云容器引擎

产品介绍

文档版本 01

发布日期 2021-08-20

华为技术有限公司

商标声明


和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目 录

1 了解云原生 2.0

作为容器最早的采用者之一，华为自2013年起就在内部多个产品落地容器技术，2014年开始广泛使用Kubernetes。在此过程中华为积累了丰富的实践经验，并在历经自身亿级用户量考验的实践后，面向企业用户提供了全栈容器服务，帮助企业轻松应对Cloud2.0时代和应用上云的挑战。

基于华为自身实践与社区的贡献积累，华为云自上线之初，就持续利用云原生技术为用户提供标准化、可移植的领先云原生基础设施服务。

云原生 2.0

随着云原生技术的成熟和市场需求的升级，云计算的发展已步入新的阶段。云原生2.0时代已经到来

从技术角度看，以容器、微服务以及动态编排为代表的云原生技术蓬勃发展，成为赋能业务创新的重要推动力，并已经应用到企业核心业务。从市场角度看，云原生技术已在金融、制造、互联网等多个行业得到广泛验证，支持的业务场景也愈加丰富，行业生态日渐繁荣。

云原生2.0，企业云化从“ON Cloud”走向“IN Cloud“，生于云、长于云且立而不破

企业新生能力基于云原生构建，使其生于云；应用、数据和AI的全生命周期云上完成，使其长于云；同时，既有能力通过立而不破的方式继承下来，并与新生能力有机协同。

智能升级新阶段，赋能“新云原生企业”

云原生2.0是企业智能升级的新阶段，企业云化从“ON Cloud”走向“IN Cloud“，成为”新云原生企业“。新生能力与既有能力立而不破、有机协同，实现资源高效、应用敏捷、业务智能，安全可信。

点此**观看视频**。

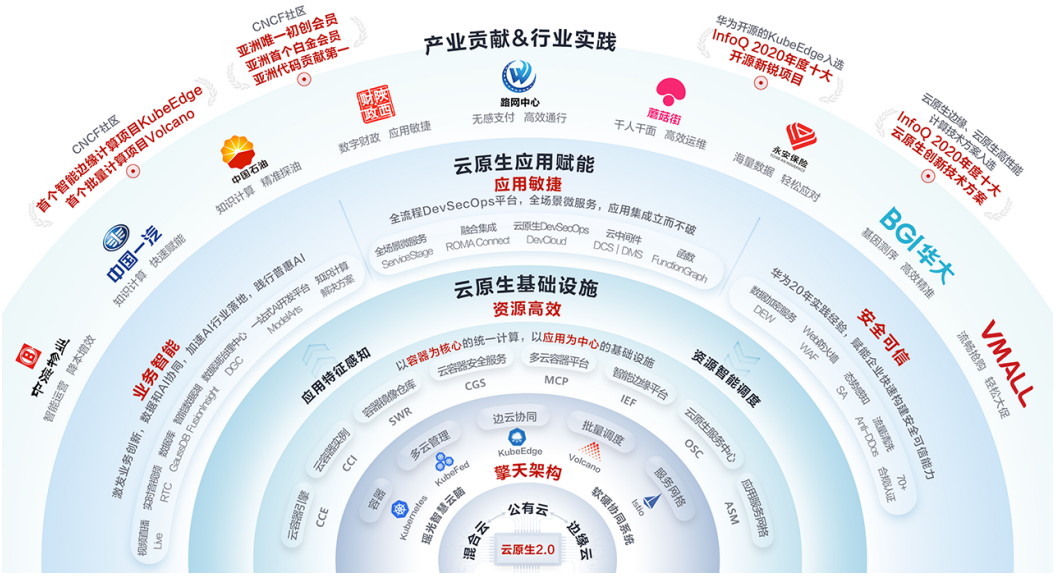
云原生 IN 基础设施

华为云基于“云原生 IN 基础设施”的理念，打造了以应用为中心的云原生基础设施。

目前，华为云云原生基础设施包含了云容器引擎CCE、云容器实例CCI、容器镜像服务SWR、智能边缘平台IEF、多云容器平台MCP、应用编排服务AOS等8大核心容器产品，并以此为基础构建了云原生裸金属、云原生高性能计算、云原生混合云、云原生

边缘计算四大解决方案，满足企业业务智能升级过程中，对高性能基础设施、分布式业务架构、完善的云原生应用生态的诉求。

图 1-1 华为云 2.0 全景图



云原生基础设施

华为云通过“重定义基础设施、新赋能泛在应用、再升级应用架构”三大创新升级，为客户提供极致体验、极致成本的云原生基础设施。使能企业构建多云、云边协同的应用架构，并统一企业应用底座，加速业务创新。

- **云容器引擎**（Cloud Container Engine，简称CCE）提供高度可扩展的、高性能的企业级Kubernetes集群，支持运行Docker容器，提供了Kubernetes集群管理、容器应用全生命周期管理、应用服务网格、Helm应用模板、插件管理、应用调度、监控与运维等容器全栈能力，为您提供一站式容器平台服务。借助云容器引擎，您可以在华为云上轻松部署、管理和扩展容器化应用程序。
- **云容器实例**（Cloud Container Instance，简称CCI）服务提供Serverless Container（无服务器容器）引擎，让您无需创建和管理服务器集群即可直接运行容器。通过CCI您只需要管理运行在Kubernetes上的容器化业务，无需管理集群和服务器即可在CCI上快速创建和运行容器负载，使容器应用零运维，使企业聚焦业务核心，为企业提供了Serverless化全新一代的体验和选择。
- **容器镜像服务**（Software Repository for Container，简称SWR），是一种支持容器镜像全生命周期管理的服务， 提供简单易用、安全可靠的镜像管理功能，帮助用户快速部署容器化服务。
- **容器安全服务**（Container Guard Service，CGS）能够扫描镜像中的漏洞与配置信息，帮助企业解决传统安全软件无法感知容器环境的问题；同时提供容器进程白名单、文件只读保护和容器逃逸检测功能，有效防止容器运行时安全风险事件的发生。
- **智能边缘平台**（Intelligent EdgeFabric，简称IEF）通过纳管用户的边缘节点，提供将云上应用延伸到边缘的能力，联动边缘和云端的数据，满足客户对边缘计算资源的远程管控、数据处理、分析决策、智能化的诉求，同时，在云端提供统一的设备/应用监控、日志采集等运维能力，为企业提供完整的边缘和云协同的一体化服务的边缘计算解决方案。

- **多云容器平台**（Multi-Cloud Container Platform，简称MCP）是华为云基于多年容器云领域实践经验和社区先进的集群联邦技术，提供的容器多云和混合云的解决方案，为您提供跨云的多集群统一管理、应用在多集群的统一部署和流量分发，为您彻底解决多云灾备问题的同时，还可以在业务流量分担、业务与数据分离、开发与生产分离、计算与业务分离等多种场景下发挥价值。
- **应用服务网格**（Application Service Mesh，简称ASM）是华为云基于开源Istio推出的服务网格平台，它深度、无缝对接了华为云的企业级Kubernetes集群服务云容器引擎（Cloud Container Engine，简称CCE），在易用性、可靠性、可视化等方面进行了一系列增强，可为客户提供开箱即用的上手体验。 应用服务网格提供非侵入式的微服务治理解决方案，支持完整的生命周期管理和流量治理，兼容Kubernetes和Istio生态，其功能包括负载均衡、熔断、限流等多种治理能力。

云原生应用赋能

华为将云原生的全栈能力赋能给客户，帮助客户应用敏捷、业务智能，安全可信，面向未来持续演进。

应用敏捷

- **软件开发平台**（DevCloud）是集华为近30年研发实践、前沿研发理念、先进研发工具为一体的一站式云端DevOps平台，面向开发者提供的云服务，即开即用，随时随地在云端进行项目管理、代码托管、流水线、代码检查、编译构建、部署、测试、发布等，让开发者快速而又轻松地开启云端开发之旅。
- **应用管理与运维平台**（ServiceStage）是一个应用托管和微服务管理平台，可以帮助企业简化部署、监控、运维和治理等应用生命周期管理工作。ServiceStage面向企业提供微服务、移动和Web类应用开发的全栈解决方案，帮助您的各类应用轻松上云，聚焦业务创新，帮助企业数字化快速转型。
- **应用与数据集成平台**（ROMA Connect）是一个全栈式的应用与数据集成平台，源自华为数字化转型集成实践，聚焦应用和数据连接，适配多种企业常见的使用场景。ROMA Connect提供轻量化消息、数据、API、设备、模型等集成能力，简化企业上云流程，支持云上云下、跨区域集成，帮助企业实现数字化转型。
- **分布式消息服务**（Kafka）是一款基于开源社区版Kafka提供的消息队列服务，向用户提供计算、存储和带宽资源独占式的Kafka专享实例。使用华为云分布式消息服务Kafka，资源按需申请，按需配置Topic的分区与副本数量，即买即用，您将有更多精力专注于业务快速开发，不用考虑部署和运维。
- **函数工作流**（FunctionGraph）是一项基于事件驱动的函数托管计算服务。使用FunctionGraph函数，只需编写业务函数代码并设置运行的条件，无需配置和管理服务器等基础设施，函数以弹性、免运维、高可靠的方式运行。

业务智能

- **知识计算解决方案**是基于一站式AI开发平台ModelArts打造的业界首个全生命周期知识计算解决方案,助力企业打造自己的知识计算平台。企业可以灵活掌控流程配置，自主完成图谱更新，适合复杂多变的企业场景。
- **AI开发平台**（ModelArts）是面向开发者的一站式AI开发平台，为机器学习与深度学习提供海量数据预处理及半自动化标注、大规模分布式Training、自动化模型生成，及端-边-云模型按需部署能力，帮助用户快速创建和部署模型，管理全周期AI工作流。
- **云数据库**（GaussDB）是华为自研的最新一代企业级高扩展海量存储分布式数据库，完全兼容MySQL。基于华为最新一代DFV存储，采用计算存储分离架构，128TB的海量存储，无需分库分表，数据0丢失，既拥有商业数据库的高可用和性能，又具备开源低成本效益。

- **数据湖治理中心**（DGC）是为了应对上述挑战、针对企业数字化运营诉求提供的数据全生命周期管理、具有智能数据管理能力的一站式治理运营平台，包含数据集成、规范设计、数据开发、数据质量监控、数据资产管理、数据服务等功能，支持行业知识库智能化建设，支持大数据存储、大数据计算分析引擎等数据底座，帮助企业快速构建从数据接入到数据分析的端到端智能数据系统，消除数据孤岛，统一数据标准，加快数据变现，实现数字化转型。

安全可信

- **数据安全中心服务**（DSC）是新一代的云化数据安全平台，提供数据分级分类、数据安全风险识别、数据水印溯源和数据静态脱敏等基础数据安全能力，通过数据安全总览整合数据安全生命周期各阶段状态，对外整体呈现云上数据安全态势。
- **企业主机安全服务**（HSS）是提升主机整体安全性的服务，通过主机管理、风险预防、入侵检测、高级防御、安全运营、网页防篡改功能，全面识别并管理主机中的信息资产，实时监测主机中的风险并阻止非法入侵行为，帮助企业构建服务器安全体系，降低当前服务器面临的主要安全风险。
- **态势感知**（SA）是华为云安全管理与态势分析平台。能够检测出超过20大类的云上安全风险，包括DDoS攻击、暴力破解、Web攻击、后门木马、僵尸主机、异常行为、漏洞攻击、命令与控制等。利用大数据分析技术，态势感知可以对攻击事件、威胁告警和攻击源头进行分类统计和综合分析，为用户呈现出全局安全攻击态势。
- **DDoS防护**（ADS）提供多种安全防护方案，您可以根据您的实际业务选择合适的防护方案。华为云DDoS防护服务（Anti-DDoS Service，简称ADS）提供了DDoS原生基础防护（Anti-DDoS流量清洗）、DDoS原生专业防护和DDoS高防三个子服务。

2 图解云容器引擎

3 什么是云容器引擎

云容器引擎（Cloud Container Engine，简称CCE）提供高度可扩展的、高性能的企业级Kubernetes集群，支持运行Docker容器。借助云容器引擎，您可以在华为云上轻松部署、管理和扩展容器化应用程序。

为什么选择云容器引擎

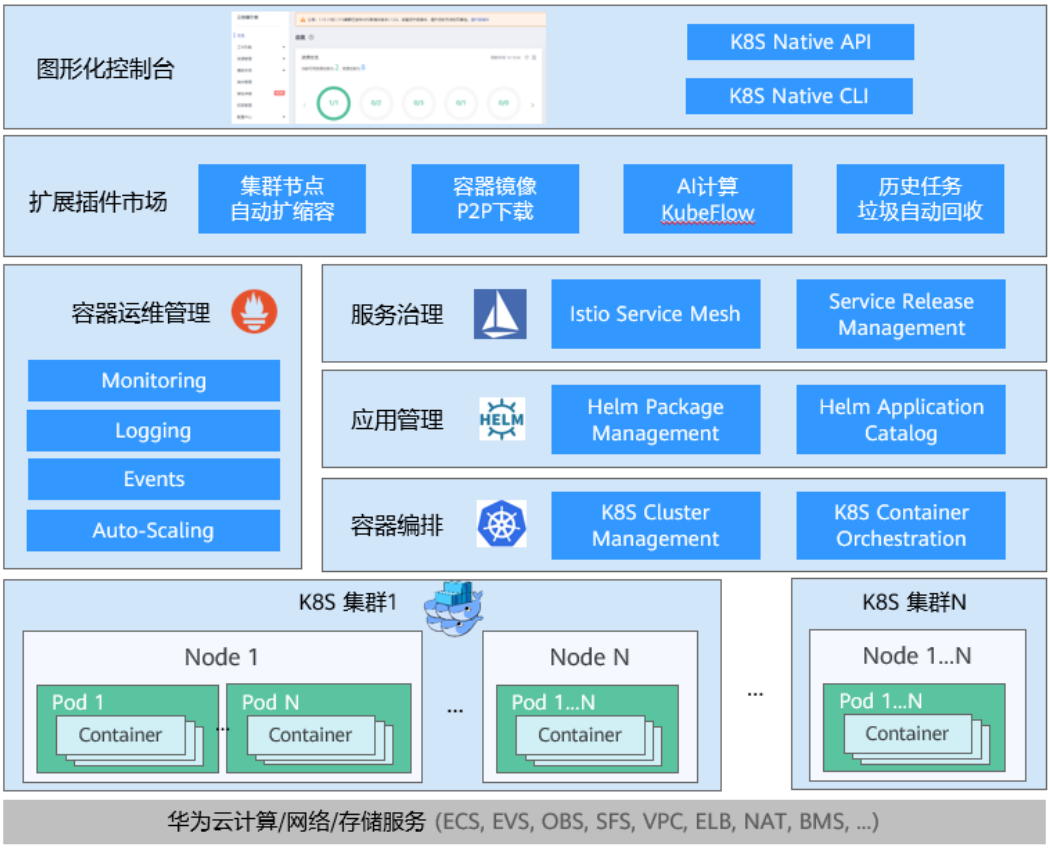
云容器引擎深度整合华为云高性能的计算（ECS/BMS）、网络（VPC/EIP/ELB）、存储（EVS/OBS/SFS）等服务，并支持GPU、NPU、ARM、FPGA等异构计算架构，支持多**可用区**（Available Zone，简称AZ）、多**区域**（Region）容灾等技术构建高可用Kubernetes集群。

华为云是全球首批Kubernetes认证服务提供商（Kubernetes Certified Service Provider，KCSP），是国内最早投入Kubernetes社区的厂商，是容器开源社区主要贡献者和容器生态领导者。华为云也是CNCF云原生计算基金会的创始成员及白金会员，云容器引擎是全球首批通过CNCF基金会Kubernetes一致性认证的容器服务。

