



(12) 发明专利申请

(10) 申请公布号 CN 118819843 A

(43) 申请公布日 2024. 10. 22

(21) 申请号 202410884313.1

(22) 申请日 2024.07.02

(71) 申请人 珠海市珠港机场管理有限公司

地址 519041 广东省珠海市珠海机场候机楼办公区

(72) 发明人 周福兴 赵兆民 叶成就

(74) 专利代理机构 北京国审领航知识产权代理有限公司 (特殊普通合伙)
16157

专利代理师 刘中航

(51) Int. Cl.

G06F 9/50 (2006.01)

G06F 11/34 (2006.01)

G06F 11/30 (2006.01)

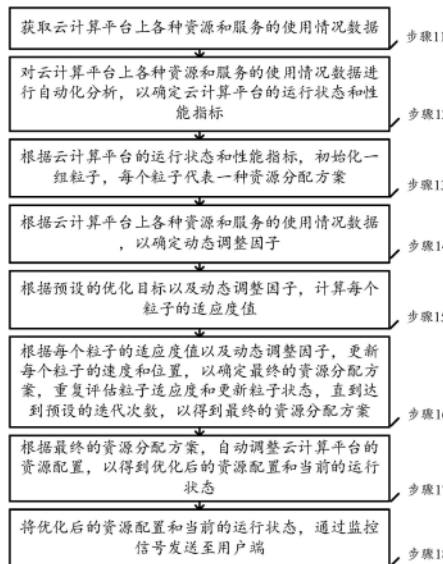
权利要求书2页 说明书9页 附图2页

(54) 发明名称

一种基于云计算平台的业务监控方法及系统

(57) 摘要

本发明涉及一种基于云计算平台的业务监控方法及系统,涉及数据处理技术领域,所述方法:根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态。本发明不仅能够实时监控云计算平台的运行状态,还能够根据实际需求进行智能的资源分配,从而提高资源利用率。



1. 一种基于云计算平台的业务监控方法,其特征在于,包括:

获取云计算平台上各种资源和服务的使用情况数据;

对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;

根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;

根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;

根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;

根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;

根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;

将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。

2. 根据权利要求1所述的基于云计算平台的业务监控方法,其特征在于,对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标,包括:

从资源和服务的使用情况数据中,筛选与云计算平台运行状态和性能指标的相关特征;

根据相关特征,通过肘部法则确定聚类的数量K;

根据聚类的数量K,对资源和服务的使用情况数据进行聚类,在聚类过程中,将资源和服务的使用情况数据划分为K个簇,每个簇代表一种相似的运行状态,直到达到预设的迭代次数,以得到聚类结果;

根据聚类结果,确定云计算平台的运行状态和性能指标,包括每个簇的平均值或中位数。

3. 根据权利要求2所述的基于云计算平台的业务监控方法,其特征在于,根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子,包括:

根据云计算平台上各种资源和服务的使用情况数据,通过计算动态调整因子。

4. 根据权利要求3所述的基于云计算平台的业务监控方法,其特征在于,资源和服务的使用情况数据包括CPU使用率、内存占用、网络带宽以及存储使用情况。

5. 一种基于云计算平台的业务监控系统,其特征在于,包括:

获取模块,用于获取云计算平台上各种资源和服务的使用情况数据;对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;

处理模块,用于根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方

案;根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。

6.一种计算设备,其特征在于,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-4中任一项所述的方法。

7.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有程序,该程序被处理器执行时实现如权利要求1-4中任一项所述的方法。

一种基于云计算平台的业务监控方法及系统

技术领域

[0001] 本发明涉及数据处理技术领域,具体地说,涉及一种基于云计算平台的业务监控方法及系统。

背景技术

[0002] 云计算平台通过虚拟化技术实现了资源的灵活分配与高效利用,为企业和个人用户提供了强大的计算能力和存储空间。然而,随着云计算平台规模的不断扩大和服务种类的日益增多,如何有效地监控和管理这些资源变得尤为重要。

