

容器服务 购买指南



腾讯云

【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

购买指南

购买指引

购买 TKE 标准集群

计费概述

购买说明

欠费说明

购买集群配额限制

集群新增资源所属项目说明

购买原生节点

原生节点产品定价

欠费说明

包年包月实例退费说明

按量计费实例转包年包月

购买超级节点

资源规格

超级节点价格说明

计费模式说明

包年包月超级节点退费说明

购买预留券

预留券概述

预留券定价

预留券使用说明

预留券管理

预留券退费说明

竞价模式说明

购买指南

购买指引

最近更新时间：2025-04-01 16:11:02

本文向您介绍如何购买腾讯云容器服务的集群及节点。

购买标准集群

容器服务（Tencent Kubernetes Engine，TKE）针对不同规格的托管集群，会收取相应的集群管理费用，以及用户实际使用的云资源费用。关于 TKE 的具体价格，请参见 [购买 TKE 标准集群](#)。

- 原生节点是由腾讯云 TKE 容器服务团队面向 Kubernetes 环境推出的全新节点类型，依托腾讯云千万核容器运维的技术沉淀，为用户提供原生化、高稳定、快响应的 K8s 节点管理能力，原生节点价格请参见 [购买原生节点](#)。
- 超级节点是腾讯云全新升级的节点产品形态，向用户提供可用区级别的、支持自定义规格的节点能力，使用超级节点类似于使用一台超大规格的 CVM，资源管理和资源扩缩容都更简单，超级节点价格请参见 [购买超级节点](#)。

购买 TKE 标准集群

计费概述

最近更新时间：2024-08-16 15:11:51

计费项

当您使用容器服务 TKE 时，产品费用由**集群管理费用**和**云产品资源费用**组成。

- 集群管理费用**
集群管理费用**仅针对托管集群收取**。TKE 托管集群提供了高可用、高性能、可伸缩、高稳定的全托管控制面，简化集群的搭建和扩容等操作，让您可专注于容器化应用的开发，而无需关心集群的管理以及维护，TKE 针对不同规格的托管集群，会收取相应的集群管理费用。具体收费细则，可参考 [集群管理费用](#)。
- 云产品资源费用**
您在使用 TKE 过程中创建的其他的云产品资源（CVM、CBS、CLB 等），将按照各自云产品的计费方式进行收费。具体收费细则，可参考 [云产品资源费用](#)。

集群管理费用

计费模式

容器服务计费方式主要为**按量计费（后付费）**。

计费项	计费模式	付款方式	计费单位
集群（个）	按量计费	购买时 冻结费用 ，每小时结算	元/小时

产品定价

⚠ 注意：

- 具体单价可能因地域调整变化，请以控制台实际展示价格为准。
- 选择规格前请仔细阅读 [购买说明](#)。
- 集群费用的计费周期自集群创建完成时开始，集群创建时间可在[容器服务控制台](#) > [集群基本信息](#)中查看。

集群规格	定价（元/小时）
L5	0.13
L20	0.40
L50	0.73

L100	1.22
L200	2.55
L500	5.11
L1000	9.38
L3000	15.60
L5000	28.04

云产品资源费用

您在使用 TKE 过程中创建的其他的云产品资源（CVM、CBS、CLB 等），将按照各自云产品的计费方式进行收费。具体收费细则，可参考各产品的收费说明。

云产品	计费说明
云服务器 CVM	云服务器计费模式
云硬盘 CBS	云硬盘价格总览
负载均衡 CLB	负载均衡计费说明

注意：

容器服务基于 Kubernetes 且为声明式服务，当不再需要容器服务创建的负载均衡 CLB、云硬盘 CBS 等 IaaS 服务资源时，请勿通过具体服务资源的管理界面删除，应该在容器服务控制台中删除相关服务资源，否则容器服务会重新创建被删除的服务资源，并继续扣除相关费用。例如您已在容器服务中创建负载均衡 CLB 服务资源，当在负载均衡控制台中删除该 CLB 实例时，容器服务会根据声明式 API 重新创建新的 CLB 实例。

购买说明

最近更新时间：2024-11-27 14:23:12

购买须知

❗ 说明：

因为不遵循本购买须知导致的服务不可用，相应服务不可用时间不属于服务不可用的计算范畴，更多请参考[《容器服务服务等级协议》](#)。

TKE 集群的可用性与集群 Pod、ConfigMap、CRD、Event 等资源数量、以及各类资源的 Get/List 读操作 QPS、Patch/Delete/Create/Update 等写操作 QPS 密切相关，应尽量避免对资源数量较多的集群发起类 List 操作，避免把 TKE 集群当数据库使用，写入过多的 ConfigMap/CRD/EndPoints 等，影响集群可用性。常见的类 List 操作如下（以 Pod 资源为例）：

- 带标签查询

```
kubectl get pod -l app=nginx
```

- 指定 namespace 查询

```
kubectl get pod -n default
```

- 查询整个集群的 pod 等

```
kubectl get pod --all-namespaces
```

- 通过 client-go 发起的 List 请求

```
k8sClient.CoreV1().Pods("").List(metav1.ListOptions{})
```

如您有类似查询集群全量资源的需求，建议使用 K8s 的 informer 机制通过本地 cache 查询。对于一些简单的场景，可以通过在 List 请求中增加 ResourceVersion 参数，在 kube-apiserver cache 中查询，如 `k8sClient.CoreV1().Pods("").List(metav1.ListOptions{ResourceVersion: "0"})`。需注意，即使从 kube-apiserver cache 查询，如果对大量资源频繁发起 List 请求，仍会对 kube-apiserver 内存造成较大压力，仅建议在请求频率较低时使用该方式。

推荐配置

请在选购集群时参考如下推荐配置，根据业务实际情况选择合适的集群规格，以免集群控制面组件负载过大导致集群不可用。

例如，如果您计划在一个集群中部署50个节点，但是计划部署2000个 Pod, 则应该选用最大管理节点规模为100（而非50）的集群规格。

❗ 说明：

- 节点指 Kubernetes Node，包含节点类型有普通节点、原生节点。**注意：节点规模计数不包含超级节点。**
- Pod 包括所有 Namespace 下，所有状态的 Pod，但不包括系统组件相关 Pod（cni-agent 等）。
- CRD 不统计 TKE 系统组件产生的 CRD（如 VPC-CNI 相关组件产生的 CRD）。
- **最大其他 K8s 资源数量**指除表格中单独列出的资源外，其他的 K8s 资源的最大数量。例如您购买了最大管理节点规模为L100的集群，为了保障集群的可用性，集群中的 ClusterRole、Service、Endpoint 等 K8s 资源的数量均不应该超过2500。
- 建议每种资源类型的所有对象总和不应超过800MiB，每个资源对象大小不超过100KB。

集群规格	最大管理节点数量	最大 Pod 数量（推荐）	最大 replicasets 数量	最大 ConfigMap 数量（推荐）	最大 CRD 数量 / 最大其他 K8s 资源数量（推荐）
L5	5	150	900	128	150
L20	20	600	3600	256	600
L50	50	1500	9000	512	1250
L100	100	3000	18000	1024	2500
L200	200	6000	36000	2048	5000
L500	500	15000	90000	4096	10000
L1000	1000	30000	180000	6144	20000
L3000	3000	90000	540000	8192	50000
L5000	5000	150000	900000	10240	100000

欠费说明

最近更新时间：2023-05-17 15:40:17

欠费说明

从您的**账户欠费**（账户余额被扣为负值时刻）起：

- 24小时内，容器服务托管集群可继续使用且继续扣费。
- 24小时后，账户内托管集群将全部进入**隔离中**状态且停止扣费。集群被隔离后，API Server 无法访问，Node 节点上的业务不受影响。

❗ 说明

如果您的账户于**收费缓冲期结束前**，即2022年4月1日上午10：00（北京时间）之前已经欠费，针对**2022年3月21日10:00（北京时间）前创建的容器服务 TKE 托管集群**，腾讯云会在缓冲期结束后（2022年4月1日上午10：00之后）隔离集群，缓冲期内您仍可以正常使用。

欠费处理

托管集群将进入**隔离中**状态后，随着隔离时间的增长，腾讯云对集群进行如下处理：

⚠ 注意

以下说明仅列出了托管集群层面的欠费处理，集群中的 **CVM**、**CLB** 等资源的欠费处理逻辑请参考对应产品的欠费说明。

隔离后时间	说明
≤ 15天	若您的账户充值至余额大于0，恢复计费，集群将自动恢复至运行中状态。
	若您的账户余额尚未充值到大于0，则集群保持为隔离中状态。
> 15天	若您的账户仍欠费，集群会被删除，集群删除后不可恢复。如集群中仍有节点，所有节点会被移出集群，TKE 不会对节点做任何其他操作。集群被删除时，我们将通过邮件及短信的方式通知到腾讯云账户的创建者以及所有协作者。

购买集群配额限制

最近更新时间：2025-06-13 09:23:42

容器服务的配额限制包括 TKE 配额限制，CVM 相关的配额限制以及托管集群的资源配额限制，详情如下：

TKE 配额限制

每个用户可购买的 TKE 配额默认如下，如果您需要更多的配额项数量，可通过 [在线咨询](#) 提出配额申请。

⚠ 注意：

2019年10月21日起，用户集群支持的最大节点配额统一调整为5000。

配额项	默认值	可查看入口	是否可提配额
单地域下集群	20	容器服务概览页右下方	是
单集群下节点	5000		
单地域下镜像命名空间	10		
单地域下镜像仓库	500		
单镜像下镜像版本	100		

CVM 配额限制

腾讯云容器服务所产生的云服务器需遵守云服务器的购买限制，详情请参见 [云服务器购买约束](#)。每个用户可购买的 CVM 配额默认如下，如果您需要更多的配额项数量，可通过 [配额申请工单](#) 提出配额申请。

配额项	默认值	可查看入口	是否可提配额
单可用区下包年包月服务器	150 台	CVM 概览页-各地域资源	是
单可用区下按量计费服务器	30 台或 60 台不等		

集群配置限制

ⓘ 说明：

集群配置限制集群规模，暂不支持修改。

配置项	地址范围	影响范围	可查看入口	是否可变更
VPC 网络-子网	用户自定义设置	该子网可添加的节点数	集群对应 VPC 子网列表页-可用 IP 数	- 不支持变更 - 可使用新子网
容器网段 CIDR	用户自定义设置	- 集群内节点上限 - 集群内 service 上限 - 每个节点 Pod 上限	集群基本信息页-容器网段	暂不支持变更

