

容器云平台构建技术研究

刘旭 戴发玉 杨玉姣 井长军 曹敏娜

(北京中电普华信息技术有限公司 北京 100192)

摘要 在企业数字化转型改造的过程中,传统企业的物理机器被虚拟化,在一定程度上降低了运维的复杂性,增加了资源的利用率,但业务应用还面临很多挑战。例如,业务应用响应效率低,业务上云部署流程复杂,运维管理体验较差,庞大的部署架构使应用的开发、测试和升级也比较复杂,让业务升级停机时间和部署成本增加。因此,亟待构建基于容器集群的云服务体系,发挥容器轻量化、敏捷性、兼容性的优势,实现自动化部署、自动伸缩、自动资源选择和调度,为企业数字化上云提供云原生解决方案。文中旨在构建具备容器集群管理、弹性伸缩、资源编排、应用管理能力的容器云平台,为企业业务应用上云提供云原生解决方案。

关键词: 容器云平台;自动化部署;弹性伸缩;资源编排;应用托管

中图分类号 TP391.9

Research on Construction Technology of Container Cloud Platform

LIU Xu, DAI Fayu, YANG Yujiao, JING Changjun and CAO Minna

(Beijing China Power Puhua Information Technology Co., Ltd., Beijing 100192, China)

Abstract In the process of enterprise Digital transformation, the physical machines of traditional enterprises are virtualized, which reduces the complexity of operation and maintenance to a certain extent and increases the utilization rate of resources., but business applications still face many challenges. For example, low response efficiency of business applications, complex cloud deployment processes on the business, poor operation and maintenance management experience, and a huge deployment architecture that makes application development, testing, and upgrading more complex, resulting in increased downtime and deployment costs for business upgrades. Therefore, it is urgent to build a cloud service system based on container clusters, leverage the advantages of container lightweight, agility, and compatibility, and achieve automated deployment Automatic scaling, automatic resource selection and scheduling, providing cloud native solutions for enterprises to digitize and cloud. The purpose of this paper is to build a container cloud platform with container cluster management, elastic scaling, resource arrangement, and application management capabilities, providing cloud native solutions for enterprise business applications on the cloud.

Key words Container cloud platform, Automated deployment, Elastic expansion and contraction, Resource allocation, Application hosting

0 引言

目前业界主流容器云平台管理技术主要基于 Kubernetes, 而其理论依据主要源自 Google 的 Borg 系统。Borg 是 Google 运行大量云应用的基础架构, 它被用来管理和运行 Google 的数以万计的容器, 而每个容器中运行着一个特定的任务。Borg 的灵感是将容器资源管理和调度自动化, 以实现资源的最优利用。传统的资源管理系统, 如传统的虚拟机, 需要定义静态的资源池。而 Borg 采用了动态的资源调度, 其根据应用需求, 能灵活地对资源进行分配。Kubernetes 把这些概念从 Google 的 Borg 中抽象出来, 并建立了一个可重用的, 可扩展的开源框架, 以支持多种应用场景。

容器云平台技术是实现多租户、自动化、自服务、动态弹性、云资源有效管控的统一云门户, 其能简化部署、监控、运维等应用生命周期管理工作, 打造一站式企业级 PaaS 运行环境, 实现“精细化服务”“即时服务”“按需服务”等关键能力。基于容器技术打造云原生 PaaS 平台, 能为企业信息化上云提供 PaaS 能力, 提供符合系统上云的云原生一体化解决方案。

1 关键技术

1.1 容器云平台集群自动化弹性扩缩以及故障自愈技术

弹性伸缩是根据应用程序的自身承载力及用户的业务需求, 自动地调整应用程序的数量, 以随时满足需求, 并利

基金项目: 容器云平台构建技术研究 (TB123017190)

作者简介: 刘旭 (1976—), 本科, 研究方向为云计算、物联网、信息技术应用创新。

用弹性伸缩功能,通过定时、周期或者监控策略,当业务量高并发时,自动完成应用程序副本数的动态增加或减少,保证业务系统的平稳运行。

目前,Kubernetes的硬件和软件出现故障已成为常态,而此时K8S集群中对应的节点会处于不稳定的状态。人工运维效率低,且容易造成误操作引发更大的风险。自愈的重点在两方面,即问题的检测和故障修复。在K8S集群中的问题主要集中在硬件和软件中。硬件包括内存板卡损坏,磁盘坏道,网卡控制器故障,机房断电跳闸及光缆故障。软件包括软件OOM,进程句柄泄露,IO hang,磁盘满,网络断链,系统负载过高,docker,kubelet故障及运行环境等。