[0003] 例如,传统的云计算平台监控方法通常侧重于对单一资源或服务的监控,而可能无法全面、实时地反映整个云计算平台的运行状态和性能指标。此外,这些方法在资源分配和优化方面也存在一定的局限性,可能无法根据云计算平台的实时状态和用户需求进行动态调整,导致资源浪费或性能瓶颈。

发明内容

[0004] 本发明要解决的技术问题在于克服现有技术的不足,提供一种基于云计算平台的业务监控方法及系统,不仅能够实时监控云计算平台的运行状态,还能够根据实际需求进行智能的资源分配,从而提高资源利用率。

[0005] 为解决上述技术问题,本发明采用技术方案的基本构思是:

[0006] 第一方面,一种基于云计算平台的业务监控方法,包括:

[0007] 获取云计算平台上各种资源和服务的使用情况数据;

[0008] 对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;

[0009] 根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;

[0010] 根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;

[0011] 根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;

[0012] 根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;

[0013] 根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;

[0014] 将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。

[0015] 优选的,对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标,包括:

[0016] 从资源和服务的使用情况数据中,筛选与云计算平台运行状态和性能指标的相关特征;

- [0017] 根据相关特征,通过肘部法则确定聚类的数量 K ;
- [0018] 根据聚类的数量 K ,对资源和服务的使用情况数据进行聚类,在聚类过程中,将资源和服务的使用情况数据划分为 K 个簇,每个簇代表一种相似的运行状态,直到达到预设的迭代次数,以得到聚类结果;
- [0019] 根据聚类结果,确定云计算平台的运行状态和性能指标,包括每个簇的平均值或中位数。
- [0020] 根据云计算平台上各种资源和服务的使用情况数据,通过计算动态调整因子。
- [0021] 资源和服务的使用情况数据包括CPU使用率、内存占用、网络带宽以及存储使用情况。
- [0022] 第二方面,一种基于云计算平台的业务监控系统,包括:
- [0023] 获取模块,用于获取云计算平台上各种资源和服务的使用情况数据;对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;
- [0024] 处理模块,用于根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。
- [0025] 第三方面,一种计算设备,包括:
- [0026] 一个或多个处理器;
- [0027] 存储装置,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现所述的方法。
- [0028] 第四方面,一种计算机可读存储介质,所述计算机可读存储介质中存储有程序,该程序被处理器执行时实现所述的方法。
- [0029] 采用上述技术方案后,本发明与现有技术相比具有以下有益效果:
- [0030] 通过获取并分析云计算平台上各种资源和服务的使用情况数据,能够实时监控云计算平台的整体运行状态和性能指标。通过设置粒子群优化算法和动态调整因子,本发明能够根据实际情况动态调整资源分配方案。这种动态优化机制能够显著提高云计算平台的资源利用率,避免资源浪费,并确保平台的高效运行。
- [0031] 每个粒子的适应度值根据预设的优化目标和动态调整因子进行计算,使得资源分配方案能够随着云计算平台状态的变化而自动调整,增强了系统的自适应性和灵活性。通过不断优化资源配置,本发明能够确保云计算平台始终保持在最佳状态,从而为用户提供更稳定、更快速的服务,极大提升了用户体验。
- [0032] 通过实时监控和动态调整,本发明能够及时发现并解决云计算平台可能出现的问题,如资源瓶颈、性能下降等,从而增强系统的稳定性和可靠性。通过自动化的资源分配和优化,减少了人工干预的需要,降低了云计算平台的运维成本。

附图说明

[0033] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。后文将参照附图以示例性而非限制性的方式详细描述本申请的一些具体实施例。附图中相同的附图标记标示了相同或类似的部件或部分,本领域技术人员应该理解的是,这些附图未必是按比例绘制的,在附图中:

[0034] 图1是本发明一种基于云计算平台的业务监控方法流程示意图。

[0035] 图2是本发明一种基于云计算平台的业务监控系统示意图。

具体实施方式

[0036] 为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述。显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本申请保护的范围。

