



(12)发明专利申请

(10)申请公布号 CN 110806954 A

(43)申请公布日 2020.02.18

(21)申请号 201910885428.1

(22)申请日 2019.09.19

(71)申请人 平安科技(深圳)有限公司

地址 518033 广东省深圳市福田区福田街
道福安社区益田路5033号平安金融中
心23楼

(72)发明人 徐锐杰

(74)专利代理机构 北京市京大律师事务所

11321

代理人 刘挽澜

(51)Int.Cl.

G06F 11/22(2006.01)

G06F 11/26(2006.01)

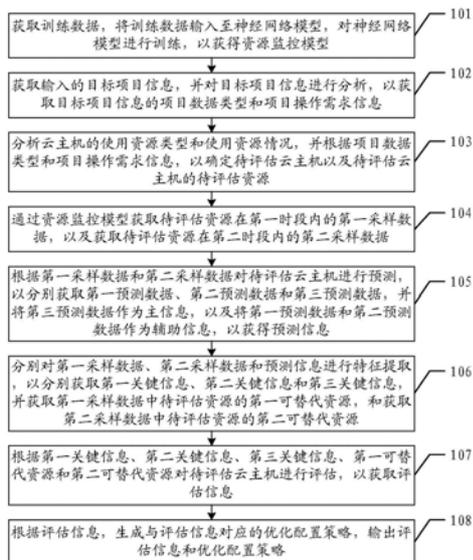
权利要求书4页 说明书17页 附图2页

(54)发明名称

评估云主机资源的方法、装置、设备及存储
介质

(57)摘要

本申请涉及云服务领域,提供一种评估云主
机资源的方法、装置、设备及存储介质,方法包
括:通过训练所得的资源监控模型对目标项目信
息、多个云主机的使用资源类型和使用资源情
况进行分析,以获取待评估云主机以及所述待评
估云主机的待评估资源,获取所述待评估资源
的第一采样数据和第二采样数据,根据所述第
一采样数据和第二采样数据获取预测信息,根
据所述第一采样数据、第二采样数据和所述预
测信息获取评估评估信息,根据所述评估信息
生成与所述评估信息对应的优化配置策略。采
用本方案,能够提高云主机的资源利用率。



1. 一种评估云主机资源的方法,其特征在于,所述方法包括:

获取训练数据,将所述训练数据输入至神经网络模型,并对所述神经网络模型进行训练,以获得资源监控模型,其中,所述训练数据包括多个云主机的多个时段的资源采样数据;

获取输入的目标项目信息,并对所述目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息;

获取云主机的使用资源类型和使用资源情况,并根据所述使用资源类型、所述使用资源情况、所述项目数据类型和所述项目操作需求信息,确定待评估云主机以及所述待评估云主机的待评估资源;

通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,其中,所述第一时段的起始时刻晚于所述第二时段的结束时刻;

根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测,以分别获取第一预测数据、第二预测数据和第三预测数据,并将所述第三预测数据作为主信息,以及将所述第一预测数据和所述第二预测数据作为辅助信息,以获取预测信息,其中,所述预测包括对所述待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,所述第三时段的起始时刻晚于所述第一时段的结束时刻;

分别对所述第一采样数据、所述第二采样数据和所述预测信息进行特征提取,以分别获取第一关键信息、第二关键信息和第三关键信息,并获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源;

根据所述第一关键信息、所述第二关键信息、所述第三关键信息、所述第一可替代资源和所述第二可替代资源对所述待评估云主机进行评估,以获取评估信息;

根据所述评估信息生成与所述评估信息对应的优化配置策略,并输出所述评估信息和所述优化配置策略。

2. 根据权利要求1所述的方法,其特征在于,所述获取输入的目标项目信息,并对所述目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息,包括:

创建项目操作需求表,其中,所述项目操作需求表包括项目的预设完成时间、主机资源需求量和所述主机资源需求量对应的最优主机资源分配量;

获取用户输入的目标项目信息,并对所述目标项目信息进行数据预处理,其中,所述数据预处理包括缺失值填补处理、去噪处理和数据标准化处理;

将经过数据预处理的目标项目信息分成N组,通过反复迭代法对分成N组的目标项目信息进行多次重新分组,以获取最优分组方案;