1.2 容器云平台资源栈可视化编辑及自动编排能力

可视化资源编排指打开可视化编辑器,通过简单拖曳图标到画布中的适当位置,连线创建资源间的依赖关系,并输入需要的参数,即可快速创建资源栈模板,高效编排云资源。

资源编排是一种简单易用的云计算资源管理和自动化运维服务。用户通过模板描述多个云计算资源的依赖关系、配置等,并自动完成所有资源的创建和配置,以达到自动化部署、运维等目的。同时,编排模板也是一种标准化的资源和应用交付方式,其可以随时编辑修改,使基础设施代码成为可能。

2 实现思路

2.1 容器云平台自动化部署及运维管理技术研究

研究容器云平台管理相关技术,需预先进行容器集群规划,通过规划容器集群部署所需的如部署区域、部署所需计算资源、存储资源、网络资源及集群规模配置等基础资源信息,通过研究自动化部署技术,按照规划内容完成容器集群的自动化部署;通过研究集群升级、扩缩容、备份等技术,实现集群维护功能;通过研究集群监控技术,实现集群资源

的监控告警等运行监控管理能力。

2.2 容器云平台资源自动弹性伸缩技术研究

研究容器云平台弹性伸缩技术,需通过预先配置的伸缩规则来触发容器集群自动弹性伸缩的机制,如当计算资源CPU使用率突破80%时,弹性伸缩根据配置的伸缩规则弹性扩张计算资源;当计算资源CPU使用率低于30%时,弹性伸缩根据伸缩规则自动收缩计算资源,由此实现对容器云平台自动弹性伸缩能力及故障自愈能力。

2.3 容器云平台业务资源编排技术研究

研究容器云平台资源编排技术,需通过编排引擎自动完成所有资源的创建和配置,以达到自动化部署、运维等目的。通过研究资源栈模板技术,实现存储、网络、数据库、中间件等资源的预先配置;通过研究可视化编排技术,实现资源栈信息的可视化编排管理功能;通过编排引擎研究,实现依据资源栈配置进行资源的创建和配置的功能。

2.4 容器云平台业务应用托管技术研究

研究容器云平台应用管理技术,需通过研究容器云平台中的应用管理机制,将业务应用信息注册、发布到云平台,实现应用在云平台上运行的功能;通过研究应用的配置管理机制,实现应用配置信息在线统一管理;通过研究应用监控机制,实现应用调用链路、日志等信息的监控能力、故障自愈能力。

3 实施方案

容器云平台构建技术主要包括容器集群管理、弹性伸缩、资源编排、应用管理等核心功能,如图1所示。研发容器集群的部署、运维、监控等集群管理功能,需依据预配置的伸缩规则实现容器云平台的自动弹性伸缩,采用资源可视化编排实现云上资源编排,依托应用管理实现应用在云平台的注册、部署、监控、运维等内容。

