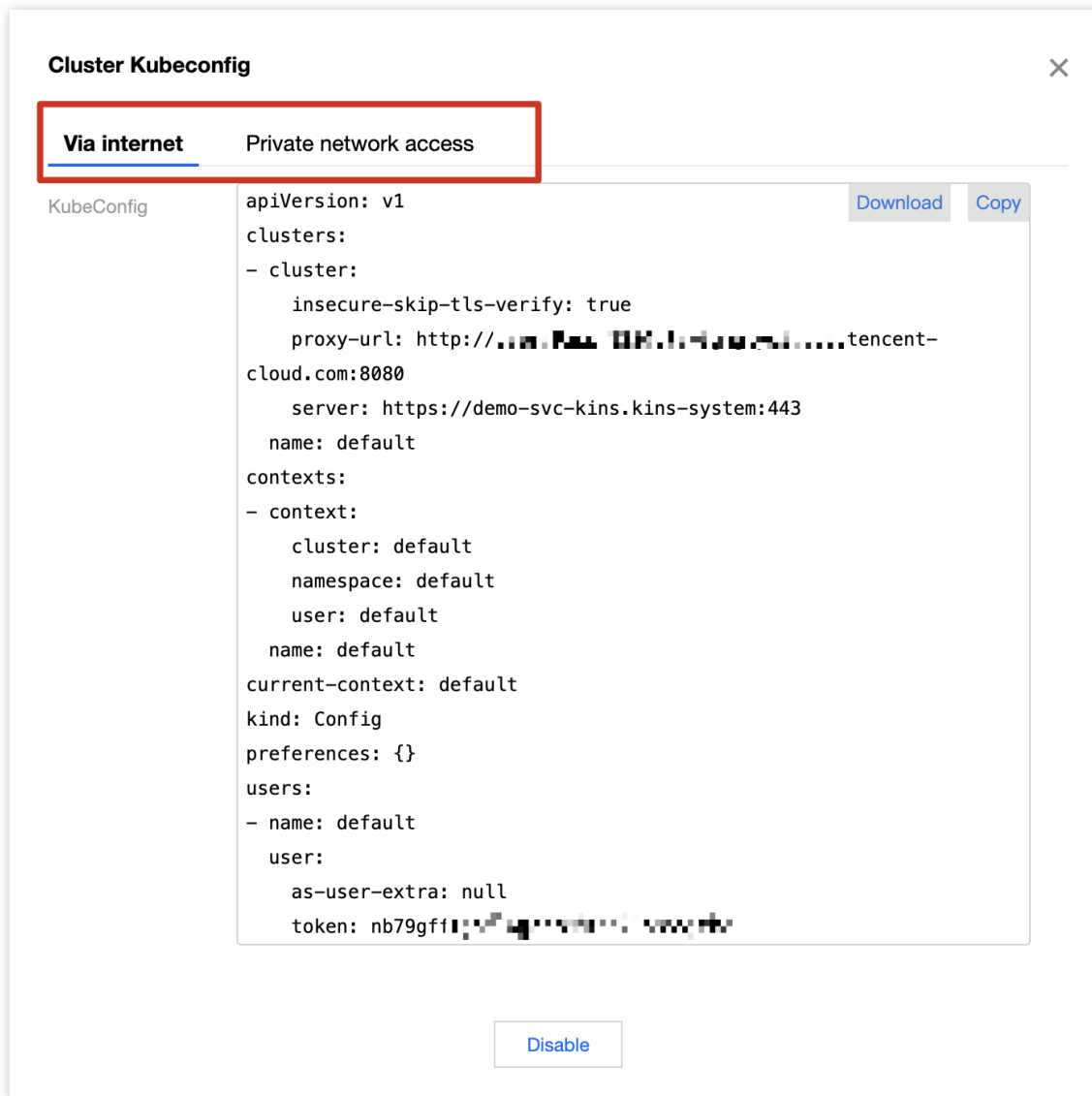
Configuration management ▼

Log

### Basic information

|                             |   |
|-----------------------------|---|
| Name                        | demo  |
| Runtime components          | containerd  |
| K8s version                 | 1.22.5  |
| Pod CIDR                    | 10.33.0.0/16  |
| Service CIDR block          | 26.0.0.0/16   |
| Kube-proxy proxy mode       | iptables  |
| Cluster Credential          | <a href="#">View cluster credential</a>   |
| Internet access             | <input checked="" type="checkbox"/><br>Enable/Disable Internet access takes effect for all K3s clusters   |
| Password of Internet Access | *****   |

And you can choose the Internet Access or Private Network Access:



If you want to access the edge K3s cluster from the public cloud side, you can download/copy the kubeconfig through "Via internet".

If you want to directly access the local K3s cluster from the edge-side node, you can download/copy the kubeconfig for "Private network access".

4. You can also directly use the console Web UI to perform some basic operations on the edge K3s cluster, as shown in the following figure:

[Cluster\(Singapore\)](#) / [cl...](#) / [demo](#)

Basic information

Node management ▾

- Node

Namespace

Workload ▾

Service and route ▾

Configuration management ▾

Log

### Node

You c

| <input type="checkbox"/> | Name | Status  | Type   | podCIDR        | kubeletVersion       |
|--------------------------|------|---------|--------|----------------|----------------------|
| <input type="checkbox"/> | gz-1 | Healthy | master | 169.254.0.0/24 | v1.22.6+k3s-416af30d |
| <input type="checkbox"/> | gz-2 | Healthy | node   | 169.254.1.0/24 | v1.22.6+k3s-416af30d |
| <input type="checkbox"/> | gz-3 | Healthy | node   | 169.254.2.0/24 | v1.22.6+k3s-416af30d |

Page 1

## Add node to edge K3s cluster

### Caution :

In the current product form, users are temporarily only supported to add new nodes to the edge K3s cluster through the Cloud console side, and independent node addition is not supported when the K3s cluster is offline. Users can directly add nodes to the NodeUnit through the cloud, and the new nodes will be automatically joined to the edge K3s cluster.