获取所述最优分组方案中各组目标项目信息的项目数据类型;

对所述目标项目信息进行分析以获取第四关键信息;

根据所述第四关键信息遍历所述项目操作需求表,以获取所述第四关键信息对应的项目操作需求信息。

3. 根据权利要求1所述的方法,其特征在于,所述待评估资源包括在第一时段内的第一任务和在第一时段内的第二任务,所述通过所述资源监控模型获取所述待评估资源在第一

时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,包括:

通过所述资源监控模型,获取所述第一任务中的第一占用率信息和第一优先级,以及获取所述第二任务中的第二占用率信息和第二优先级;

根据所述第一优先级,对所述第一任务进行分类,并标识第一类别标签,以及根据所述第二优先级,对所述第二任务进行分类,并标识第二类别标签;

根据所述第一占用率信息,对标识所述第一类别标签的第一任务进行分类,并标识第三类别标签,以及所述第二占用率信息,对标识所述第二类别标签的第二任务进行分类,并标识第四类别标签;

判断所述第一优先级是否符合第一预设采样条件和/或所述第一占用率信息是否符合第二预设采样条件,以及判断所述第二优先级是否符合第三预设采样条件和/或所述第二占用率信息是否符合第四预设采样条件;

若所述第一优先级符合第一预设采样条件和/或所述第一占用率信息符合第二预设采样条件,则按照预设第一采样频率,对符合第一预设采样条件和/或第二预设采样条件的同一类别标签的任务进行采样,以获得第一采样数据;

若所述第二优先级符合第三预设采样条件和/或所述第二占用率信息符合第四预设采样条件,则按照预设第二采样频率,对符合第三预设采样条件和/或第四预设采样条件的同一类别标签的任务进行采样,以获得第二采样数据。

4. 根据权利要求1所述的方法,其特征在于,所述获取预测信息之后,所述方法还包括:

获取所述预测信息的第一时间序列数据;

对所述第一时间序列数据进行滑动窗口处理,以生成预设数量的具备预设长度的时序子序列;

分析所述时序子序列的统计指标以获取统计特征信息,并以所述统计特征信息作为更新后的预测信息,其中,所述统计特征信息包括最大值、最小值、中位数、第一四分位数、第三四分位数、方差和标准差。

5. 根据权利要求1所述的方法,其特征在于,所述根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测之前,所述方法还包括:

获取所述第一采样数据的第二时间序列数据,以及获取所述第二采样数据的第三时间序列数据;

通过所述数据处理子模型中的指数加权移动平均法EWMA,分别对所述第二时间序列数据和所述第三时间序列数据进行求值,以获得第一平滑处理数据和第一平滑处理数据,对所述第二时间序列数据和所述第三时间序列数据的求值计算如下:

$$V_t = \frac{\alpha V_{t-1} + (1-\alpha)x_t}{1-\alpha^t},$$

其中, x_t 为时刻t的实际第二时间序列数据或实际第三时间序列数据,系数 α 为加权下降的速率, V_t 为t时刻的EWMA值;

通过所述数据处理子模型中的马氏距离Mahalanobis算法,分别对所述第一平滑处理数据和所述第二平滑处理数据中的极度异常点进行检测与标识,并删除所标识的极度异常点,检测与标识所述极度异常点的计算如下:

$$MDist(\mathbf{b}, \bar{\mathbf{b}}) = \sqrt{(\mathbf{b}, \bar{\mathbf{b}})^T \mathbf{S}^{-1} (\mathbf{b} - \bar{\mathbf{b}})},$$

其中, $MDist(\mathbf{b}, \bar{\mathbf{b}})$ 是 \mathbf{b} 与 $\bar{\mathbf{b}}$ 的距离, $\bar{\mathbf{b}}$ 是所述平滑处理数据的均值向量, \mathbf{b} 是为所述平滑处理数据中的其他对象, \mathbf{S} 是协方差矩阵。

6. 根据权利要求1所述的方法, 其特征在于, 所述方法包括资源数据库, 所述资源数据库包括收集的多种数据类型对应的资源数据, 所述获取所述第一采样数据中待评估资源的第一可替代资源, 和获取所述第二采样数据中待评估资源的第二可替代资源, 包括:

获取所述第一采样数据中待评估资源的第一特征信息, 以及获取所述第二采样数据中待评估资源的第二特征信息, 其中, 所述第一特征信息和所述第二特征信息均包括数据类型、资源总容量、各数据类型对应的资源的使用占比、资源的性能和特性;

根据所述第一特征信息在所述资源数据库中获取与所述第一特征信息对应的第一资源数据, 以及根据所述第二特征信息在所述资源数据库中获取与所述第二特征信息对应的第二资源数据;

通过预设替换条件对所述第一资源数据进行计算和筛选以获取第一可替代资源, 以及通过所述预设替换条件对所述第二资源数据进行计算和筛选以获取第二可替代资源;

根据所述第一可替代资源, 在所述待评估云主机之外的云主机中进行分析 and 匹配, 以获取第一可替代云主机, 并获取所述第一可替代云主机的第一资源使用信息, 并将所述第一可替代云主机和所述第一资源使用信息标记在所述第一可替代资源上, 以获取最终的第一可替代资源, 以及根据所述第二可替代资源, 在所述待评估云主机之外的云主机中进行分析 and 匹配, 以获取第二可替代云主机, 并获取所述第二可替代云主机的第二资源使用信息, 并将所述第二可替代云主机和所述第二资源使用信息标记在所述第二可替代资源上, 以获取最终的第二可替代资源。

7. 根据权利要求1所述的方法, 其特征在于, 所述获取训练数据, 将所述训练数据输入至神经网络模型, 对所述神经网络模型进行训练, 以获得资源监控模型, 包括:

获取训练数据, 对所述训练数据进行数据预处理;

将经过平滑处理和异常数据处理的训练数据存储于训练数据库中, 并设置配置文件, 其中, 所述配置文件包括网络结构、训练时长、训练与测试的比例安排、输出内容、优化学习率的设定、优化参数和存档规则设定;

根据所述配置文件, 对所述训练数据进行平滑处理, 以得到预测信息;

根据预设的综合评估规则对所述预测信息进行评估, 以获得评估信息, 其中, 所述评估信息包括评分范围以及分析信息;

根据所述评估信息生成与所述评估信息对应的优化配置策略, 获得资源监控模型;

通过已创建的检测脚本对所述资源监控模型进行准确性检测与性能测试;

若准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值, 则将所述资源监控模型作为最终的资源监控模型;

若准确性检测的结果未达到第一预设阈值和/或性能测试的结果未达到第二预设阈值, 则通过不断更新所述训练数据和修改所述预设的综合评估规则, 并对所述资源监控模型进行再训练, 直至准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值。

8. 一种用于评估云主机资源的装置,其特征在于,所述装置包括:

输入输出模块,用于获取训练数据,用于获取输入的目标项目信息;

处理模块,用于将所述输入输出模块获取的训练数据输入至神经网络模型,并对所述神经网络模型进行训练,以获得资源监控模型,其中,所述训练数据包括多个云主机的多个时段的资源采样数据;并对所述输入输出模块获取的输入的目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息;获取云主机的使用资源类型和使用资源情况,并根据所述使用资源类型、所述使用资源情况、所述项目数据类型和所述项目操作需求信息,确定待评估云主机以及所述待评估云主机的待评估资源;通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,其中,所述第一时段的起始时刻晚于所述第二时段的结束时刻;根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测,以分别获取第一预测数据、第二预测数据和第三预测数据,并将所述第三预测数据作为主信息,以及将所述第一预测数据和所述第二预测数据作为辅助信息,以获取预测信息,其中,所述预测包括对所述待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,所述第三时段的起始时刻晚于所述第一时段的结束时刻;分别对所述第一采样数据、所述第二采样数据和所述预测信息进行特征提取,以分别获取第一关键信息、第二关键信息和第三关键信息,并获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源;根据所述第一关键信息、所述第二关键信息、所述第三关键信息、所述第一可替代资源和所述第二可替代资源对所述待评估云主机进行评估,以获取评估信息;根据所述评估信息生成与所述评估信息对应的优化配置策略,将所述评估信息和所述优化配置策略至显示模块,并通过所述显示模块输出所述评估信息和所述优化配置策略;