K8s 资源配额说明

❗ 说明：

下列配额自2022年4月30日开始生效，该配额无法被移除。配额不足时，您可通过升高集群规格提高各资源配额。

如有特殊场景需要调整配额，可 [提交工单](#) 联系我们，并说明需要调整的原因。

如需检查此配额，您可以执行以下命令：

```
kubectl get resourcequota tke-default-quota -o yaml
```

如需查看给定命名空间的 `tke-default-quota` 对象，请添加 `--namespace` 选项以指定命名空间。

❗ 说明：

- 其他 K8s 资源限制指集群中除了 Pod，Node，ConfigMap 的其他 K8s 资源均不能超过该数值。例如 L100 的集群，集群中的 ClusterRole，Service，Endpoint 等 K8s 资源的数量均不能超过 10000。
- CRD 数量限制指集群中所有 CRD 的总和不应该超过该限制。某类 CRD 的增多，会占用 CRD 配额，导致其他 CRD 能创建的数量变少。

集群规格	Pod 数量限制	ConfigMap 数量限制	CRD / 其他 K8s 资源数量限制
L5	600	256	1250
L20	1500	512	2500

L50	3000	1024	5000
L100	6000	2048	10000
L200	15000	4096	20000
L500	30000	6144	50000
L1000	90000	8192	100000
L3000	150000	10240	150000
L5000	200000	20480	200000

集群规格升配

1. 集群开启自动升配功能后，当集群控制面组件负载过高或者节点规模超过当前规格管理上限时，会自动升至更高规格。满足以下任意一项条件将触发集群自动升配（监控来源平于 metrics-server 获取的实时数据）：
 - 1.1 节点数（指 Kubernetes Node，包含节点类型有普通节点、原生节点）大于当前集群等级规格就触发升配。
 - 1.2 过去1天内，kube-apiserver OOM 重启次数大于2，或 kube-controller-manager/kube-scheduler OOM 重启次数大于3将触发升配。
 - 1.3 在过去2小时内，kube-apiserver OOM 重启次数大于1，且实时持续10分钟内负载大于70%将触发升配。
 - 1.4 kube-apiserver/kube-controller-manager/kube-scheduler 任意一组件连续10分钟以上，CPU 和内存实时负载持续95%将触发升配。
2. 升配过程中控制面（Master 节点）组件需要进行滚动更新，可能会有短暂中断，建议升配期间停止其他操作（如创建工作负载等）。
3. 集群升配后不会自动降配，您需手动下调集群规格。

针对命名空间分配配额

默认情况下，任意一个命名空间余量都相同（余量 = 当前集群等级的配额 - 整个集群已使用的量）。如果您在一个命名空间创建了资源，会导致余量减少，即其他命名空间的可使用量也会在一定时间后相应减少。

如果您有自定义分配比例的需求，可以在 `kube-system` 下创建一个 `tke-quota-config` 的 ConfigMap，指定余量在各个命名空间的分配比例。

例如：下面的例子代表在 `default` 的命名空间分配50%的余量，在 `kube-system` 的命令空间分配40%的余量，其余命名空间分配到10%的余量。如果设置的百分比之和超过100%，则 TKE 认为比例分配无效，会采用默认的分配策略。

```
apiVersion: v1
data:
```

```
default: "50"  
kube-system: "40"  
kind: ConfigMap  
metadata:  
  name: tke-quota-config  
  namespace: kube-system
```

集群新增资源所属项目说明

最近更新时间：2023-09-04 14:39:01

总述

如需要通过分项目进行财务核算等，请先阅读以下内容：

1. 集群无项目属性，集群内云服务器、负载均衡器等资源有项目属性。
2. 集群新增资源所属项目：仅将新增到该集群下的资源归属到该项目下。

建议

1. 建议集群内的所有资源在同一个项目。
2. 如若需要集群内云服务器分布在不同的项目，请自行前往云服务器控制台迁移项目。操作详情见 [调整项目配置](#)。

购买原生节点

原生节点产品定价

最近更新时间：2025-10-29 17:04:11

产品定价

说明：
原生节点为容器服务 TKE 单独计费云产品，其定价与云服务器 CVM 存在差异，不同规格的计费信息以控制台展示为准。

TKE 原生节点定价分为资源费和增值服务费。其中原生节点的资源费定价和普通节点定价基本一致，具体包含 CPU、内存、系统盘和 GPU。增值费定价区分机型种类，按照CPU、内存和 GPU 的总资源费的一定比例收取，默认溢价比例在 5%-20%，具体可能因商务优惠有所调整。

机型种类	增值费种类	默认溢价比例
CPU 虚拟机	按照 CPU 和内存资源增值	20%
GPU 虚拟机	按照 CPU、内存和 GPU 资源增值	10%
裸金属	按照整体资源增值	5%

计费公式：费用 = 原生节点实例资源的资源费和增值服务费单价 × 运行时间

计费项	描述
实例类型	当前原生节点支持的实例类型，详细介绍可参考 CVM 实例 ： - 标准型：S2、S4、S5、SA2、SA3、S6、SA5、SA4、S7、S8、SA9 - 计算型：C3、C4、C5、C6 - 内存型：M3、MA3、M5、M6、MA4、MA2、MA5、MA8 - 高 IO 型：IT5、ITA5 - 蜂驰型：BF1 - GPU 型：GN7、GNV4、PNV4、GN10X、GN10Xp、GT4、PNV5、PNV5b、PNV5i、GC49d - 裸金属：BMS5、HCCPNV4sne、HCCPNV4sn、HCCPNV5、HCCPNV5v、HCCPNV4ane
资源规格	当前原生节点支持的资源及规格，以控制台实际展示为主： CPU：最低支持2核，实际可选规格以控制台展示为准。

- 内存：最低支持2GB，实际可选规格以控制台展示为准。
- 系统盘：支持高性能云盘、SSD 云盘、通用型 SSD、增强型 SSD，系统盘大小为 20GB–2048GB。
- GPU：仅 GPU 型原生节点实例含 GPU 资源，仅支持整卡 GPU 规格。

计费模式

TKE 提供两种不同的原生节点实例购买模式：包年包月、按量计费，分别适用于不同场景下的用户需求，详情如下：

计费模式	包年包月	按量计费
付款方式	预付费	后付费（购买时冻结费用，每小时结算）
计费单位	元/月	元/秒
单价	单价较低	单价较高
最少使用时长	至少一个月	按秒计费，按小时结算，随时购买随时释放
使用场景	适用于节点资源需求量长期稳定的成熟业务。	适用于转码、大数据、电商抢购等周期性计算场景或开启自动伸缩（HPA）的潮汐型在线服务场景。

包年包月

计费公式：支付金额 = 申请的原生节点实例数 × 原生节点实例资源配置单价（包月） × 包月时长

包年包月是 TKE 原生节点实例的一种预付费模式，提前一次性支付一个月或多月甚至多年的费用，价格相较于按量计费模式更低廉。在购买包月原生节点实例前，可先查看云账户中的可用余额，如果余额小于系统将扣除的金额，请先充值后再购买。

按量计费

计费公式：支付金额 = 申请的原生节点实例数 × 原生节点实例资源配置单价（按量） × 按量时长

开通按量计费原生节点实例时，会预先冻结该实例一个小时的资源配置单价费用（含 CPU、内存、GPU 和系统盘），并在每个整点（北京时间）进行一次结算，根据您在上一个小时的实际使用时长进行扣费。购买时实例单价按小时呈现，结算时按实际使用秒数计价，费用四舍五入到小数点后2位。计费的起点以每个原生节点实例资源购买成功的时间点为准，终点以每个节点销毁（删除）操作完成的时间点为准。实例销毁时，系统将会对冻结的费用进行解冻。

相关文档

- [欠费说明](#)
- [包年包月实例退费说明](#)

欠费说明

最近更新时间：2023-12-20 19:19:01

当您的账号发生欠费时，系统将为您推送欠费提醒，请在收到欠费通知后，及时前往控制台 [充值中心](#) 进行充值，以免影响您的业务。本文向您介绍原生节点欠费的相关说明。

预警说明

预警类型	说明
到期预警	集群中的资源会在到期前第7天内，将向用户推送到期预警消息。预警消息将通过邮件及短信的方式通知到腾讯云账户的创建者以及全局资源协作者、财务协作者。
欠费预警	集群中的资源到期当天及以后，将向用户推送欠费预警消息。预警消息将通过邮件及短信的方式通知到腾讯云账户的创建者以及所有协作者。

回收机制

说明：

- 停服：指资源已经对用户停止服务，用户不可继续使用；但数据仍未清除，用户可对资源进行恢复。
- 释放：指资源的数据被删除，无法恢复。

包年包月回收机制

- 原生节点资源到期前7天内，系统会给您发送续费提醒通知。
- 账户余额充足的情况下，若您已设置自动续费，设备在到期当日会执行自动续费。
- 若您的原生节点实例在到期前（包括到期当天）未进行续费，系统将在到期**24小时内**对其作停服并释放处理（节点资源将被销毁且数据不可恢复）。
- 若您在欠费后的24小时内完成手动续费，续费周期的起始时间为上一周期的到期日。

按量计费回收机制

- 从您的账户余额被扣为负值时刻起，原生节点实例在**2小时内**可继续使用且继续扣费，2小时后若您的账户余额未充值到>0，系统对该实例作停服并释放处理（节点资源将被销毁且数据不可恢复）。
- 若您在账户欠费后的2小时内完成账户余额已充值到>0，原生节点实例将正常按照按量计费规则进行扣费。

包年包月实例退费说明

最近更新时间：2023-08-25 14:26:21

为了方便您使用 TKE 原生节点模式，如果您在购买包年包月原生节点后有任何不满意，我们支持自助退货退款。每个主体可享受1次1台原生节点实例五天无理由退还，您支付的有效金额将返还至您的腾讯云账户。除此之外，您还可享受普通自助退还，扣除您已使用的费用，将按购买支付使用的现金和赠送金支付比例退还至您的腾讯云账户的方式返还到您的账户。以上操作均可在容器服务 TKE 控制台自助进行，具体操作详情请参见 [删除原生节点](#)。

五天无理由自助退还

TKE 原生节点遵守腾讯云 [云服务退货说明](#)，如果您在购买原生节点后有任何不满意，我们支持五天内无理由自助退还，具体规则如下：

- 每个主体下，包年包月预付费原生节点自**新购之日起五天之内（含五天）**，默认享受1台原生节点五天无理由退还。
- 符合五天无理由退还场景的订单，退款金额为购买时花费的全部消耗金额，包括现金账户金额、收益转入账户金额以及赠送账户金额。
- 如需退还参与活动购买的原生节点，且退款规则与活动规则有冲突，则以活动规则为准。
- 如出现疑似异常/恶意退货，腾讯云有权拒绝您的退货申请。

