

大数据智能运维在云计算环境下的资源调度与优化研究

王贾琳

中国电子科技集团公司第十五研究所, 中国·北京 100083

摘要: 在云计算环境下, 大数据分析已经成为企业有效地运用海量的数据进行管理与优化的重要手段。云计算的资源调度与优化通过对云计算环境中的计算资源进行有效的配置与管理, 从而提升云计算系统的运行效率。论文阐述了大数据智能运维概念, 分析了在云计算环境下资源调度的挑战, 并对大数据智能运维在云计算环境下的资源调度与优化进行了探讨。

关键词: 大数据智能运维; 云计算环境; 资源调度

Research on Resource Scheduling and Optimization of Big Data Intelligent Operations in Cloud Computing Environment

Jialin Wang

The 15th Research Institute of China Electronics Technology Group Corporation, Beijing, 100083, China

Abstract: In the cloud computing environment, big data analysis has become an important means for enterprises to effectively utilize massive amounts of data for management and optimization. The resource scheduling and optimization of cloud computing improve the operational efficiency of cloud computing systems by effectively configuring and managing computing resources in the cloud computing environment. This paper elaborates on the concept of big data intelligent operation and maintenance, analyzes the challenges of resource scheduling in cloud computing environments, and explores the resource scheduling and optimization of big data intelligent operation and maintenance in cloud computing environments.

Keywords: big data intelligent operation and maintenance; cloud computing environment; resource scheduling

0 前言

云计算是一种基于需求的计算模式, 它可以为用户提供灵活而高效的计算资源, 是一种新兴的计算方式。但是, 由于服务需求的动态变化以及资源种类的多样性, 现有的资源调度算法已经很难满足其动态、高效的要求。大数据智能化运维以数据驱动的智能算法为基础, 自动优化资源配置, 可以大幅提高云环境中的资源利用效率。

1 大数据智能运维概念

大数据智能运维是一种基于大数据分析、机器学习等现代技术手段, 可以进行实时监测、自动处理、预测分析和优化决策等。AIOps 的核心是利用智能化的方法, 使运维工作由被动反应向主动防范转变, 从而达到自动化和智能化的目的。采用 Agentless 等方法, 对信息系统中的监测数据、网络流量等数据进行综合采集。该数据向 AIOps 平台提供完整的系统操作视图。大数据分析通过对大量运行和运行数据进行处理, 为辅助决策提供支撑。例如, 采用时序分析等方法, 可以对系统的负荷、资源利用率等做出预测。机器学习法是通过分析数据等, 对可能出现的问题进行预测, 并及时进行干预。AIOps 与机器人过程自动化 (RPA) 相结合, 实现日志分析、配置管理等重复运行任务的自动化。这样的

自动化不仅提升了运行效率, 而且也减少人为的误差^[1]。

2 在云计算环境下资源调度的挑战

2.1 复杂的资源调度

在云计算环境下, 包括计算资源、网络资源等维度, 又常常分散于不同的数据中心, 构成了巨大的资源池。在这样的环境下, 传统的静态配置方法已经不能满足企业的需求。计算资源配置的复杂程度表现为不同的计算任务对计算能力及并行处理能力等要求存在差异。云计算平台要求在多个应用场景下, 根据任务特点, 灵活配置不同类型的计算资源。但是, 由于计算资源具有动态等特点, 如何高效地进行资源配置具有较大的挑战。同时, 在存储资源的配置上, 也遇到许多困难。大数据时代来临, 对数据的存储要求呈爆炸性增长。云计算环境下, 如何对海量存储资源进行有效的管理, 保证数据安全、高效的存取成为亟待解决的问题。但是, 由于存储资源的异构性, 数据分布和数据频繁访问, 使系统的资源配置变得更加复杂。

2.2 需求动态变动

业务流量的高峰与低谷如潮水一般不可预测, 也无法避免, 使资源需求具有较强的波动性与不确定性。这一波动性既表现在时间上, 也表现在不同时期的资源需求、业务类型和外部环境等方面。从时间上看, 电商促销季等流量高峰

期将快速推动计算、网络等资源需求。但在低谷期，却有可能出现资源闲置，浪费大量的资源。这就需要云计算服务提供商具有快速响应的资源配置机制，保证高峰资源充分，低谷期有效回收与重新使用。不同种类的服务也使资源调度变得更加复杂。不同服务对资源有不同的要求，如数据密集的服务对带宽等要求较高，而计算密集服务要求更高的性能。云服务提供商必须在保证资源高效调度的前提下，持续地对各种服务进行优化，以适应各种服务的具体要求。同时，由于用户行为的不确定性，也给资源调度带来了巨大的挑战。由于用户的访问习惯和偏好的改变及突发事件的发生，使得网络资源的需求发生突变^[2]。