所述显示模块,用于显示所述评估信息和所述优化配置策略。

9. 一种计算机设备,其特征在于,所述计算机设备包括:

至少一个处理器、存储器、显示器和输入输出单元;

其中,所述存储器用于存储程序代码,所述处理器用于调用所述存储器中存储的程序代码来执行如权利要求1-7中任一项所述的方法。

10. 一种计算机存储介质,其特征在于,其包括指令,当其在计算机上运行时,使得计算机执行如权利要求1-7中任一项所述的方法。

评估云主机资源的方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及云监控领域,尤其涉及评估云主机资源的方法、装置、设备及存储介质。

背景技术

[0002] 云主机作为新一代的主机租用服务,它整合了高性能服务器与优质网络带宽,有效解决了传统主机租用价格偏高和服务品质参差不齐的缺点,能全面满足中小企业、个人站长用户对主机租用服务低成本、高可靠、易管理的需求,因而,云主机被广泛使用。但是,随着时间的发展、人员的变动或者项目的搁置等情况变化,导致云主机资源的不合理利用和浪费,诸如离职人员所归属的云主机由于可能存在项目的关联而在离职过程中并未被释放,从而导致许多服务器资源使用量极低却配备高配置云主机,进而造成云主机资源的浪费。

[0003] 目前的云监控中,通过获取云主机中当前的资源使用信息,对所述资源使用信息进行分析与评估,以获得评估信息,根据评估信息确定云主机的使用状态,并根据所述使用状态确定优化配置方案。

[0004] 由于云主机的资源使用量会随着用户人数和使用时间的增加而增长,用户人数和使用时间会影响到云主机的资源使用量,在该情况下,若继续采用对云主机当前的资源使用量进行分析和评估的方案,则会导致所分析和评估的资源使用信息单一,致使所确定的优化配置策略不严谨且不全面,从而,导致云主机的资源利用率低。

发明内容

[0005] 本申请提供了一种评估云主机资源的方法、装置、设备及存储介质,能够解决云主机的资源利用率低的问题。

[0006] 第一方面,本申请提供一种评估云主机资源的方法,所述方法包括:

[0007] 获取训练数据,将所述训练数据输入至神经网络模型,并对所述神经网络模型进行训练,以获得资源监控模型,其中,所述训练数据包括多个云主机的多个时段的资源采样数据;

[0008] 获取输入的目标项目信息,并对所述目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息;

[0009] 获取云主机的使用资源类型和使用资源情况,并根据所述使用资源类型、所述使用资源情况、所述项目数据类型和所述项目操作需求信息,确定待评估云主机以及所述待评估云主机的待评估资源;

[0010] 通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,其中,所述第一时段的起始时刻晚于所述第二时段的结束时刻;

[0011] 根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测,以分

别获取第一预测数据、第二预测数据和第三预测数据,并将所述第三预测数据作为主信息,以及将所述第一预测数据和所述第二预测数据作为辅助信息,以获取预测信息,其中,所述预测包括对所述待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,所述第三时段的起始时刻晚于所述第一时段的结束时刻;

[0012] 分别对所述第一采样数据、所述第二采样数据和所述预测信息进行特征提取,以分别获取第一关键信息、第二关键信息和第三关键信息,并获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源;

[0013] 根据所述第一关键信息、所述第二关键信息、所述第三关键信息、所述第一可替代资源和所述第二可替代资源对所述待评估云主机进行评估,以获取评估信息;

[0014] 根据所述评估信息生成与所述评估信息对应的优化配置策略,并输出所述评估信息和所述优化配置策略。

[0015] 一种可能的设计中,所述获取输入的目标项目信息,并对所述目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息,包括:

[0016] 创建项目操作需求表,其中,所述项目操作需求表包括项目的预设完成时间、主机资源需求量和所述主机资源需求量对应的最优主机资源分配量;

[0017] 获取用户输入的目标项目信息,并对所述目标项目信息进行数据预处理,其中,所述数据预处理包括缺失值填补处理、去噪处理和数据标准化处理;