❗ 说明

- 抵扣的代金券不予以退还。
- 退还金额将全部退还到腾讯云账号余额。

普通自助退还

普通自助退还将按购买支付使用的现金和赠送金支付比例退还至您的腾讯云账户。

普通自助退还限制

不满足五天无理由退货退款规则的退货订单，退款策略如下：每个账户每年支持**199台**包年包月原生节点控制台自助退还，普通自助退还将按购买支付使用的现金和赠送金支付比例退还至您的腾讯云账户。

普通自助退还规则

退款金额 = 当前有效订单金额 + 未开始订单金额 - 资源已使用价值

- 当前有效订单金额**：指生效中订单的付款金额，不包含折扣和代金券。
- 未开始订单金额**：将来生效订单的付款金额，不包含代金券。
- 资源已使用价值**：用户发起退费的当天，原生节点使用已满整月的，按已使用整月的包年包月价格及对应折扣进行扣除；不满整月的，按已使用时长（单位精确至秒）的按量计费价格进行扣除。

说明：

退款金额 ≤ 0 ，按0计算并清退资源。

原生节点包年包月实例自助退还说明

- 包年包月实例自助退还后，挂载在实例上**系统盘均随实例一起回收**，保存在这些存储上的数据将丢失。挂载在该实例上的数据盘会继续保留，数据不受影响。
- 包年包月实例自助退还后，实例上运行的服务彻底中断，相关资源彻底释放，**不支持续费恢复**。
- 若包年包月实例在创建时绑定了弹性公网 IP，则**实例销毁后 IP 资源随实例一起释放**。
- 如出现疑似异常/恶意退货，腾讯云有权拒绝您的退货申请。

原生节点自助退还计费规则示例

说明

以下价格仅供示例，非官网实际价格。具体单价可能因活动或策略等调整变化，请以控制台实际单价为准。

案例背景

广州二区标准型 S2 1核1GB，10GB 高性能云硬盘（系统盘），不含带宽，85元/月，使用200元代金券，购买1年，包年享7折。

折扣金额为 $85 \times 12 \times 0.7 = 714$ （元）

支付金额为 $714 - 200 = 514$ （元）

五天无理由退还场景

购买五天内发现不满意，想要退还，且为该主体首次退还。

退费金额 = 真实支付价514元

普通自助退还场景

案例1

购买五天内，且为该账户本年度首次退还。

退还现金金额 = 实际支付金额514元

案例2

购买五天内，且不是该账户本年度首次退还，总使用时长48小时。

退还现金金额 = $514 - 48 \times 0.29 = 500.08$ （元）

说明：

0.29元为同样配置按量计费实例的单价。

案例3

购买五天内，且不是该账户本年度首次退还，已使用时长48小时；其中又续费1年，包年官网7折，续费实际支付金额714元。

退还现金金额 = $514 - 48 \times 0.29 + 714 = 1214.08$ （元）

 说明：

48×0.29 为生效订单退款金额，714元为未开始订单金额。

按量计费实例转包年包月

最近更新时间：2025-04-08 15:04:12

操作场景

腾讯云容器服务 TKE 开放了原生节点按量计费转换成包年包月功能，您可以在原生节点控制台及云 API 进行转换操作。本文主要介绍在控制台进行按量计费实例转换成包年包月实例的功能操作。

转换规则

在原生节点控制台提供计费模式转换功能，具体规则如下：

- 转换以节点池维度进行操作，节点池内的原生节点将一次性完成按量到包月模式的切换。
- 按量计费节点转换包年包月时会产生一个续费订单，必须完成该订单的支付流程，计费方式的变更才能生效。若未支付或未成功支付，该订单可在您的 [订单管理](#) 页面查看和处理。
- 计费模式由按量计费转换为包年包月的原生节点不支持五天内无理由退还。
- 计费方式转换成功及支付成功后，节点会即刻按照包年包月计费，新包年包月节点的起始时间为转换成功时间。
- 在未成功支付前，不可对该节点池进行重复计费模式转换。
- 按量计费转包年包月功能支持同步转换节点和磁盘（包括系统盘和数据盘）的计费方式。

⚠ 注意：

- 非按量计费实例不支持转换。
- 按量计费转包年包月为不可逆操作。

操作限制

- 所有节点状态为正常时才能进行计费模式的转换操作。
- 节点池在转换前需关闭弹性伸缩功能才能进行计费模式的转换操作。

操作步骤

1. 登录 [容器服务控制台](#)，选择左侧导航栏中的**集群**。
2. 在集群列表页中，单击集群 ID，进入该集群详情页。
3. 选择**节点管理**，在**节点池**中选择目标节点池，进入节点池的详情页。
4. 在**节点启动配置信息**模块中修改计费模式。如下图所示：

节点启动配置信息

操作系统 TencentOS Server 3.1 (TK4)

计费模式 按量计费

支持子网 subnet

安全组 sg-

关联ssh密钥 skey-

机型 ① SA2.MEDIUM2(主)
[配置多机型](#)

系统盘 高性能云硬盘 50GB

数据盘 -

弹性公网 IP 不访问外网

主机名 ① 自动命名

5. 在弹出的按量计费转包年包月窗口中，根据实际需求，设置续费时长以及是否自动续费。如下图所示：

按量计费转包年包月

重要声明

- 转换以节点池维度进行操作，节点池内的原生节点将一次性完成按量到包月模式的切换。
- 计费方式转换成功后，节点会即刻按照包年包月计费，新包年包月节点的起始时间为转换成功时间。
- 按量计费转包年包月功能支持同步转换节点和磁盘（包括系统盘和数据盘）的计费方式。
- 按量计费转包年包月为不可逆操作，请谨慎操作。

计费模式

包年包月

购买时长

1个月

2

3

6

1年

2年

3年

其他时长

自动续费

☒ 账户余额足够时，设备到期后按月自动续费

节点列表

ID/节点名	状态	配置	IP	创建时间	预计价格
np- tke-np- worker	健康	SA2.MEDIUM2 2核, 2GB, -Mbps 系统盘: 50GB	-		

存储

同时转为包年包月模式

费用

服务条款

☒ 我已阅读并同意协议

确定

取消

- 续费时长：选择转换为包年包月后的购买时长。
- 自动续费：根据您的需求，选择自动续费。
- 存储：节点系统盘和数据盘将一同转换为包年包月模式。

6. 勾选已阅读并同意按量计费转包年包月规则，单击确定。

常见问题

计费模式转换过程中，是否会影响业务的正常运行？

不会，计费模式转换仅涉及订单流程，不会对节点进行资源变更或操作。

为什么计费模式转换按钮不可点击？

为保障转换的一致性，转换操作需要满足上述操作限制。请确保所有节点的状态为正常，并关闭节点池的弹性伸缩功能。只有在满足这些限制条件后，计费模式转换按钮才会变为可点击状态。

购买超级节点 资源规格

最近更新时间：2025-06-05 18:30:12

概述

使用 TKE Serverless，您无需关心集群节点，但为了合理分配资源和准确核算，您需在部署工作负载时指定 Pod 申请的资源规格。腾讯云会根据您指定的规格为工作负载分配计算资源，也会根据此规格进行费用核算。当您使用 Kubernetes API 或 Kubectl 创建 TKE Serverless 工作负载时，可以通过 Annotation 指定资源规格。如果不指定，TKE Serverless 会根据工作负载设置的容器 Request、Limit 等参数进行计算。详情请参考 [指定资源规格](#)。

⚠ 注意：

- 资源规格是 Pod 内容器可使用的最大资源量（含操作系统、容器运行时、日志采集等系统组件资源占用）。
- 目前支持分配的 CPU、GPU 规格见下表，分配时请勿超出所支持的规格。
- Pod 内所有 Container 设置的 Request 之和不可大于最大 Pod 规格。
- Pod 内任何 Container 设置的 Limit 不可大于最大 Pod 规格。

CPU 规格

TKE Serverless 在所有支持 CPU 资源类型的地域提供以下 CPU Pod 规格。TKE Serverless 提供一系列 CPU 选项，不同的 CPU 大小会对应不同的内存选择区间，请在创建工作负载时根据您的实际需求选择最合适规格，并进行资源分配。

Intel

CPU/核	内存区间/GiB	内存区间粒度/GiB
0.25	0.5、1、2	—
0.5	0.5、1、2、3、4	—
1	1 – 8	1
2	2、4 – 16	1
4	8 – 32	1
8	16 – 64	1

12	24 – 96	1
16	16 – 128	1
32	64、128、256	–
64	128、192、256、512	–

星星海 AMD

基于腾讯云自研星星海服务器，提供可靠、安全、稳定的高性能。详情请参见 [云服务器标准型 SA2 介绍](#)。

CPU/核	内存区间/GiB	内存区间粒度/GiB
1	1 – 4	1
2	2 – 8	1
4	4 – 16,32	1
8	8 – 32,64	1
16	16 – 64,128	1
32	64、128、256	–
64	128、192、256	–

GPU 规格

TKE Serverless 提供以下型号 GPU Pod 规格，不同的 GPU 卡型号和大小会对应不同的 CPU、内存选项，请在创建工作负载时根据您的实际需求选择最合适规格，并进行资源分配。

⚠ 注意：

如果您需要通过 yaml 来创建、管理和使用 GPU 的工作负载，请参考如下 Annotation 信息和 [Annotation 说明](#)。

1/4 NVIDIA T4 和 1/2 NVIDIA T4 的规格有更新，详情请参见 [关于 Serverless 容器服务 T4 vGPU Pod 规格变化的公告](#)。

GPU 型号	GPU/卡	GPU 显存	支持驱动	CUDA 版本	CPU/核	内存/GiB
--------	-------	--------	------	---------	-------	--------

NVIDIA Tesla V100 – 1 颗	1	1 * 32GB	470	11.4	8	40
NVIDIA Tesla V100 – 2 颗	2	2 * 32GB	470	11.4	18	80
NVIDIA Tesla V100 – 4 颗	4	4 * 32GB	470	11.4	36	160
NVIDIA Tesla V100 – 8 颗	8	8 * 32GB	470	11.4	72	320
1/4 NVIDIA T4 – 1/4 颗	1	4GB vGPU	450	11.0	4	16
1/2 NVIDIA T4 – 1/2 颗	1	8GB vGPU	450	11.0	8	32
NVIDIA T4 – 1颗	1	1 * 16GB	470	11.4	8	32

NVIDIA T4 – 1颗	1	1 * 16GB	470	11.4	20	80	
NVIDIA T4 – 1颗	1	1 * 16GB	470	11.4	32	128	
NVIDIA T4 – 2颗	2	2 * 16GB	470	11.4	40	160	
NVIDIA T4 – 4颗	4	4 * 16GB	470	11.4	80	320	
NVIDIA A10 – PNV4 – 1颗	1	1 * 24GB	470	11.4	28	116	
NVIDIA A10 – PNV4 – 2颗	2	2 * 24GB	470	11.4	56	232	