[0037] 本申请下述实施例以基于云计算平台的业务监控方法为例进行详细说明本申请的方案,但是此实施例并不能限制本申请保护范围。

[0038] 如图1所示,本发明提供了一种基于云计算平台的业务监控方法,所述方法包括以下步骤:

[0039] 步骤11,获取云计算平台上各种资源和服务的使用情况数据;资源和服务的使用情况数据包括CPU使用率、内存占用、网络带宽以及存储使用情况;

[0040] 步骤12,对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;

[0041] 步骤13,根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;

[0042] 步骤14,根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;

[0043] 步骤15,根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;

[0044] 步骤16,根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;

[0045] 步骤17,根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;

[0046] 步骤18,将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。

[0047] 在本发明实施例中,通过获取并分析云计算平台上各种资源和服务的使用情况数据,能够实时监控云计算平台的整体运行状态和性能指标。通过设置粒子群优化算法和动态调整因子,本发明能够根据实际情况动态调整资源分配方案。这种动态优化机制能够显著提高云计算平台的资源利用率,避免资源浪费,并确保平台的高效运行。每个粒子的适应度值根据预设的优化目标和动态调整因子进行计算,使得资源分配方案能够随着云计算平台状态的变化而自动调整,增强了系统的自适应性和灵活性。通过不断优化资源配置,本发

明能够确保云计算平台始终保持在最佳状态,从而为用户提供更稳定、更快速的服务,极大提升了用户体验。通过实时监控和动态调整,本发明能够及时发现并解决云计算平台可能出现的问题,如资源瓶颈、性能下降等,从而增强系统的稳定性和可靠性。通过自动化的资源分配和优化,减少了人工干预的需要,降低了云计算平台的运维成本。

[0048] 在本发明另一优选的实施例中,设计一套数据收集接口或利用现有的监控工具(如Zabbix、Prometheus等)来收集所需的数据;这些接口应与云计算平台的各个组件(如虚拟机、容器、物理服务器等)紧密集成,以便能够实时、准确地获取资源使用情况。通过监控工具或接口定期(如每秒或每分钟)采集每个计算节点的CPU使用率,CPU使用率通常表示为某个时间段内CPU被占用的时间与该时间段总时间的比例;收集的数据可以包括用户态CPU时间、系统态CPU时间以及空闲时间等,以便更全面地分析CPU的使用情况。监控工具会定期记录每个计算节点的内存使用情况,包括已用内存、空闲内存、缓存内存等,这些数据有助于分析内存的使用效率,以及是否存在内存泄漏或过度使用的情况。

[0049] 通过网络接口卡(NIC)或其他网络监控工具来收集网络带宽的使用情况,这包括入站和出站的数据流量、网络延迟、丢包率等关键指标,对网络带宽的监控有助于识别网络瓶颈和优化数据传输效率。监控工具会记录存储设备的容量、已用空间、剩余空间以及I/O性能(如读写速度、IOPS等)。收集到的资源使用情况数据会被整合并存储到中央数据库或时间序列数据库中,这些数据可以按照时间戳进行索引,以便后续的数据分析和可视化,在收集和使用这些数据时,需要确保数据的安全性和隐私性,应实施适当的访问控制和加密措施,以防止数据泄露或未经授权的访问。

[0050] 在本发明一优选的实施例中,对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标,包括:

[0051] 从资源和服务的使用情况数据中,筛选与云计算平台运行状态和性能指标的相关特征;

[0052] 根据相关特征,通过肘部法则确定聚类的数量K;

[0053] 根据聚类的数量K,对资源和服务的使用情况数据进行聚类,在聚类过程中,将资源和服务的使用情况数据划分为K个簇,每个簇代表一种相似的运行状态,直到达到预设的迭代次数,以得到聚类结果;