2.3 故障处理

云计算平台因其规模大、组件多等特性，在运行过程中难免会有硬件故障、网络中断等意外事件。为了保证服务的持续与稳定，保证服务的持续与稳定，不会因为服务中断而给顾客及服务带来负面影响。在资源调度中，必须构建有效的故障监测机制。通过对云平台各节点的运行状态、网络通信状态等进行实时监测，对异常情况进行检测，并启动相应的故障处理流程。这一监测并不局限于硬件，而是包括软件应用和数据存储等方面，要能够及时发现问题。在故障处理过程中，迅速找到故障原因非常重要。通过对所采集到的故障信息进行深入分析，从而快速地确定出故障的确切地点及原因，缩短故障诊断的时间，并为随后的故障恢复争取时间。当发生故障时，需要对资源调度系统进行自动故障恢复，使故障对企业运营造成的影响降到最低。此外，该系统还提供人工操作的功能，让管理人员通过手工方式处理复杂的故障状况。

3 大数据智能运维在云计算环境下资源调度的应用

3.1 预测建模

云计算大数据智能化运营管理的重要应用是对其进行预测建模。利用机器学习等方法，对历史数据进行分析与建模，实现对未来资源需求变化的预测，为优化配置资源提供科学依据。这样可以提高系统的资源利用效率，有效地减少运行维护费用。机器学习是一种新兴的数学模型。在云计算环境下，通过对不同类型数据进行分析，如 CPU 利用率、磁盘 I/O 等。采用森随机森林回归模型。随机森林是一种融合学习的方法，通过多个决策的构建，提升预测精度和稳定性。利用测试集对其进行预测，并利用均方误差等方法对模型的性能进行评估。例如，某家大型电商平台采用智能维护系统，运用 AIOps 平台对常见的问题进行自动化处理，并对资源需求进行大数据的分析与预测，使系统的故障响应时间减少 70%，年维护费用减少 30%^[3]。云计算环境下，利用机器学习模型对资源的需求进行预测，并对其进行优化，可以有效地提升系统的资源利用效率，提升系统的整体性能。

3.2 智能调度算法

基于人工智能、深度学习等理论，采取智能调度算法。该方法可以在一定程度上实现对网络资源的有效利用。利用深度学习等方法，对运行数据进行分析，并对异常情况进行预测。例如，通过对历史和监测信息的分析，可以对未来的资源需求进行预先预测，实现对资源的优化配置。深度学习（神经网络等）可以应对复杂运行环境，为用户提供智能解决方案。又如，利用 LSTM 对时间序列进行预测，可以对未来的资源需要进行精确的预测，并可以提前进行资源配置。Q-Learning 通过强化学习算法，对状态—行为的最佳值函数进行学习，对资源调配进行动态调整。再如，某云平台利用 Q-Learning 算法对资源进行动态调整，从而有效地提升系统的资源利用率。在云环境中，Q-Learning 算法通过模拟不同的负载情况下的资源配置策略，对资源的配置进行优化，可以有效地提高网络资源利用率。多智能体强化学习（MARL）利用多智能体协同优化资源调度，可以有效提高系统的运行效率。在动态的云计算环境下，通过多 Agent 协作来实现对资源的最优配置，提高系统的整体效率。例如，在某云平台上，利用多 Agent 增强学习方法，通过多方主体协同优化资源配置，可以缩短网络的响应时间。

3.3 自动化执行

大数据智能运维在云计算环境中，特别是在资源调度中，具有较好的灵活性。将其与智能调度技术相结合，不仅可以对资源进行动态配置和调整，而且可以大大提高系统的资源利用效率，保证服务质量和用户体验度。智能调度系统是大数据智能运维的关键部件，其主要功能是对云平台资源（CPU、内存等）进行智能化管理与优化配置。其中，人工智能算法可以实时调度现有资源，并对未来的资源需求进行预测，实现对资源的精确调度。例如，某家大型电商平台在促销期间，由于后台服务资源的需求量急剧增加。传统的人工调度方法很难对突发事件进行快速响应，从而造成服务延误。在引入大数据的大数据智能运维后，该系统可以对一些重要的数据进行实时监测，如数据库的查询次数等。采用深度学习方法，对未来数个小时内的资源需求量进行预测，并根据实际情况，自动启动资源扩充策略，提高服务器与数据库的连接数量，保证服务的稳定运行。通过智能化运营，平均响应时间减少了 30%，消费者的满意度也得到了明显的提高^[4]。

4 大数据智能运维在云计算环境下的资源调度优化

4.1 基于机器学习的自适应调度算法

云计算环境下大数据智能化运维管理，特别是以机器学习为基础的自适应调度算法，是提高云服务质量和响应能力的重要手段。该方法是通过 CPU、内存、任务执行时间等数据进行深入的挖掘与分析，建立高精度的机器学习模