更多选择理由，请参见**产品功能**、**产品优势**和**应用场景**。

产品架构

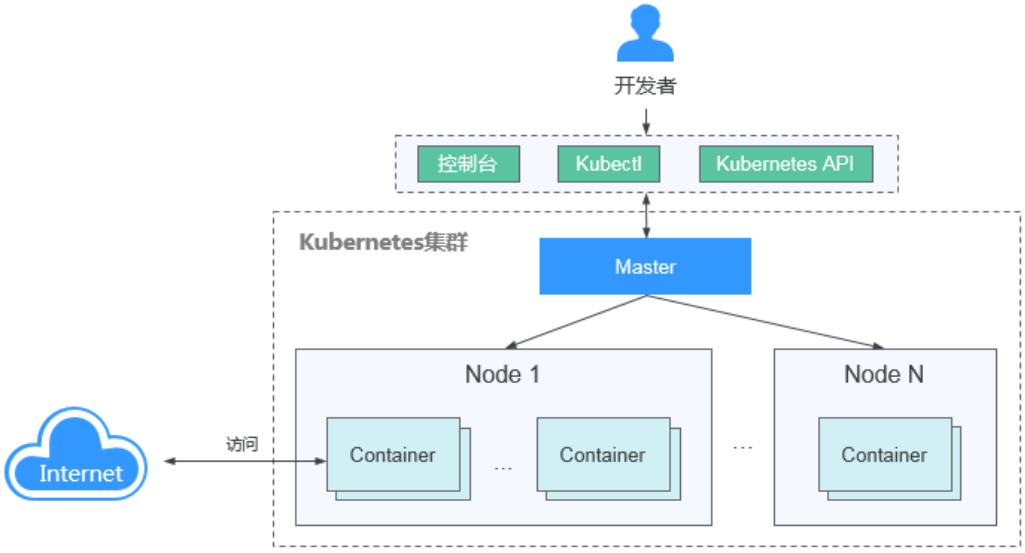
图 3-1 CCE 产品架构



访问方式

您可以通过**CCE控制台**、**Kubectl命令行**、**Kubernetes API**使用云容器引擎服务。具体请参见**图3-2**。

图 3-2 使用云容器引擎



云容器引擎学习路径

您可以借助云容器引擎**成长地图**，快速了解产品，由浅入深学习使用和运维CCE。

4 CCE Turbo 集群

以容器为核心的云原生基础设施，不仅让资源利用率更高，还能解放运维人员，聚焦应用和业务创新。但是，容器的规模化应用，也对性能、弹性、调度能力提出了更高要求。

没有容器化，就没有应用现代化。以容器为核心的云原生基础设施，不仅提供更高效的资源，还能把开发运维人员从资源的调配和运维中解放出来，聚焦于应用和业务创新。容器全面规模化应用的同时也对性能、弹性、调度能力提出了更高的要求。

华为云已在**华为开发者大会2021**正式发布革命性的容器集群——**CCE Turbo**，点此**观看发布会现场视频**。

华为云CCE Turbo容器集群在计算、网络和调度上全方位加速，让容器真正成为企业应用创新的强劲引擎。

集群优势

- **计算加速**
  软硬协同基础设施，更少资源可实现更好的性能。
- **网络加速**
  云原生网络2.0，无损网络让应用间通信更流畅。
- **调度加速**
  智能混合调度，简单高效地管理应用与资源。
- **安全隔离**
  安全容器引擎，使应用拥有虚机级安全隔离。

CCE Turbo 集群与 CCE 集群对比

CCE支持多种类型的集群创建，以满足您各种业务需求，如下为CCE Turbo集群与CCE集群区别：

表 4-1 集群类型对比

维度	子维度	CCE Turbo集群	CCE集群
集群	定位	面向云原生2.0的新一代容器集群产品，计算、网络、调度全面加速	标准版本集群，提供商用级的容器集群服务
	节点形态	支持虚拟机和裸金属服务器混合	支持虚拟机和裸金属服务器混合
	支持机型	基于擎天软硬件协同架构的机型	通用机型
网络	网络模型	云原生网络2.0：面向大规模和高性能的场景。 组网规模2000节点	**云原生网络1.0**：面向性能和规模要求不高的场景。 • 隧道网络模式 • VPC网络模式
	网络性能	VPC网络和容器网络融合，性能无损耗	VPC网络叠加容器网络，性能有一定损耗
	容器网络隔离	基于安全组的隔离策略，支持集群内外部统一的安全隔离，支持SecurityGroup。	• 隧道网络模式：集群内部网络隔离策略，支持Networkpolicy。 • VPC网络模式：不支持
安全	隔离性	• 裸金属服务器：安全容器，支持虚机级别的隔离 • 虚拟机：普通容器	普通容器，Cgroups隔离

如何购买

点此可了解如何购买并使用CCE Turbo集群。

5 功能总览

云容器引擎提供高度可扩展的、高性能的企业级Kubernetes集群，包括集群管理、节点管理、节点池管理、工作负载、亲和/反亲和性调度、容器网络、容器存储、插件管理、模板市场、弹性伸缩、权限管理、系统管家等功能，为您提供一站式容器平台服务。

集群管理

云容器引擎CCE是一种托管的Kubernetes服务，可进一步简化基于容器的应用程序部署和管理，您可以在CCE中方便的创建Kubernetes集群、部署您的容器化应用，以及方便的管理和维护。

- **一站式部署和运维：**使用云容器引擎，您可以一键创建Kubernetes容器集群，无需自行搭建Docker和Kubernetes集群。您可以通过云容器引擎自动化部署和一站式运维容器应用，使得应用的整个生命周期都在云容器引擎内高效完成。
- **支持多类型容器集群：**通过云容器引擎您可以直接使用华为云高性能的弹性云服务器、裸金属服务器、GPU加速云服务器等多种异构基础设施，您可以根据业务需要在云容器引擎中快速创建CCE集群、鲲鹏集群、CCE Turbo集群，并通过云容器引擎对创建的集群进行统一管理。

表 5-1 集群管理功能介绍

功能模块	功能概述
CCE Turbo集群	支持第二代裸金属容器，同时支持传统虚拟机节点，基于华为云新一代高性能基础设施提供极致的性能体验。
CCE集群	CCE集群支持虚拟机与裸金属服务器混合、支持GPU、NPU等异构节点的混合部署，基于高性能网络模型提供全方位、多场景、安全稳定的容器运行环境。
鲲鹏集群	鲲鹏容器集群（ARM指令集）提供了容器在鲲鹏（ARM架构）服务器上的运行能力，提供与X86服务器相同的调度伸缩，快速部署能力，并具有大幅降低成本的潜力。
集群弹性扩容	根据实际业务需要对CCE集群的工作节点进行扩容和缩容，当集群中出现由于资源不足而无法调度的工作负载时自动触发扩容，从而减少人力成本。

功能模块	功能概述
集群升级	通过云容器引擎管理控制台快速升级到Kubernetes最新版本或者bugfix版本，以支持新特性的使用。
集群监控	实时查看每个集群控制节点的资源使用情况，了解CCE集群控制节点的监控指标，及时收到异常告警并做出反应，保证业务顺畅运行。

节点管理

节点是容器集群组成的基本元素。节点取决于业务，既可以是虚拟机，也可以是物理机。每个节点都包含运行Pod所需要的基本组件，包括 Kubelet、Kube-proxy 、Container Runtime等。在云容器引擎CCE中，主要采用高性能的弹性云服务器ECS或裸金属服务器BMS作为节点来构建高可用的Kubernetes集群。

表 5-2 节点管理功能介绍

功能模块	功能概述
添加节点	支持两种添加节点的方式：购买节点和纳管节点，纳管节点是指将“已购买的弹性云服务器（ ECS ）加入到CCE集群中”。 支持虚拟机、裸金属服务器、GPU、NPU等异构节点的购买添加。
节点监控	CCE通过云监控服务（ Cloud Eye ）为您提供节点的监控，每个节点对应一台弹性云服务器。
重置节点	在CCE集群中重置节点会将该节点以及节点内运行的业务都销毁，重置前请确认您的正常业务运行不受影响，请谨慎操作。该功能支持v1.13及以上版本的集群。
删除节点	在CCE集群中删除节点会将该节点以及节点内运行的业务都销毁，删除前请确认您的正常业务运行不受影响，请谨慎操作。

节点池管理

支持创建新的自定义节点池，借助节点池基本功能方便快捷地创建、管理和销毁节点，而不会影响整个集群。新节点池中所有节点的参数和类型都彼此相同，您无法在节点池中配置单个节点，任何配置更改都会影响节点池中的所有节点。

表 5-3 节点池管理功能介绍

功能模块	功能概述
创建节点池	创建节点池、查看节点池
管理节点池	编辑节点池、删除节点池、拷贝节点池、迁移节点

工作负载

工作负载是在Kubernetes上运行的应用程序。无论您的工作负载是单个组件还是协同工作的多个组件，您都可以在Kubernetes上的一组Pod中运行它。在Kubernetes中，工作负载是对一组Pod的抽象模型，用于描述业务的运行载体，包括Deployment、Statefulset、Daemonset、Job、CronJob等多种类型。

CCE提供基于Kubernetes原生类型的容器部署和管理能力，支持容器工作负载部署、配置、监控、扩容、升级、卸载、服务发现及负载均衡等生命周期管理。

表 5-4 工作负载功能介绍

功能模块	功能概述
设置容器规格	支持在创建工作负载时为添加的容器设置资源限制。可以对工作负载中每个实例所用的CPU配额、内存配额进行申请和限制，对每个实例所用的GPU和昇腾 310配额设置使用或不使用。
设置容器生命周期	提供了回调函数，在容器的生命周期的特定阶段执行调用，比如容器在停止前希望执行某项操作，就可以注册相应的钩子函数。
设置容器启动命令	创建工作负载或任务时，通常通过镜像指定容器中运行的进程。在默认情况下，镜像会运行默认命令，如果想运行特定命令或重写镜像默认值，需要用到以下设置： ● 工作目录：指定运行命令的工作目录。若镜像中未指定工作目录，且在界面中也未指定，默认是“/”。 ● 运行命令：控制镜像运行的实际命令。 ● 运行参数：传递给运行命令的参数。
设置容器健康检查	健康检查是指容器运行过程中，根据用户需要，定时检查容器健康状况。若不配置健康检查，如果服务出现业务异常，pod将无法感知，也不会自动重启去恢复业务。最终导致虽然pod状态显示正常，但pod中的业务异常的情况。 提供了两种健康检查的探针： 1. 工作负载存活探针：用于检测容器是否正常，类似于我们执行 ps 命令检查进程是否存在。如果容器的存活检查失败，集群会对该容器执行重启操作；若容器的存活检查成功则不执行任何操作。 2. 工作负载业务探针：用于检查用户业务是否就绪，如果未就绪，则不转发流量到当前实例。一些程序的启动时间可能很长，比如要加载磁盘数据或者要依赖外部的某个模块启动完成才能提供服务。这时候程序进程在，但是并不能对外提供服务。这种场景下该检查方式就非常有用。如果容器的就绪检查失败，集群会屏蔽请求访问该容器；若检查成功，则会开放对该容器的访问。
设置环境变量	环境变量是指容器运行环境中设定的一个变量，环境变量可以在工作负载部署后修改，为工作负载提供极大的灵活性。
采集容器日志	支持配置工作负载日志策略，便于日志的统一收集、管理和分析，以及按周期防爆处理。

亲和/反亲和性调度

云容器引擎提供工作负载和可用区、工作负载和节点以及工作负载间的亲和性/反亲和调度。您可根据业务需求设置亲和性，实现工作负载的就近部署，容器间通信就近路由，减少网络消耗；您也可以对同个工作负载的多个实例设置反亲和部署，减少宕机影响，对互相干扰的应用反亲和部署，避免干扰。

表 5-5 调度策略功能介绍

功能模块	功能概述
自定义调度策略	开放节点亲和、工作负载亲和以及工作负载反亲和调度策略的配置，以满足用户的更高需求。在自定义调度策略中用户可以设置“节点亲和性”、“工作负载亲和性”和“工作负载反亲和性”。
简易调度策略	提供简单便捷以及足够功能的调度方式。简易调度策略提供工作负载和可用区的亲和性、工作负载和节点的亲和性以及工作负载间的亲和性调度，用户可根据业务需求进行相应的设置部署工作负载。

网络访问方式

云容器引擎通过将Kubernetes网络和华为云VPC深度集成，提供了稳定高性能的网络访问方式，能够满足多种复杂场景下工作负载间的互相访问。

表 5-6 网络管理功能介绍

功能模块	功能概述
Service	Service是一种资源，提供了我们访问单个或多个容器应用的能力。每个服务在其生命周期内，都拥有一个固定的IP地址和端口。每个服务对应了后台的一个或多个Pod，通过这种方式，客户端就不需要关心Pod所在的位置，方便后端进行Pod扩容、缩容等操作。 支持的Service类型包括： ● 集群内访问（ClusterIP）：仅在集群内访问服务。 ● 节点访问（NodePort）：使用节点私有IP或弹性公网IP访问。 ● 负载均衡（LoadBalancer）：使用弹性负载均衡器访问服务。 ● DNAT网关（DNAT）：通过DNAT网关访问服务。
七层负载均衡（Ingress）	七层负载均衡（Ingress）是采用了共享型弹性负载均衡和独享型弹性负载均衡，在四层负载均衡访问方式的基础上支持了URI配置，通过对应的URI将访问流量分发到对应的服务。同时，服务根据不同URI实现不同的功能。

功能模块	功能概述
网络策略（NetworkPolicy）	基于Kubernetes的网络策略功能进行了加强，通过配置网络策略，允许在同个集群内实现网络的隔离，也就是可以在某些实例（Pod）之间架起防火墙。使用场景例如：某个用户有支付系统，且严格要求只能某几个组件能访问该支付系统，否则有被攻破的安全风险，通过配置网络策略可免除该风险。
网络平面(NetworkAttachmentDefinition)	网络平面（NetworkAttachmentDefinition）是集群的一种crd资源，为容器对接ENI（Elastic Network Interface，弹性网络接口）提供配置项，如VPC，子网等。关联网络平面的工作负载支持对接弹性网卡服务，容器能直接绑定弹性网卡，并对外提供服务。

持久化存储卷

云容器引擎除支持本地磁盘存储外，还支持将工作负载数据存储在华为云的云存储上，当前支持的云存储包括：云硬盘存储卷（EVS）、文件存储卷（SFS）、对象存储卷（OBS）和极速文件存储卷（SFS Turbo）。

表 5-7 存储管理功能介绍

功能模块	功能概述
本地磁盘存储	通过本地磁盘存储将容器所在宿主机的文件目录挂载到容器的指定路径中（对应Kubernetes的HostPath），也可以不填写源路径（对应Kubernetes的EmptyDir），不填写时将分配主机的临时目录挂载到容器的挂载点，指定源路径的本地硬盘数据卷适用于将数据持久化存储到容器所在宿主机，EmptyDir（不填写源路径）适用于容器的临时存储。
云硬盘存储卷	支持将云硬盘（EVS）挂载到容器中。通过云硬盘，可以将存储系统的远端文件目录挂载到容器中，数据卷中的数据将被永久保存，即使删除了容器，数据卷中的数据依然保存在存储系统中。
文件存储卷	支持创建SFS存储卷并挂载到容器的某一路径下，也可以使用底层SFS服务创建的文件存储卷，SFS存储卷适用于多读多写的持久化存储，适用于多种工作负载场景，包括媒体处理、内容管理、大数据分析和分析工作负载程序等场景。
对象存储卷	支持创建OBS对象存储卷并挂载到容器的某一路径下，对象存储适用于云工作负载、数据分析、内容分析和热点对象等场景。
极速文件存储卷	CCE支持创建SFS Turbo极速文件存储卷并挂载到容器的某一路径下，极速文件存储具有按需申请，快速供给，弹性扩展，方便灵活等特点，适用于DevOps、容器微服务、企业办公等应用场景。