NVIDIA A10 – PNV4 – 4颗	4	4 * 24GB	470	11.4	112	466
NVIDIA A10 – PNV4 – 8颗	8	8 * 24GB	470	11.4	224	932
NVIDIA A10 – GNV4 – 1颗	1	24GB	470	11.4	12	44
NVIDIA A10 – GNV4v – 1/4颗	1	6GB vGPU	470	11.4	6	24
NVIDIA A10 – GNV4v – 1/2颗	1	12GB vGPU	470	11.4	14	58
NVIDIA A10 – GNV4v – 1颗	1	1 * 24GB	470	11.4	28	116

--	--	--	--	--	--	--	--

超级节点价格说明

计费模式说明

最近更新时间：2025-07-09 18:44:52

超级节点的新版产品定价于2023年7月1日00:00开始生效。

计费模式说明

TKE Serverless 以超级节点维度承载资源，提供以下几种计费模式：按量计费、包年包月、预留券和竞价模式。下表列出了这些计费模式的区别：

计费模式	超级节点包年包月模式	超级节点按量计费模式	超级节点预留券模式	竞价模式
付款方式	预付费	购买时 冻结费用 ，每小时结算	预付费	购买时冻结费用，每小时结算
计费单位	元/月	元/秒	元/月	元/秒
单价	单价较低	单价较高	单价较低	单价最低，固定折扣比例，同规格按量计费模式的20%（2折）
最少使用时长	至少使用一个月	按秒计费，按小时结算，随时购买随时释放	至少使用一个月	按秒计费，按小时结算，随时购买随时释放
最小购买规格	10C20G	无限制	与 Pod 规格一致	无限制
使用场景	适用于算力需求量长期稳定的成熟在线小规格（8C及以下）业务	适用于算力需求量瞬间大幅波动的场景	适用于算力需求量长期稳定、需要保障资源的业务，推荐单个 Pod 超过8C时使用预留券模式	适用于短时长突发、周期任务，可切分的计算任务及无状态或者支持断点续传能力的计算任务。
计费模式维度	超级节点	Pod 维度	一张预留券对应抵扣一个 Pod	Pod 维度

按量计费定价

在按量计费场景下，服务会根据用户选择的容器资源类型计算相关费用。计费公式为：**费用 = 相关计费项配置 × 资源单位时间价格 × 运行时间（精确到秒）**。

目前，超级节点支持的计费项配置详情请参见 [资源规格](#)。如果您想了解如何选择计费项配置，请参见 [指定资源规格](#)。

Intel Pod 按量计费

地域	计费项			
	CPU（元/核/小时）	内存（元/GiB/小时）	高性能云硬盘（元/GB/小时） 超过20GB部分收费	SSD 云硬盘（元/GB/小时） 超过20GB部分收费
上海、北京、广州、南京、重庆、成都、武汉、长沙、郑州、西安、沈阳、福州、合肥、济南、杭州、石家庄、上海金融、深圳金融、北京金融、上海自动驾驶云、日本、泰国、美国硅谷、弗吉尼亚、圣保罗	0.12	0.05	0.0009	0.0025
中国香港、中国台北、新加坡	0.12	0.06	0.0005	0.0024
首尔、法兰克福	0.126	0.063	0.0005	0.0024
雅加达	0.14	0.07	0.0005	0.0024

AMD Pod 按量计费

地域	计费项			
	CPU（元/核/小时）	内存（元/GiB/小时）	高性能云硬盘（元/GB/小时） 超过20GB部分收费	SSD 云硬盘（元/GB/小时） 超过20GB部分收费
上海、北京	0.065	0.032	0.0009	0.0025
重庆、成都、南京	0.055	0.032	0.0005	0.0024

广州、武汉、长沙、郑州、西安、沈阳、福州、合肥、济南、杭州、石家庄	0.061	0.032	0.0005	0.0024
上海金融、深圳金融、北京金融、上海自动驾驶云	0.099	0.049	0.0005	0.0024
中国香港、中国台北、日本、新加坡、泰国、弗吉尼亚、圣保罗、法兰克福	0.087	0.044	0.0005	0.0024
美国硅谷	0.072	0.036	0.0005	0.0024

GPU V100 Pod 按量计费

计量	价格（秒）	价格（小时）	备注
GPU	每卡 0.003193 元	每卡 11.5 元	同样适用于 vGPU，vGPU 以虚拟化系数（0.25、0.5）乘以此价格
CPU	每核 0.00005778 元	每核 0.208 元	—
内存	每 GiB 0.00003389 元	每 GiB 0.122 元	—

GPU T4 Pod 按量计费

计量	价格（秒）	价格（小时）	备注
GPU	每卡 0.001447 元	每卡 5.21 元	同样适用于 vGPU，vGPU 以虚拟化系数（0.25、0.5）乘以此价格
CPU	每核 0.00002411 元	每核 0.0868 元	—
内存	每 GiB 0.00002411 元	每 GiB 0.0868 元	—

	元		
--	---	--	--

GPU A10 Pod 按量计费

GNV4v

计量	价格（秒）	价格（小时）	备注
GPU	每卡 0.00170718 元	每卡 6.15 元	同样适用于 vGPU，vGPU 以虚拟化系数（0.25、0.5）乘以此价格
CPU	每核 0.0000185 元	每核 0.0666 元	—
内存	每 GiB 0.00000926 元	每 GiB 0.0333 元	—
license	每个 0.00012733 元	每个 0.458 元	—

GNV4

计量	价格（秒）	价格（小时）	备注
GPU	每卡 0.00170718 元	每卡 6.15 元	同样适用于 vGPU，vGPU 以虚拟化系数（0.25、0.5）乘以此价格
CPU	每核 0.00002894 元	每核 0.1042 元	—
内存	每 GiB 0.00001446 元	每 GiB 0.0521 元	—
license	每个 0.00012733 元	每个 0.458 元	—

PNV4

计量	价格（秒）	价格（小时）	备注
----	-------	--------	----

GPU	每卡 0.00170718 元	每卡 6.15 元	同样适用于 vGPU，vGPU 以虚拟化系数（0.25、0.5）乘以此价格
CPU	每核 0.0000185 元	每核 0.0666 元	—
内存	每 GiB 0.00000926 元	每 GiB 0.0333 元	—

按量计费 Pod 运行时间说明

按量计费 Pod 运行时间是指从 Pod 资源分配完成，准备拉取首个容器的镜像开始计算，到 Pod 运行终止结束的时间段。这段时间以秒为单位计算，作为 Pod 的计费时间。

包年包月定价

地域	可用区	计费项	
		CPU（元/核/月）	内存（元/GiB/月）
上海	上海四区、上海五区、上海七区	38	20
北京	北京七区	38	20
上海、北京、广州、南京	—	36	18
中国香港、中国台北	—	48	24
新加坡、法兰克福、弗吉尼亚、美国硅谷	—	46	23

说明：

- 未指定可用区通常指除了指定的可用区之外的其他可用区。
- 包年包月模式暂不支持 GPU 类型。
- 包年包月模式下支持购买的最小规格为10C20G。
- 包年包月模式仅抵扣8C及以内规格的 Pod。
- 包年包月模式支持调度 Intel 和 AMD 类型的 Pod，如无明确指定将优先调度 intel 机型。
- 包年包月退费说明请参见 [包年包月超级节点退费说明](#)。

计费示例

包年包月

示例 1

包年包月购买广州地域总规格为 200C400G 的超级节点 1 个月，总费用为 $36 \times 200 + 18 \times 400 = 14400$ 元。该节点可累计调度算力总和为 200C400G 的符合包年包月超级节点规则的 Pod 资源。

按量计费

示例 1

某 Deployment 的 Pod 资源规格为 2 核 4GB，指定 Intel，调度到广州六区，Pod 数固定为 2。假设该 Deployment 从启动到终止，共耗时 5 分钟，即要计算 300 秒的费用。则该 Deployment 的运行费用为 $2 \times (2 \times 0.00003334 + 4 \times 0.00001389) \times 300 = 0.073344$ 元，匹配优惠折扣后，最终计费为 $0.073344 \times 0.675 = 0.0495072$ 元。

示例 2

某 CronJob 需要每次启动 10 个 Pod，每个 Pod 资源规格为 4 核 8GB，指定 Intel，调度到上海五区，运行 10 分钟后结束。假设该 CronJob 每天执行 2 次，使用 TKE Serverless 服务托管该任务。则该任务每天收费为 $2 \times 10 \times (4 \times 0.00003334 + 8 \times 0.00001389) \times 600 = 2.93376$ 元，匹配优惠折扣后，最终计费为 2.93376 元。

预留券定价

TKE Serverless 服务预留券是一种预付费产品。支持按月购买，购买后可以对同地域、同资源规格类型、同资源规格的容器（Serverless 集群 Pod 或容器实例）进行计算费用抵扣。详情请参见 [预留券概述](#)，定价请参见 [预留券定价](#)。

竞价模式

TKE Serverless 竞价模式是一种低成本资源购置模式，其核心特点为价格低于按量付费实例，在一些场景可以极大降低运行容器的成本，但资源可能会被腾讯云中断回收。您在使用竞价模式运行时仅需支付较少的费用，并运行到容器资源被回收为止。详情请参见 [竞价模式说明](#)。

包年包月超级节点退费说明

最近更新时间：2023-05-17 15:40:25

为了方便您使用超级节点，如果您在购买包年包月超级节点后有任何不满意，我们支持自助退货退款。每个主体可享受1次1个超级节点五天无理由退还，您支付的有效金额将返还至您的腾讯云账户。除此之外，您还可享受普通自助退还，扣除您已使用的费用，将按购买支付使用的现金和赠送金支付比例退还至您的腾讯云账户。您可在容器服务控制台自助进行。

五天无理由自助退还

TKE 超级节点遵守 [云服务退货说明](#)，若您在购买超级节点后有任何不满意，我们支持五天内无理由自助退还，具体策略如下：

- 每个主体下，包年包月预付费超级节点自**新购之日起五天之内（含五天）**，默认享受1个超级节点五天无理由退还。
- 符合五天无理由退还场景的订单，退款金额为购买时花费的全部消耗金额，包括现金账户金额、收益转入账户金额以及赠送账户金额。
- 如需退还参与活动购买的超级节点，且退款规则与活动规则有冲突，则以活动规则为准。
- 具体退款规则请参见 [五天内无理由全额退款](#)。
- 如出现疑似异常/恶意退货，腾讯云有权拒绝您的退货申请。