型,从而准确地预测未来的任务负荷。通过对业务流程的提前预测,实现对业务资源的有效使用,并对业务服务质量进行不断优化。例如,某云计算平台需要面对大量的用户需求及高动态性的业务负荷。基于深度学习的资源调度方法。在前期工作基础上,通过采集一年不同的运行数据,对不同类型的应用程序进行分析。采用神经网络模型对所采集的数据进行深度学习与模式识别,并对其在不同时期的计算、网络等方面的特殊要求进行预测,可以明显提高系统的资源利用效率。使 CPU 的平均使用效率从 60% 提升到 85%,而存储器的使用率从 55% 增加到 78%^[5]。同时,该系统还能使其响应速度降低 20%,大大提高了用户的满意度与企业的市场竞争力。

4.2 算智能调度算法

算力网络融合“算力+网络”多因素智能计算,根据需求动态制定服务调度方案,实现资源利用率的最大化。该智能调度框架以“算法+知识+数据+算力”为基础,可以满足复杂任务的调度要求。算力网络通过对计算资源的实时感知,对计算任务进行有效的调度与调度,从而构建具有感知能力和可调度能力的计算资源网络。基于资源调度算法,该方法综合考虑网络时延及用户多样化等因素。智能调度架构可以针对不同的客户提供个性化的资源解决方案,从而更好地满足客户的需要,提高客户的使用体验与满意度。对计算、存储等资源进行合理配置,提高了资源配置的效率。利用机器学习等方法,对用户行为与资源利用规律进行实时分析,并对未来的需求进行预测,进而进行动态调度。在物联网环境下,算力网络融合了云计算的可伸缩性和分布式存储的高可用性,为大规模感知数据提供了新的解决方案。例如,某汽车制造企业应用分布式对象存储技术和云原生算力调度,将冲压过程模拟数据预处理时间由 3.2h 缩短到 18min,内存开销减少 57%。在智慧安保系统中,利用先进的算法模型,对大量的监测数据进行实时分析,利用算力网络,可以快速地发现隐患,从而保证系统的安全性。采用多因素的智能运算与动态调度,极大地提高系统的资源利用率。某家企业在对资源调度进行优化后,其存储成本下降了 57%^[6]。

4.3 多模态资源一体化管控

多模态资源一体化管控是其中的核心策略,已逐渐成为提高资源使用效率与稳定运行的重要途径。云计算中存在着大量的资源,其中包括计算、存储等多种资源。这些资源在使用与管理方面也有较大的差异。传统的资源管理模式很难有效地处理这一复杂问题,多模态资源一体化管控通过建

立多维度的资源知识地图,对资源进行分级管理与动态配置。多维度资源知识地图就是通过对资源及其联系的可视化展示,对资源的属性、使用等信息进行采集与分析,建立起全面的资源知识地图。该地图有助于运营人员对资源的状况进行全面的认识,并对其进行优化调度。例如,某云计算平台推出的 OpenLake,其采用 DLF 元数据管理平台对结构化、无结构化的数据进行统一管理,建立起“大数据—搜索—人工智能”集成的能力体系。该方法不仅可以对多模态数据进行统一管理,而且可以使不同类型的计算引擎进行协作,从而提升系统的资源利用效率。基于多维度的资源知识地图,对多种资源进行分类管理与动态调度。将资源划分为不同类型,并据此设计更合理的调度方案。同时,通过动态组态技术,实现了对网络资源的合理配置与利用,保证整个系统处于最佳状态。例如,某企业通过运用多模态资源一体化管控方法,对其进行有效的资源管理与动态配置,保证资源的有效使用。通过对多模态资源一体化管控,可以使资源利用效率提高 30%、维护费用减少 15%。

5 结语

总之,在云计算环境下,对资源调度与优化进行研究具有重要的意义。大数据智能运维是云计算资源调度和优化的重要手段。对其进行深入的分析与预测,可以实现对资源的动态调度与优化,从而提升系统的整体效能。但是,在大数据的规模与复杂度等方面仍存在诸多挑战。随着科技的不断进步与发展,云计算中的大数据智能运维将会不断拓展,为云计算的优化与创新提供新的动力。

参考文献:

- [1] 于雅娴.大数据驱动的云计算资源动态调度算法研究[J].信息与电脑(理论版),2024,36(18):160-163.
- [2] 都繁杰,李静,王亮,等.MEF:基于在线容错的云计算资源调度方法[J].计算机应用与软件,2024,41(3):336-344.
- [3] 张桂兰.云计算环境下的资源调度与优化算法研究[J].信息系统工程,2023(12):64-67.
- [4] 石学鹏.云计算环境下的虚拟机迁移算法优化与资源调度研究[J].电子元器件与信息技术,2023,7(8):105-109.
- [5] 王婷.人工智能技术的云计算资源调度优化研究[J].自动化技术与应用,2023,42(8):157-160.
- [6] 李伟彦,董宝良,王凯,等.基于金豺优化算法的云计算资源调度研究[J].电子设计工程,2023,31(15):41-45.

作者简介:王贾琳(1990-),女,中国北京人,硕士,高级工程师,从事信息系统管理与网络安全研究。