[0054] 根据聚类结果,确定云计算平台的运行状态和性能指标,包括每个簇的平均值或中位数。

[0055] 在本发明实施例中,对收集到的资源和服务使用情况数据进行预处理,包括数据清洗、去重、异常值处理等;接着,从这些数据中筛选出与云计算平台运行状态和性能指标相关的特征,特征包括CPU使用率、内存占用率、网络带宽利用率、存储I/O性能等;通过绘制不同K值的聚类成本(如WSS, Within-Cluster Sum of Squares)曲线图,找到曲线的“肘点”(即成本下降速度突然变缓的点),该点对应的K值即为较佳的聚类数量,根据上一步确定的K值,采用K-means聚类算法对资源和服务的使用情况数据进行聚类,在聚类过程中,算法会随机选择K个初始质心,然后将每个数据点分配给最近的质心,形成K个簇;重新计算每个簇的质心,并重复数据点的分配过程,直到质心的位置不再发生显著变化或达到预设的迭代次数,聚类完成后,每个簇代表了一种相似的运行状态,对于每个簇,可以计算其内部数据点的平均值或中位数,以此来代表该运行状态下的性能指标。例如,可以计算每个簇的平均

CPU使用率、平均内存占用率等,从而得到云计算平台在不同运行状态下的性能表现。根据聚类结果,可以确定云计算平台的几种典型运行状态(如空闲状态、轻负载状态、重负载状态等),针对每种运行状态,通过计算簇内数据点的统计特征(如平均值、中位数等),得到该状态下的具体性能指标,这些性能指标可以用于评估云计算平台的性能表现。通过上述步骤,可以实现对云计算平台上各种资源和服务的使用情况数据的自动化分析,进而确定云计算平台的运行状态和性能指标。

[0056] 在本发明一优选的实施例中,根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案,包括:

[0057] 每个粒子将包含位置(代表资源分配方案)和速度两个主要属性,位置可以是一个多维向量,每个维度代表一种资源的分配量,如CPU核心数、内存容量、网络带宽等,速度也是一个多维向量,表示粒子在各个资源维度上的调整方向和速率。

[0058] 根据问题的复杂性和所需的搜索精度,确定初始化的粒子数量。粒子数量越多,搜索空间越广泛,但计算成本也越高;基于云计算平台的当前运行状态和性能指标,为每个粒子的位置向量赋予初始值,这些初始值可以根据历史数据、经验规则或随机生成,但要确保它们在合理的资源分配范围内。例如,如果当前CPU使用率高,那么初始化的粒子位置可能在分配更多CPU资源的区域。粒子的初始速度可以设置为零或根据某种分布(如均匀分布或正态分布)随机生成,速度的初始化应考虑到搜索空间的规模和问题的特性,以避免粒子过早收敛或偏离搜索空间。对于每个初始化的粒子(即资源分配方案),使用预定的评价标准(如成本、性能、资源利用率等)进行评估,这个评估将作为后续迭代中粒子更新和选择的基础,完成粒子的初始化和评估后,系统准备进入迭代过程,通过不断更新粒子的位置和速度来搜索最优的资源分配方案。通过这个步骤,为后续的粒子群优化算法提供了一个初始的解空间,每个粒子都代表了一个可能的资源分配方案。在接下来的迭代中,算法将根据每个粒子的性能和适应度来调整它们的位置和速度,以找到最优的资源分配策略。

[0059] 在本发明一优选的实施例中,根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子,包括:

[0060] 根据云计算平台上各种资源和服务的使用情况数据,通过 $DAF = \gamma \times$

$$\left(\alpha \times \frac{\sum_{i=1}^n (w_i \times U_i \times \frac{R_i}{R_{max}})}{\sum_{i=1}^n w_i} + \beta \times \frac{1}{m} \sum_{j=1}^m \left(L_j \times \frac{S_j}{S_{max}} \right) \right) + (1 - \gamma) \times \left(\delta \times \frac{1}{p} \sum_{k=1}^p T_k + \right.$$

[0061] $\left. \epsilon \times \frac{1}{q} \sum_{l=1}^q V_l \right)$ 计算动态调整因子;