普通自助退还

不满足五天无理由退货退款规则的退货订单，退款策略如下：

普通自助退还将按购买支付使用的现金和赠送金支付比例退还至您的腾讯云账户。

普通自助退还限制

每个账户最多支持**199**个超级节点控制台自助退还。

普通自助退还规则

退款金额 = 当前有效订单金额 + 未开始订单金额 - 资源已使用价值

- 当前有效订单金额：指生效中订单的付款金额，不包含折扣和代金券。
- 未开始订单金额：将来生效订单的付款金额，不包含代金券。
- 资源已使用价值按照如下策略计算：已使用部分 = 已使用超级节点费用（用户发起退费的当天，超级节点使用已满整月的，按已使用整月的包年包月价格及对应的官网折扣进行扣除；不满整月的，按已使用时长内调度 Pod 的按量计费价格进行扣除。）
- 退款金额 ≤ 0 ，按0计算并清退资源。
- 抵扣的代金券不予以退还。
- 退还金额将按购买支付使用的现金和赠送金支付比例退还至用户的腾讯云账户。

超级节点自助退还后，超级节点的状态从生效中变为销毁中或已销毁，此时，不再产生该超级节点相关的费用，包月超级节点上也不可再调度 Pod。

包月超级节点自助退还后，对用户实际运行的 Pod 不产生影响，而是转换为按量计费模式。

升配费用规则

超级节点升配遵循按天补差价，升配费用 = 按月升配差价 × 升配月数 × 适用折扣。

- 按月升配差价：新老配置原价按月的单价差。
- 升配月数 = 升配天数 / (365/12)
- 升配的费用按天计算：升配天数 = 资源到期时间 - 当前时间
- 适用折扣：根据升配月数匹配官网适用折扣，其中折扣为官网生效的折扣。
- 升配不影响资源到期时间。
- 升配可以使用代金券抵扣费用。

降配费用规则

超级节点降配遵循按天退还差价，降配费用 = 剩余整月部分退还费用 + 剩余不足整月部分退还费用

- 剩余整月部分退还费用 = 按月降配差价 × 降配月数 × 适用折扣
- 剩余不足整月部分退还费用 = 按月降配差价 - 降配资源数 × 已使用时间 × 按量计费刊例价
- 升配不影响资源到期时间。
- 退款金额 ≤ 0 ，按0计算并清退资源。

购买预留券

预留券概述

最近更新时间：2023-05-31 17:56:18

TKE Serverless 服务预留券是一种预付费产品。支持按月购买，购买后可以同时对同地域、同资源规格类型、同资源规格的容器（Serverless 集群 Pod 或容器实例）进行计算费用抵扣。

预留券特点

- 单位时间价格便宜**
预留券仅支持按月购买，在生效期内，每小时/秒的单位时间价格相比 TKE Serverless 服务按量计费模式更便宜，具体可参考 [产品定价](#)。
- 生命周期独立**
预留券和容器的生命周期互相独立，您可以先创建容器，再根据容器的资源规格购买匹配的预留券。也可以提前购买预留券，再创建与其资源规格一致的容器。
- 即时抵扣**
成功购买预留券后，可在本账单周期（通常是一小时）内抵扣与之匹配的容器的费用。抵扣原则请参考 [预留券使用说明](#)。
- 预保留资源**
当指定可用区/超级节点时，可选择保留指定配置的资源，保证在此可用区/节点下创建出资源。
- 抵扣计算费用**
抵扣计算费用，例如 CPU、内存费用，不包括网络、存储等其他资源或服务费用。

相关概念

资源规格类型

由于弹性容器本身不同算力资源（Intel CPU、AMD CPU、GPU）的价格不同，预留券也分为 CPU 通用、AMD 专用以及其他资源专用等类型。其中：

预留券类型	使用场景
CPU 通用类型	可以抵扣包括 Intel、AMD 型号 CPU Pod 产生的费用，并且会优先抵扣 Intel 型号。该类型适合您在混合使用 Intel、AMD 资源的时候采购。
AMD 专用类型	仅可以抵扣 AMD 型号 CPU Pod 产生的费用，但单位时长的费用更低。该类型适合您仅使用 AMD 资源的时候采购。

资源规格

资源规格是指预留券的配置，例如8核 CPU、16GiB内存等。为了保证预留券能适配抵扣容器费用，TKE Serverless 服务仅支持购买弹性容器支持规格的预留券，支持规格详情请参见 [资源规格](#)。

抵扣范围

- 地域级：预留券只指定地域，不预留资源，可抵扣该地域下的全量 Pod。
- 可用区级：预留券勾选保留资源，并指定可用区，只可抵扣该可用区下的 Pod。
- 节点级：预留券勾选保留资源，并选择绑定到某个超级节点上，只可抵扣该节点上的 Pod。

计费公式

预留券支持预付费按月购买，将根据用户选择的预留券类型、资源规格以及购买时长，计算相关费用。计费公式如下：

费用 = 预留券配置 × 资源单位时间价格 × 购买时长

其中，

- 预留券支持配置详情请参考 [资源规格](#)。
- 资源单位时间具体价格请参考 [产品定价](#)。
- 购买时长单位为月，至少购买一个月。

适用场景

- 需要长期运行，负载稳定的业务。
- 需要长期运行，但需要弹性的业务，建议将最低保留容器负载转为预留券模式，动态扩缩的容器负载依然使用传统按量计费模式。
- 需要长期持有资源，保证资源供给的情况，建议选择可用区并预留资源。

预留券定价

最近更新时间：2025-11-13 18:20:41

Common 通用预留券

CPU 通用类型预留券可以优先抵扣 Intel CPU 容器的费用，详情请参见 [预留券使用说明](#)。

说明：
未指定可用区通常指除了指定的可用区之外的其他可用区。

地域	可用区	计费项	
		CPU（元/核/月）	内存（元/GiB/月）
上海、北京、广州、成都、重庆、南京	—	49	22
武汉、长沙、郑州、西安、沈阳、福州、合肥、济南、杭州、石家庄	—	49	24
上海金融、深圳金融、北京金融、上海自动驾驶云	—	72	36
中国香港	—	58	30
中国台北、首尔、东京	—	66	33
新加坡	—	56	33
泰国、美国硅谷、弗吉尼亚、圣保罗、法兰克福	—	50	25

AMD 预留券

AMD 专用类型预留券可以抵扣 AMD CPU 容器的费用，详情请参见 [预留券使用说明](#)。

说明：
未指定可用区通常指除了指定的可用区之外的其他可用区。

地域	可用区	计费项	
		CPU（元/核/月）	内存（元/GiB/月）

上海、北京	—	36	17
广州、成都、重庆、武汉、长沙、郑州、西安、沈阳、福州、合肥、济南、杭州、石家庄	—	35	17
南京	—	35	16
上海金融、深圳金融、北京金融	—	52	27
上海自动驾驶云	—	54	24
中国香港、中国台北	—	45	23
首尔、东京	—	47	23
新加坡、法兰克福、多伦多	—	43	22
美国硅谷	—	36	18
曼谷、雅加达、孟买、弗吉尼亚、圣保罗	—	43	23

GPU V100 预留券

GPU V100 预留券可以用于抵扣 V100 型号的 GPU 容器的费用，详情请参见 [预留券使用说明](#)。

计量	价格（月）	价格（小时）
GPU	每卡 3310 元	约每卡 4.597 元
CPU	每核 60 元	约每核 0.083 元
内存	每 GiB 35 元	约每 GiB 0.069 元

GPU T4 预留券

GPU T4 预留券可以用于抵扣 T4 型号的 GPU 容器的费用，详情请参见 [预留券使用说明](#)。

计量	价格（月）	价格（小时）
GPU	每卡 1500 元	约每卡 2.083 元
CPU	每核 25 元	约每核 0.035 元
内存	每 GiB 25 元	约每 GiB 0.035 元

GPU A10 预留券

GNV4v 预留券

计量	价格（月）	价格（小时）
GPU	每卡 2950 元	约每卡 4.097 元
CPU	每核 32 元	约每核 0.0444 元
内存	每 GiB 16 元	约每 GiB 0.0222 元
License	每个 220 元	约每个 0.3056 元

GNV4 预留券

计量	价格（月）	价格（小时）
GPU	每卡 2950 元	约每卡 4.097 元
CPU	每核 50 元	约每核 0.0694 元
内存	每 GiB 25 元	约每 GiB 0.0347 元
License	每个 220 元	约每个 0.3056 元

PNV4 预留券

计量	价格（月）	价格（小时）
GPU	每卡 2950 元	约每卡 4.097 元
CPU	每核 32 元	约每核 0.0444 元
内存	每 GiB 16 元	约每 GiB 0.0222 元

预留券使用说明

最近更新时间：2024-05-23 15:10:51

使用说明

预留券在生效时间内，尽量匹配并抵扣每个容器（Serverless 集群 Pod 或容器实例）结算周期产生的容器计算费用。

生效时间

预留券购买即时生效，并会抵扣当前结算周期账单。购买月数个自然月之后日期的23点59分59秒到期，到期后停止抵扣行为。

例如，在4月15日16点购买一个月预留券，则在5月15日23点59分59秒到期。

容器结算周期及预留券抵扣状态

容器结算周期为1小时，预留券生效时会默认覆盖本结算周期。

控制台抵扣状态显示为两个计费周期前的抵扣状态，并在整点更新。

示例

用户在13点30分购买预留券，可以在14点的抵扣周期抵扣13点到14点的按量计费 Pod，此时控制台预留券抵扣状态属于11点到12点这一计费周期；14点时控制台预留券抵扣状态更新为12点到13点的计费周期抵扣状态。