功能模块	功能概述
快照与备份	通过EVS服务为您提供快照功能，云硬盘快照简称快照，指云硬盘数据在某个时刻的完整拷贝或镜像，是一种重要的数据容灾手段，当数据丢失时，可通过快照将数据完整的恢复到快照时间点。

插件扩展

CCE提供了多种类型的系统插件，用于管理集群的扩展功能，以支持选择性扩展满足特性需求的功能。

- 提供OpenAPI和社区原生API。
- 提供Kubectl插件和社区原生Kubectl工具。

生态工具

云容器引擎深度集成应用服务网格和Kubernetes Helm标准模板。

表 5-8 Kubernetes 生态

功能模块	功能概述
应用服务网格	提供非侵入式的微服务治理解决方案，支持完整的生命周期管理和流量治理能力，兼容Kubernetes和Istio生态。您无需修改任何服务代码，也无需手动安装代理，只需开启应用服务网格功能，即可实现丰富的服务治理能力。
模板市场	模板市场是CCE基于Kubernetes Helm标准的模板提供统一的资源管理与调度，高效地实现了模板的快速部署与后期管理，大幅简化了Kubernetes资源的安装管理过程。CCE提供的模板市场功能包括：示例模板和我的模板。 • 示例模板：使用社区Helm开源镜像，提供基础的容器集群体验与模板体验功能，当前支持redis、etcd、mysql-ndb、mongodb、istio、zookeeper、elasticsearch、kibana等示例模板。 • 我的模板：通过自定义Helm模板来简化工作负载部署的服务。

弹性伸缩

CCE支持集群节点、工作负载的弹性伸缩，支持手动伸缩和自动弹性伸缩，并可以自由组合多种弹性策略以应对业务高峰期的突发流量浪涌。

表 5-9 弹性伸缩功能介绍

功能模块	功能概述
工作负载伸缩	提供HPA策略和CustomedHPA策略两种创建方式。 ● HPA策略：即Horizontal Pod Autoscaling，是Kubernetes中实现POD水平自动伸缩的功能。该策略在kubernetes社区HPA功能的基础上，增加了HPA级别的冷却时间窗和扩缩容阈值等功能。 ● CustomedHPA策略：华为云自研的弹性伸缩增强能力，能够基于指标（CPU利用率、内存利用率）或周期（每天、每周、每月或每年的具体时间点），对无状态工作负载进行弹性扩缩容。
节点伸缩	通过节点自动伸缩组件autoscaler实现的，可以按需弹出节点实例，支持多可用区、多实例规格、多种伸缩模式，满足不同的节点伸缩场景。

权限管理

权限管理是在统一身份认证服务（IAM）与Kubernetes的角色访问控制（RBAC）的能力基础上，打造的细粒度权限管理功能，支持基于IAM的细粒度权限控制和IAM Token认证，支持集群级别、命名空间级别的权限控制，帮助用户便捷灵活的对租户下的IAM用户、用户组设定不同的操作权限。

表 5-10 权限管理功能介绍

功能模块	功能概述
集群权限	CCE集群权限是基于IAM**系统策略**和**自定义策略**的授权，可以通过用户组功能实现IAM用户的授权
命名空间权限	基于Kubernetes RBAC能力的授权，通过权限设置可以让不同的用户或用户组拥有操作不同Kubernetes资源的权限。同时CCE基于开源能力进行了增强，可以支持基于IAM用户或用户组粒度进行RBAC授权、IAM token直接访问API进行RBAC认证鉴权。

系统管家

CCE提供的系统管家功能包括：系统体检和系统加固。

表 5-11 系统管家功能介绍

功能模块	功能概述
系统体检	主要用于实时检测并发现节点上的一些故障或者异常情况。

功能模块	功能概述
系统加固	主要用于对一些系统组件（如coredns插件）进行加固。当前已支持coredns自动水平伸缩，根据coredns的请求量情况自动伸缩其实例个数，以防止请求量过大导致coredns解析性能下降或者解析超时失败。

容器 DevOps 能力

配合**容器镜像服务**提供容器自动化交付流水线，您无需编写Dockerfile与Kubernetes Manifest，基于ContainerOps流水线模板可以自定义企业级容器DevOps流程，大幅提升容器交付效率。

6 产品优势

云容器引擎的优势

云容器引擎是基于业界主流的Docker和Kubernetes开源技术构建的容器服务，提供众多契合企业大规模容器集群场景的功能，在系统可靠性、高性能、开源社区兼容性等多个方面具有独特的优势，满足企业在构建容器云方面的各种需求。

简单易用

- 通过WEB界面一键创建Kubernetes集群，支持管理虚拟机节点或裸金属节点，支持虚拟机与物理机混用场景。
- 一站式自动化部署和运维容器应用，整个生命周期都在容器服务内一站式完成。
- 通过Web界面轻松实现集群节点和工作负载的扩容和缩容，自由组合策略以应对多变的突发浪涌。
- 通过Web界面一键完成Kubernetes集群的升级。
- 深度集成应用服务网格和Helm标准模板，真正实现开箱即用。

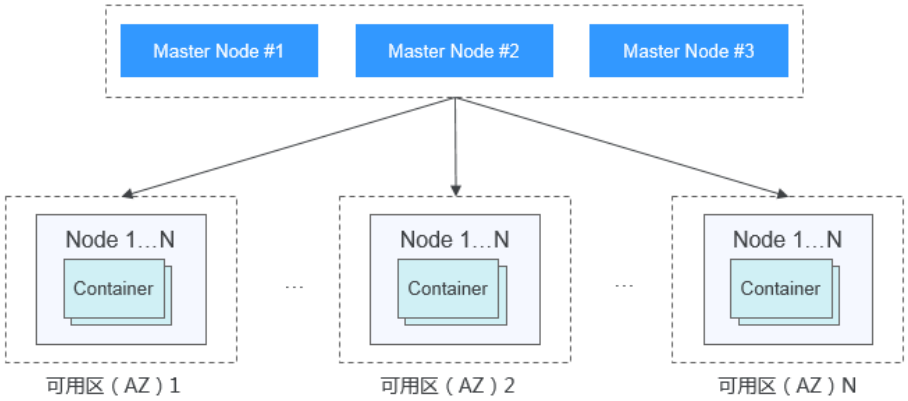
高性能

- 基于华为在计算、网络、存储、异构等方面多年的行业技术积累，提供业界领先的高性能云容器引擎，支撑您业务的高并发、大规模场景。
- 采用高性能裸金属NUMA架构和高速IB网卡，AI计算性能提升3-5倍以上。

安全可靠

- 高可靠：集群控制面支持3 Master HA高可用，当其中某个或者两个控制节点故障时，集群依然可用，从而保障您的业务高可用。集群内节点和工作负载支持跨可用区（AZ）部署，帮助您轻松构建多活业务架构，保证业务系统在主机故障、机房中断、自然灾害等情况下可持续运行，获得生产环境的高稳定性，实现业务系统零中断。

图 6-1 集群高可用



- 高安全：私有集群，完全由用户掌控，并深度整合华为云帐号和Kubernetes RBAC能力，支持用户在界面为子用户设置不同的RBAC权限。

开放兼容

- 云容器引擎在Docker技术的基础上，为容器化的应用提供部署运行、资源调度、服务发现和动态伸缩等一系列完整功能，提高了大规模容器集群管理的便捷性。
- 云容器引擎基于业界主流的Kubernetes实现，完全兼容Kubernetes/Docker社区原生版本，与社区最新版本保持紧密同步，完全兼容Kubernetes API和Kubectl。

云容器引擎对比自建 Kubernetes 集群

表 6-1 云容器引擎和自建 kubernetes 集群对比

对比项	自建kubernetes集群	云容器引擎
易用性	自建kubernetes集群管理基础设施通常涉及安装、操作、扩展自己的集群管理软件、配置管理系统和监控解决方案，管理复杂。每次升级集群的过程都是巨大的调整，带来繁重的运维负担。	**简化集群管理，简单易用** 借助云容器引擎，您可以一键创建和升级Kubernetes容器集群，无需自行搭建Docker和Kubernetes集群。您可以通过云容器引擎自动化部署和一站式运维容器应用，使得应用的整个生命周期都在容器服务内高效完成。 您可以通过云容器引擎轻松使用深度集成的应用服务网格和Helm标准模板，真正实现开箱即用。 您只需启动容器集群，并指定想要运行的任务，云容器引擎帮您完成所有的集群管理工作，让您可以集中精力开发容器化的应用程序。
可扩展性	自建kubernetes集群需要根据业务流量情况和健康情况人工确定容器服务的部署，可扩展性差。	**灵活集群托管，轻松实现扩缩容** 云容器引擎可以根据资源使用情况轻松实现集群节点和工作负载的自动扩容和缩容，并可以自由组合多种弹性策略，以应对业务高峰期的突发流量浪涌。

对比项	自建kubernetes集群	云容器引擎
可靠性	自建kubernetes集群多采用单控制节点，一旦出现故障，集群和业务将不可使用。	**服务高可用** 创建集群时若“高可用”选项配置为“是”，集群将创建3个Master节点，在单个控制节点发生故障时，集群仍然可用，从而保障您的业务高可用。
高效性	自建kubernetes集群需要自行搭建镜像仓库或使用第三方镜像仓库，镜像拉取方式多采用串行传输，效率低。	**镜像快速部署，业务持续集成** 云容器引擎配合**容器镜像服务**提供容器自动化交付流水线，您无需编写Dockerfile与Kubernetes Manifests，基于ContainerOps流水线模板可以自定义企业级容器DevOps流程，镜像拉取方式采用并行传输，大幅提升容器交付效率。
成本	自建kubernetes集群需要投入资金构建、安装、运维、扩展自己的集群管理基础设施，成本开销大。	**云容器引擎成本低** 您只需支付用于存储和运行应用程序的基础设施资源（例如云服务器、云硬盘、弹性IP/带宽、负载均衡等）费用和容器集群控制节点费用。

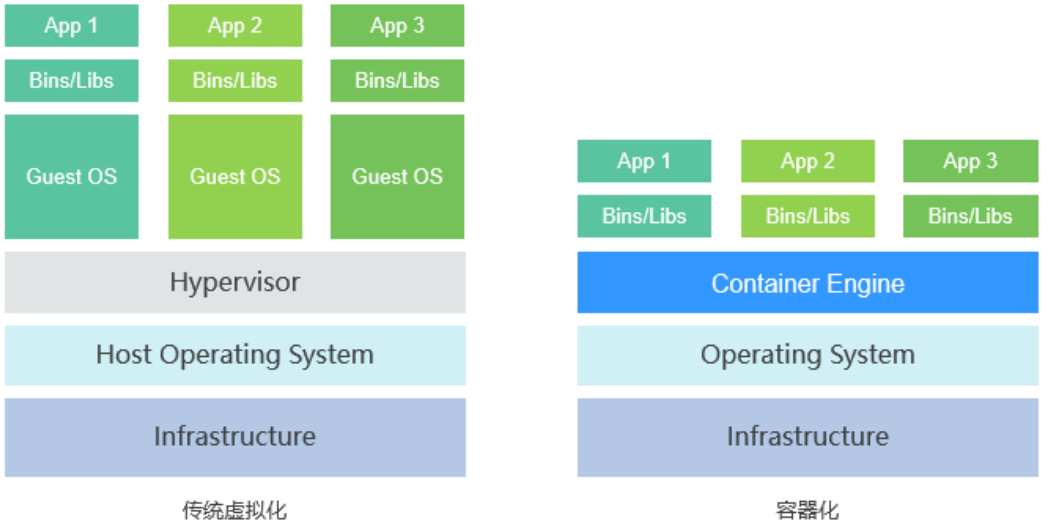
容器的优势

Docker使用Google公司推出的Go语言进行开发实现，基于Linux内核的cgroup，namespace，以及AUFS类的Union FS等技术，对进程进行封装隔离，属于操作系统层面的虚拟化技术。由于隔离的进程独立于宿主和其它的隔离的进程，因此也称其为容器。

Docker在容器的基础上，进行了进一步的封装，从文件系统、网络互联到进程隔离等，极大的简化了容器的创建和维护。

传统虚拟机技术是虚拟出一套硬件后，在其上运行一个完整操作系统，在该系统上再运行所需应用进程；而容器内的应用进程直接运行于宿主的内核，容器内没有自己的内核，而且也没有进行硬件虚拟。因此使得Docker技术比虚拟机技术更为轻便、快捷。

图 6-2 传统虚拟化和容器化方式的对比



作为一种新兴的虚拟化方式，Docker跟虚拟机相比具有众多的优势：

更高效的利用系统资源

由于容器不需要进行硬件虚拟以及运行完整操作系统等额外开销，Docker对系统资源的利用率更高。无论是应用执行速度、内存损耗或者文件存储速度，都要比传统虚拟机技术更高效。因此，相比虚拟机技术，一个相同配置的主机，往往可以运行更多数量的应用。

更快速的启动时间

传统的虚拟机技术启动应用服务往往需要数分钟，而Docker容器应用，由于直接运行于宿主内核，无需启动完整的操作系统，因此可以做到秒级、甚至毫秒级的启动时间。大大的节约了开发、测试、部署的时间。

一致的运行环境

开发过程中一个常见的问题是环境一致性问题。由于开发环境、测试环境、生产环境不一致，导致有些bug并未在开发过程中被发现。而Docker的镜像提供了除内核外完整的运行时环境，确保了应用运行环境一致性。

持续交付和部署

对开发和运维（DevOps）人员来说，最希望的就是一次创建或配置，可以在任意地方正常运行。

使用Docker可以通过定制应用镜像来实现持续集成、持续交付、部署。开发人员可以通过Dockerfile来进行镜像构建，并结合持续集成（Continuous Integration）系统进行集成测试，而运维人员则可以直接在生产环境中快速部署该镜像，甚至结合持续部署（Continuous Delivery/Deployment） 系统进行自动部署。

而且使用Dockerfile使镜像构建透明化，不仅开发团队可以理解应用运行环境，也方便运维团队理解应用运行所需条件，帮助更好的生产环境中部署该镜像。

更轻松的迁移

由于Docker确保了执行环境的一致性，使得应用的迁移更加容易。Docker可以在很多平台上运行，无论是物理机、虚拟机、公有云、私有云，甚至是笔记本，其运行结果

是一致的。因此用户可以很轻易的将在一个平台上运行的应用，迁移到另一个平台上，而不用担心运行环境的变化导致应用无法正常运行的情况。

更轻松的维护和扩展

Docker使用的分层存储以及镜像的技术，使得应用重复部分的复用更为容易，也使得应用的维护更新更加简单，基于基础镜像进一步扩展镜像也变得非常简单。此外，Docker团队同各个开源项目团队一起维护了一大批高质量的官方镜像，既可以直接在生产环境使用，又可以作为基础进一步定制，大大的降低了应用服务的镜像制作成本。

