

- Serverless 容器从入门到精通: Serverless K8s -

第 02 讲

Serverless Kubernetes 应用部署及扩缩容

轻零 阿里巴巴技术专家



关注“Serverless”公众号
获取第一手技术资料



Serverless Kubernetes 集群创建

创建 Kubernetes 集群

Kubernetes 专有版

Kubernetes 托管版

Serverless Kubernetes

Kubernetes 边缘托管版（公测）

创建集群注册代理（公测）

集群模板

返回集群列表

Serverless Kubernetes 所调度的 ECI 容器实例为按量收费资源，请您在创建应用前务必了解 ECI 的计费标准，详情请参见 ECI 计费标准

容器 ECI 免费试用3个月，限量100元代金券，立即领取

* 集群名称

名称为1~63个字符，可包含数字、汉字、英文字符，或"-"

资源组

未选择

地域

华北 2（北京）

华北 3（张家口）

华北 5（呼和浩特）

华东 1（杭州）

华东 2（上海）

华南 1（深圳）

西南 1（成都）

中国（香港）

新加坡

澳大利亚（悉尼）

印度尼西亚（雅加达）

印度（孟买）

美国（弗吉尼亚）

美国（硅谷）

英国（伦敦）

德国（法兰克福）

Kubernetes 版本

1.16.9-aliyun.1

1.14.8-aliyun.1

专有网络

自动创建

使用已有

可用区

华北 2 可用区 H

NAT 网关

☐ 在 VPC 中自动创建 NAT 网关并配置 SNAT 规则

推荐

自动创建 VPC 时必须配置 SNAT

当前配置

地域：华北2（北京）

创建集群

开通按量付费的云服务器，现金账户余额不得少于100.00元，如账户金额少于100.00元，需充值后方可开通，否则创建集群会失败。去充值

开通 云服务器、负载均衡 需要先进行实名认证，马上认证>>

您目前可以创建 20 个集群，每个集群最多可以添加 100 个节点。如需更高配额，请提交工单申请

如需更多实例规格，请提交工单申请

<https://cs.console.aliyun.com/#/k8s/cluster/createV2/serverless>

应用部署

<

集群: serverlesee-demo / default / 无状态

帮助文档

无状态有状态任务定时任务容器组

应用基本信息容器配置高级配置创建完成

应用名称

名称为1-63个字符，可包含数字、小写英文字母以及短划线（-），且不能以短划线（-）开头

副本数量

2

类型

无状态

标签

+ 添加

注解

+ 添加

返回

下一步

基本信息

存储卷

命名空间

工作负载

服务

路由

发布

配置管理

运维管理

事件列表

集群审计

应用部署



1

集群创建及应用部署

2

常用功能介绍

3

应用扩缩容

Serverless Kubernetes 控制台

<

集群: serverlesee-demo

刷新

通过 CloudShell 管理集群

基本信息

存储卷

命名空间

工作负载

服务

路由

发布

配置管理

▼ 运维管理

事件列表

集群审计

基本信息

集群 ID: c079636dfa647474f94df3a6f8f8a7c26

虚拟专有网络

● 运行中

地域: 华北2 (北京)

删除保护: ☐

集群信息

API Server 公网连接端点

https://59.110.164.25:6443

API Server 内网连接端点

https://192.168.7.33:6443

集群资源

虚拟专有网络 VPC

vpc-2zewb358rv3qursycneyx

节点虚拟交换机

vsw-2ze29treiscw3mn7tevy3

安全组

sg-2zehdb9gj0vmhp406xm5

通过 kubectl 连接 Kubernetes 集群

通过 CloudShell 管理集群

1. 下载最新的 kubectl 客户端, 请参见 [Kubernetes 版本页面](#)。

2. 安装和设置 kubectl 客户端。有关详细信息请参见 [安装和设置 kubectl](#)。

3. 配置集群凭据:

公网访问

内网访问

吊销 KubeConfig

阿里云

Kubectl

```
shell@alicloud:~$ kubectl -h
kubectl controls the Kubernetes cluster manager.

Find more information at: https://kubernetes.io/docs/reference/kubectl/overview/

Basic Commands (Beginner):
  create      Create a resource from a file or from stdin.
  expose      Take a replication controller, service, deployment or pod and expose it as a new Kubernetes Service
  run         Run a particular image on the cluster
  set         Set specific features on objects

Basic Commands (Intermediate):
  explain     Documentation of resources
  get         Display one or many resources
  edit        Edit a resource on the server
  delete      Delete resources by filenames, stdin, resources and names, or by resources and label selector

Deploy Commands:
  rollout     Manage the rollout of a resource
  scale       Set a new size for a Deployment, ReplicaSet or Replication Controller
  autoscale   Auto-scale a Deployment, ReplicaSet, or ReplicationController

Cluster Management Commands:
  certificate Modify certificate resources.
  cluster-info Display cluster info
  top         Display Resource (CPU/Memory/Storage) usage.
  cordon      Mark node as unschedulable
  uncordon    Mark node as schedulable
  drain       Drain node in preparation for maintenance
  taint       Update the taints on one or more nodes

Troubleshooting and Debugging Commands:
  describe    Show details of a specific resource or group of resources
  logs        Print the logs for a container in a pod
  attach      Attach to a running container
  exec        Execute a command in a container
  port-forward Forward one or more local ports to a pod
  proxy       Run a proxy to the Kubernetes API server
  cp          Copy files and directories to and from containers.
  auth        Inspect authorization

Advanced Commands:
  diff        Diff live version against would-be applied version
  apply       Apply a configuration to a resource by filename or stdin
  patch       Update field(s) of a resource using strategic merge patch
  replace     Replace a resource by filename or stdin
```


1

集群创建及应用部署

2

常用功能介绍

3

应用扩缩容

应用扩缩容方式一： 人工扩缩容

伸缩

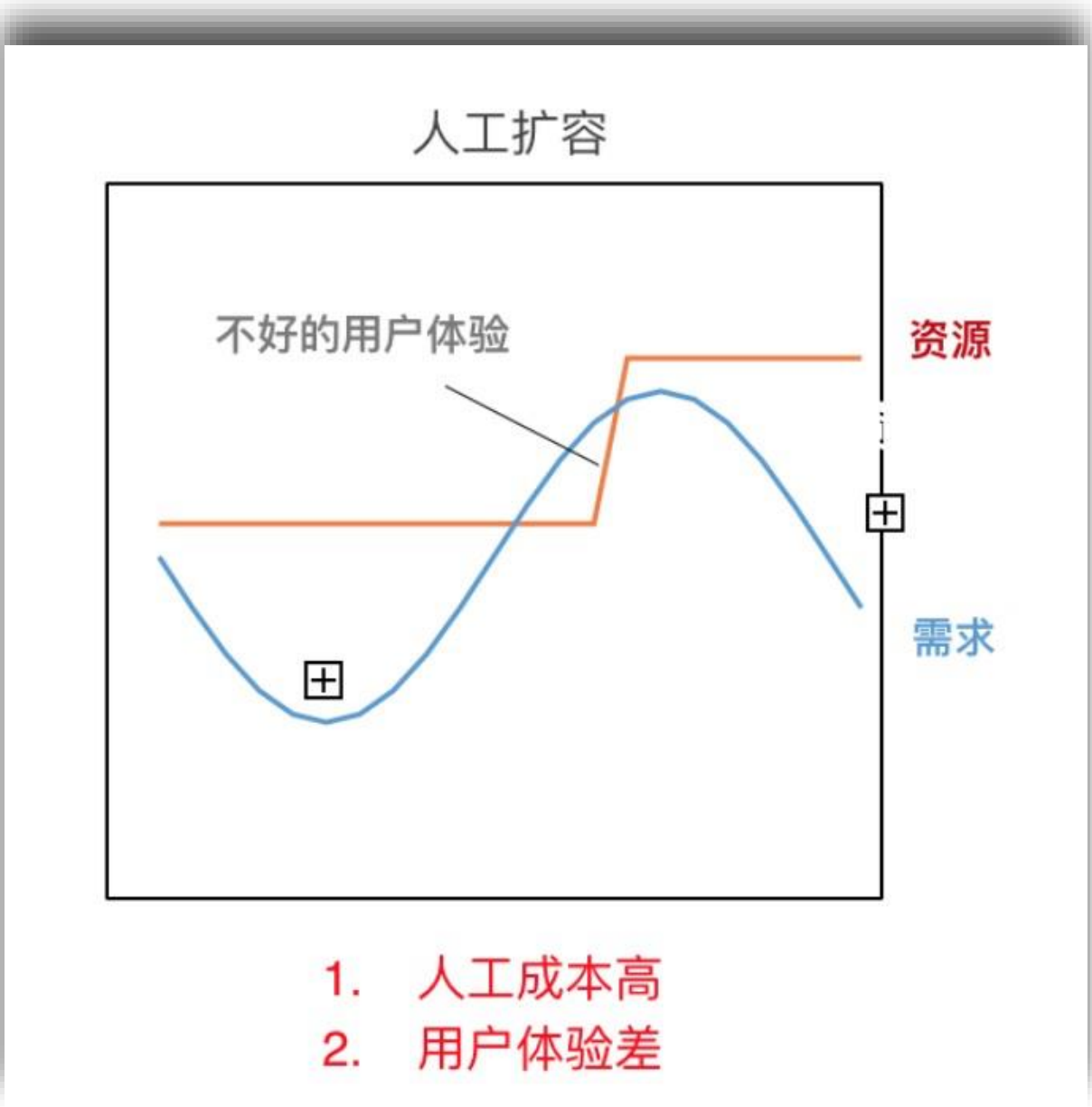
所需容器组数量：

2

更新资源 nginx 的容器数量
已创建 2 个， 总共需要 2 个

确定

取消



应用扩缩容方式二：HPA

创建

名称:

指标:

+ 添加

监控项	阈值
CPU使用量	<div></div> % <div>-</div>
内存使用量	<div></div> % <div>-</div>

最大容器数量:

10

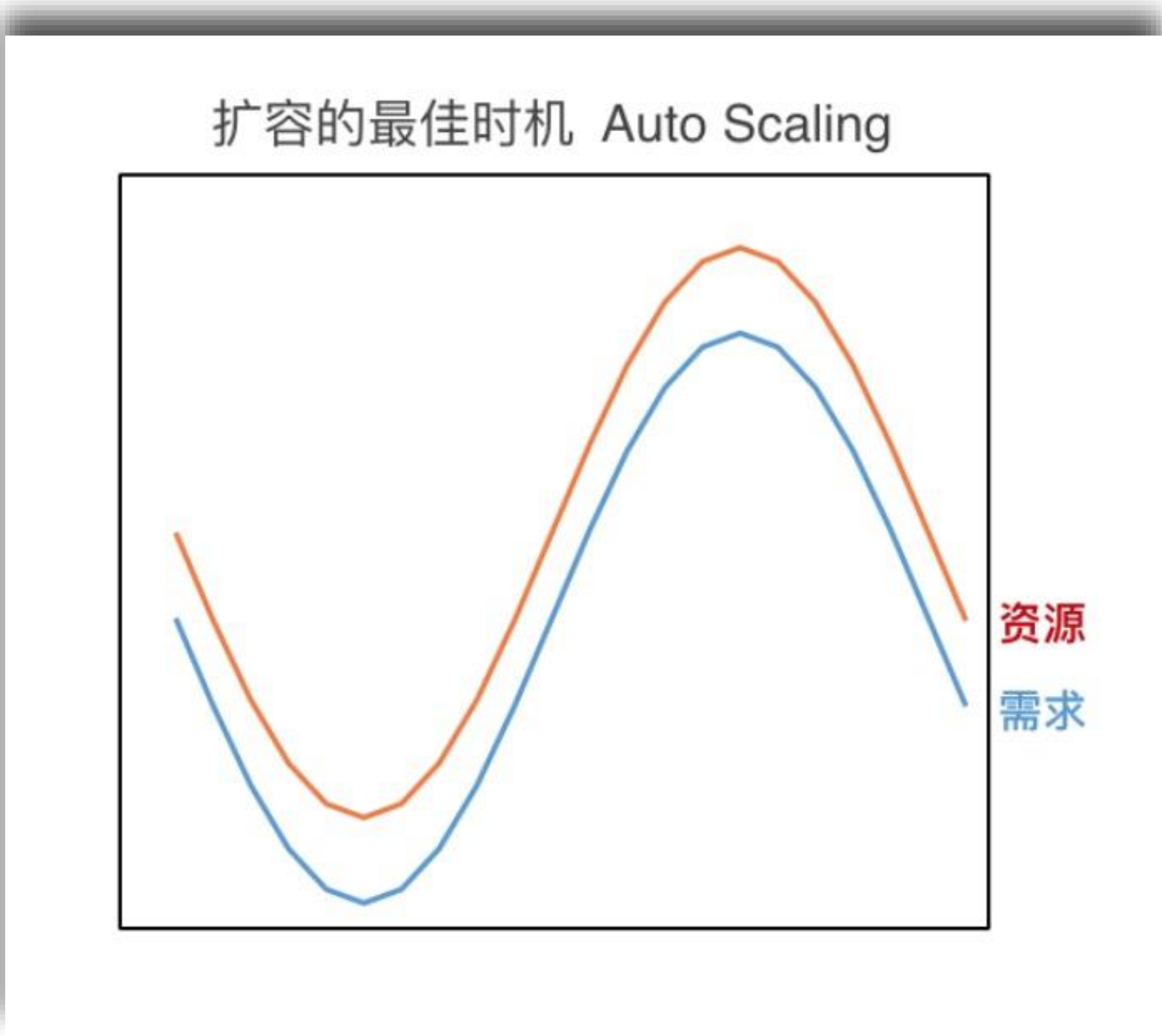
最小容器数量:

1

使用模板创建

确定

取消



应用扩缩容更多方式

- **Cron HPA**
- **External Metrics (alibaba-cloud-metrics-adapter)**

谢谢观看

THANK YOU



关注“Serverless”公众号
获取第一手技术资料