匹配抵扣原则

● 匹配抵扣范围

- 地域级预留券，可抵扣该地域下的全量 Pod。
- 可用区级，只可抵扣该可用区下的 Pod。
- 节点级预留券，只可抵扣该节点上的 Pod。

● 同地域

预留券与容器必须在同一地域才能进行费用抵扣。

例如，广州地域的预留券不可抵扣上海地域的容器费用。

● 同资源规格类型

预留券与容器必须资源类型相同才能进行费用抵扣。

例如，AMD 专用预留券不可抵扣 Intel CPU 容器费用。

● 同资源规格

预留券与容器必须配置相同才能进行费用抵扣。

例如，8核16GiB 的预留券不可抵扣4核8GiB 的容器费用。

● 同时长

预留券最多在一个结算周期内抵扣本周期内（相同算力 × 结算周期时长）的容器费用。

计算费用

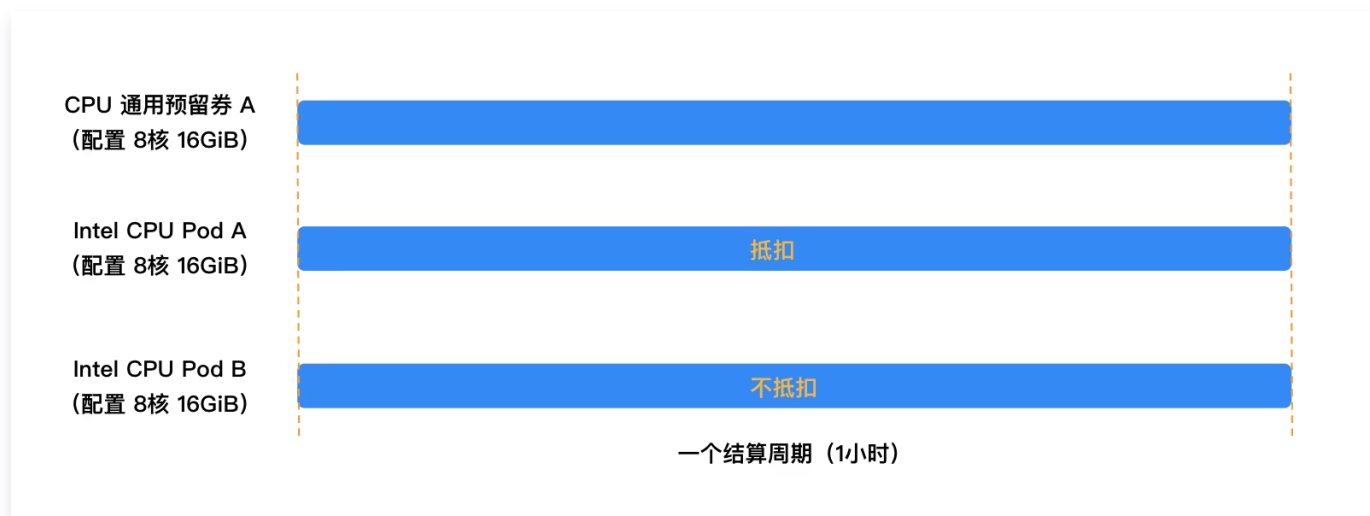
计算费用包括 CPU、内存、GPU 费用，不包括网络、存储等其他资源和服务费用。

匹配抵扣示例

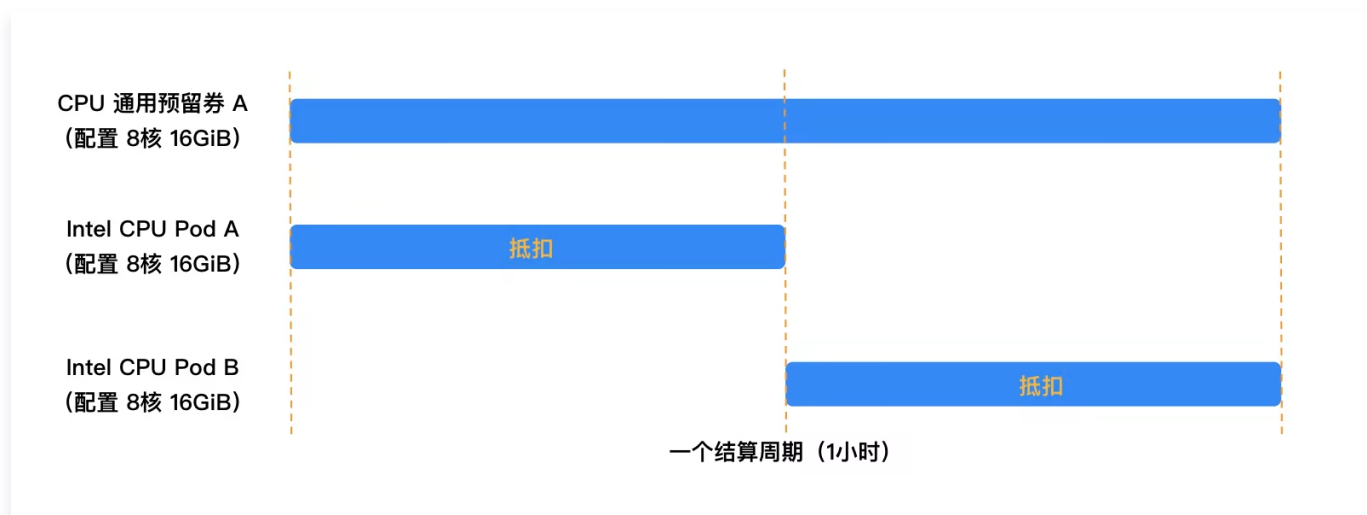
同时长匹配

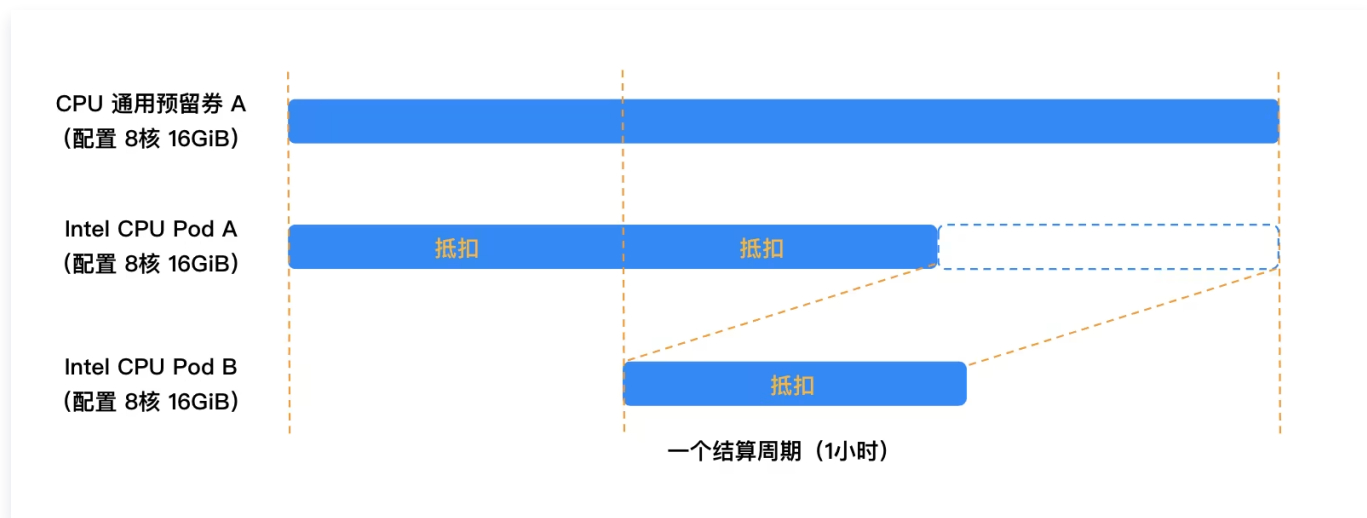
同样资源规格配置的预留券 A、Pod A、Pod B，根据结算周期内的两个容器的运行时长不同，存在以下三种抵扣情况：