表 6-2 容器对比传统虚拟机总结

特性	容器	虚拟机
启动	秒级	分钟级
硬盘使用	一般为MB	一般为GB
性能	接近原生	弱
系统支持量	单机支持上千个容器	一般几十个

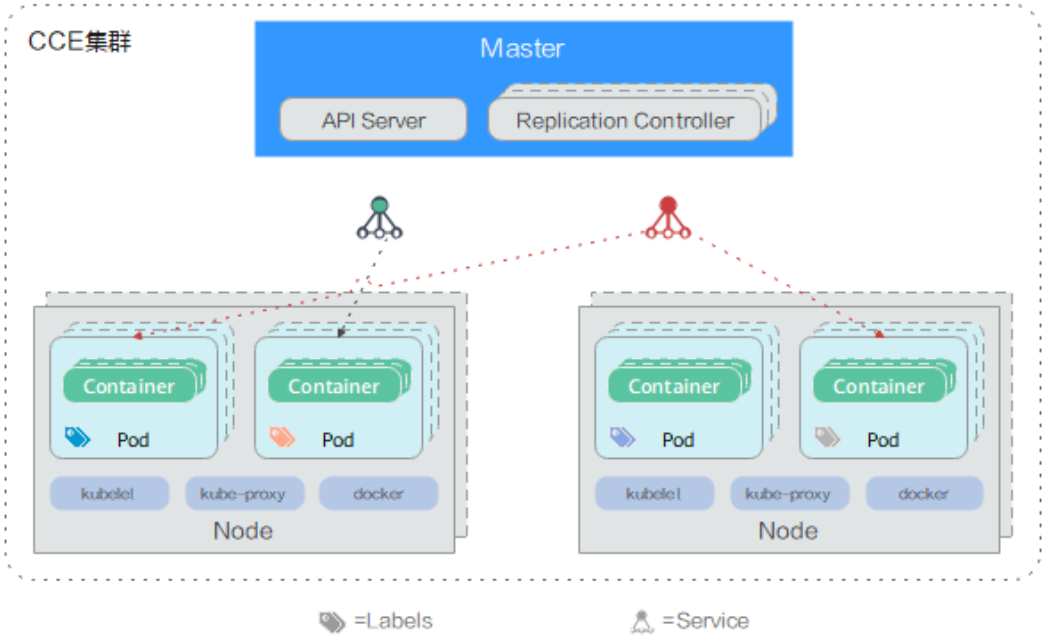
7 应用场景

7.1 基础设施与容器应用管理

应用场景

CCE集群支持管理X86资源池和ARM资源池，能方便的创建Kubernetes集群、部署您的容器化应用，以及方便的管理和维护。

图 7-1 CCE 集群



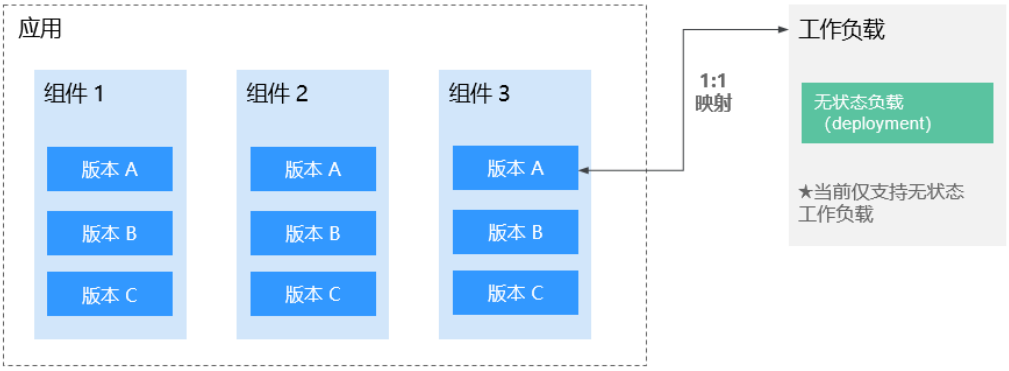
价值

通过容器化改造，使应用部署资源成本降低，提升应用的部署效率和升级效率，可以实现升级时业务不中断以及统一的自动化运维。

优势

- 多种类型的容器部署
  支持部署无状态工作负载、有状态工作负载、守护进程集、普通任务、定时任务等。
- 应用升级
  支持替换升级、滚动升级（按比例、实例个数进行滚动升级）；支持升级回滚。
- 弹性伸缩
  支持节点和工作负载的弹性伸缩。

图 7-2 工作负载



7.2 秒级弹性伸缩

应用场景

- 电商客户遇到促销、限时秒杀等活动期间，访问量激增，需及时、自动扩展云计算资源。
- 视频直播客户业务负载变化难以预测，需要根据CPU/内存使用率进行实时扩缩容。
- 游戏客户每天中午12点及晚上18:00-23:00间需求增长，需要定时扩容。

价值

云容器引擎可根据用户的业务需求预设策略自动调整计算资源，使云服务器或容器数量自动随业务负载增长而增加，随业务负载降低而减少，保证业务平稳健康运行，节省成本。

优势

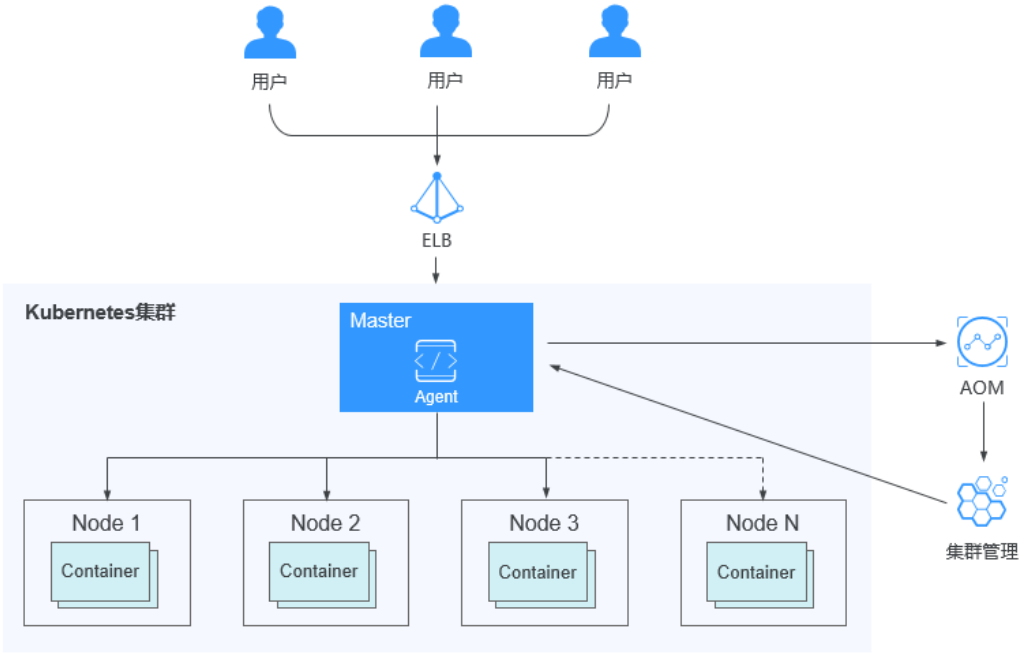
- 自由灵活
  支持多种策略配置，业务流量达到扩容指标，秒级触发容器扩容操作。
- 高可用
  自动检测伸缩组中实例运行状况，启用新实例替换不健康实例，保证业务健康可用。
- 低成本

只按照实际用量收取云服务器费用。

建议搭配使用

autoscaler插件（集群自动扩缩容）+应用运维管理AOM（工作负载伸缩）

图 7-3 弹性伸缩场景



7.3 微服务流量治理

应用场景

伴随着互联网技术的不断发展，各大企业的系统越来越复杂，传统的系统架构越来越不能满足业务的需求，取而代之的是微服务架构。微服务是将复杂的应用切分为若干服务，每个服务均可以独立开发、部署和伸缩；微服务和容器组合使用，可进一步简化微服务的交付，提升应用的可靠性和可伸缩性。

随着微服务的大量应用，其构成的分布式应用架构在运维、调试、和安全管理等维度变得更加复杂，在管理微服务时，往往需要在业务代码中添加微服务治理相关的代码，导致开发人员不能专注于业务开发，还需要考虑微服务治理的解决方案，并且将解决方案融合到其业务系统中。

价值

云容器引擎深度集成应用服务网格，提供开箱即用的应用服务网格流量治理能力，用户无需修改代码，即可实现灰度发布、流量治理和流量监控能力。

优势

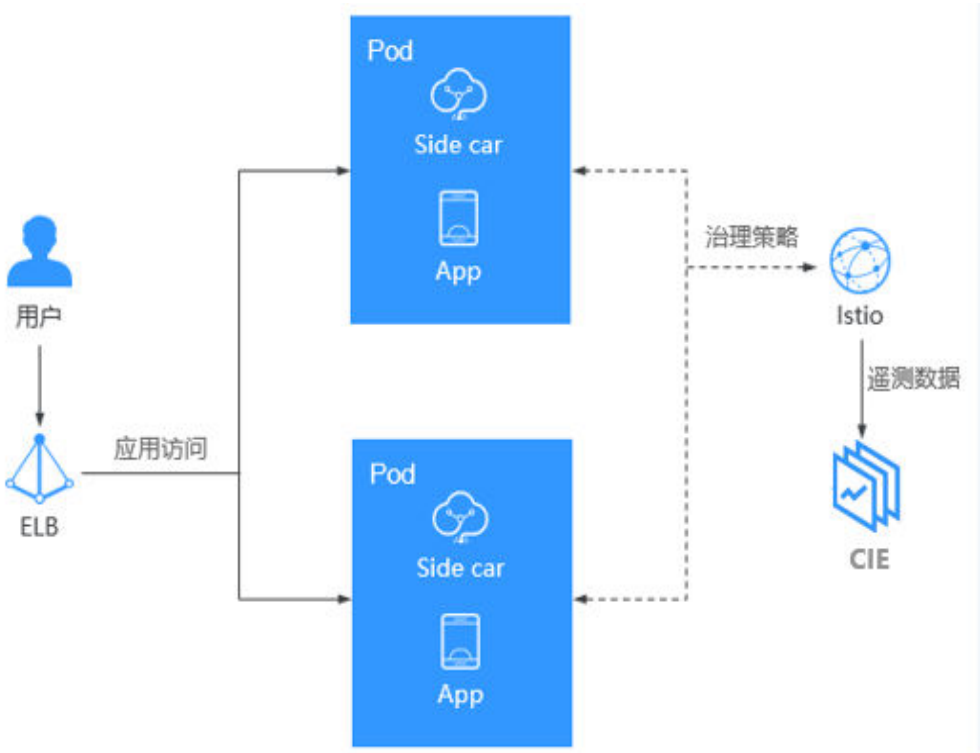
- 开箱即用
  与云容器引擎无缝对接，一键开启后即可提供非侵入的智能流量治理解决方案。

- 策略化智能路由
  无需修改代码，即可实现HTTP、TCP等服务连接策略和安全策略。
- 流量治理可视化
  基于无侵入的监控数据采集，深度整合华为云APM能力，提供实时流量拓扑、调用链等服务性能监控和运行诊断，构建全景的服务运行视图，可实时、一站式观测服务流量健康和性能状态。

建议搭配使用

弹性负载均衡ELB + 应用性能管理APM + 应用运维管理AOM

图 7-4 微服务治理场景



7.4 DevOps 持续交付

应用场景

当前IT行业发展日益快速，面对海量需求必须具备快速集成的能力。经过快速持续集成，才能保证不间断的补全用户体验，提升服务质量，为业务创新提供源源不断的动力。大量交付实践表明，不仅传统企业，甚至互联网企业都可能在持续集成方面存在研发效率低、工具落后、发布频率低等方面的问题，需要通过持续交付提高效率，降低发布风险。

价值

云容器引擎搭配容器镜像服务提供DevOps持续交付能力，能够基于代码源自动完成代码编译、镜像构建、灰度发布、容器化部署，实现一站式容器化交付流程，并可对接已有CI/CD，完成传统应用的容器化改造和部署。

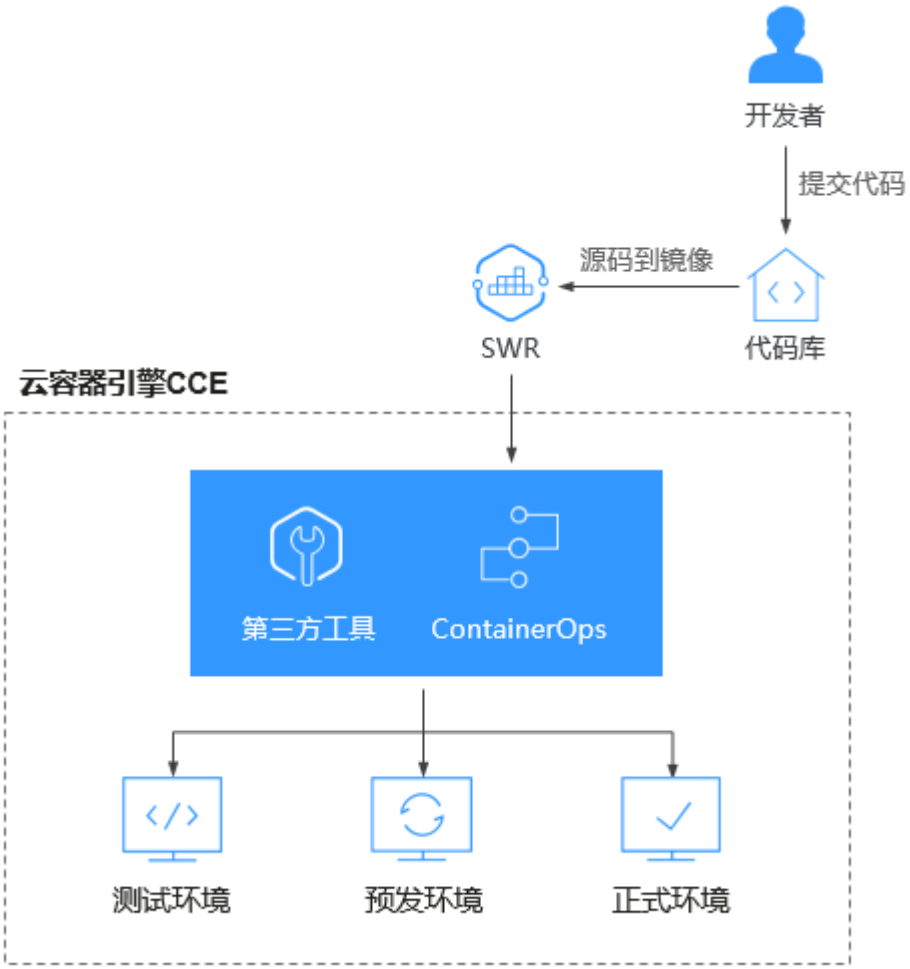
优势

- 高效流程管理
  更优的流程交互设计，脚本编写量较传统CI/CD流水线减少80%以上，让CI/CD管理更高效。
- 灵活的集成方式
  提供丰富的接口便于与企业已有CI/CD系统进行集成，灵活适配企业的个性化诉求。
- 高性能
  全容器化架构设计，任务调度更灵活，执行效率更高。

建议搭配使用

容器镜像服务SWR + 对象存储服务OBS + 虚拟专用网络VPN

图 7-5 DevOps 持续交付场景



7.5 混合云架构

应用场景

- 多云部署、容灾备份

  为保证业务高可用，需要将业务同时部署在多个云的容器服务上，在某个云出现事故时，通过统一流量分发的机制，自动的将业务流量切换到其他云上。

- 流量分发、弹性伸缩

  大型企业客户需要将业务同时部署在不同地域的云机房中，并能自动弹性扩容和缩容，以节约成本。

- 业务上云、数据库托管

  对于金融、安全等行业用户，业务数据的敏感性要求将数据业务保留在本地的IDC中而将一般业务部署在云上，并需要进行统一管理。

- 开发与部署分离

  出于IP安全的考虑，用户希望将生产环境部署在公有云上，而将开发环境部署在本地的IDC。

价值

云容器引擎利用容器环境无关的特性，将私有云和公有云容器服务实现网络互通和统一管理，应用和数据可在云上云下无缝迁移，并可统一运维多个云端资源，从而实现资源的灵活使用以及业务容灾等目的。