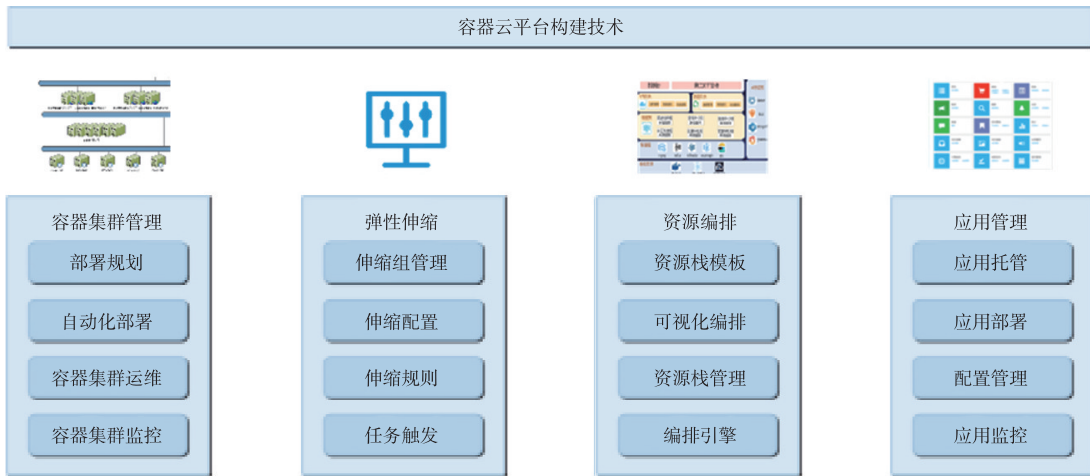


图1 容器云平台构建技术

3.1 容器云平台自动化部署及运维管理技术研究

研究容器云平台的部署规划、自动化部署、集群运维、

集群监控等容器云平台管理技术,需通过规划容器集群,定义集群所需的诸如部署区域、部署所需计算资源、存储资源

源、网络资源及集群规模配置等基础资源信息,完成容器集群资源栈配置工作。同时,依据部署规划的资源栈配置,进行容器集群的自动化部署,通过集群运维进行集群升级、扩

缩容、备份等集群运行维护。通过集群监控实现集群资源的监控告警、故障自愈等管理能力。容器云平台管理技术如图2所示。

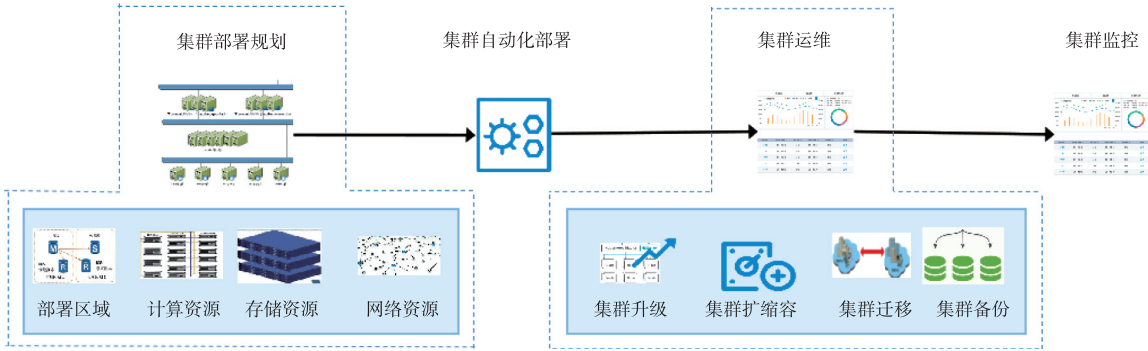


图2 容器云平台管理技术

容器集群创建过程解析如下。(1)在容器集群部署规划阶段进行容器集群资源配置信息的收集。(2)自动化部署服务通过 ansible,发送配置信息至 master 节点、worker 节点。(3)自动化部署服务在各节点安装 docker 环境。(4)自动化部署服务部署缓存服务至 master 节点,并配置启动,用于存储集群配置信息。(5)自动化部署服务部署 master 节点控制服务,并配置启动,用于接受来自 worker 节点 agent 的注册及心跳请求。(6)自动化部署服务部署 master 节点 api 服务,并配置启动,用于向外提供本容器集群 restful 接口。(7)自动化部署服务部署 master 节点调度服务,并配置启动,用于进行 worker 节点工作负载调度。(8)自动化部署服务部署并启动 worker 节点 agent 服务。(9)自动化部署服务部署并启动 worker 节点负载均衡服务。

容器集群监控过程如下。(1)监控服务从集群 master 节点获取集群中所有的节点信息。(2)监控服务通过这些节点上的代理服务获取监控数据。(3)监控服务将获取到的数据推到后端存储,同时支持数据可视化展示。

3.2 容器云平台资源自动弹性伸缩技术研究

弹性伸缩主要包括伸缩组管理、伸缩配置、伸缩规则、

任务触发,通过容器云平台弹性伸缩中的伸缩组管理功能,创建伸缩组,将伸缩规则,配置到伸缩组中,完成伸缩组配置。

任务触发服务通过监控虚拟机 CPU 利用率、内存利用率、网络入/出流量、存储读/写流量、VGPU/显存利用率等指标,依据事先配置好的伸缩规则实现容器云平台的自动扩缩容及故障自愈能力。

3.3 容器云平台业务资源编排技术研究

资源编排包括资源栈模板、可视化编辑、资源栈管理、编排引擎,通过在资源栈可视化编排界面定义上线云平台业务系统涉及的应用、存储、网络、数据库、缓存等云资源以及资源间的依赖关系,通过自动化编排引擎,实现依据资源栈配置进行云资源的自动化部署和配置实现标准化的版本控制和资源变化跟踪,统一对上线业务系统所需云资源的管理,简化云应用交付过程,也可以用于从研发环境快速复制整套资源至测试环境、生产环境等。

3.4 容器云平台业务应用托管技术研究

容器云平台应用管理流程如图3所示。

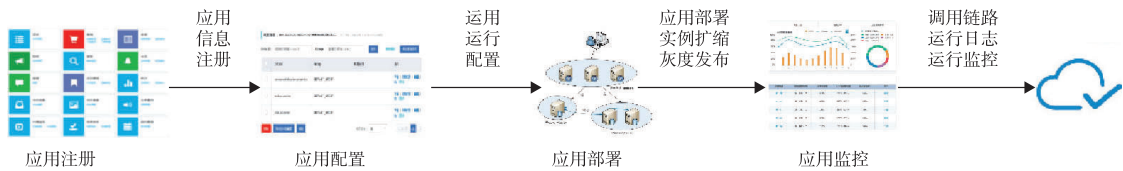


图3 容器云平台应用管理流程

应用管理包括应用注册、应用部署、配置管理、应用监控,通过容器云平台应用管理,提供应用在容器云平台全生命周期管理服务,包括应用部署、扩容、缩容、停止、删除等,同时应用管理集成了弹性伸缩、限流降级、流量管理、监控检查等服务治理组件;也提供了应用配置在线管理的能力,并提供了服务拓扑、服务报表、调用链查询等功能。通过应用管理可以实时监控应用中资源和服务的状态,及时发现问题,且通过日志和诊断组件快速定位,

