Alibaba Cloud
Container Service for Kubernetes

problem

Issue: 20180906

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, special, or punitive damages, including lost profits arising from the use or trust in this document, even if Alibaba Cloud has been notified of the possibility of such a loss.
5. By law, all the content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade

secrets. No part of the Alibaba Cloud website, product programs, or content shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates).

6. Please contact Alibaba Cloud directly if you discover any errors in this document.

Generic conventions

Table -1: Style conventions

Style	Description	Example
	This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 **Danger:** Resetting will result in the loss of user configuration data.
	This warning information indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 **Warning:** Restarting will cause business interruption. About 10 minutes are required to restore business.
	This indicates warning information, supplementary instructions, and other content that the user must understand.	 **Note:** Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	 **Note:** You can use **Ctrl** + **A** to select all files.
>	Multi-level menu cascade.	**Settings** > **Network** > **Set network type**
Bold	It is used for buttons, menus, page names, and other UI elements.	Click **OK**.
`Courier font`	It is used for commands.	Run the `cd /d C:/windows` command to enter the Windows system folder.
Italics	It is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	It indicates that it is a optional value, and only one item can be selected.	`ipconfig` *[-all|-t]*
{} or {a|b}	It indicates that it is a required value, and only one item can be selected.	`swich` *{stand | slave}*

Contents

1 如何选择Kubernetes集群网络插件

容器服务在Kubernetes集群创建时提供了两种网络插件选择，Terway和Flannel，那么创建集群时您需要选择哪个呢？下面看一下两种插件的功能：

- Flannel: 这个网络插件使用的是简单稳定的社区的*Flannel* cni插件，配合阿里云的VPC的高速网络，能给集群高性能和稳定的容器网络体验，但功能偏简单，支持的特性少，目前建议在生产环境中选择Flannel的网络插件。
- Terway: 这个网络插件是阿里云容器服务自研的网络插件，支持将阿里云的弹性网卡分配给容器，支持Kubernetes的NetworkPolicy来定义容器间的访问策略，支持对单个容器做带宽的限流，但目前该插件还在公测阶段，对于测试环境或者希望体验这些新特性的用户可以尝试这个网络插件。
  - 使用弹性网卡：

    预先给Worker节点绑定好弹性网卡，并重启节点，然后在pod的`annotation`中配置：`k8s.aliyun.com/eni: "true"`，在容器创建时就会使用预先绑定的弹性网卡了。
  - 使用NetworkPolicy:

    参见*https://kubernetes.io/docs/concepts/services-networking/network-policies/*。
  - 对Pod做带宽限制：

    在Pod的**`annotation`**中分别通过`k8s.aliyun.com/ingress-bandwidth`, `k8s.aliyun.com/egress-bandwidth` 可以指定Pod的最大入网带宽和出网带宽， 例如：`k8s.aliyun.com/ingress-bandwidth: 1m`，`k8s.aliyun.com/egress-bandwidth: 1m`

2 如何支持私有镜像

```
kubectl create secret docker-registry regsecret --docker-server=
registry-internal.cn-hangzhou.aliyuncs.com --docker-username=abc@
aliyun.com --docker-password=xxxxxx --docker-email=abc@aliyun.com
```

其中：

- **regsecret**： 指定密钥的键名称，可自行定义。
- **—docker-server**：指定 Docker 仓库地址。
- **—docker-username**: 指定 Docker 仓库用户名。
- **—docker-password**：指定 Docker 仓库登录密码。
- **—docker-email**：指定邮件地址（选填）。

yml 文件加入密钥参数。

```
containers:
    - name: foo
      image: registry-internal.cn-hangzhou.aliyuncs.com/abc/test:1.0
imagePullSecrets:
    - name: regsecret
```

其中：

- `imagePullSecrets` 是声明拉取镜像时需要指定密钥。
- `regsecret` 必须和上面生成密钥的键名一致。
- `image` 中的 Docker 仓库名称必须和 `--docker-server` 中的 Docker 仓库名一致。

详情信息参见官方文档 ######。

3 FAQ about storage volumes

Storage volumes cannot be mounted

Check if flexvolume is installed.