优势

- 云上容灾

  通过云容器引擎，可以将业务系统同时部署在多个云的容器服务上，统一流量分发，单云故障后能够自动将业务流量切换到其他云上，并能快速自动解决现网事故。

- 流量自动分发

  通过云容器引擎的统一流量分发机制，实现应用访问流量的地域亲和，降低业务访问时延，并需要能够将线下IDC中的业务在云上扩展，可根据业务流量峰值情况，自动弹性扩容和缩容。

- 计算与数据分离，能力共享

  通过华为云容器引擎，用户可以实现敏感业务数据与一般业务数据的分离，可以实现开发环境和生产环境分离，可以实现特殊计算能力与一般业务的分离，并能够实现弹性扩展和集群的统一管理，达到云上云下资源和能力的共享。

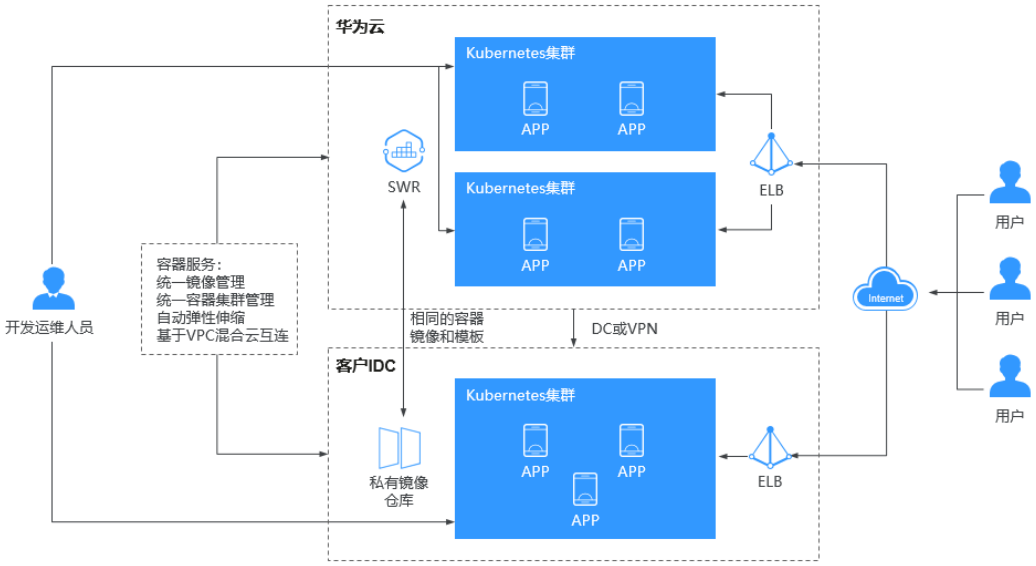
- 降低成本

  业务高峰时，利用公有云资源池快速扩容，用户不再需要根据流量峰值始终保持和维护大量资源，节约成本。

建议搭配使用

弹性云服务器ECS + 云专线DC + 虚拟专用网络VPN + 容器镜像服务SWR

图 7-6 混合云场景



7.6 高性能调度

CCE通过集成Volcano提供高性能计算能力。

Volcano是基于Kubernetes的批处理系统，源自于华为云AI容器。Volcano提供了一个针对BigData和AI场景下，通用、可扩展、高性能、稳定的原生批量计算平台，方便AI、大数据、基因、渲染等诸多行业通用计算框架接入，提供高性能任务调度引擎，高性能异构芯片管理，高性能任务运行管理等能力。

应用场景 1：多类型作业混合部署

随着各行各业的发展，涌现出越来越多的领域框架来支持业务的发展，这些框架都在相应的业务领域有着不可替代的作用，例如Spark，Tensorflow，Flink等。在业务复杂性能不断增加的情况下，单一的领域框架很难应对现在复杂的业务场景，因此现在普遍使用多种框架达成业务目标。但随着各个领域框架集群的不断扩大，以及单个业务的波动性，各个子集群的资源浪费比较严重，越来越多的用户希望通过统一调度系统来解决资源共享的问题。

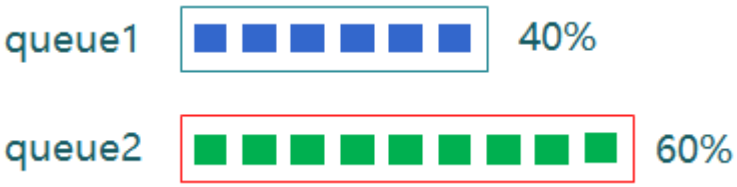
Volcano在Kubernetes之上抽象了一个批量计算的通用基础层，向下弥补Kubernetes调度能力的不足，向上提供灵活通用的Job抽象。Volcano通过提供多任务模板功能实现了利用Volcano Job描述多种作业类型（Tensorflow、Spark、MPI、PyTorch等），并通过Volcano统一调度系统实现多种作业混合部署，解决集群资源共享问题。

应用场景 2：多队列场景调度优化

用户在使用集群资源的时候通常会涉及到资源隔离与资源共享，Kubernetes中没有队列的支持，所以它在多个用户或多个部门共享一个机器时无法做资源共享。但不管在HPC还是大数据领域中，通过队列进行资源共享都是基本的需求。

在通过队列做资源共享时，我们提供了多种机制。可以为队列设置weight值，集群通过计算该队列weight值占所有weight总和的比例来给队列划分资源；另外也可以为队列设置资源的Capability值，来确定该队列能够使用的资源上限。

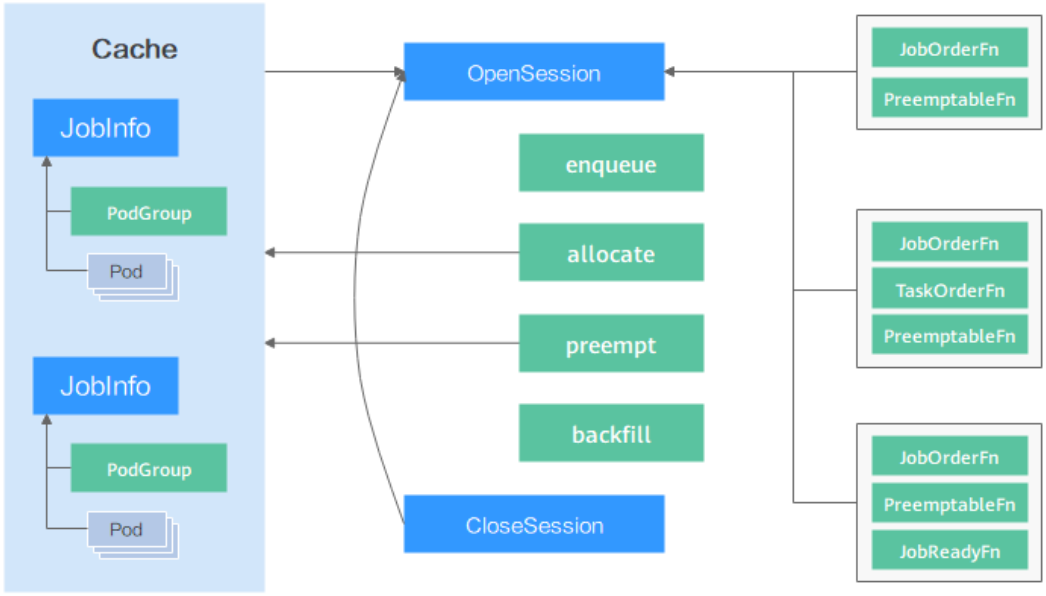
例如下图中，通过这两个队列去共享整个集群的资源，一个队列获得40%的资源，另一个队列获得60%的资源，这样可以把两个不同的队列映射到不同的部门或者是不同的项目中。并且在一个队列里如果有多余的空闲资源，可以把这些空闲资源分配给另外一个队列里面的作业去使用。



应用场景 3：多种高级调度策略

当用户向Kubernetes申请容器所需的计算资源（如CPU、Memory、GPU等）时，调度器负责挑选出满足各项规格要求的节点来部署这些容器。通常，满足各项要求的节点并非唯一，且水位（节点已有负载）各不相同，不同的分配方式最终得到的分配率存在差异，因此，调度器的一项核心任务就是以最终资源利用率最优的目标从众多候选机器中挑出最合适的节点。

下图为Volcano scheduler调度流程，首先将API server中的Pod、PodGroup信息加载到scheduler cache中。Scheduler周期被称为session，每个scheduler周期会经历OpenSession，调用Action，CloseSession三个阶段。其中OpenSession阶段加载用户配置的scheduler plugin中实现的调度策略；调用Action阶段逐一调用配置的action以及在OpenSession阶段加载的调度策略；CloseSession为清理阶段。



Volcano scheduler通过插件方式提供了多种调度Action（例如enqueue，allocate，preempt，reclaim，backfill）以及调度策略（例如gang，priority，drf，proportion，binpack等），用户可以根据实际业务需求进行配置。通过实现Scheduler提供的接口也可以方便灵活的进行定制化开发。

应用场景 4：高精度资源调度

Volcano 在支持AI，大数据等作业的时候提供了高精度的资源调度策略，例如在深度学习场景下计算效率非常重要。以TensorFlow计算为例，配置“ps”和“worker”之间的亲和性，以及“ps”与“ps”之间的反亲和性，可使“ps”和“worker”尽量调度到同一台节点上，从而提升“ps”和“worker”之间进行网络和数据交互的效率，进而提升计算效率。然而Kubernetes默认调度器在调度Pod过程中，仅会检查Pod与现有集群下所有已经处于运行状态Pod的亲和性和反亲和性配置是否冲突或吻合，并不会考虑接下来可能会调度的Pod造成的影响。

Volcano提供的Task-topology算法是一种根据Job内task之间亲和性和反亲和性配置计算task优先级和Node优先级的算法。通过在Job内配置task之间的亲和性和反亲和性策略，并使用task-topology算法，可优先将具有亲和性配置的task调度到同一个节点上，将具有反亲和性配置的Pod调度到不同的节点上。同样是处理亲和性和反亲和性配置对Pod调度的影响，task-topology算法与Kubernetes默认调度器处理的不同点在于，task-topology将待调度的Pods作为一个整体进行亲和性和反亲和性考虑，在批量调度Pod时，考虑未调度Pod之间的亲和性和反亲和性影响，并通过优先级施加到Pod的调度进程中。

价值

面向AI计算的容器服务，采用高性能GPU计算实例，并支持多容器共享GPU资源，在AI计算性能上比通用方案提升3~5倍以上，并大幅降低了AI计算的成本，同时帮助数据工程师在集群上轻松部署计算应用，您无需关心复杂的部署运维，专注核心业务，快速实现从0到1快速上线。

优势

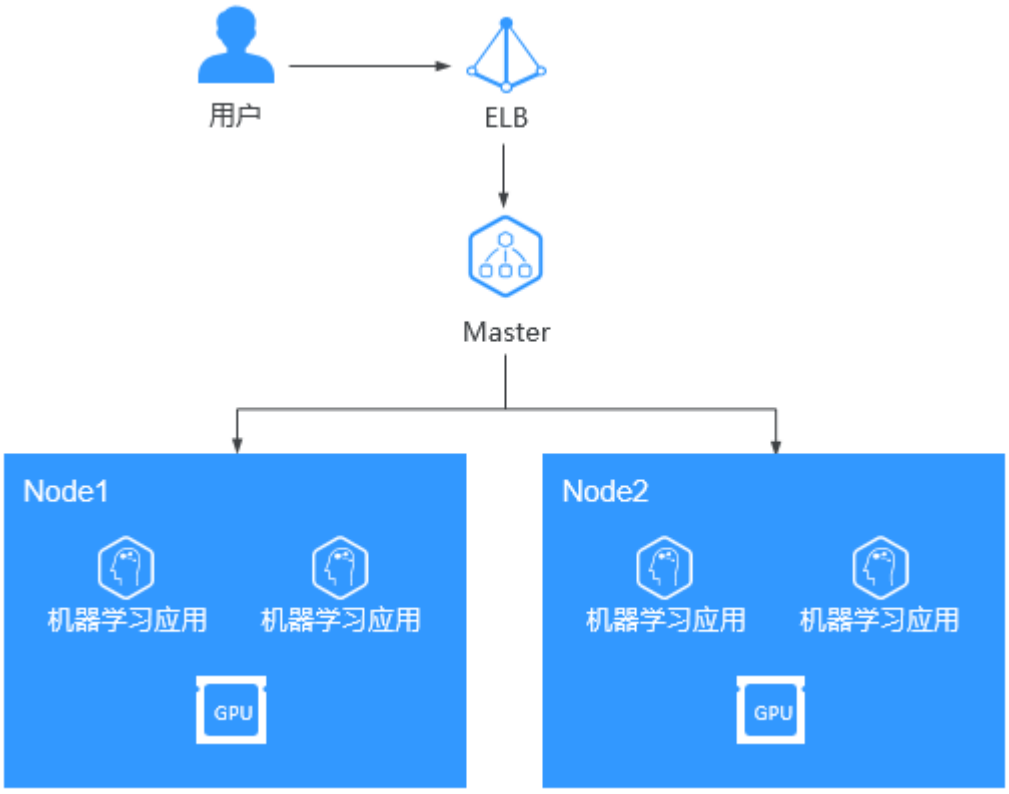
CCE通过集成Volcano，在高性能计算、大数据、AI等领域有如下优势：

- 多种类型作业混合部署：支持AI、大数据、HPC作业类型混合部署。
- 多队列场景调度优化：支持多队列用于多租资源共享与分组规划，支持优先级与分时复用。
- 多种高级调度策略：支持gang-scheduling、公平调度、资源抢占、GPU拓扑等高级调度策略。
- 多任务模板：支持单一Job多任务模板定义，打破Kubernetes原生资源束缚，Volcano Job描述多种作业类型（Tensorflow、MPI、PyTorch等）。
- 作业扩展插件配置：在提交作业、创建Pod等多个阶段，Controller支持配置插件用来执行自定义的环境准备和清理的工作，比如常见的MPI作业，在提交前就需要配置SSH插件，用来完成Pod资源的SSH信息配置。

建议搭配使用

GPU加速云服务器 + 弹性负载均衡ELB + 对象存储服务OBS

图 7-7 AI 计算



8 约束与限制

本文主要为您介绍华为云云容器引擎（CCE）集群使用过程中的一些限制。

集群/节点限制

- 集群一旦创建以后，不支持变更以下项：
  - 变更集群类型，例如“鲲鹏集群”更换为“CCE集群”。
  - 变更集群的控制节点数量。
  - 变更控制节点可用区。
  - 变更集群的网络配置，如所在的虚拟私有云VPC、子网、容器网段、服务网段、IPv6、kubeproxy代理（转发）模式。
  - 变更网络模型，例如“容器隧道网络”更换为“VPC网络”。
- 不支持应用在不同命名空间下迁移。
- 创建的ECS实例（节点）目前支持“按需计费”和“包年/包月”，其他资源（例如负载均衡）为按需计费，您可以通过管理控制台将按需计费实例转换成包年/包月实例。
- 集群创建过程中创建的节点支持“按需计费”和“包年包月”，但有如下限制：
  - 如果创建的集群是“按需计费”，那么在该集群下创建的节点只能为“按需计费”。
  - 如果创建的集群是“包年包月”，那么该集群下的节点可以为“按需计费”或者“包年包月”。
  - 纳管的节点在“包年包月”场景下，无法通过集群为其续费，需用户单独续费。

  备注：集群创建完成后再购买节点时，节点的付费方式不受集群的付费方式限制。
- 由于ECS（节点）等底层依赖产品配额及库存限制，创建集群、扩容集群或者自动弹性扩容时，可能只有部分节点创建成功。
- ECS（节点）规格要求：CPU > 2核且内存 > 4GiB。
- 通过搭建VPN方式访问CCE集群，需要注意VPN网络和集群所在的VPC网段、容器使用网段不能冲突。