1. On the NodeUnit list page, select the desired NodeUnit and click Update Configuration, as shown in the following figure:



← Cluster(Singapore) / 

- Node
- Edge node pool
- Node group

### Add-on management

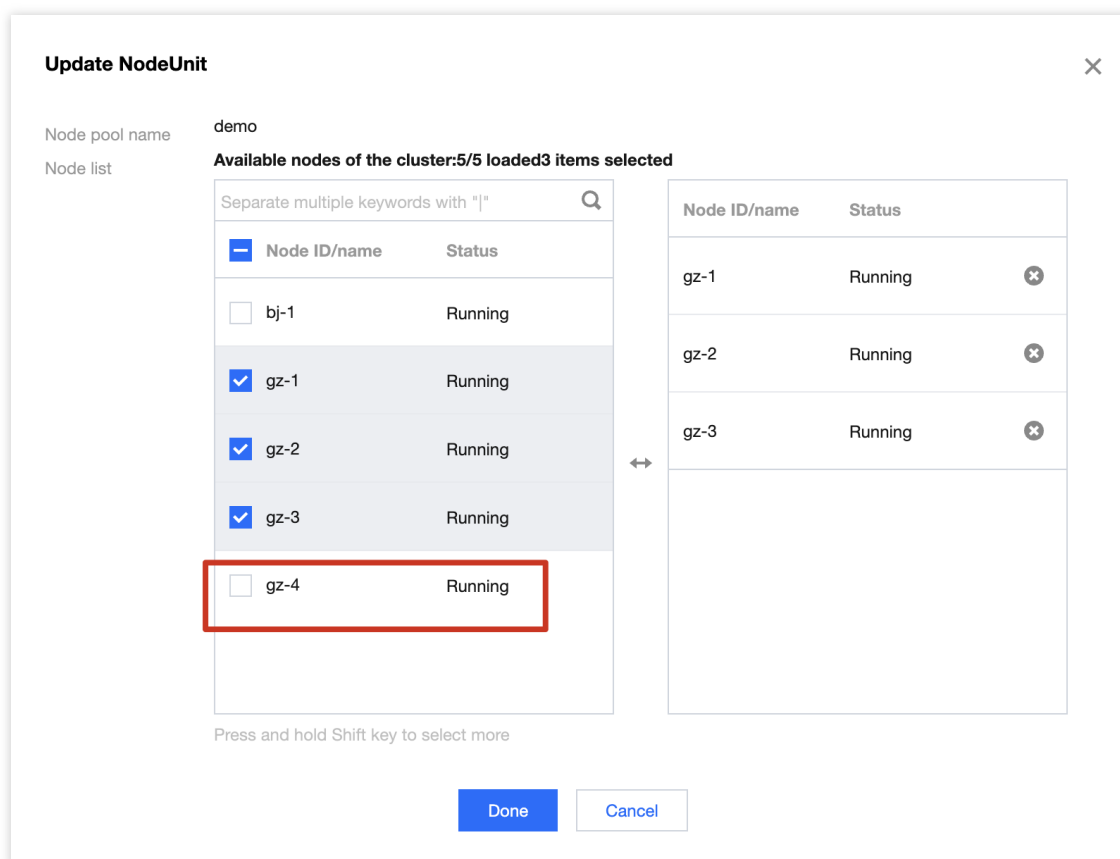
## NodeUnit

You c

| Name          | Node list | Type   | Running nodes/desir... | Enable                              |
|---------------|-----------|--|------------------------|-------------------------------------|
| demo          | gz-1      | Standalone   | 3/3                    | <input checked="" type="checkbox"/> |
|               | gz-2      |  |                        |                                     |
|               | gz-3      |  |                        |                                     |
| unit-node-all | bj-1      | Standard   | 4/4                    | <input type="checkbox"/>            |
|               | gz-1      |  |                        |                                     |
|               | gz-2      |  |                        |                                     |
|               | ...       |  |                        |                                     |

Page 1

Page 113 of 136



3. Click "Done" and you'll see "gz-4" has already been in the NodeUnit "demo".

4. After waiting for a while, click "Access NodeUnit" button and enter edge K3s console page. Observe the cluster node information, and you will find that the added node has joined the edge K3s cluster, as shown in the following figure:

Cluster(Singapore) / demo

Basic information

Node management

Node

Namespace

Workload

Service and route

Configuration management

Log

### Node

| <input type="checkbox"/> Name | Status  | Type   | podCIDR        | kubeletVersion       |
|-------------------------------|---------|--------|----------------|----------------------|
| <input type="checkbox"/> gz-1 | Healthy | master | 169.254.0.0/24 | v1.22.6+k3s-416af30d |
| <input type="checkbox"/> gz-2 | Healthy | node   | 169.254.1.0/24 | v1.22.6+k3s-416af30d |
| <input type="checkbox"/> gz-3 | Healthy | node   | 169.254.2.0/24 | v1.22.6+k3s-416af30d |
| <input type="checkbox"/> gz-4 | Healthy | node   | 169.254.3.0/24 | v1.22.6+k3s-416af30d |

Page 1

5. Through interactive operation with NodeUnit on the cloud, you can now freely add computing nodes to the edge K3s cluster.