Execute the following command on the master node:

```
# kubectl get pod -n kube-system | grep flexvolume

flexvolume-4wh8s            1/1       Running   0          8d
flexvolume-65z49            1/1       Running   0          8d
flexvolume-bpc6s            1/1       Running   0          8d
flexvolume-l8pml            1/1       Running   0          8d
flexvolume-mzkpv            1/1       Running   0          8d
flexvolume-wbfhv            1/1       Running   0          8d
flexvolume-xf5cs            1/1       Running   0          8d
```

Check if the flexvolume pod status is Running and if the number of running flexvolume pods is the same as the number of nodes.

If not, see *#unique_6*.

If the flexvolume pod status is not running, see the running log analysis of the plug-in.

Check if the dynamic storage plug-in is installed

To use the dynamic storage function of a cloud disk, execute the following command to verify the dynamic storage plug-in is installed:

```
# kubectl get pod -n kube-system | grep alicloud-disk

alicloud-disk-controller-8679c9fc76-lq6zb     1/1 Running   0   7d
```

If not, see *#unique_6*.

If the dynamic storage plug-in status is not running, see the running log analysis of the plug-in.

How to view types of storage logs?

View flexvolume logs by executing commands on the master1 node

Execute the following get command to view the error pod:

```
# kubectl get pod -n kube-system | grep flexvolume
```

Execute the following log command to view the log for the error pod:

```
# kubectl logs flexvolume-4wh8s -n kube-system
# kubectl describe pod flexvolume-4wh8s -n kube-system
```

```
# The last several lines in the pod description are the descriptio
ns of pod running status. You can analyze pod errors based on the
descriptions.
```

View drive logs of the cloud disk, Network Attached Storage (NAS), and Object Storage Service (OSS):

```
# View the persistent logs on the host node;
# If a pod mount fails, view the address of the node on which the pod
resides:

# kubectl describe pod nginx-97dc96f7b-xbx8t | grep Node
Node: cn-hangzhou.i-bp19myla3uvnt6zihejb/192.168.247.85
Node-Selectors:  <none>

# Log on to the node to view logs:

# ssh 192.168.247.85
# ls /var/log/alicloud/flexvolume*
flexvolume_disk.log  flexvolume_nas.log  flexvolume_o#ss.log

You can see logs mounted on the cloud disk, NAS, and OSS;
```

View provsioner plug-in logs by executing commands on the master1 node

Execute the following get command to view the error pod:

```
# kubectl get pod -n kube-system | grep alicloud-disk
```

Execute the log command to view the log for the error pod:

```
# kubectl logs alicloud-disk-controller-8679c9fc76-lq6zb -n kube-
system
# kubectl describe pod alicloud-disk-controller-8679c9fc76-lq6zb -n
kube-system

# The last several lines in the pod description are the descriptio
ns of pod running status. You can analyze pod errors based on the
descriptions.
```

View Kubelet logs

```
# If a pod mount fails, view the address of the node on which the pod
resides:

# kubectl describe pod nginx-97dc96f7b-xbx8t | grep Node
Node: cn-hangzhou.i-bp19myla3uvnt6zihejb/192.168.247.85
Node-Selectors: <none>

# Log on to the node to view kubelet logs:

# ssh 192.168.247.85
# journalctl -u kubelet -r -n 1000 &> kubelet.log
```

```
# The value of -n indicates the number of log lines that you expect to
 see;
```

The above are methods to obtain error logs of flexvolume, provsioner, and kubelet. If the logs cannot help you to repair the status, contact Alibaba Cloud technical support with the logs.

FAQ about cloud disks

Cloud disk mount fails with timeout errors

If the node is added manually, the failure may be caused by problem about Security Token Service (STS) permissions. You need to manually configure Resource Access Management (RAM) permissions: *Use the instance RAM role in the console*.