- Pod A、Pod B 运行时长均覆盖整个结算周期，预留券 A 仅可抵扣其中一个容器的费用。如下图所示：

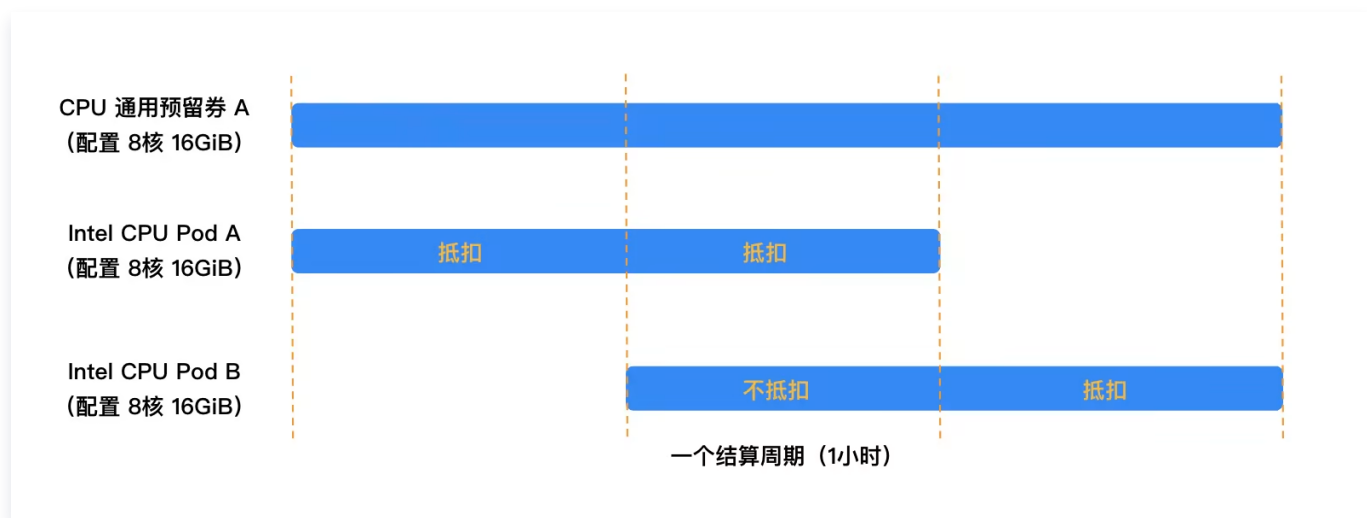


- Pod A、Pod B 运行时长总和正好覆盖整个结算周期，预留券 A 可抵扣两个容器的费用。如下图所示：



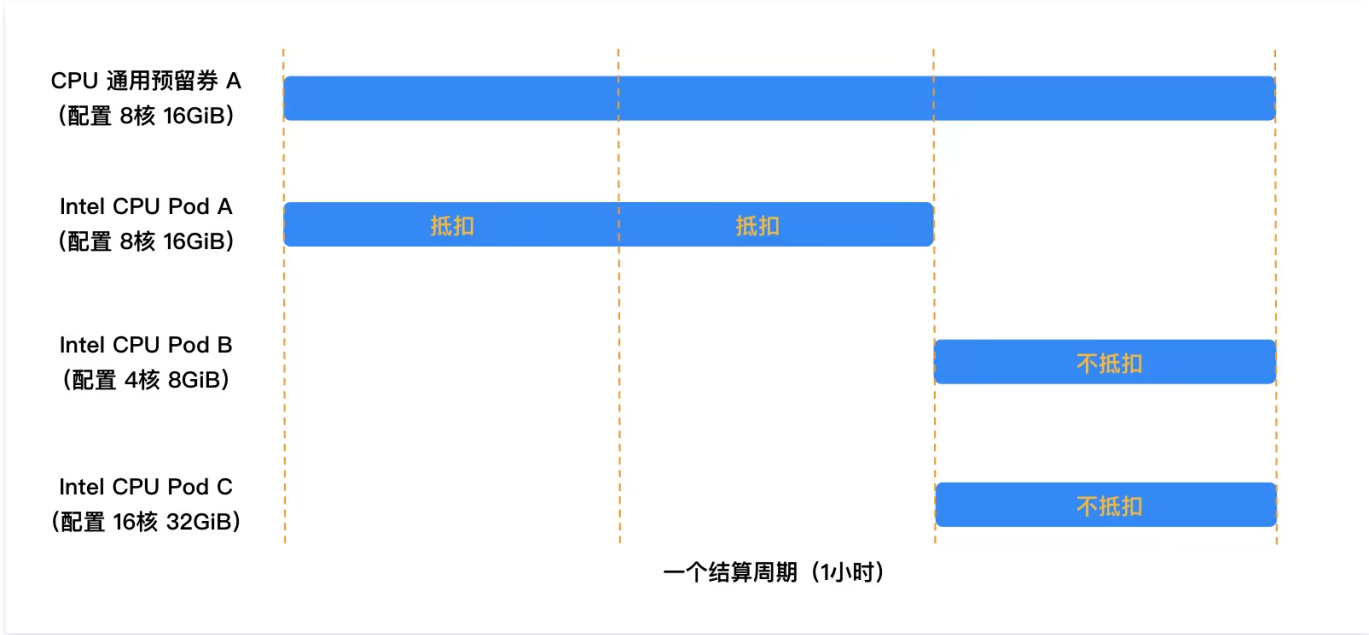


- Pod A、Pod B 运行时长总和超过整个结算周期，预留券 A 最多可抵扣该资源规格容器一个周期时长的费用。如下图所示：

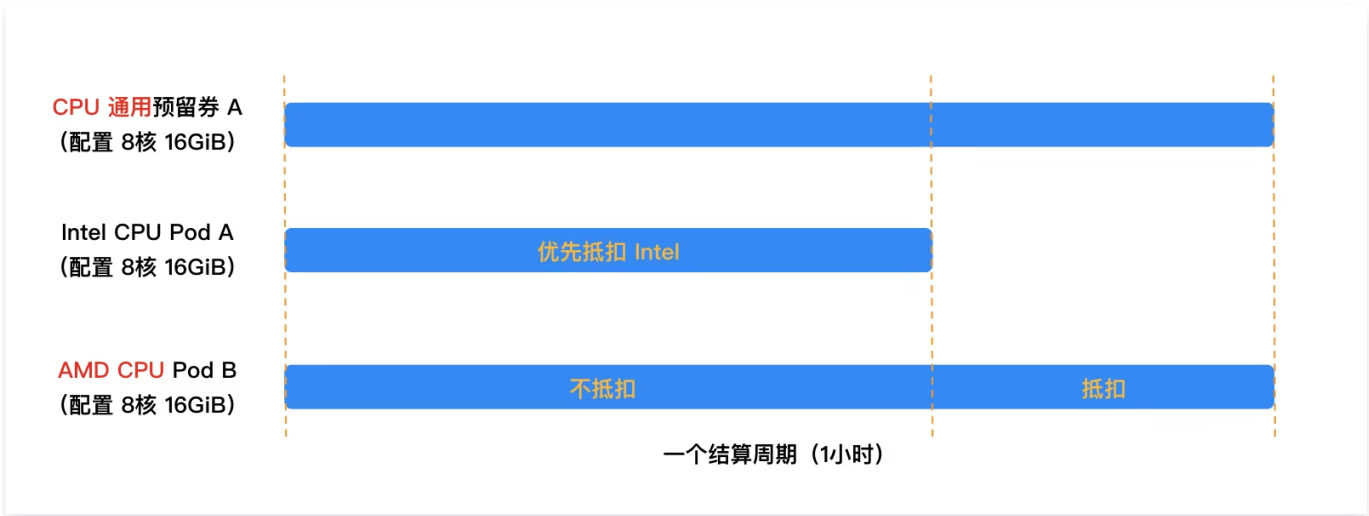


同资源规格匹配

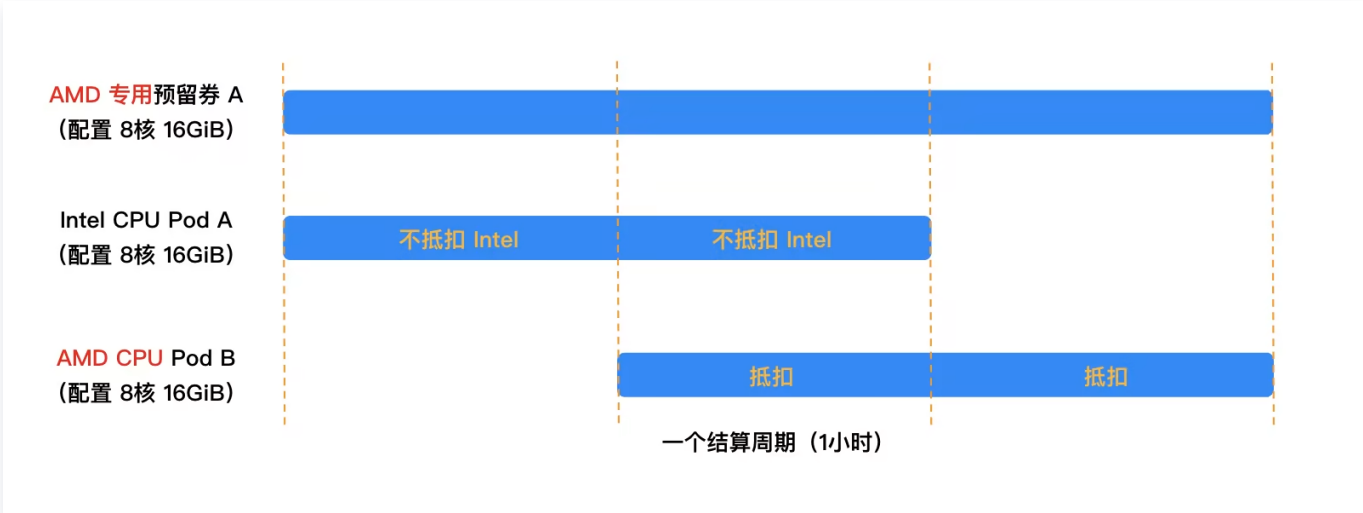
预留券 A 只能抵扣同样资源规格配置的 Pod A。如下图所示：



CPU 通用类型，优先抵扣 Intel



AMD 专用类型，仅抵扣 AMD



预留券管理

最近更新时间：2025-07-15 11:28:52

操作场景

本文介绍如何创建、管理 TKE Serverless 服务预留券，包括如何购买、退还预留券和查看预留券抵扣信息。

前提条件

在创建 TKE Serverless 服务预留券前，您需要完成以下工作：

- 注册[腾讯云账号](#)，并完成[实名认证](#)。
- 预留券是预付费产品，请确保账号余额充足。

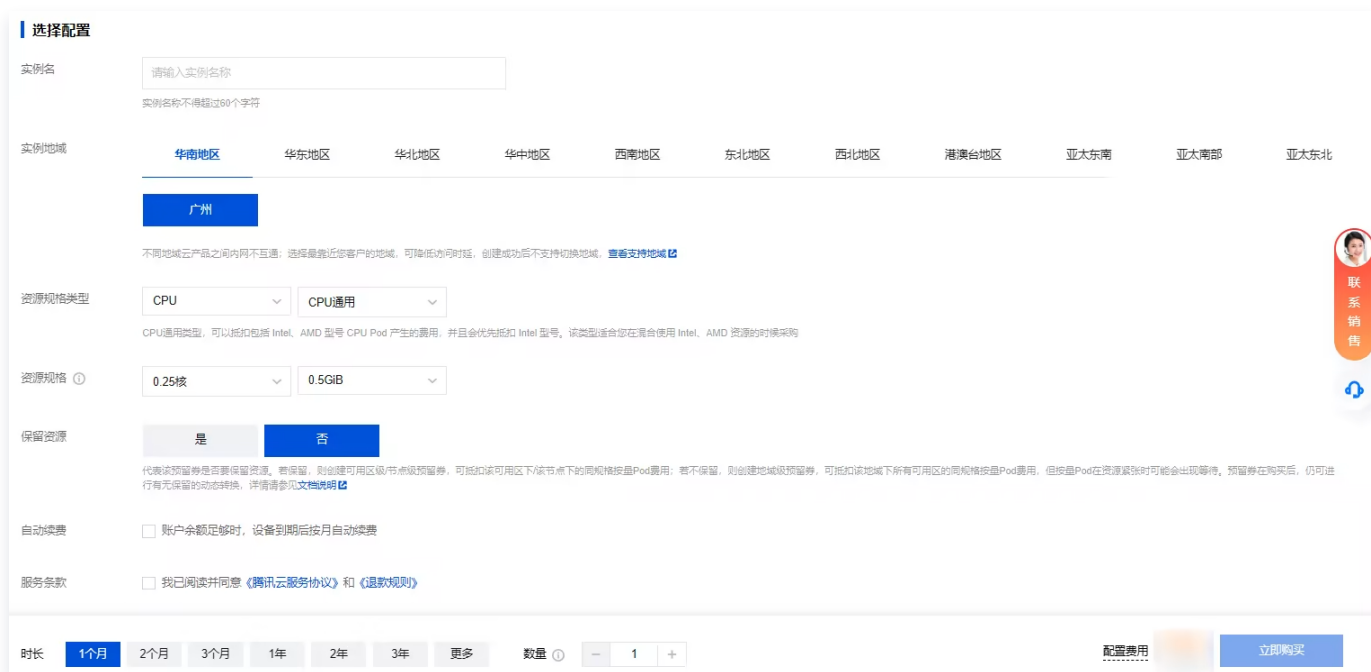
操作步骤

创建预留券

- 登录[容器服务控制台](#)，选择左侧导航栏的[运维中心](#) > [预留券](#)。
- 在[资源预留券](#)页面，单击[新建](#)，进入[Serverless 容器服务预留券购买页](#)。如下图所示：