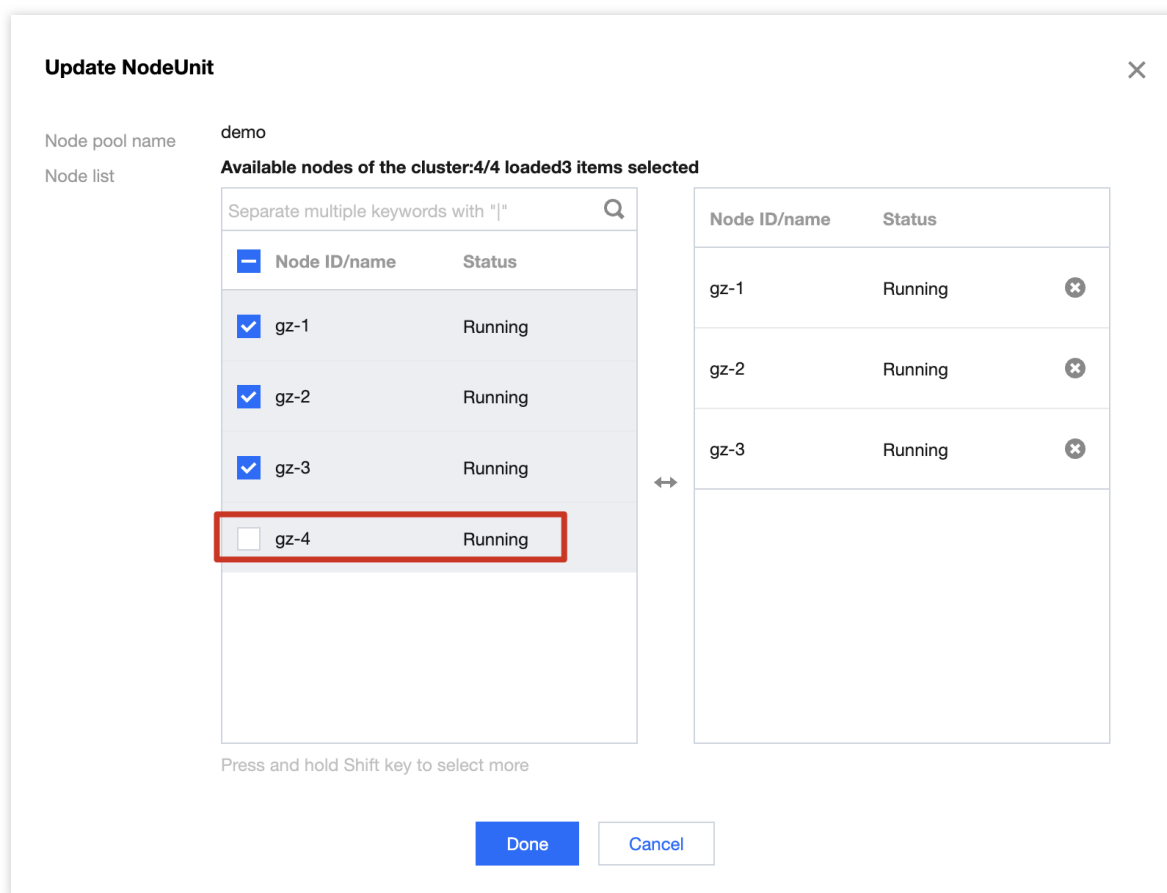
## Delete node from edge K3s cluster

At present, this product only temporarily supports users to delete edge nodes in the K3s cluster **via the cloud**. There are two ways for users to remove nodes from the edge K3s cluster.

### Method 1

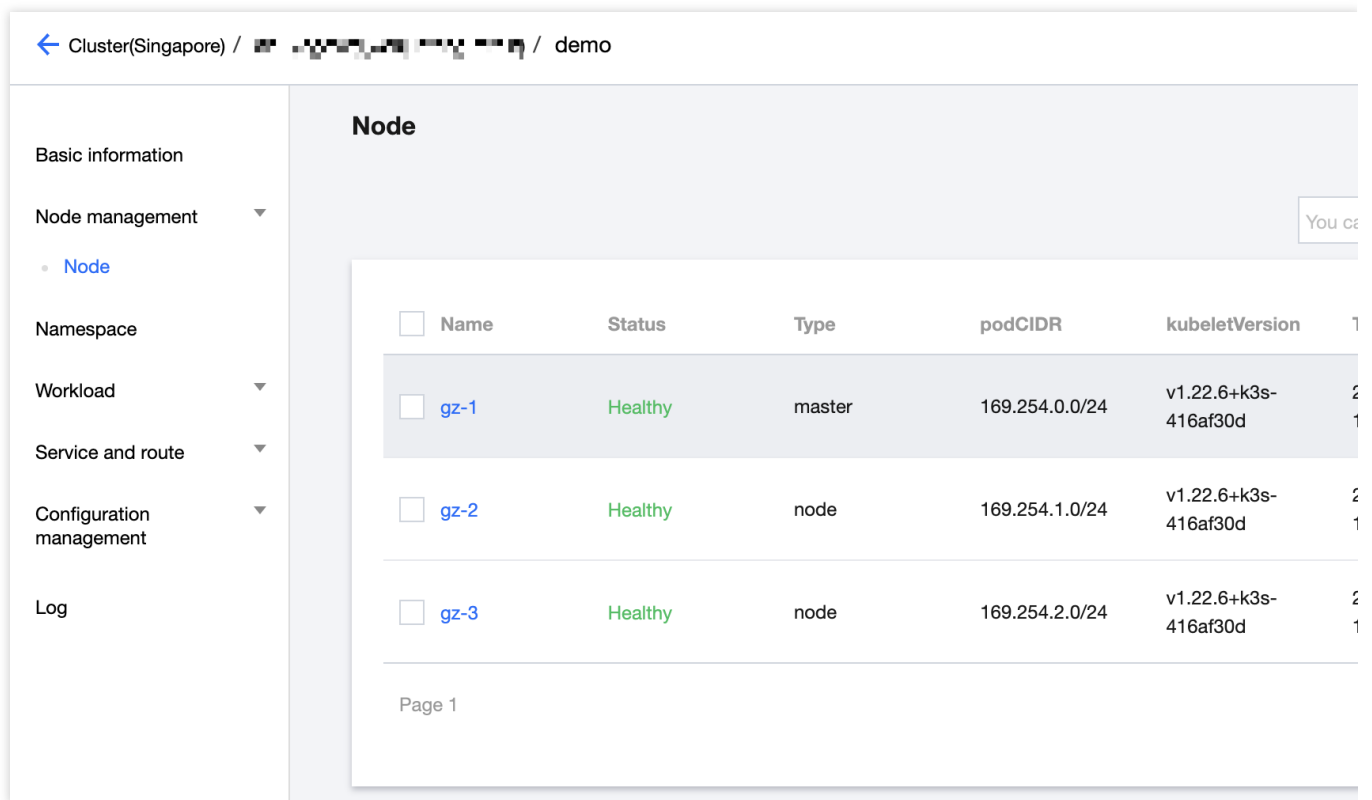
If users only want to clear the corresponding node from NodeUnit while not needing to remove it from the TKE-Edge cluster, they just need to modify the NodeUnit's node configuration. As shown below, simply deselect the corresponding node:

1. On the NodeUnit list page, select the specified NodeUnit and click on **"Update configuration"**.
2. On "Update NodeUnit" page, deselect the specified node as shown in the image below:



3. Click on "Done", and you can observe the edge K3s cluster node status. The corresponding node will be cleared, as shown in the following image:





## Method 2

If a user wants to delete a node on TKE-Edge and no longer use it, they can directly remove the corresponding node from the TKE-Edge node page. The corresponding node in the edge K3s cluster will also be cleaned up accordingly.