Cloud disk mount fails with size errors

The following are size requirements for creating a cloud disk:

Note:

- Basic cloud disk: Minimum 5Gi
- Ultra cloud disk: Minimum 20Gi
- SSD cloud disk: Minimum 20Gi

Cloud disk mount fails with zone errors

When the ECS mounts a cloud disk, they must be in the same zone under the same region. Otherwise, the cloud disk cannot be mounted successfully.

After your system is upgraded, the cloud disk sometimes reports input/output error

1. Upgrade flexvolume to v1.9.7-42e8198 or later.
2. Rebuild pods that have already gone wrong.

Upgrading command:

```
# kubectl set image daemonset/flexvolume acs-flexvolume=registry.cn-
hangzhou.aliyuncs.com/acs/flexvolume:v1.9.7-42e8198 -n kube-system
```

Flexvolume version information: To obtain the latest version of flexvolume, log on to the container image service console, click **Image search** in the left-side navigation pane, and search for acs/flexvolume.

FAQ about NAS

NAS mount time is too long

If the NAS volume contains a large amount of files and the chmod parameter is configured in the mount template, the mount time may be too long. To solve this problem, remove the chmod parameter.

NAS mount fails with the timeout error

Check if the NAS mount point and the cluster are within the same Virtual Private Cloud (VPC). If not, NAS cannot be mounted.

FAQ about OSS

OSS mount fails

Check if the AK used is correct.

4 收集 Kubernetes 诊断信息

1. 在 master 节点下载诊断脚本，并增加运行权限。

```
curl -o /usr/local/bin/diagnose_k8s.sh http://aliacs-k8s-cn-hangzhou
.oss-cn-hangzhou.aliyuncs.com/public/diagnose/diagnose_k8s.sh
chmod u+x /usr/local/bin/diagnose_k8s.sh
```

2. 执行诊断脚本。

```
diagnose_k8s.sh
 ......
 + echo 'please get diagnose_1514939155.tar.gz for diagnostics'   ##
每次执行诊断脚本，产生的日志文件的名称不同
 please get diagnose_1514939155.tar.gz for diagnostics
 + echo '请上传 diagnose_1514939155.tar.gz'
 请上传 diagnose_1514939155.tar.gz
```

3. 列出产生的日志文件。

```
cd /usr/local/bin
ls -ltr|grep diagnose_1514939155.tar.gz           ##注意替换为生成的日
志文件名
```

5 Failed to create a Kubernetes cluster

Check the cause of failure

You can check the cause of cluster creation failure by viewing the cluster creation events.

Log on to the *Resource Orchestration Service (ROS) console*.

Select the region in which the cluster resides. Click **Manage** at the right of the cluster. Click **Event** in the left-side navigation pane. Move the cursor over the failed event to view the specific error message of the failure.

If the preceding error message is displayed, it means that the cluster creation failed because the number of Virtual Private Cloud (VPC) instances has reached the quota.

Failure codes and solutions

- **Code: QuotaExceeded.Eip, Message: Elastic IP address quota exceeded**

  Solution: Release unused EIPs, or open a ticket to raise the EIP quota.
- **The maximum number of SLB instances is exceeded. Code: ORDER.QUANTITY_INVALID**

  Solution: Release unused SLB instances, or open a ticket to raise the SLB quota.
- **Resource CREATE failed: ResponseException: resources.k8s_vpc: VPC quota exceeded . Code: QuotaExceeded.Vpc**

  Solution: Release unused VPCs, or open a ticket to raise the VPC quota.
- **Resource CREATE failed: ResponseException: resources.k8s_master_1: The specified image does not support cloud-init. Code: ImageNotSupportCloudInit**

  Solution: When using custom image to create a cluster, the custom image used must be developed based on the lastest Centos public cloud image.
- **Status Code: 403 Code: InvalidResourceType.NotSupported Message: This resource type is not supported;**

  Solution: ECS is out of stock or the type of ECS instances you selected are not supported.

6 Failed to delete Kubernetes clusters: ROS stack cannot be deleted

Root cause

Some resources are manually added (for example, manually add a VSwitch under the Virtual Private Cloud (VPC) created by Resource Orchestration Service (ROS)) under the resources created by ROS. ROS does not have permissions to delete those resources. This causes ROS to fail to process the VPC when deleting the Kubernetes resources and then the cluster fails to be deleted.