网络

- 节点访问(NodePort)的使用约束：默认为VPC内网访问，如果需要使用弹性IP通过公网访问该服务，请提前在集群的节点上绑定弹性IP。
- CCE中的负载均衡（LoadBalancer）访问类型使用弹性负载均衡 ELB提供网络访问，存在如下产品约束：
  - 自动创建的ELB实例建议不要被其他资源使用，否则会在删除时被占用，导致资源残留。
  - 正在使用的ELB实例请不要修改监听器名称，否则可能导致无法正常访问。
- 网络策略(NetworkPolicy)，存在如下产品约束：
  - 当前仅容器隧道网络模式的集群支持网络策略（NetworkPolicy）。
  - 网络策略（NetworkPolicy）暂不支持设置出方向（egress）。
  - 不支持对IPv6地址网络隔离。
  - v1.13及v1.15版本的容器隧道网络类型的集群，节点操作系统内核为Centos时，如果使用NetworkPolicy请升级openvswitch的版本，升级方法请参考**操作系统内核升级**。
- 网络平面(NetworkAttachmentDefinition)：
  - VPC网络模型的集群使用ENI为受限功能，未全面开放，如有使用ENI需求请创建CCE Turbo集群。
  - 仅网络模型为VPC网络（且未开启IPv6）的集群支持创建网络平面；网络模型为容器隧道网络时列表中仅显示“default-network”，不能新增或修改。
  - 需v1.13.7-r0及以上版本的集群才能启用，v1.13.7-r0以下版本集群需要升级到最新版本后才能启用。

存储卷

- 云硬盘存储卷使用约束：
  - CCE默认创建计费模式为“按需计费”的云硬盘。如需使用包周期的云硬盘，请参考**云硬盘包周期**。
  - 云硬盘不支持跨可用区挂载，且暂时不支持被多个工作负载、同一个工作负载的多个实例或多个任务使用。
  - 由于CCE集群各节点之间暂不支持共享盘的数据共享功能，多个节点挂载使用同一个云硬盘可能会出现读写冲突、数据缓存冲突等问题，所以创建无状态工作负载时，若使用了EVS云硬盘，建议工作负载只选择一个实例。
  - 创建有状态工作负载并添加云存储时，云硬盘暂不支持使用已有存储。
  - 不支持导入分区过或者具有非ext4文件系统的云硬盘。
  - CCE集群中的容器存储目前已支持加密（Kubernetes 1.13版本及以上），当前仅在部分区域（Region）提供端到端支持。
  - 存储不支持选择企业项目，新创建的存储卷默认创建到default企业项目下。
- 文件存储卷使用约束：
  - CCE集群中的容器存储目前已支持加密（Kubernetes 1.13版本及以上），当前仅在部分区域（Region）提供端到端支持。
  - 存储不支持选择企业项目，新创建的存储卷默认创建到default企业项目下。
- 对象存储卷使用约束如下：
  - CCE v1.7.3-r8及以下版本集群不支持创建对象存储服务，请参照界面要求创建新版本集群，再使用对象存储服务。

  - 目前鲲鹏集群暂时不支持obsfs，无法挂载并行文件系统。
  - 存储不支持选择企业项目，新创建的存储卷默认创建到default企业项目下。
- 极速文件存储卷使用约束如下：

  暂不支持直接创建极速文件存储卷，您可以参照界面要求前往**SFS Turbo控制台**创建后，再使用极速文件存储服务。
- 快照与备份使用约束：
  - 快照功能**仅支持v1.15及以上版本**的集群，且需要安装基于CSI的Everest插件才可以使用。
  - 基于快照创建的云硬盘，其子类型（普通IO/高IO/超高IO）、是否加密、磁盘模式（VBD/SCSI）、共享性(非共享/共享)、容量等都要与快照关联母盘保持一致，这些属性查询和设置出来后不能够修改。

服务（Service）数量

此处的服务对应kubernetes的service资源，即工作负载所添加的服务。

每个命名空间下，创建的服务数量不能超过6000个。

CCE 集群配额限制

针对每个用户，华为云云容器引擎的集群在每个地域分配了固定配额。

表 8-1

限制项	普通用户限制	例外申请方式
实名认证	实名认证	没有例外
单Region下集群总数	50	**到“我的配额”提交申请**
单集群下节点（集群管理规模）	可选择50节点、200节点、1000节点或2000节点多种管理规模，最大支持5000节点。	**到“我的配额”提交申请**
每个worker节点创建容器实例最大数	创建集群时界面可设置。 VPC网络：最大256。	没有例外

依赖底层云产品配额限制

表 8-2

限制大类	限制项	普通用户限制	例外申请方式
计算	实例数	1000	**提交工单**
	核心数	8000核	**提交工单**
	RAM容量（MB）	16384000	**提交工单**

限制大类	限制项	普通用户限制	例外申请方式
网络	一个用户创建虚拟私有云的数量	5	提交工单
	一个用户创建子网的数量	100	提交工单
	一个用户拥有的安全组数量	100	提交工单
	一个用户拥有的安全组规则数量	5000	提交工单
	一个路由表里拥有的路由数量	100	没有例外
	一个虚拟私有云拥有路由数量	100	没有例外
	一个区域下的对等连接数量	50	没有例外
	一个用户拥有网络ACL数量	200	提交工单
	一个用户创建二层连接网关的数量	5	提交工单
负载均衡	弹性负载均衡	50	提交工单
	弹性负载均衡监听器	100	提交工单
	弹性负载均衡证书	120	提交工单
	弹性负载均衡转发策略	500	提交工单
	弹性负载均衡后端主机组	500	提交工单
	弹性负载均衡后端服务器	500	提交工单

9 计费说明

计费项

云容器引擎（CCE）本身不收取任何费用，但在使用过程中会创建相关资源（如节点、带宽等），您需要为您使用的这些资源付费。CCE相关资源的计费项分为如下两部分：

1. **集群：**控制节点资源费用，按照每个集群的类型（虚拟机或裸金属、控制节点数）、集群规模（最大支持的节点数）的差异收取不同的费用。

   说明

   集群规模是指用户在集群下创建和购买的云主机或者裸金属服务器的数量。

   控制节点资源的价格目录请参见：**云容器引擎价格目录**。

2. **IaaS基础设施：**集群工作节点所使用的IaaS基础设施费用，包括集群创建使用过程中自动创建或手动加入的相关资源，如云服务器、云硬盘、弹性IP/带宽、负载均衡等，价格参照相应产品价格表。

   更多价格目录请参见：**产品价格详情**。

计费模式

CCE支持按需计费、包年/包月两种计费模式，供您灵活选择。

- 按需计费：一种先使用后付费的方式，从“开通”开启计费到“删除”结束计费，按实际购买时长计费。这种购买方式比较灵活，您可以按需取用资源，随时开启和释放，无需提前购买大量资源。

  说明

  关于CCE集群休眠或节点关机后的收费说明：

  - **集群休眠：**集群休眠后，控制节点资源费用将停止收费。
  - **节点关机：**集群休眠后，集群中的工作节点（即ECS）并不会自动关机，节点关机后不再收费。如需关机请登录ECS控制台操作，具体请参见**节点关机**。

    ECS关机不收费： 对于按需购买的普通ECS（**不含本地硬盘，FPGA卡**），用户关机后，ECS实例本身（vCPU，内存，镜像）不计费，其它所挂载的资源如云硬盘、公网IP、带宽则正常计费。实例的vCPU和内存将不再保留，再次启动时会重新申请vCPU和内存，在资源不足时会有启动失败的风险，您可以通过稍后启动或更改实例规格的方式来恢复。**包含本地硬盘（如磁盘增强型，GPU加速型等）和包含FPGA卡的实例，关机后仍然正常收费，同时vCPU和内存等资源也会保留。**具体请参见**ECS计费模式**。

- 包年/包月：先购买再使用的方式。这种购买方式相对于按需计费能够提供更大的折扣，对于长期使用者，推荐该方式。用户在购买时，系统会根据用户所选的机型对用户云账户中的金额进行扣除。
- 计费模式更改：计费周期内暂不支持计费模式更改。

须知

- 以集群作为计费量纲，根据集群类型和规模大小，按阶梯计费。
- 华为云提供给客户进行续费与充值的时间，当您的包周期资源到期未续订或按需资源欠费时提供宽限期和保留期，详情请参见**宽限期保留期**。

变更配置

按需计费：您可以将集群的“按需计费”方式变更为包周期（包年/包月）计费。按需计费变更为包周期后，控制节点、工作节点及附属资源（云硬盘，弹性公网IP）将转为包周期资源，并会生成新的订单，用户支付订单后，包周期资源立即生效。

包年/包月：包年/包月计费方式的集群在计费周期内不支持变更配置。创建后不能删除，如需停止使用，请到费用中心执行**退订**操作。

变更须知：

- 集群中使用代金券购买的云服务器降低规格时，系统不会退还代金券。
- 升级规格配置后需按照与原规格的价差，结合已使用的时间周期，补上差价。
- 集群中的云服务器规格（CPU或内存）变小，会影响云服务器的性能。
- 降低规格配置后，如需重新升级至原规格，可能需要补交费用。

10 权限管理

CCE权限管理是在**统一身份认证服务（IAM）**与**Kubernetes的角色访问控制（RBAC）**的能力基础上，打造的细粒度权限管理功能，支持基于IAM的细粒度权限控制和IAM Token认证，支持集群级别、命名空间级别的权限控制，帮助用户便捷灵活的对租户下的IAM用户、用户组设定不同的操作权限。

CCE的权限管理包括“集群权限”和“命名空间权限”两种能力，能够从集群和命名空间层面对用户组或用户进行细粒度授权，具体解释如下：

- **集群权限**：是基于IAM系统策略的授权，可以通过用户组功能实现IAM用户的授权。用户组是用户的集合，通过集群权限设置可以让某些用户组操作集群（如创建/删除集群、节点、节点池、模板、插件等），而让某些用户组仅能查看集群。

  集群权限涉及CCE非Kubernetes API，支持IAM细粒度策略、企业项目管理相关能力。
- **命名空间权限**：是基于Kubernetes RBAC能力的授权，通过权限设置可以让不同的用户或用户组拥有操作不同Kubernetes资源的权限（**如工作负载、任务、服务等Kubernetes原生资源**）。同时CCE基于开源能力进行了增强，可以支持基于IAM用户或用户组粒度进行RBAC授权、IAM token直接访问API进行RBAC认证鉴权。

  命名空间权限涉及CCE Kubernetes API，基于Kubernetes RBAC能力进行增强，支持对接IAM用户/用户组进行授权和认证鉴权，但与IAM细粒度策略独立，详见**Kubernetes RBAC**。

⚠ 注意

- 集群权限仅针对与集群相关的资源（如集群、节点等）有效，您必须确保同时配置了**命名空间权限**，才能有操作Kubernetes资源（如工作负载、任务、Service等）的权限。
- 任何用户创建v1.11.7-r2或以上版本集群后，CCE会自动为该用户添加该集群的所有命名空间的cluster-admin权限，也就是说该用户允许对集群以及所有命名空间中的全部资源进行完全控制。

集群权限（IAM 系统策略授权）

默认情况下，管理员创建的IAM用户没有任何权限，需要将其加入用户组，并给用户组授予策略或角色，才能使得用户组中的用户获得对应的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

CCE部署时通过物理区域划分，为项目级服务。授权时，“作用范围”需要选择“区域级项目”，然后在指定区域（如华北-北京四）对应的项目（cn-north-4）中设置相关权限，并且该权限仅对此项目生效；如果在“所有项目”中设置权限，则该权限在所有区域项目中都生效。访问CCE时，需要先切换至授权区域。

权限根据授权精细程度分为角色和策略。

- 角色：IAM最初提供的一种根据用户的工作职能定义权限的粗粒度授权机制。该机制以服务为粒度，提供有限的服务相关角色用于授权。由于华为云各服务之间存在业务依赖关系，因此给用户授予角色时，可能需要一并授予依赖的其他角色，才能正确完成业务。角色并不能满足用户对精细化授权的要求，无法完全达到企业对权限最小化的安全管控要求。
- 策略：IAM最新提供的一种细粒度授权的能力，可以精确到具体服务的操作、资源以及请求条件等。基于策略的授权是一种更加灵活的授权方式，能够满足企业对权限最小化的安全管控要求。例如：针对CCE服务，租户（Domain）能够控制用户仅能对某一类集群和节点资源进行指定的管理操作。多数细粒度策略以API接口为粒度进行权限拆分，CCE支持的API授权项请参见**权限策略和授权项**。

如**表10-1**所示，包括了CCE的所有系统权限。

表 10-1 CCE 系统权限

系统角色/策略名称	描述	类别	依赖关系
CCE Administrator	具有CCE集群及集群下所有资源（包含集群、节点、工作负载、任务、服务等）的读写权限。	系统角色	拥有该权限的用户必须同时拥有以下权限： **全局服务：** OBS Buckets Viewer、OBS Administrator。 **区域级项目：** Tenant Guest、Server Administrator、ELB Administrator、SFS Administrator、SWR Admin、APM FullAccess。 **说明** 如果同时拥有NAT Gateway Administrator权限，则可以在集群中使用NAT网关的相关功能。
CCE FullAccess	CCE服务集群相关资源的普通操作权限，不包括集群（启用Kubernetes RBAC鉴权）的命名空间权限，不包括委托授权、生成集群证书等管理员角色的特权操作。	策略	无

系统角色/策略名称	描述	类别	依赖关系
CCE ReadOnly Access	CCE服务集群相关资源的查看权限，不包括集群（启用Kubernetes RBAC鉴权）的命名空间权限。	策略	无

表 10-2 CCE 常用操作与系统权限的关系

操作	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
创建集群	x	√	√
删除集群	x	√	√
更新集群，如后续允许集群支持RBAC，调度参数更新等	x	√	√
升级集群	x	√	√
唤醒集群	x	√	√
休眠集群	x	√	√
查询集群列表	√	√	√
查询集群详情	√	√	√
添加节点	x	√	√
删除节点/批量删除节点	x	√	√
更新节点，如更新节点名称	x	√	√
查询节点详情	√	√	√
查询节点列表	√	√	√
查询任务列表（集群层面的job）	√	√	√
删除任务/批量删除任务（集群层面的job）	x	√	√
查询任务详情（集群层面的job）	√	√	√
创建存储	x	√	√

操作	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
删除存储	x	√	√
操作所有kubernetes资源（具体权限请在**命名空间权限**中配置）。	√	√	√
ECS（弹性云服务器）服务的所有权限。	x	√	√
EVS（云硬盘）的所有权限。 可以将云硬盘挂载到云服务器，并可以随时扩容云硬盘容量	x	√	√
VPC（虚拟私有云，包含二代ELB）的所有权限。 创建的集群需要运行在虚拟私有云中，创建命名空间时，需要创建或关联VPC，创建在命名空间的容器都运行在VPC之内。	x	√	√
ECS（弹性云服务器）所有资源详情的查看权限。 CCE中的一个节点就是具有多个云硬盘的一台弹性云服务器	√	√	√
ECS（弹性云服务器）所有资源列表的查看权限。	√	√	√
EVS（云硬盘）所有资源详情的查看权限。可以将云硬盘挂载到云服务器，并可以随时扩容云硬盘容量	√	√	√
EVS（云硬盘）所有资源列表的查看权限。	√	√	√
VPC（虚拟私有云，包含二代ELB）所有资源详情的查看权限。 创建的集群需要运行在虚拟私有云中，创建命名空间时，需要创建或关联VPC，创建在命名空间的容器都运行在VPC之内	√	√	√

操作	CCE ReadOnlyAccess	CCE FullAccess	CCE Administrator
VPC（虚拟私有云，包含二代ELB）所有资源列表的查看权限。	√	√	√
ELB（弹性负载均衡）服务所有资源详情的查看权限。	x	x	√
ELB（弹性负载均衡）服务所有资源列表的查看权限。	x	x	√
SFS（弹性文件服务）服务所有资源详情的查看权限。	√	√	√
SFS（弹性文件服务）服务所有资源列表额查看权限。	√	√	√
AOM（应用运维管理）服务所有资源详情的查看权限。	√	√	√
AOM（应用运维管理）服务所有资源列表的查看权限。	√	√	√
AOM（应用运维管理）服务自动扩缩容规则的所有操作权限。	√	√	√