## Degrade edge K3s cluster

If a user no longer needs the edge K3s standalone cluster and wishes to revert to using the standard TKE-Edge edge NodeUnit, follow these steps:

1. On the NodeUnit list page, select the NodeUnit that needs to be downgraded, and click on **"Degrade to normal NodeUnit"**. As shown in the image below:

Cluster(Singapore) / [Cluster Name]

Basic information

Node management

- Node
- NodeUnit**
- NodeGroup

Namespace

Service Group

Workload

Auto scaling

Service

Configuration management

Add-on management

### NodeUnit

Create

| Name          | Node list                   | Type       | Running nodes |
|---------------|-----------------------------|------------|---------------|
| demo          | gz-1<br>gz-2<br>gz-3        | Standalone | 3/3           |
| unit-node-all | gz-1<br>gz-2<br>gz-3<br>... | Standard   | 4/4           |

Page 1

2. After being degraded to a standard node pool, its default behavior will be consistent with that of a standard NodeUnit.

## Delete edge K3s cluster

### Caution :

By default, the TKE-Edge product only allows users to delete standard NodeUnit from the interface. The delete button for independent node pools is disabled by default, so users are not permitted to directly delete "Standalone" NodeUnit. If you wish to delete an independent node pool, you must first perform the edge K3s cluster degrade operation to downgrade the standalone NodeUnit to a standard NodeUnit. After that, you can delete it from the NodeUnit page.

# Edge Platform Capability Monitor

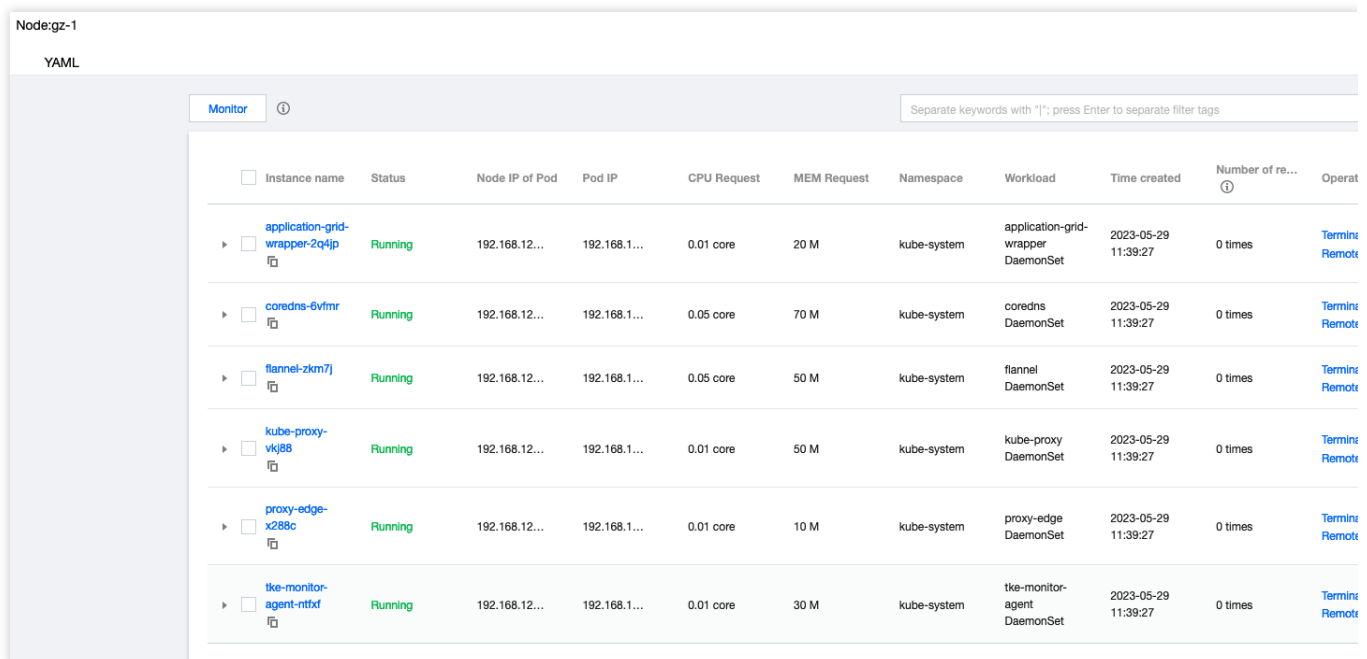
Last updated : 2023-06-01 11:22:54

The monitoring component capability is a basic supporting component of a platform system. The edge container reuses Tencent Cloud's cloud monitoring capabilities, which can upload the edge-side monitoring metrics to the cloud through the edge container's "Tunnel Cloud-Edge Channel" providing basic monitoring capabilities. You can view the cluster monitoring information according to the following operations:

## Operation Steps

### Basic Monitoring Information

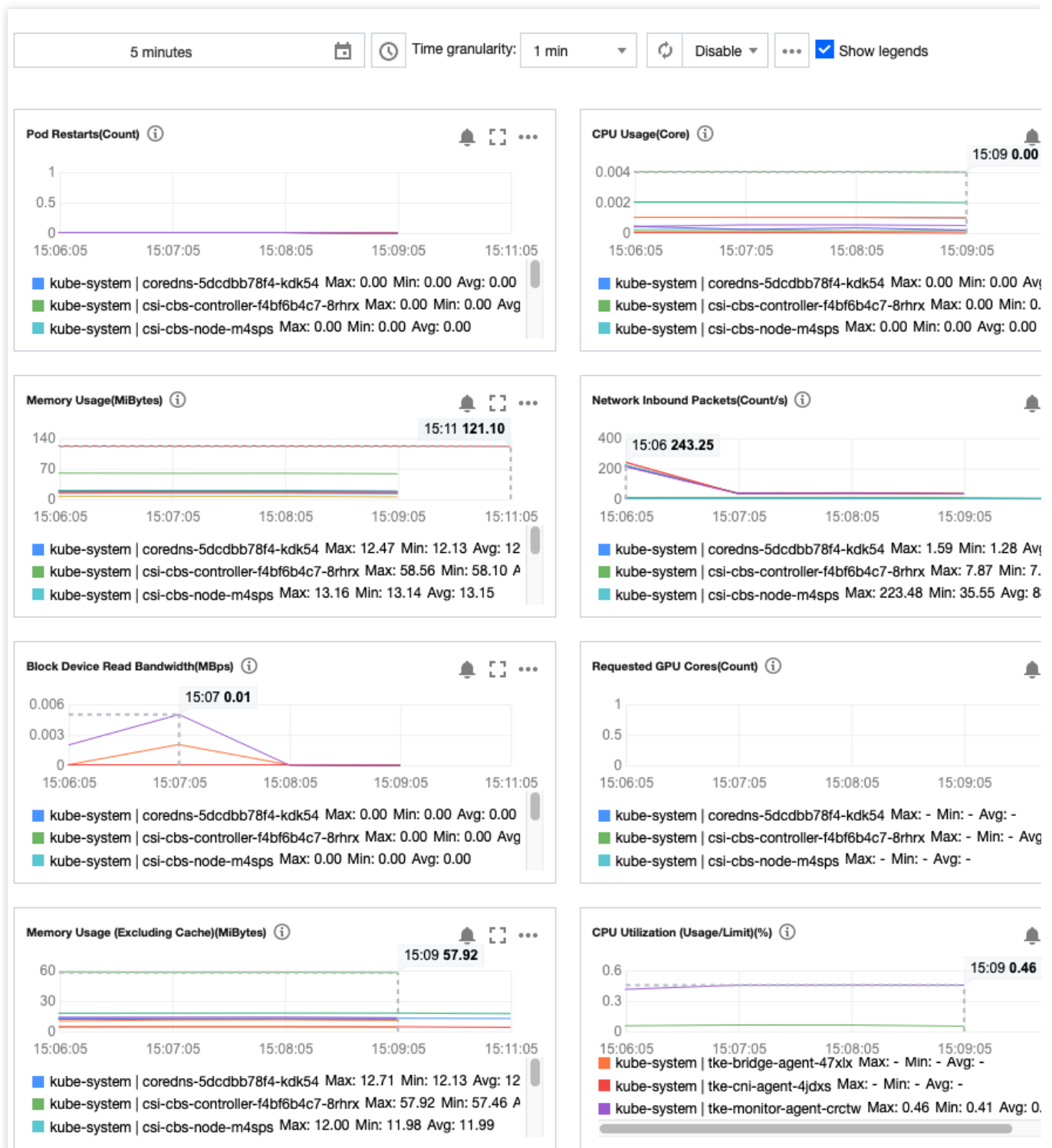
1. Log in to the [Tencent Kubernetes Engine](#), and select the Cluster option in the left-side navigation menu.
2. On the Cluster Management page, click on the Edge Cluster ID to enter the cluster details page.
3. Select Node Management > Nodes on the left side of the page to enter the node list page, and click on Monitoring to view specific monitoring information, as shown in the figure below:



The screenshot shows the 'Node:gz-1' monitoring page. A 'Monitor' tab is selected, displaying a table of pods. The table has columns for Instance name, Status, Node IP of Pod, Pod IP, CPU Request, MEM Request, Namespace, Workload, Time created, Number of re..., and Operat. The pods listed are all in a 'Running' state.

| Instance name                  | Status  | Node IP of Pod | Pod IP       | CPU Request | MEM Request | Namespace   | Workload                           | Time created        | Number of re... | Operat        |
|--------------------------------|---------|----------------|--------------|-------------|-------------|-------------|------------------------------------|---------------------|-----------------|---------------|
| application-grid-wrapper-2q4jp | Running | 192.168.12...  | 192.168.1... | 0.01 core   | 20 M        | kube-system | application-grid-wrapper DaemonSet | 2023-05-29 11:39:27 | 0 times         | Termin Remote |
| coredns-6vfmr                  | Running | 192.168.12...  | 192.168.1... | 0.05 core   | 70 M        | kube-system | coredns DaemonSet                  | 2023-05-29 11:39:27 | 0 times         | Termin Remote |
| flannel-zkm7l                  | Running | 192.168.12...  | 192.168.1... | 0.05 core   | 50 M        | kube-system | flannel DaemonSet                  | 2023-05-29 11:39:27 | 0 times         | Termin Remote |
| kube-proxy-vkj88               | Running | 192.168.12...  | 192.168.1... | 0.01 core   | 50 M        | kube-system | kube-proxy DaemonSet               | 2023-05-29 11:39:27 | 0 times         | Termin Remote |
| proxy-edge-x288c               | Running | 192.168.12...  | 192.168.1... | 0.01 core   | 10 M        | kube-system | proxy-edge DaemonSet               | 2023-05-29 11:39:27 | 0 times         | Termin Remote |
| tke-monitor-agent-ntxf         | Running | 192.168.12...  | 192.168.1... | 0.01 core   | 30 M        | kube-system | tke-monitor-agent DaemonSet        | 2023-05-29 11:39:27 | 0 times         | Termin Remote |

4. View monitoring information, as shown in the figure below:



The specific monitoring Metrics include:

**Events:** Pod restart count, Pod Ready status.

**CPU:** CPU usage (cores), CPU utilization (percentage of limit), CPU utilization (percentage of node), CPU utilization (percentage of request).

**Memory:** Memory usage (MBytes), memory usage (excluding cache), memory utilization (percentage of node, excluding cache), memory utilization (percentage of request), memory utilization (percentage of request, excluding cache), memory utilization (percentage of limit), memory utilization (percentage of limit, excluding cache), memory

utilization (percentage of node), memory usage\_working\_set, memory utilization\_working\_set percentage of request, memory utilization\_working\_set percentage of limit, memory utilization\_working\_set percentage of node.

**Network:** Network input packet rate (packets/s), network output traffic (MBytes), network input traffic (MBytes), network output bandwidth (Mbps), network input bandwidth (Mbps), and network output packet rate (packets/s).

**GPU (if included):** GPU memory usage rate (%), GPU encoding resource usage rate (%), GPU decoding resource usage rate (%), and GPU stream processor usage rate (%).

**Storage:** Rootfs space size (MiB).

### Note

The "Basic Monitoring" capability provided by the edge container is currently available to users for free.

## Cloud-native Monitoring

There are some shortcomings in the platform's built-in monitoring information:

Metrics are limited, and custom metrics cannot be defined.

It lacks monitoring capabilities for cloud-native related Metrics.