[0062] 其中,DAF是动态调整因子; γ 、 α 、 β 、 δ 和 ϵ 是权重系数,满足 $0 \leq \gamma, \alpha, \beta, \delta, \epsilon \leq 1$,且 $\alpha + \beta = 1$, $\delta + \epsilon = 1$;n是资源类型数量; w_i 是第i种资源的权重; U_i 是第i种资源的使用率; R_i 是第i种资源的当前请求量或需求量; R_{max} 是第i种资源的最大请求量或容量;m是考虑的服务类型数量; L_j 是第j种服务的负载指标; S_j 是第j种服务的当前并发量或会话数; S_{max} 是第j种服务的最大并发量或会话数;p是时间段数量; T_k 是第k个时间段内的平均响应时间或延迟;q是用户反馈数量; V_l 是第l个用户反馈的评分或满意度;i、j、l和k为索引。

[0063] 在本发明实施例中,该动态调整因子(DAF)综合考虑了多种资源和服务的使用情况,包括资源使用率、服务负载、响应时间以及用户反馈等,从而能够更全面、准确地反映云

计算平台的实际运行状态和需求。由于DAF是基于实时的资源和服务使用数据计算得出的,因此它能够动态地适应云计算平台的变化。当资源需求、服务负载或用户反馈发生变化时,DAF会相应地调整,为资源分配提供及时的指导。公式中的权重系数(如 γ 、 α 、 β 、 δ 和 ε)可以根据实际情况进行调整,这使得DAF具有更高的灵活性和可定制性。通过纳入用户反馈(如评分或满意度),DAF能够更直接地反映用户对云计算服务的感受。通过动态调整资源分配方案,基于DAF的优化策略可以帮助云计算平台在保障性能的同时,提高资源利用效率,减少资源浪费,并增强系统的稳定性。

[0064] 在本发明一优选的实施例中,根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值,包括:

[0065] 根据预设的优化目标以及动态调整因子,通过 $Fit_i = \left(\frac{1}{1 + \left| \frac{DAF_i - DAF_t}{DAF_t} \right|} \right) \times \left(\frac{1}{1 + Cost_i} \right)$ 计算每个粒子的适应度值,其中, Fit_i 是第i个粒子的适应度值; DAF_i 是根据第i个粒子的位置计算出的动态调整因子; DAF_t 是预设的优化目标中期望达到的动态调整因子值; $Cost_i$ 是与第i个粒子相关的成本。

[0066] 在本发明实施例中,适应度函数的设计考虑了预设的优化目标(即期望达到的动态调整因子值)。这确保了优化过程是有目标导向的,能够专注于提升那些对达到优化目标最为关键的粒子。适应度值的计算不仅考虑了动态调整因子与目标的接近程度,还通过成本项的引入,平衡了性能提升与资源成本之间的关系。这有助于避免过度投入资源以追求性能提升,从而实现经济效益与性能优化的平衡。由于动态调整因子是根据每个粒子的位置实时计算得出的,因此适应度值也能够动态地反映粒子在当前环境下的优劣。这种动态适应性有助于算法在变化的云计算环境中持续有效地进行资源分配优化。适应度函数采用了平滑的过渡形式(通过分数形式实现),这避免了在优化过程中出现突变或跳跃,有助于算法的稳定收敛。该适应度值的计算方案不仅适用于特定的云计算环境或应用场景,而且通过调整预设的优化目标和成本函数,可以灵活地应用于多种不同的云计算资源分配问题。通过综合考虑动态调整因子和成本,该适应度函数有助于引导粒子群算法在全局范围内搜索最优解,而不仅仅是陷入局部最优。

[0067] 在本发明一优选的实施例中,在根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置当中,速度的更新公式为:

[0068]
$$v_i^{(t+1)} = w \times v_i^{(t)} + c_1 \times r_1 \times (p_i - x_i^{(t)}) + c_2 \times r_2 \times (gb - x_i^{(t)}) + c_3 \times (DAF_i - DAF_t);$$

[0069] 其中, $v_i^{(t)}$ 和 $v_i^{(t+1)}$ 分别是第i个粒子在时间t和t+1的速度; w 是惯性权重, c_1 和 c_2 是学习因子, r_1 和 r_2 是介于0和1之间的随机数; p_i 是第i个粒子的最优位置; gb 是所有粒子的全局最优位置。