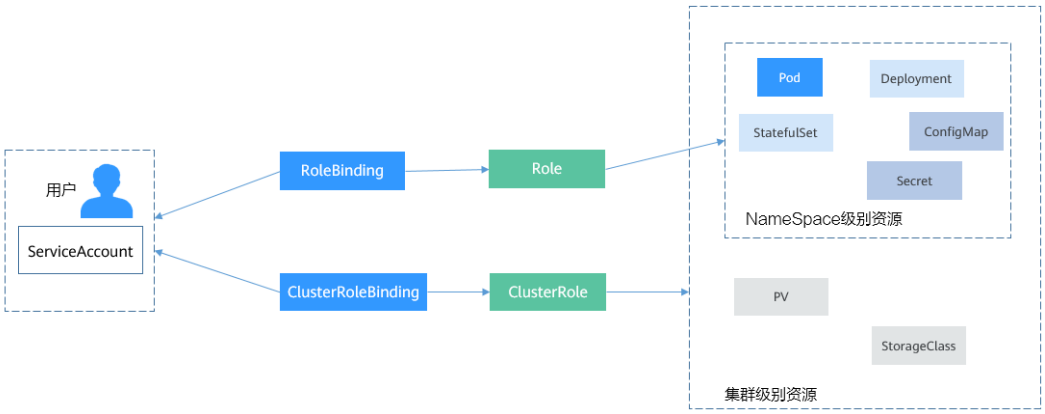
命名空间权限（kubernetes RBAC 授权）

命名空间权限是基于Kubernetes RBAC能力的授权，通过权限设置可以让不同的用户或用户组拥有操作不同Kubernetes资源的权限。Kubernetes RBAC API定义了四种类型：Role、ClusterRole、RoleBinding与ClusterRoleBinding，这四种类型之间的关系和简要说明如下：

- Role：角色，其实是定义一组对Kubernetes资源（命名空间级别）的访问规则。
- RoleBinding：角色绑定，定义了用户和角色的关系。
- ClusterRole：集群角色，其实是定义一组对Kubernetes资源（集群级别，包含全部命名空间）的访问规则。
- ClusterRoleBinding：集群角色绑定，定义了用户和集群角色的关系。

Role和ClusterRole指定了可以对哪些资源做哪些动作，RoleBinding和ClusterRoleBinding将角色绑定到特定的用户、用户组或**ServiceAccount**上。如下图所示。

图 10-1 角色绑定



在CCE控制台中可以授予用户或用户组命名空间权限，可以对某一个命名空间或全部命名空间授权，CCE控制台中默认提供如下5个ClusterRole。

- view：拥有查看命名空间资源的权限
- edit：拥有修改命名空间资源的权限
- admin：拥有命名空间全部权限
- cluster-admin：拥有集群的全部权限
- psp-global：是集群级别的资源，它能够控制Pod规约中与安全性相关的各个方面。具体请参见**Pod安全策略配置**。

除了使用cluster-admin、admin、edit、view这4个最常用的clusterrole外，您还可以通过定义Role和RoleBinding来进一步对命名空间中不同类别资源（如Pod、Deployment、Service等）的增删改查权限进行配置，从而做到更加精细化的权限控制。具体请参见**命名空间权限**。

相关链接

- **IAM产品介绍**
- **创建用户组、用户并授权CCE权限**
- **策略支持的授权项**

11 基本概念

11.1 基本概念

云容器引擎（Cloud Container Engine，简称CCE）提供高度可扩展的、高性能的企业级Kubernetes集群，支持运行Docker容器。借助云容器引擎，您可以在华为云上轻松部署、管理和扩展容器化应用程序。

云容器引擎提供Kubernetes原生API，支持使用kubectl，且提供图形化控制台，让您能够拥有完整的端到端使用体验，使用云容器引擎前，建议您先了解相关的基本概念。

集群（Cluster）

集群指容器运行所需要的云资源组合，关联了若干云服务器节点、负载均衡等云资源。您可以理解为集群是“同一个子网中一个或多个弹性云服务器（又称：节点）”通过相关技术组合而成的计算机群体，为容器运行提供了计算资源池。

节点（Node）

每一个节点对应一台服务器（可以是虚拟机实例或者物理服务器），容器应用运行在节点上。节点上运行着Agent代理程序（kubelet），用于管理节点上运行的容器实例。集群中的节点数量可以伸缩。

节点池（NodePool）

节点池是集群中具有相同配置的一组节点，一个节点池包含一个节点或多个节点。

虚拟私有云（VPC）

虚拟私有云是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。您可以在VPC中定义与传统网络无差别的虚拟网络，同时提供弹性IP、安全组等高级网络服务。

安全组

安全组是一个逻辑上的分组，为同一个VPC内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。安全组创建后，用户可以在安全组中定义各种访问规则，当弹性云服务器加入该安全组后，即受到这些访问规则的保护。

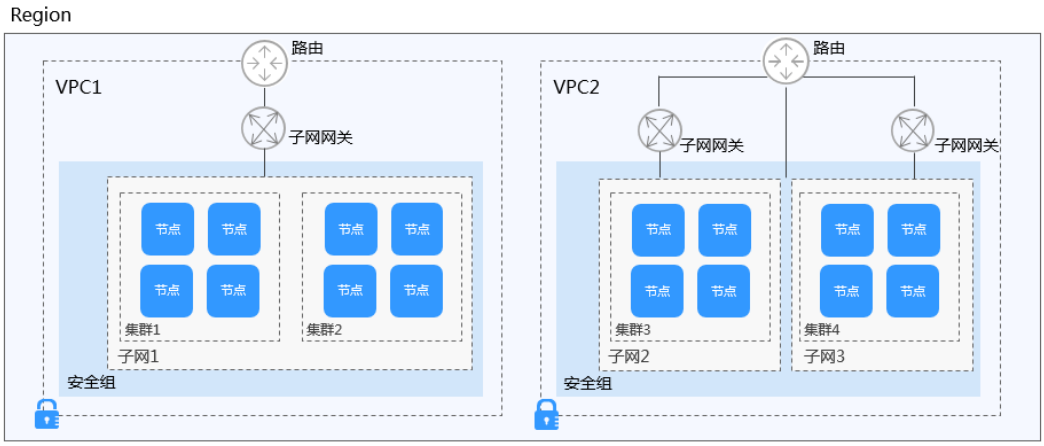
详细介绍请参见**安全组**。

集群、虚拟私有云、安全组和节点的关系

如图11-1，同一个Region下可以有多个虚拟私有云（VPC）。虚拟私有云由一个个子网组成，子网与子网之间的网络交互通过子网网关完成，而集群就是建立在某个子网中。因此，存在以下三种场景：

- 不同集群可以创建在不同的虚拟私有云中。
- 不同集群可以创建在同一个子网中。
- 不同集群可以创建在不同的子网中。

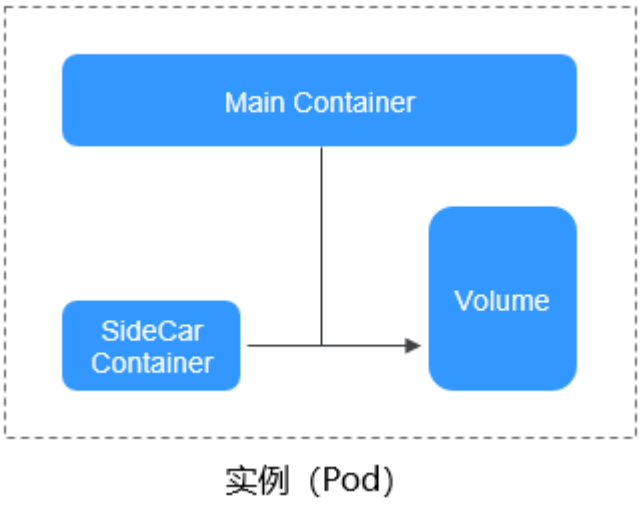
图 11-1 集群、VPC、安全组和节点的关系



实例（Pod）

实例（Pod）是 Kubernetes 部署应用或服务的最小的基本单位。一个Pod 封装多个应用容器（也可以只有一个容器）、存储资源、一个独立的网络 IP 以及管理控制容器运行方式的策略选项。

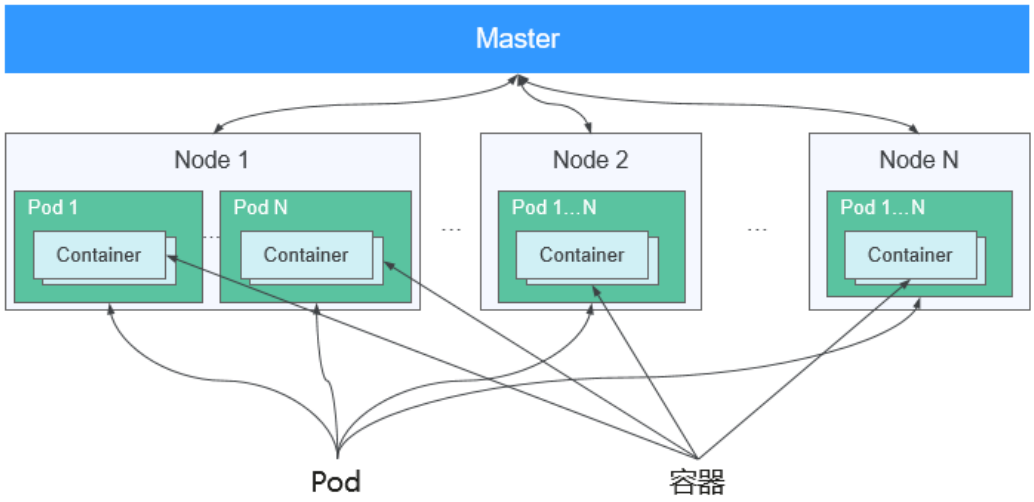
图 11-2 实例（Pod）



容器（Container）

一个通过 Docker 镜像创建的运行实例，一个节点可运行多个容器。容器的实质是进程，但与直接在宿主执行的进程不同，容器进程运行于属于自己的独立的命名空间。

图 11-3 实例 Pod、容器 Container、节点 Node 的关系

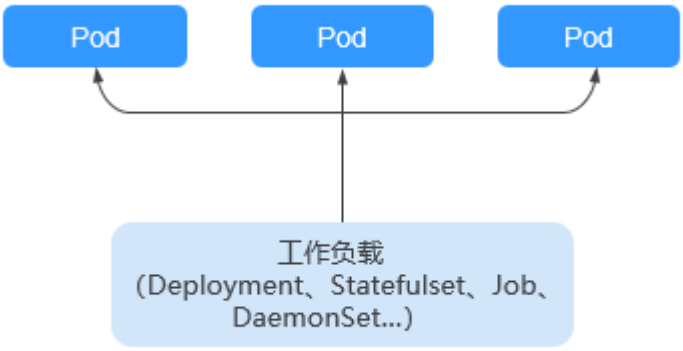


工作负载

工作负载是在Kubernetes上运行的应用程序。无论您的工作负载是单个组件还是协同工作的多个组件，您都可以在Kubernetes上的一组Pod中运行它。在Kubernetes中，工作负载是对一组Pod的抽象模型，用于描述业务的运行载体，包括Deployment、Statefulset、Daemonset、Job、CronJob等多种类型。

- **无状态工作负载：**即kubernetes中的“Deployment”，无状态工作负载支持弹性伸缩与滚动升级，适用于实例完全独立、功能相同的场景，如：nginx、wordpress等。
- **有状态工作负载：**即kubernetes中的“StatefulSet”，有状态工作负载支持实例有序部署和删除，支持持久化存储，适用于实例间存在互访的场景，如ETCD、mysql-HA等。
- **创建守护进程集：**即kubernetes中的“DaemonSet”，守护进程集确保全部（或者某些）节点都运行一个Pod实例，支持实例动态添加到新节点，适用于实例在每个节点上都需要运行的场景，如ceph、fluentd、Prometheus Node Exporter等。
- **普通任务：**即kubernetes中的“Job”，普通任务是一次性运行的短任务，部署完成后即可执行。使用场景为在创建工作负载前，执行普通任务，将镜像上传至镜像仓库。
- **定时任务：**即kubernetes中的“CronJob”，定时任务是按照指定时间周期运行的短任务。使用场景为在某个固定时间点，为所有运行中的节点做时间同步。

图 11-4 工作负载与 Pod 的关系



编排模板

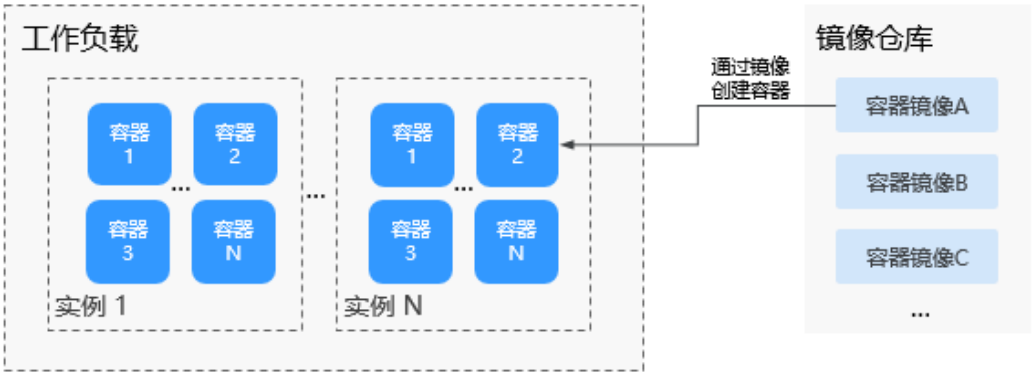
编排模板包含了一组容器服务的定义和其相互关联，可以用于多容器应用的部署和管理。

镜像（Image）

Docker镜像是一个模板，是容器应用打包的标准格式，用于创建Docker容器。或者说，Docker镜像是一个特殊的文件系统，除了提供容器运行时所需的程序、库、资源、配置等文件外，还包含了一些为运行时准备的配置参数（如匿名卷、环境变量、用户等）。镜像不包含任何动态数据，其内容在构建之后也不会被改变。在部署容器化应用时可以指定镜像，镜像可以来自于 Docker Hub、华为云**容器镜像服务**或者用户的私有 Registry。例如一个Docker镜像可以包含一个完整的Ubuntu操作系统环境，里面仅安装了用户需要的应用程序及其依赖文件。

镜像（Image）和容器（Container）的关系，就像是面向对象程序设计中的类和实例一样，镜像是静态的定义，容器是镜像运行时的实体。容器可以被创建、启动、停止、删除、暂停等。

图 11-5 镜像、容器、工作负载的关系



命名空间（Namespace）

命名空间是对一组资源和对象的抽象整合。在同一个集群内可创建不同的命名空间，不同命名空间中的数据彼此隔离。使得它们既可以共享同一个集群的服务，也能够互不干扰。例如：

- 可以将开发环境、测试环境的业务分别放在不同的命名空间。
- 常见的pods，services，replication controllers和deployments等都是属于某一个namespace的（默认是default），而node，persistentVolumes等则不属于任何namespace。

服务（Service）

Service是将运行在一组 Pods 上的应用程序公开为网络服务的抽象方法。

使用Kubernetes，您无需修改应用程序即可使用不熟悉的服务发现机制。 Kubernetes为Pods提供自己的IP地址和一组Pod的单个DNS名称，并且可以在它们之间进行负载平衡。