Therefore, the edge container service also supports the integration of Prometheus cloud-native monitoring capabilities. With this product, the edge container platform can support custom metrics and comes with free node metrics (node-exporter), kubelet metrics, and cluster-level kube-state-metrics information. This provides a complete cloud-native monitoring capability for edge clusters. For more information, refer to [Prometheus Monitoring](#).

### Note

This product is a paid product. Please understand the detailed features and purchase it as needed.

# Log

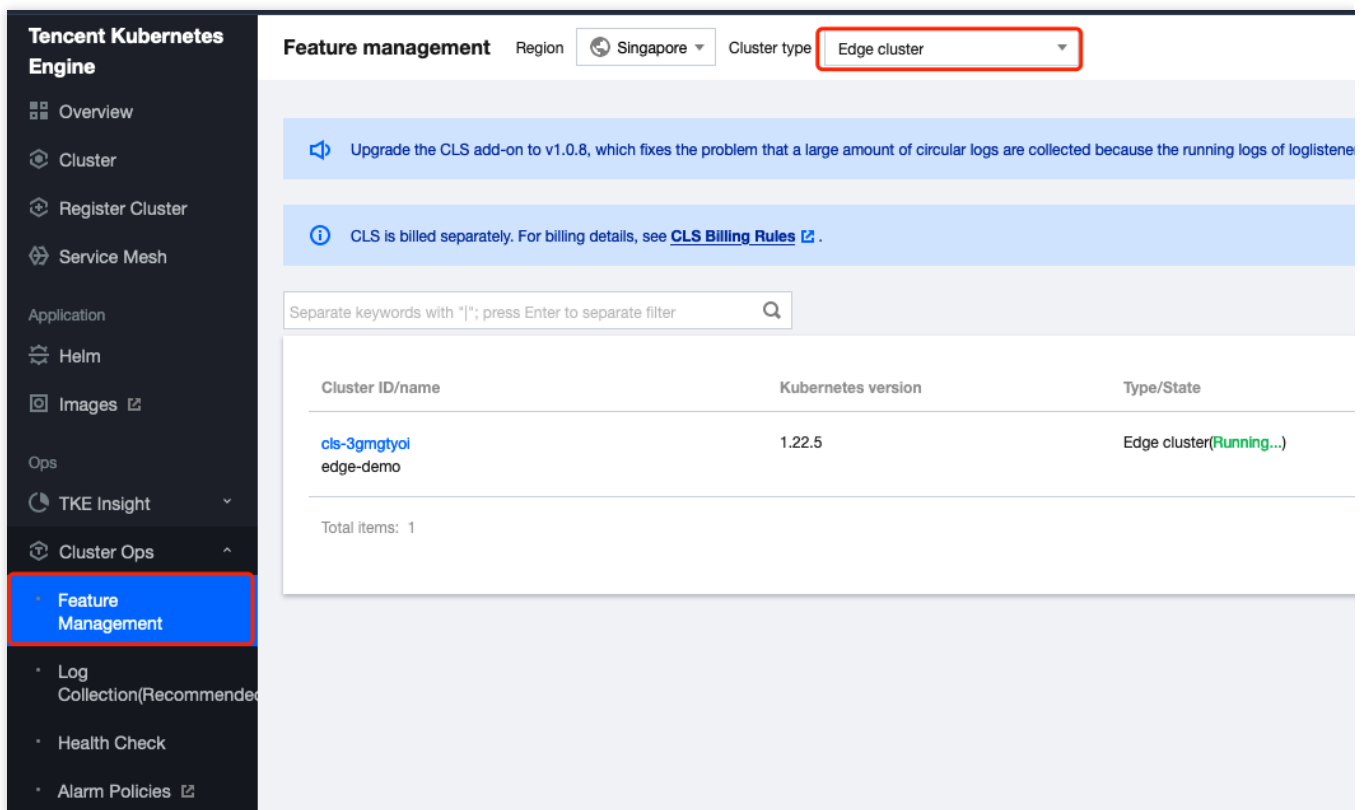
Last updated : 2023-06-01 11:22:54

The logging component capability is a basic supporting component of a platform system. The edge container reuses Tencent Cloud's logging service, which can upload the edge-side log data to the cloud, providing a complete log collection and retrieval capability. You can enable the log collection function according to the following operations:

## Operation Steps

### Enable Edge Container Logging Capability

1. Log in to the [Tencent Kubernetes Engine](#), and select Operation Management > Feature Management in the left-side navigation menu.
2. In the "Feature Management" page, select the region and cluster type at the top. As shown in the figure below:



3. Click the settings on the right side of the cluster for which you want to enable log collection.
4. On the "Set Features" page, click Edit on the right side of Log Collection. As shown in the figure below:

## Configure features

### Log collection

Log collection

Disabled

### Cluster Auditing

 CLS is billed separately. Total amount = Traffic fees + Storage fees + Other fees. For details, see [CLS Billing Rules](#)

Cluster Auditing

Disabled

### Event storage

 CLS is billed separately. Total amount = Traffic fees + Storage fees + Other fees. For details, see [CLS Billing Rules](#)

Event storage

Disabled

### Master logging

kube-apiserver

Disabled

kube-controller-manager


Disabled

kube-scheduler

Disabled

Check the "Enable log collection" and then click Confirm. As shown in the figure below:

**Configure features****Log collection**☒ Enable log collection

 If the current cluster does not have a logging rule, please enable Log Collection and go to [Log Collection Rule](#) page to edit the collection rule.

When log collection is enabled, the log collection component tke-log-agent (DaemonSet) will be deployed to the cluster system (namespace). Please reserve at least **0.1 core and 16 MiB** on each node.

**Confirm**

Cancel

5. After enabling, you can view or retrieve logs in the Log Service Console. For more details, please refer to the [Cloud Log Service](#).



# O&M Management

## Edgectl script manual

Last updated : 2023-06-01 11:53:42

Edgectl is an edge node management tool used for remotely adding edge nodes. You can obtain this tool by going to Node Management > [Add Nodes via Script](#).