[0070] 在本发明实施例中,惯性权重 w 有助于平衡全局搜索和局部搜索的能力。一个较大的 w 值有利于全局搜索,而较小的 w 值则更倾向于局部搜索。通过调整 w 的大小,可以控制粒子在搜索空间中的探索和开发能力。学习因子 c_1 它决定了粒子向其自身历史最优位置学习

的程度。随机数 r_1 它为这种学习引入了随机性,有助于粒子在搜索空间中更广泛地探索。学习因子 c_2 决定了粒子向全局最优位置学习的程度。随机数 r_2 同样为这种学习引入了随机性,有助于粒子之间共享信息,促进整个粒子群向全局最优解收敛。系数 c_3 引入了动态调整因子与预设目标之间的差异来影响粒子的速度更新,这使得粒子能够根据当前的动态调整因子值与目标值之间的差距来调整其搜索方向和速度,从而更加直接地朝着优化目标前进。

[0071] 在本发明一优选的实施例中,在根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置当中,位置的更新公式为:

$$[0072] \quad x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)};$$

[0073] 其中, $x_i^{(t)}$ 和 $x_i^{(t+1)}$ 分别是第 i 个粒子在时间 t 和 $t+1$ 的位置。

[0074] 在本发明实施例中,由于速度 $v_i^{(t+1)}$ 是根据粒子的个体最优位置、全局最优位置以及动态调整因子等多个因素计算得出的,因此它反映了粒子在搜索空间中的动态行为,这使得位置更新能够适应搜索过程的进展,并有助于粒子向更优的位置移动。位置更新公式提供了粒子在搜索空间中的连续移动。这种连续性有助于算法在搜索过程中保持稳定性和平滑性,避免突变或跳跃,从而有助于找到更优的解。通过考虑粒子的个体最优和全局最优位置,以及动态调整因子的影响,位置更新公式使得粒子能够根据搜索过程的进展自适应地调整其移动方向和步长。这种自适应性有助于算法在不同的搜索阶段采取不同的策略,从而更好地平衡探索和开发能力。

[0075] 步骤17,经过前面的优化算法(如粒子群优化算法)运行后,会得到一个最终的资源分配方案。这个方案详细说明了各个计算节点或服务应该分配多少资源(如CPU核心数、内存大小、存储容量、网络带宽等);在实际应用资源分配方案之前,系统会进行一系列的验证检查,确保方案的合理性和可行性。例如,检查资源总量是否超出云平台的实际承载能力,以及是否满足各种服务级别协议(SLA)的要求;一旦方案通过验证,云计算平台的管理系统(如Kubernetes、OpenStack等)会开始执行资源调整命令,这包括启动或关闭虚拟机、调整容器的资源配额、重新分配网络带宽等;在资源调整过程中,系统会实时监控各项指标的变化,如资源利用率、服务响应时间等,以确保调整过程顺利进行且没有对服务造成不良影响;调整完成后,系统会记录新的资源配置情况,并更新云计算平台的运行状态。

[0076] 步骤18,系统会收集优化后的资源配置详细信息,包括每个节点或服务分配到的具体资源量,以及云计算平台当前的运行状态数据(如CPU使用率、内存占用率、网络吞吐量等);收集到的数据会被格式化成用户易于理解的报告或图表形式。这有助于用户直观地了解资源优化后的效果和云平台的运行状态;根据用户的偏好和系统的配置,选择合适的通信协议和方式将数据发送给用户端,这可能是通过电子邮件、短信通知、Web界面更新、API调用或其他自定义的通信方式;将格式化后的数据作为监控信号发送给用户端。用户可以根据这些信息评估资源优化的效果,以及云平台当前的性能和健康状况;用户接收到监控信号后,可以根据实际情况提供反馈,并可能需要对资源分配方案进行微调,以更好地满足业务需求或性能要求。这些反馈和调整会作为新一轮优化的输入,持续改进云计算平台的资源配置。

[0077] 图2所示,本发明的实施例还提供一种基于云计算平台的业务监控系统20,包括:

[0078] 获取模块21,用于获取云计算平台上各种资源和服务的使用情况数据;对云计算平台上各种资源和服务的使用情况数据进行自动化分析,以确定云计算平台的运行状态和性能指标;根据云计算平台的运行状态和性能指标,初始化一组粒子,每个粒子代表一种资源分配方案;

[0079] 处理模块22,用于根据云计算平台上各种资源和服务的使用情况数据,以确定动态调整因子;根据预设的优化目标以及动态调整因子,计算每个粒子的适应度值;根据每个粒子的适应度值以及动态调整因子,更新每个粒子的速度和位置,以确定最终的资源分配方案,重复评估粒子适应度和更新粒子状态,直到达到预设的迭代次数,以得到最终的资源分配方案;根据最终的资源分配方案,自动调整云计算平台的资源配置,以得到优化后的资源配置和当前的运行状态;将优化后的资源配置和当前的运行状态,通过监控信号发送至用户端。

[0080] 需要说明的是,该系统是与上述方法相对应的系统,上述方法实施例中的所有实现方式均适用于该实施例中,也能达到相同的技术效果。

[0081] 本发明的实施例还提供一种计算设备,包括:处理器、存储有计算机程序的存储器,所述计算机程序被处理器运行时,执行如上所述的方法。上述方法实施例中的所有实现方式均适用于该实施例中,也能达到相同的技术效果。

[0082] 本发明的实施例还提供一种计算机可读存储介质,存储指令,当所述指令在计算机上运行时,使得计算机执行如上所述的方法。上述方法实施例中的所有实现方式均适用于该实施例中,也能达到相同的技术效果。

[0083] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0084] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0085] 在本发明所提供的实施例中,应该理解到,所揭露的装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0086] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0087] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0088] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以

存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0089] 此外,需要指出的是,在本发明的装置和方法中,显然,各部件或各步骤是可以分解和/或重新组合的。这些分解和/或重新组合应视为本发明的等效方案。并且,执行上述系列处理的步骤可以自然地按照说明的顺序按时间顺序执行,但是并不需要一定按照时间顺序执行,某些步骤可以并行或彼此独立地执行。对本领域的普通技术人员而言,能够理解本发明的方法和装置的全部或者任何步骤或者部件,可以在任何计算装置(包括处理器、存储介质等)或者计算装置的网络中,以硬件、固件、软件或者它们的组合加以实现,这是本领域普通技术人员在阅读了本发明的说明的情况下运用的基本编程技能就能实现的。

[0090] 因此,本发明的目的还可以通过在任何计算装置上运行一个程序或者一组程序来实现。所述计算装置可以是公知的通用装置。因此,本发明的目的也可以仅仅通过提供包含实现所述方法或者装置的程序代码的程序产品来实现。也就是说,这样的程序产品也构成本发明,并且存储有这样的程序产品的存储介质也构成本发明。显然,所述存储介质可以是任何公知的存储介质或者将来所开发出来的任何存储介质。还需要指出的是,在本发明的装置和方法中,显然,各部件或各步骤是可以分解和/或重新组合的。这些分解和/或重新组合应视为本发明的等效方案。并且,执行上述系列处理的步骤可以自然地按照说明的顺序按时间顺序执行,但是并不需要一定按照时间顺序执行。某些步骤可以并行或彼此独立地执行。

[0091] 以上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明所述原理的前提下,还可以作出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