[0018] 将经过数据预处理的目标项目信息分成N组,通过反复迭代法对分成N组的目标项目信息进行多次重新分组,以获取最优分组方案;

[0019] 获取所述最优分组方案中各组目标项目信息的项目数据类型;

[0020] 对所述目标项目信息进行分析以获取第四关键信息;

[0021] 根据所述第四关键信息遍历所述项目操作需求表,以获取所述第四关键信息对应的项目操作需求信息。

[0022] 一种可能的设计中,所述待评估资源包括在第一时段内的第一任务和在第一时段内的第二任务,所述通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,包括:

[0023] 通过所述资源监控模型,获取所述第一任务中的第一占用率信息和第一优先级,以及获取所述第二任务中的第二占用率信息和第二优先级;

[0024] 根据所述第一优先级,对所述第一任务进行分类,并标识第一类别标签,以及根据所述第二优先级,对所述第二任务进行分类,并标识第二类别标签;

[0025] 根据所述第一占用率信息,对标识所述第一类别标签的第一任务进行分类,并标识第三类别标签,以及所述第二占用率信息,对标识所述第二类别标签的第二任务进行分类,并标识第四类别标签;

[0026] 判断所述第一优先级是否符合第一预设采样条件和/或所述第一占用率信息是否符合第二预设采样条件,以及判断所述第二优先级是否符合第三预设采样条件和/或所述第二占用率信息是否符合第四预设采样条件;

[0027] 若所述第一优先级符合第一预设采样条件和/或所述第一占用率信息符合第二预设采样条件,则按照预设第一采样频率,对符合第一预设采样条件和/或第二预设采样条件的同一类别标签的任务进行采样,以获得第一采样数据;

[0028] 若所述第二优先级符合第三预设采样条件和/或所述第二占用率信息符合第四预设采样条件,则按照预设第二采样频率,对符合第三预设采样条件和/或第四预设采样条件的同一类别标签的任务进行采样,以获得第二采样数据。

[0029] 一种可能的设计中,所述获取预测信息之后,所述方法还包括:

[0030] 获取所述预测信息的第一时间序列数据;

[0031] 对所述第一时间序列数据进行滑动窗口处理,以生成预设数量的具备预设长度的时序子序列;

[0032] 分析所述时序子序列的统计指标以获取统计特征信息,并以所述统计特征信息作为更新后的预测信息,其中,所述统计特征信息包括最大值、最小值、中位数、第一四分位数、第三四分位数、方差和标准差。

[0033] 一种可能的设计中,所述根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测之前,所述方法还包括:

[0034] 获取所述第一采样数据的第二时间序列数据,以及获取所述第二采样数据的第三时间序列数据;

[0035] 通过所述数据处理子模型中的指数加权移动平均法EWMA,分别对所述第二时间序列数据和所述第三时间序列数据进行求值,以获得第一平滑处理数据和第一平滑处理数据,对所述第二时间序列数据和所述第三时间序列数据的求值计算如下:

$$[0036] \quad V_t = \frac{\alpha V_{t-1} + (1-\alpha)x_t}{1-\alpha^t},$$

[0037] 其中, x_t 为时刻t的实际第二时间序列数据或实际第三时间序列数据,系数 α 为加权下降的速率, V_t 为t时刻的EWMA值;

[0038] 通过所述数据处理子模型中的马氏距离Mahalanobis算法,分别对所述第一平滑处理数据和所述第二平滑处理数据中的极度异常点进行检测与标识,并删除所标识的极度异常点,检测与标识所述极度异常点的计算如下:

$$[0039] \quad \text{MDist}(\mathbf{b}, \bar{\mathbf{b}}) = \sqrt{(\mathbf{b} - \bar{\mathbf{b}})^T \mathbf{S}^{-1} (\mathbf{b} - \bar{\mathbf{b}})},$$

其中, $\text{MDist}(\mathbf{b}, \bar{\mathbf{b}})$ 是 \mathbf{b} 与 $\bar{\mathbf{b}}$ 的距离, $\bar{\mathbf{b}}$ 是所述平滑处理数据的均值向量, \mathbf{b} 是为所述平滑处理数据中的其他对象, \mathbf{S} 是协方差矩阵。