Edgectl provides the following three commands:

| Command                         | Description   |
|---------------------------------|---|
| <a href="#">edgectl check</a>   | Check if the node meets the conditions for installing an edge node. |
| <a href="#">edgectl clear</a>   | Clean up the edge node.   |
| <a href="#">edgectl install</a> | Install the edge node.  |

The sample code is shown below:



```
# ./edgectl -h
Usage:
  edgectl command [flags]
Available Commands:
  check      Check the edge node if to be add to clusters
  install    Install components to edge node
  clear      Clear edge node and recovery as usual

Flags:
  -h, --help  Help for edgectl
```

# edgectl check command

## Meaning

Check if the node meets the conditions for installing an edge node. The check project includes the following:

Check if the user is the root user.

Check if the system is within the supported range.

Check if the swap area is turned off.

Check if the firewall is turned off.

Check if ufw is turned off.

Check if the ports are occupied (the ports checked are 1443 and the ranges 10249-10259 and 51000-51020).

Check if cgroup memory is enabled.

Check if the node has previously installed kubeadm, docker, kubelet, and kubectl.

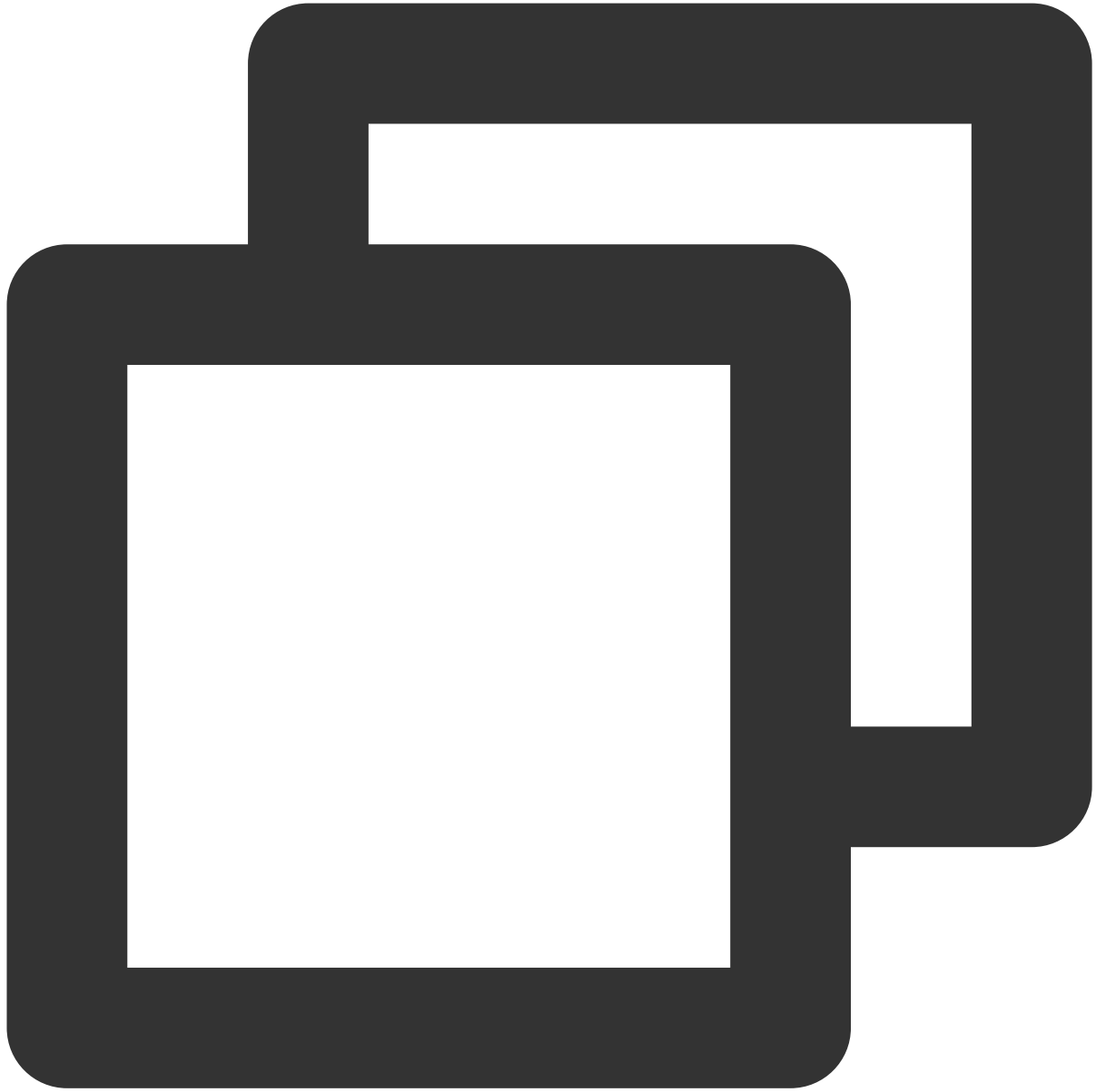
## Usage example



```
# ./edgectl check
Unit firewalld.service could not be found.
WARN >> Port: 1443 occupied. Please turn off port service.
...
WARN >> The machine is not clean. Please reinstall the system.
/usr/bin/kubelet
...
>> Check Environment Finish! <<
```

For instance, after executing the sample code this time, the user is prompted with the following 2 risk items:

The edge node's port 1443 is already in use, please shut down its service.



```
WARN >> Port: 1443 occupied. Please turn off port service.
```

The node has kubelet remnants. It is recommended to reinstall the node system, or execute the `edgectl clear` command to perform the cleanup operation.



```
WARN >> The machine is not clean. Please reinstall the system.  
/usr/bin/kubelet ## The kubelet has been installed previously
```

## edgectl clear command

### Meaning

Clean up the edge node. The clear command will remove the following information:

Remove all containers and Pods running on the edge node.

Stop kubelet, lite-apiserver, and docker.

Remove created network information and routing information.

Delete the following folders or files:

/etc/kubernetes

/etc/docker

/root/.kube/config

/var/lib/kubelet >/dev/null 2>&1

/var/lib/cni

/etc/cni

/etc/sysconfig/kubelet/

/etc/sysconfig/lite-apiserver

/data/lite-apiserver >/dev/null 2>&1

/usr/lib/systemd/system/{kubelet, docker, lite-apiserver}.service

Note :

`edgectl clear` will remove all containers and Pods on the node, please execute with caution.

`edgectl clear` will remove relevant folders or files, please back up important materials in advance.