- 在购买页中，参考以下信息购买预留券。如下图所示：



- **实例名：**预留券的名称。不得超过60个字符。
- **实例地域：**建议选择与您的客户最近的地域，可降低访问时延、提高访问速度。不同地域云产品之间内网不互通。详情可参见 [地域和可用区](#)。
- **资源规格类型：**
 - **CPU 通用类型：**可以抵扣包括 Intel、AMD 型号 CPU Pod 产生的费用，并且会优先抵扣 Intel 型号。该类型适合您在混合使用 Intel、AMD 资源的时候采购。
 - **AMD 专用类型：**仅可以抵扣 AMD 型号 CPU Pod 产生的费用，但单位时长的费用更低。该类型适合您仅使用 AMD 资源的时候采购。
- **资源规格：**资源规格是容器（TKE Serverless 集群 Pod）可使用的最大资源量。预留券的资源规格需要和容器资源规格完全一致，以保证匹配抵扣和预留资源。
- **保留资源（可选）：**代表该预留券是否要保留资源。若不勾选，则创建地域级预留券，可抵扣该地域下所有可用区的同规格按量 Pod 费用，不保留资源，按量 Pod 在资源紧张时可能会出现等待；若勾选，则创建可用区级/节点级预留券，可抵扣该可用区下/该节点下的同规格按量 Pod 费用，保留资源。（预留券在购买后，仍可进行有无保留资源的动态转换。）

保留资源

是

否

代表该预留券是否要保留资源。若保留，则创建可用区级/节点级预留券，可抵扣该可用区下/该节点下的同规格按量Pod费用；若不保留，则创建地域级行有无保留的动态转换，详情请参见[文档说明](#)

可用区

广州三区

广州四区

广州六区

广州七区

绑定超级节点

☐ 开启

- **绑定超级节点（可选）：**代表该预留券是否要绑定到某个具体的超级节点，只可抵扣该节点上的 Pod。若勾选，则创建节点级预留券。如下图所示：

保留资源

是

否

代表该预留券是否要保留资源。若保留，则创建可用区级/节点级预留券，可抵扣该可用区下/该节点下的同规格按量Pod费用；若不保留，则创建地域级行有无保留的动态转换，详情请参见[文档说明](#)

绑定超级节点

☒ 开启

集群

请选择

超级节点

请选择

- **数量：**一次购买数量，请尽量保证和希望抵扣的容器（TKE Serverless 集群 Pod）数量一致。

注意：

单次购买数量上限为 300。

- 选择购买时长，并勾选服务条款。请在配置完成基本信息后确认已选规格及配置费用，计费详情请参见 [产品定价](#)。
- 单击**立即购买**购买已选配置的预留券。
- 您可在[资源预留券](#)页查看已购买的预留券。

查看抵扣范围

在[资源预留券](#)页，可查看预留券的抵扣范围，如下所示：



资源预留券 地域 广州

新建 概览 退还 续费 设为自动续费 设为手动续费 修改抵扣范围

多个过滤标签用回车键分隔

ID/名称	状态	抵扣状态	抵扣范围	时长	规格类型	规格	生效时间	到期时间	购买时间	操作
	生效中	未抵扣	广州	1个月	CPU通用 (含Intel和AMD)	CPU:0.25核 内存:0.5GiB				续费 退还 更多

修改抵扣范围

当需要在地域级、可用区级、节点级预留券之间进行转换时，选择**更多 > 修改抵扣范围**，如下图所示：



资源预留券 地域 广州

新建 概览 退还 续费 设为自动续费 设为手动续费 修改抵扣范围

多个过滤标签用回车键分隔

ID/名称	状态	抵扣状态	抵扣范围	时长	规格类型	规格	生效时间	到期时间	购买时间	操作
	生效中	未抵扣	广州	1个月	CPU通用 (含Intel和AMD)	CPU:0.25核 内存:0.5GiB				续费 退还 更多

共 1 条

20 条 / 页

1

设置成自动续费

修改抵扣范围

在修改以下预留券的抵扣范围页中，选择想要修改的抵扣范围，单击**确定**，如下图所示：

修改以下预留券的抵扣范围



您已选择1张预留券

ID/名称	抵扣范围	时长	规格类型	规格
	广州		CPU通用 (含Intel和AMD)	CPU: 0.25核 内存: 0.5GiB

抵扣范围

地域范围抵扣

可用区范围抵扣

节点范围抵扣

不支持保留资源，可抵扣该地域下所有可用区的同规格按量pod费用

确定

取消

查看预留券抵扣记录

1. 登录 [容器服务控制台](#)，选择左侧导航栏的**运维中心 > 预留券**。
2. 在**资源预留券**页面，选择需要查看的预留券名称，进入预留券详情页。
3. 在**抵扣历史**页签中，查看该预留券的费用抵扣记录。

退还预留券

注意：

退还预留券属于费用有损的限制性操作，请勿随意操作。退费详情可参见 [退费规则](#)。

1. 登录 [容器服务控制台](#)，选择左侧导航栏的**运维中心 > 预留券**。
2. 在**资源预留券**页面，单击需要退还的预留券所在行右侧的**退还**。
3. 在退还预留券页，勾选“已阅读并同意退费规则”，单击**确定**即可退还当前所选预留券。如下图所示：

退还预留券



您已选择1张预留券

ID/名称	时长	状态	规格类型	规格
		生效中	CPU通用 (含Intel和AMD)	CPU: 0.25核 内存: 0.5GiB

- 资源销毁后，退款金额将按购买支付使用的现金和赠送金比例退还至您的腾讯云账户，[详情说明](#)
- 若购买时享有折扣或代金券，折扣和代金券**不予退还**

退费规则

☒ 已阅读并同意[退费规则](#)

确定

取消

4. 您可在[资源预留券](#)页面查看到该预留券的状态为“已退还”，且不再进行费用抵扣和资源预留。

预留券退费说明

最近更新时间：2025-03-27 11:23:23

为了方便您使用 Serverless 容器服务预留券，如果您在购买预留券后有任何不满意，我们支持自助退货退款。您可在容器服务控制台自助进行，具体操作详情请参见 [退还预留券](#)。

说明

- 每个账户可享受**1次1张**预留券的**五天无理由**退还，您支付的有效金额将返还至您的腾讯云账户。
- 您可享受普通自助退还，扣除您已使用的费用，将按**购买支付使用的现金和赠送金支付比例**退还至您的**腾讯云账户**的方式返还到您的账户。

五天无理由自助退还

弹性容器服务预留券遵守腾讯云 [云服务退货说明](#)，若您在购买预留券后有任何不满意，我们支持**五天**内无理由自助退还，具体策略如下：

- 弹性容器服务预留券**新购时长** ≤ **五天**，默认支持五天无理由退款。
- 对于每个主体，每个产品**只能退还1个实例**。例如，同一订单包含20张弹性容器服务预留券，最多只能退还其中1张预留券。
- 符合五天无理由退还场景的订单，退款金额为**购买时花费的全部消耗金额**，包括现金账户金额、收益转入账户金额以及赠送账户金额。
- 具体退款规则请参见 [五天内无理由全额退款](#)。

普通自助退还

不满足五天无理由退货退款规则的退货订单，退款策略如下：

普通自助退还将按购买支付使用的**现金和赠送金支付比例**退还至您的腾讯云账户。

普通自助退还限制

每个账户最多支持**199**张弹性容器服务预留券控制台自助退还。

普通自助退还规则

退费金额 = 实际支付金额 - (已使用时长 / 总时长) × 订单原价 × 订单折扣

- 已使用时长精确到天，不满一天按照一天计算。
- 退款金额 ≤ 0，按0计算并清退资源。

预留券自助退还计费规则示例

案例背景

在广州购买1张 CPU 通用型1核2GiB 资源规格的预留券。购买时长1年原价为1068元，包年享83折。购买时使用300元的代金券。

折后需付金额为 $1068 \times 0.83 = 886.44$ (元)

实际支付金额为 $886.44 - 300 = 586.44$ (元)

五天无理由退还场景

案例

购买五天内发现不满意，想要退还，且该账户首次退还。

退费金额 = 实际支付金额 586.44 (元)

普通自助退还场景

案例 1

购买五天内，且该账户首次退还。

退还金额为 实际支付金额 586.44 (元)

案例 2

购买五天内，且不是该账户首次退还，总使用时长48小时，按照2天计算。

退还金额为 $586.44 - (2 / 365) \times 1068 \times 0.83 = 581.58$ (元)

案例 3

购买超过五天，总使用时长132小时，按照6天计算。

退还金额为 $586.44 - (6 / 365) \times 1068 \times 0.83 = 571.87$ (元)

案例 4

购买超过五天，总使用时长260天。

退还金额为 $586.44 - (260 / 365) \times 1068 \times 0.83 < 0$ (元)

由于购买时使用了代金券，而普通退还时不支持退还代金券，此时的退还金额为0，不建议进行退还操作。

竞价模式说明

最近更新时间：2025-06-09 17:21:12

竞价模式概述

TKE Serverless 竞价模式是一种低成本资源购置模式，其核心特点为价格低于按量付费实例，在一些场景可以极大降低运行容器的成本，但资源可能会被腾讯云中断回收。您在使用竞价模式运行时仅需支付较少的费用，并运行到容器资源被回收为止。

使用竞价模式时，您可以像使用普通按量计费资源一样在容器部署工作负载，且同样具备普通按量计费模式下的所有功能。

竞价模式策略

价格策略

目前 TKE Serverless 采用的竞价模式为**固定折扣比例（2折）**，即所有规格的竞价模式将以原规格 **产品定价** 的固定折扣出售。

⚠ 注意：

该折扣仅对 Pod 资源规格计费项（CPU、内存、GPU）生效。不包括网络带宽、网络流量、持久化存储等资源的费用。

回收中断机制

竞价模式下的容器会因为腾讯云计算资源池库存不足而产生回收中断。当库存不足时，会从已分配的竞价模式容器里随机回收，容器缓存数据不会保留。

回收中断时将会产生以下中断事件：

```
EVENT REASON : "SpotPodInterruption"
EVENT MESSAGE : "Spot pod was interrupted, it will be killed and re-
created"
```

适用场景

适合短时长突发、周期任务

适用于不需要长期运行的突发性、周期性短时长工作负载。例如视频转码、视频渲染、服务压测、批量计算、爬虫等。

适合可切分的计算任务

适用于可以将长时间作业按作业对象切分为细粒度任务进行计算的系统。例如 EMR 等大数据套件。

适合无状态或者支持断点续传能力的计算任务

- 适用于将计算中间结果放到持久化存储上，可接受 Pod 被回收重启后继续运算的工作负载。
- 适用于支持自动负载均衡和服务发现的无状态工作负载，可接受 Pod 被回收重启的工作负载。

竞价模式开启

您可以通过在工作负载 YAML 中定义如下 Pod template annotation 方式，为工作负载开启竞价模式。

```
eks.tke.cloud.tencent.com/spot-pod: "true"
```

其他支持 Annotation 以及范例请参见 [Annotation 说明](#)。