Kubernetes允许指定一个需要的类型的Service，类型 的取值以及行为如下：

- ClusterIP：集群内访问。通过集群的内部 IP 暴露服务，选择该值，服务只能够在集群内部可以访问，这也是默认的 ServiceType。
- NodePort：节点访问。通过每个Node上的 IP 和静态端口（NodePort）暴露服务。NodePort服务会路由到ClusterIP服务，这个ClusterIP服务会自动创建。通过请求 <NodeIP>:<NodePort>，可以从集群的外部访问一个 NodePort 服务。
- LoadBalancer：负载均衡。使用云提供商的负载均衡器，可以向外部暴露服务。外部的负载均衡器可以路由到NodePort服务和ClusterIP服务。
- DNAT：DNAT网关。可以为集群节点提供网络地址转换服务，使多个节点可以共享使用弹性IP。与弹性IP方式相比增强了可靠性，弹性IP无需与单个节点绑定，任何节点状态的异常不影响其访问。

七层负载均衡（Ingress）

Ingress是为进入集群的请求提供路由规则的集合，可以给service提供集群外部访问的URL、负载均衡、SSL终止、HTTP路由等。

网络策略（NetworkPolicy）

NetworkPolicy提供了基于策略的网络控制，用于隔离应用并减少攻击面。它使用标签选择器模拟传统的分段网络，并通过策略控制它们之间的流量以及来自外部的流量。

配置项（Configmap）

ConfigMap用于保存配置数据的键值对，可以用来保存单个属性，也可以用来保存配置文件。ConfigMap跟secret很类似，但它可以更方便地处理不包含敏感信息的字符串。

密钥（Secret）

Secret解决了密码、token、密钥等敏感数据的配置问题，而不需要把这些敏感数据暴露到镜像或者Pod Spec中。Secret可以以Volume或者环境变量的方式使用。

标签（Label）

标签其实就一对 key/value，被关联到对象上，比如Pod。标签的使用我们倾向于能够标示对象的特殊特点，并且对用户而言是有意义的，但是标签对内核系统是没有直接意义的。

选择器（LabelSelector）

Label selector是Kubernetes核心的分组机制，通过label selector客户端/用户能够识别一组有共同特征或属性的资源对象。

注解（Annotation）

Annotation与Label类似，也使用key/value键值对的形式进行定义。

Label具有严格的命名规则，它定义的是Kubernetes对象的元数据（Metadata），并且用于Label Selector。

Annotation则是用户任意定义的“附加”信息，以便于外部工具进行查找。

存储卷（PersistentVolume）

PersistentVolume（PV）是集群之中的一块网络存储。跟 Node 一样，也是集群的资源。

存储声明（PersistentVolumeClaim）

PV 是存储资源，而 PersistentVolumeClaim (PVC) 是对 PV 的请求。PVC 跟 Pod 类似：Pod 消费 Node 资源，而 PVC 消费 PV 资源；Pod 能够请求 CPU 和内存资源，而 PVC 请求特定大小和访问模式的数据卷。

弹性伸缩（HPA）

Horizontal Pod Autoscaling，简称HPA，是Kubernetes中实现POD水平自动伸缩的功能。Kubernetes集群可以通过Replication Controller的scale机制完成服务的扩容或缩容，实现具有伸缩性的服务。

亲和性与反亲和性

在应用没有容器化之前，原先一个虚机上会装多个组件，进程间会有通信。但在做容器化拆分的时候，往往直接按进程拆分容器，比如业务进程一个容器，监控日志处理或者本地数据放在另一个容器，并且有独立的生命周期。这时如果他们分布在网络中两个较远的点，请求经过多次转发，性能会很差。

- 亲和性：可以实现就近部署，增强网络能力实现通信上的就近路由，减少网络的损耗。如：应用A与应用B两个应用频繁交互，所以有必要利用亲和性让两个应用的尽可能的靠近，甚至在一个节点上，以减少因网络通信而带来的性能损耗。
- 反亲和性：主要是出于高可靠性考虑，尽量分散实例，某个节点故障的时候，对应用的影响只是 N 分之一或者只是一个实例。如：当应用采用多副本部署时，有必要采用反亲和性让各个应用实例打散分布在各个节点上，以提高HA。

节点亲和性（NodeAffinity）

通过选择标签的方式，可以限制pod被调度到特定的节点上。

节点反亲和性（NodeAntiAffinity）

通过选择标签的方式，可以限制pod不被调度到特定的节点上。

工作负载亲和性（PodAffinity）

指定工作负载部署在相同节点。用户可根据业务需求进行工作负载的就近部署，容器间通信就近路由，减少网络消耗。

工作负载反亲和性（PodAntiAffinity）

指定工作负载部署在不同节点。同个工作负载的多个实例反亲和部署，减少宕机影响；互相干扰的应用反亲和部署，避免干扰。

资源配额（Resource Quota）

资源配额（Resource Quotas）是用来限制用户资源用量的一种机制。

资源限制（Limit Range）

默认情况下，K8S中所有容器都没有任何CPU和内存限制。LimitRange(简称limits)用来给Namespace增加一个资源限制，包括最小、最大和默认资源。在pod创建时，强制执行使用limits的参数分配资源。

环境变量

环境变量是指容器运行环境中设定的一个变量，您可以在创建容器模板时设定不超过30个的环境变量。环境变量可以在工作负载部署后修改，为工作负载提供了极大的灵活性。

在CCE中设置环境变量与Dockerfile中的“ENV”效果相同。

应用服务网格（Istio）

Istio是一个提供连接、保护、控制以及观测功能的开放平台。

云容器引擎深度集成了应用服务网格，提供非侵入式的微服务治理解决方案，支持完整的生命周期管理和流量治理能力，兼容Kubernetes和Istio生态。一键开启应用服务网格后即可提供非侵入的智能流量治理解决方案，其功能包括负载均衡、熔断、限流等多种治理能力。应用服务网格内置金丝雀、蓝绿等多种灰度发布流程，提供一站式自动化的发布管理。基于无侵入的监控数据采集，深度整合华为云**应用性能管理**（APM）能力，提供实时流量拓扑、调用链等服务性能监控和运行诊断，构建全景的服务运行视图。

11.2 CCE 与原生 Kubernetes 名词对照

Kubernetes，简称K8S，是开源的容器集群管理系统，可以实现容器集群的自动化部署、自动扩缩容、维护等功能。它既是一款容器编排工具，也是全新的基于容器技术的分布式架构领先方案。在Docker技术的基础上，为容器化的应用提供部署运行、资源调度、服务发现和动态伸缩等功能，提高了大规模容器集群管理的便捷性。

本文主要为您介绍云容器引擎CCE与原生Kubernetes名词对照情况和简单解释。

表 11-1 CCE 与原生 Kubernetes 名词对照

云容器引擎CCE	原生Kubernetes
集群	Cluster
节点	Node
节点池	NodePool
容器	Container
镜像	Image
命名空间	Namespace
无状态工作负载	Deployment
有状态工作负载	StatefulSet
守护进程集	DaemonSet
普通任务	Job
定时任务	CronJob
实例（容器组）	Pod
服务（Service）	Service
虚拟集群IP	Cluster IP
节点端口	NodePort
负载均衡	LoadBalancer
七层负载均衡	Ingress
网络策略	NetworkPolicy
模板	Template
配置项	ConfigMap
密钥	Secret
标签	Label
选择器	LabelSelector
注解	Annotation
存储卷	PersistentVolume
存储声明	PersistentVolumeClaim
弹性伸缩	HPA
节点亲和性	NodeAffinity
节点反亲和性	NodeAntiAffinity

云容器引擎CCE	原生Kubernetes
工作负载亲和性	PodAffinity
工作负载反亲和性	PodAntiAffinity
触发器	Webhook
终端节点	Endpoint
资源配额	Resource Quota
资源限制	Limit Range

11.3 区域与可用区

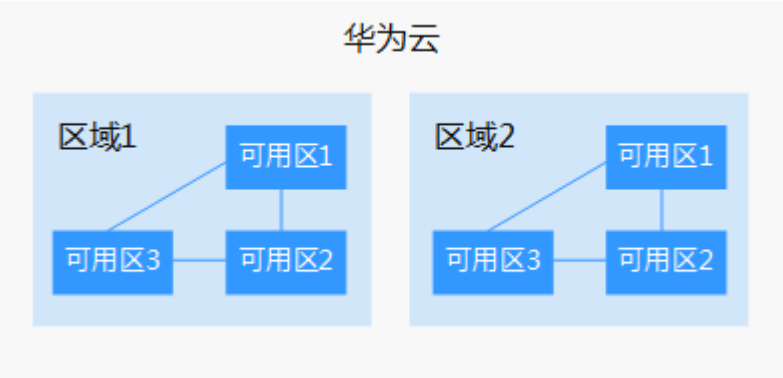
什么是区域、可用区?

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。
- 可用区（AZ，Availability Zone）：一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

图11-6阐明了区域和可用区之间的关系：

图 11-6 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。更多信息请参见华为云全球站点。

如何选择区域?

选择区域时，您需要考虑以下几个因素：

- 地理位置

  一般情况下，建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。不过，在基础设施、BGP网络品质、资源的操作与配置等方面，中国大陆各个区域间区别不大，如果您或者您的目标用户在中国大陆，可以不用考虑不同区域造成的网络时延问题。

  - 在除中国大陆以外的亚太地区有业务的用户，可以选择“中国-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
  - 在非洲地区有业务的用户，可以选择“南非-约翰内斯堡”区域。
  - 在欧洲地区有业务的用户，可以选择“欧洲-巴黎”区域。

- 资源的价格

  不同区域的资源价格可能有差异，请参见**华为云服务价格详情**。

如何选择可用区?

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

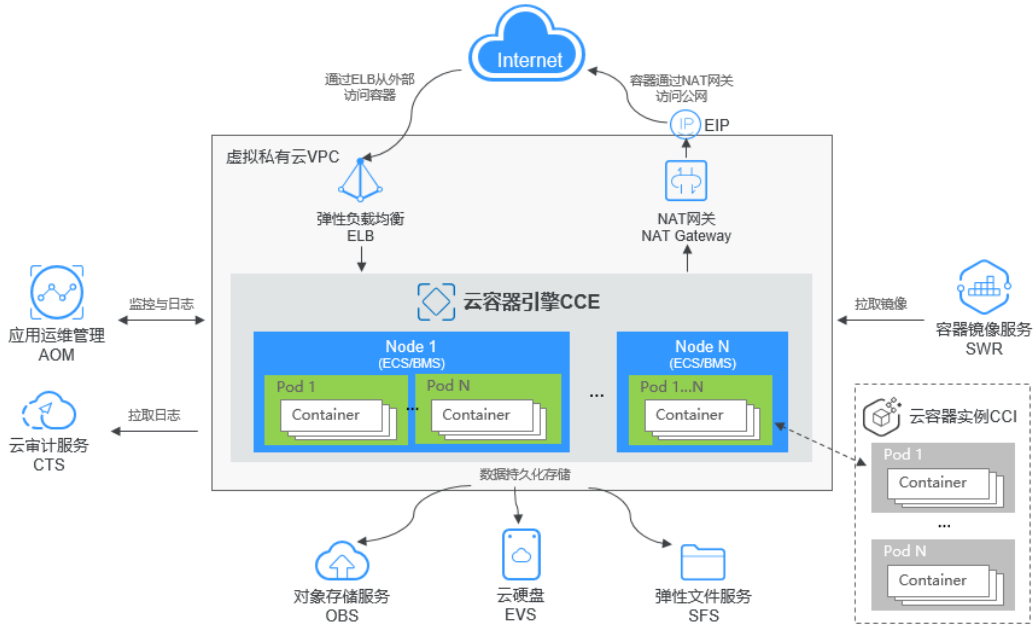
区域和终端节点

当您通过API使用资源时，您必须指定其区域终端节点。有关华为云的区域和终端节点的更多信息，请参阅**地区和终端节点**。

12 与其它云服务的关系

云容器引擎需要与其他云服务协同工作，云容器引擎需要获取如下云服务资源的权限。

图 12-1 云容器引擎与其他服务的关系示意图



云容器引擎与其他服务的关系

表 12-1 云容器引擎与其他服务的关系

服务名称	云容器引擎与其他服务的关系	主要交互功能
弹性云服务器 ECS	在云容器引擎中具有多个云硬盘的一台弹性云服务器就是一个节点，您可以在创建节点时指定弹性云服务器的规格。	• 购买节点 • 纳管已有节点到集群

服务名称	云容器引擎与其他服务的关系	主要交互功能
虚拟私有云 VPC	在云容器引擎中创建的集群需要运行在**虚拟私有云**中，您创建命名空间时，需要创建或关联VPC，创建在命名空间的容器都运行在VPC之内，从而保障网络安全。	**购买CCE集群**
弹性负载均衡 ELB	云容器引擎支持将创建的应用对接到弹性负载均衡，从而提高应用系统对外的服务能力，提高应用程序容错能力。 您可以通过**弹性负载均衡**，从外部网络访问容器负载。	• **创建无状态负载(Deployment)** • **创建有状态负载(StatefulSet)** • **负载均衡(LoadBalancer)**
NAT网关	NAT网关能够为VPC内的容器实例提供网络地址转换（Network Address Translation）服务，SNAT功能通过绑定弹性公网IP，实现私有IP向公有IP的转换，可实现VPC内的容器实例共享弹性公网IP访问Internet。 您可以通过**NAT网关**设置SNAT规则，使得容器能够访问Internet。	• **创建无状态负载(Deployment)** • **创建有状态负载(StatefulSet)** • **DNAT网关(DNAT)**
容器镜像服务 SWR	容器镜像服务提供的镜像仓库是用于存储、管理docker容器镜像的场所，可以让使用人员轻松存储、管理、部署docker容器镜像。 您可以使用**容器镜像服务**中的镜像创建负载。	• **创建无状态负载(Deployment)** • **创建有状态负载(StatefulSet)**
云容器实例 CCI	云容器实例的Serverless Container是从使用角度，无需创建、管理Kubernetes集群，也就是从使用的角度看不见服务器（Serverless），直接通过控制台、kubectl、Kubernetes API创建和使用容器负载，且只需为容器所使用的资源付费。 当CCE集群资源不足时，支持将Pod弹性部署到CCI集群。	• **virtual kubelet插件** • **华为云CCE弹性至CCI**
云硬盘 EVS	可以将云硬盘挂载到云服务器，并可以随时扩容云硬盘容量。 在云容器引擎中一个节点就是具有多个云硬盘的一台弹性云服务器，您可以在创建节点时指定云硬盘的大小。	**使用云硬盘存储卷**

服务名称	云容器引擎与其他服务的关系	主要交互功能
对象存储服务 OBS	对象存储服务是一个基于对象的海量存储服务，为客户提供海量、安全、高可靠、低成本的数据存储能力，包括：创建、修改、删除桶，上传、下载、删除对象等。 云容器引擎支持创建OBS对象存储卷并挂载到容器的某一路径下。	**使用对象存储卷**
弹性文件服务 SFS	弹性文件服务提供托管的共享文件存储，符合标准文件协议（NFS），能够弹性伸缩至PB规模，具备可扩展的性能，为海量数据、高带宽型应用提供有力支持。 您可以使用弹性文件服务作为容器的持久化存储，在创建任务负载的时候挂载到容器上。	**使用文件存储卷**
应用运维管理 AOM	云容器引擎对接了AOM，AOM会采集容器日志存储中的“.log”等格式日志文件，转储到AOM中，方便您查看和检索；并且云容器引擎基于AOM进行资源监控，为您提供弹性伸缩能力。	**采集容器内路径日志**
云审计服务 CTS	云审计服务提供云服务资源的操作记录，记录内容包括您从公有云管理控制台或者开放API发起的云服务资源操作请求以及每次请求的结果，供您查询、审计和回溯使用。	**云审计服务支持的CCE操作列表**