The `edgectl install` command will execute `edgectl clear` by default. Before executing `edgectl install`, please consider the risk items removed by `edgectl clear`.

## Usage example



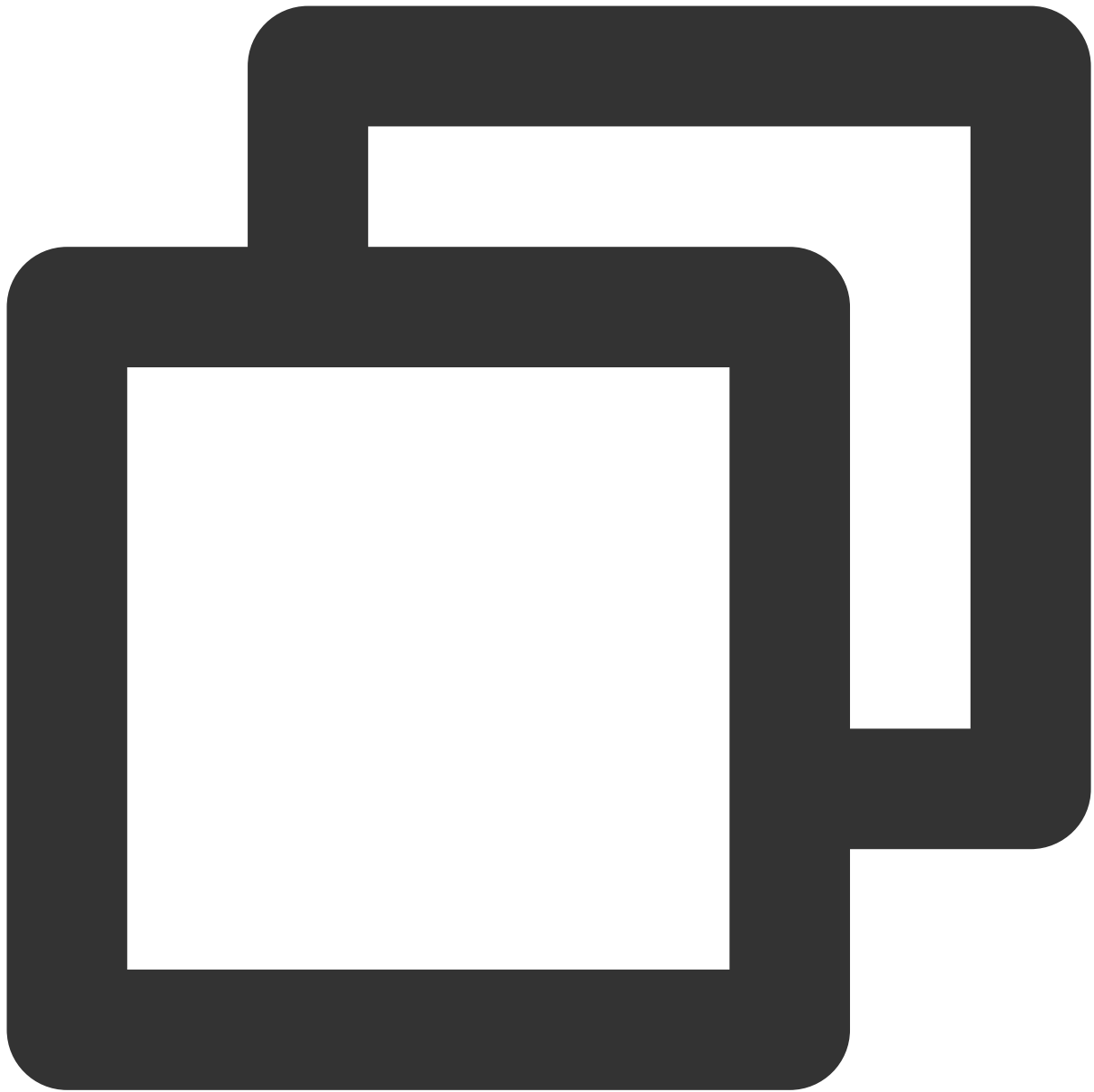
```
# ./edgectl clear  
removed '/etc/kubernetes/cluster-ca.crt'  
...  
>> Clear Node Complete! <<
```

## edgectl install command

### Meaning



Install the edge node. Execute the following command to install the edge node.



```
# ./edgectl install -h
Usage:
  edgectl install [flags]
Flags:
  -n, --node-name      Node name in edge cluster. Must!## Node name, required
  -i, --interface      Default network interface name. ## Node's default network i
```

**--node-name** : Edge node name, abbreviated as -n.

The name is a required option and must strictly follow the [Kubernetes node naming conventions](#).

Please ensure that the node name is unique within the cluster being added, otherwise, it will cause a problem of continuous switching of nodes with the same node name.

Node names do not support using IP addresses directly, otherwise it will cause the `kubect` log and `kubect` exec commands to become invalid.

**--interface** : The default network interface name for the edge node, abbreviated -i.

Optional, the default network interface name of the edge node. Entering it incorrectly will cause the flannel and coredns components to malfunction.

The default value is the specified network interface name entered from the Tencent Cloud page, and specifying `--interface` will override the original default value.

## Usage example



```
# ./edgectl install --node-name node-192.168.67.91 --interface eth0
NOTE:
    input: [ edgectl install --node-name node-192.168.67.91 --interface eth0 ] ##
    logPath: /tmp/tke-edge-install.log ## The log location of this installation
    success-message: Install Edge Node: node-192.168.67.91 Success! ## Message af

Start Install Edge Node node-192.168.67.91, Please Waiting...
Waiting Running of the base service
Dockerd kubelet lite-apiserver has Running! ## Dockerd kubelet lite-apiserver succ
Install Edge Node: node-192.168.67.91 Success! ## Indicates that the edge node i
```

After the edge node installation is completed, you can execute the command `kubectl -n kube-system get pod` to check whether all Pods are in the Running state.

The `edgectl install` command is valid for one hour and can be used to add different nodes multiple times within the valid period. If it expires, please perform the Script Add Nodes operation again to request a new `edgectl`.

If there is no prompt to exit during the installation process, please refer to the installation log location `logPath: /tmp/tke-edge-install.log` provided in the NOTE for specific errors.

If there is no "Install Edge Node: node-192.168.67.91 Success!" after execution, the installation is considered failed.

Please check the installation log for troubleshooting.