及故障自愈。

4 应用效益分析

基于容器云平台的构建,可以简化业务应用运行环境的运维管理工作,优化应用程序部署过程,实现弹性伸缩、故障自愈,降低运维团队运维成本,提高运维效率。同时,减少基础环境管理方面的重复投入,按每个项目30万计算,每年20个项目推广应用即可节约600万。

5 结语

通过自主可控技术路线的容器云平台关键技术的研究,推动行业自主可控云平台技术突破,可以有效降低技术创新和产品创新成本,提高创新能力,提升公司在能源行业信息通信技术、产品及服务的社会影响力和品牌形象,促进社会效益的持续提升。

参考文献

[1] 唐燕. 开源技术支持下的企业容器云平台构建与应用研究

[J], 电子技术与软件工程, 2023(8):198-201.

[2] 王军波. 大规模容器云平台海量软件数据资源实时监控方法[J]. 自动化与仪器仪表, 2023(1):10-14.

[3] 刘宏娟, 黄炜, 李文吉, 等. 基于 Kubernetes 的卫星遥感数据容器云平台[J]. 计算机测量与控制, 2022(1):209-214.

[4] 吕志强, 孙超, 王艳蓉, 等. 基于 Kubernetes 的电网调度云应用容器管理系统的实现[J]. 工业控制计算机, 2022(12):130-132.

[5] 张有帅, 余霞, 尹雪龙. 基于 Kubernetes 的容器云平台研究与设计[J]. 电子设计工程, 2021(22):180-183.

(上接第 284 页)

基础。对用户行为数据进行交叉验证后,对模型训练中的数据进行分类与回归,并将任意用户行为数据作为输入变量,形成输出标签,经回归数据处理后,即可获得最终的处理结果。其计算式如式(2)所示:

$$Y = f(X) \quad (2)$$

其中, Y 表示输出变量; X 为输入变量。

通过对用户行为标签进行深层次挖掘,可分析其中的异常数据情况,如通过读者登录系统的行为与频次来判断是否存在数据流失风险;通过大数据感知用户行为数据,预测用户的潜在兴趣。在设定预测标签时,可通过大数据识别读者的阅读行为及兴趣范围,考虑读者在图书馆和系统内的行为类型是有限的,某些画像数据的缺失对画像的精准度影响较大,需处理标签预测中的数据平衡的问题。

为解决上述问题,本文引入了随机森林方法。该方法的核心是通过多棵树完成样本的训练与分类,整个数据处理过程如下。

第 1 步,输入 N 个训练样本及 M 个特征数目。

第 2 步,在输入 m 个特征数目后,计算决策树上的节点计算结果,此时 $m < M$ 。

第 3 步,通过回放抽样的方式,从训练样本中连续取样 N 次,并将上述数据整合到训练集中之后,在数据集中抽取样本进行预测,估算数据间的误差。

第 4 步,在选择 m 个数据后确定特征数据,并进行数据分类。

4 结语

基于大数据的信息行为领域技术具有广阔的应用空间,随着相关技术的不断成熟,其应用范围不断拓展,信息处理效果和技术优势更加显著。本文针对图书馆信息系统建设中存在的问题,应用信息行为领域相关技术,解决了图书馆在当前服务功能实现中存在的不足,对于类似的信息系统建设具有一定的借鉴与指导作用。

参考文献

[1] 卢新元, 张进澳, 雷晓鹏. 人工智能生成内容环境下用户信息行为研究——以对话式搜索引擎为例[J]. 情报理论与实践, 2023, 46(12):84-92.

[2] 古婷骅, 陈忆金, 曹树金. 信息行为领域中情感的核心概念及其演化路径分析[J]. 情报理论与实践, 2021, 44(12):114-121.

[3] 王晰巍, 孟盈, 张响, 等. 信息素养视角下虚拟现实用户信息获取效果影响因素研究[J]. 情报理论与实践, 2021, 44(9):82-89.

[4] 许桂敏, 张转. 非法获取公民个人信息行为的智能化、解读与规制——基于技术的多维面向[J]. 中国人民公安大学学报(社会科学版), 2020, 36(6):130-142.

[5] 王晰巍, 李玥琪, 王铎, 等. 虚拟现实环境下用户信息行为研究动态及趋势分析[J]. 图书情报工作, 2020, 64(5):12-21.