Note:
For more information about the resources automatically created by ROS when the Kubernetes cluster is created, see *#unique_11*.

Solutions

1. If the cluster fails to be deleted (the cluster status is **Failed to delete**), go to the *ROS console*.
2. Select the region in which the cluster resides and find the stack `k8s-for-cs-{cluster-id}` corresponding to the cluster. You can see the status is **Failed to delete**.
3. Click the stack name to go to the stack details page. Click **Resource** in the left-side navigation pane.

   You can see what resources failed to be deleted. In this example, the VSwitch under Server Load Balancer failed to be deleted.
4. Go to the console in which the resource that failed to be deleted resides and find that resource.

   In this example, log on to the VPC console and find the VPC in which the cluster resides. Find the VSwitch that failed to be deleted under that VPC.
5. Click **Delete** at the right of the VSwtich to manually delete it.

   In this example, the VSwitch has resources to release and cannot be deleted.

   Manually release the resources under this VSwitch and try to delete this VSwitch again.

6. Manually delete all the resources that failed to be deleted under the Kubernetes cluster in this way and try to delete the Kubernetes cluster again.

7 Upgrade Helm manually

Log on to the master node of the Kubernetes cluster, see *#unique_13*.

Execute the following command:

```
helm init --tiller-image registry.cn-hangzhou.aliyuncs.com/acs/tiller:
v2.9.1 --upgrade
```

The image address can use the VPC domain name of the region corresponding to the image. For example, the image address of a machine in the Hangzhou region can be replaced by registry-vpc .cn-hangzhou.aliyuncs.com/acs/tiller:v2.9.1.

Wait for **`tiller`** passing through health check. Then you can execute `helm version` to view the upgraded version.

Note:
Only the Helm server version is upgraded here. To use the Helm client, download the corresponding client binary.

Helm 2.9.1 client download address: *https://github.com/kubernetes/helm/releases/tag/v2.9.1*。
Currently, the latest version of Helm supported by Alibaba Cloud is 2. 9.1.

After the Helm client and server are both upgraded, you can see the following information by executing the **`helm version`** command:

```
#helm version
Client: &version.Version{SemVer:"v2.9.1", GitCommit:"a80231648a
1473929271764b920a8e346f6de844", GitTreeState:"clean" }
Server: &version.Version{SemVer:"v2.9.1", GitCommit:"a80231648a
1473929271764b920a8e346f6de844", GitTreeState:"clean "}
```

8 How to manually install alicloud-application-controller

By default, alicloud-application-controller is installed in Alibaba Cloud Container Service in version 1.10.4 and later to provide the release based on custom resource definition (CRD).

Note:

In the Kubernetes cluster of the latest version, alicloud-application-controller is installed by default. In Kubernetes clusters of old versions, manually install alicloud-application-controller and the oldest version of Kubernetes cluster must be 1.9.3.

Use the `kubectl create -f alicloud-application-controller.yml` command to deploy alicloud-application-controller. In *`alicloud-application-controller.yml`*, enter the following orchestration template:

```
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  name: alicloud-application-controller
  labels:
    owner: aliyun
    app: alicloud-application-controller
  namespace: kube-system
spec:
  replicas: 1
  selector:
    matchLabels:
      owner: aliyun
      app: alicloud-application-controller
  template:
    metadata:
      labels:
        owner: aliyun
        app: alicloud-application-controller
      annotations:
        scheduler.alpha.kubernetes.io/critical-pod: ''
    spec:
      tolerations:
      - effect: NoSchedule
        operator: Exists
        key: node-role.kubernetes.io/master
      - effect: NoSchedule
        operator: Exists
        key: node.cloudprovider.kubernetes.io/uninitialized
      containers:
        - name: alicloud-application-controller
          image: registry.cn-hangzhou.aliyuncs.com/acs/aliyun-app-
lifecycle-manager:0.1-c8d5da8
          imagePullPolicy: IfNotPresent
      serviceAccount: admin
```