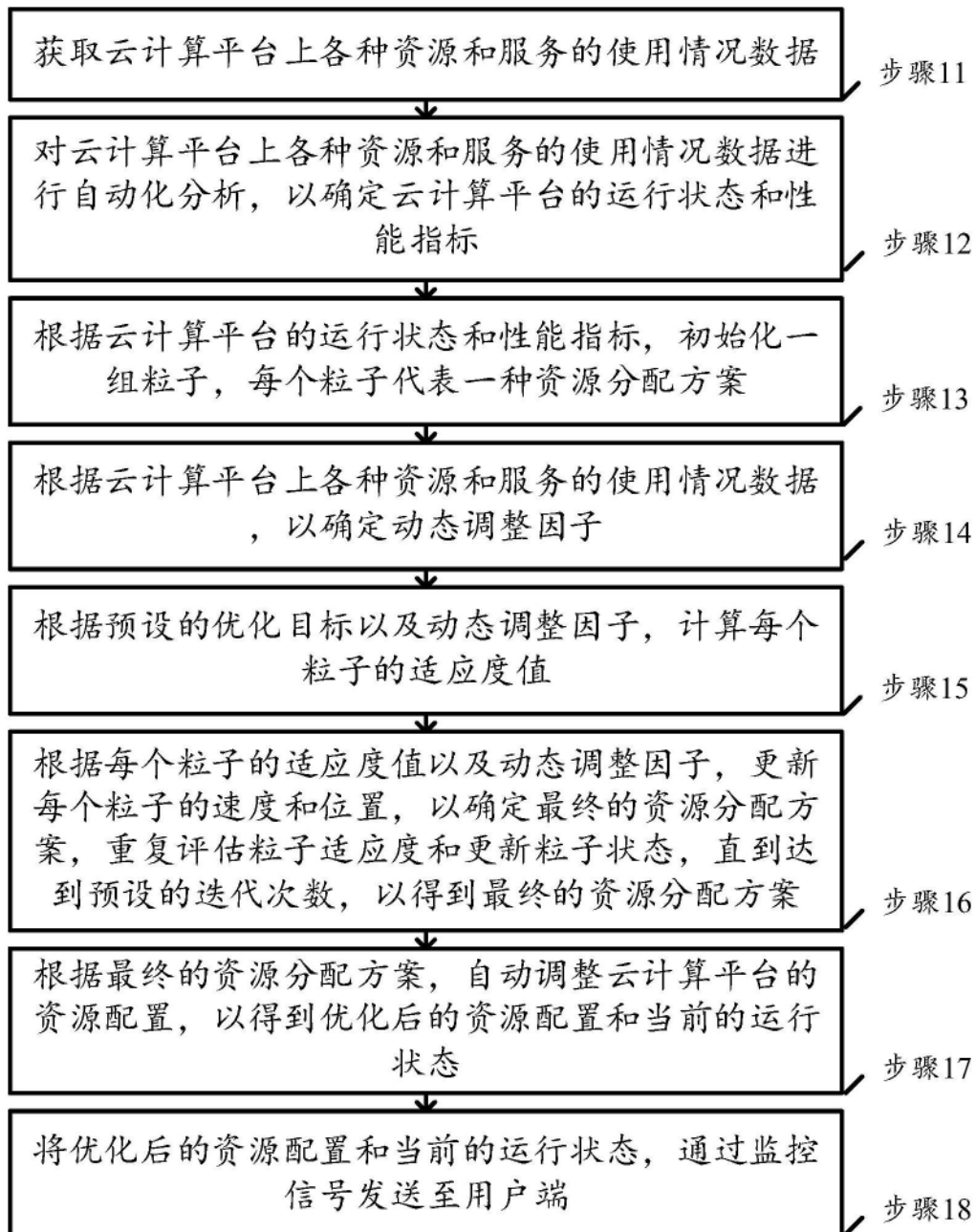


图1

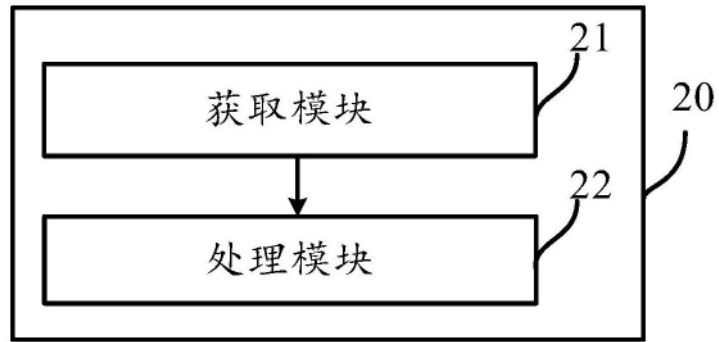


图2