[0040] 一种可能的设计中,所述方法包括资源数据库,所述资源数据库包括收集的多种数据类型对应的资源数据,所述获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源,包括:

[0041] 获取所述第一采样数据中待评估资源的第一特征信息,以及获取所述第二采样数据中待评估资源的第二特征信息,其中,所述第一特征信息和所述第二特征信息均包括数据类型、资源总容量、各数据类型对应的资源的使用占比、资源的性能和特性;

[0042] 根据所述第一特征信息在所述资源数据库中获取与所述第一特征信息对应的第一资源数据,以及根据所述第二特征信息在所述资源数据库中获取与所述第二特征信息对应的第二资源数据;

[0043] 通过预设替换条件对所述第一资源数据进行计算和筛选以获取第一可替代资源,以及通过所述预设替换条件对所述第二资源数据进行计算和筛选以获取第二可替代资源;

[0044] 根据所述第一可替换资源,在所述待评估云主机之外的云主机中进行分析 and 匹配,以获取第一可替代云主机,并获取所述第一可替代云主机的第一资源使用信息,并将所述第一可替代云主机和所述第一资源使用信息标记在所述第一可替换资源上,以获取最终的第一可替换资源,以及根据所述第二可替换资源,在所述待评估云主机之外的云主机中进行分析 and 匹配,以获取第二可替代云主机,并获取所述第二可替代云主机的第二资源使用信息,并将所述第二可替代云主机和所述第二资源使用信息标记在所述第二可替换资源上,以获取最终的第二可替换资源。

[0045] 一种可能的设计中,所述获取训练数据,将所述训练数据输入至神经网络模型,对所述神经网络模型进行训练,以获得资源监控模型,包括:

[0046] 获取训练数据,对所述训练数据进行数据预处理;

[0047] 将经过平滑处理和异常数据处理的训练数据存储于训练数据库中,并设置配置文件,其中,所述配置文件包括网络结构、训练时长、训练与测试的比例安排、输出内容、优化学习率的设定、优化参数和存档规则设定;

[0048] 根据所述配置文件,对所述训练数据进行平滑处理,以得到预测信息;

[0049] 根据预设的综合评估规则对所述预测信息进行评估,以获得评估信息,其中,所述评估信息包括评分范围以及分析信息;

[0050] 根据所述评估信息生成与所述评估信息对应的优化配置策略,获得资源监控模型;

[0051] 通过已创建的检测脚本对所述资源监控模型进行准确性检测与性能测试;

[0052] 若准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值,则将所述资源监控模型作为最终的资源监控模型;

[0053] 若准确性检测的结果未达到第一预设阈值和/或性能测试的结果未达到第二预设阈值,则通过不断更新所述训练数据和修改所述预设的综合评估规则,并对所述资源监控模型进行再训练,直至准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值。

[0054] 第二方面,本申请提供一种用于评估云主机资源的装置,具有实现对应于上述第一方面提供的评估云主机资源的方法的功能。所述功能可以通过硬件实现,也可以通过硬件执行相应的软件实现。硬件或软件包括一个或多个与上述功能相对应的模块,所述模块可以是软件和/或硬件。

[0055] 一种可能的设计中,所述装置包括:

[0056] 输入输出模块,用于获取训练数据,用于获取输入的目标项目信息;

[0057] 处理模块,用于将所述输入输出模块获取的训练数据输入至神经网络模型,并对所述神经网络模型进行训练,以获得资源监控模型,其中,所述训练数据包括多个云主机的多个时段的资源采样数据;并对所述输入输出模块获取的输入的目标项目信息进行分析,以获取所述目标项目信息的项目数据类型和项目操作需求信息;获取云主机的使用资源类型和使用资源情况,并根据所述使用资源类型、所述使用资源情况、所述项目数据类型和所述项目操作需求信息,确定待评估云主机以及所述待评估云主机的待评估资源;通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据,其中,所述第一时段的起始时刻晚于所述第二时段的结

束时刻;根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测,以分别获取第一预测数据、第二预测数据和第三预测数据,并将所述第三预测数据作为主信息,以及将所述第一预测数据和所述第二预测数据作为辅助信息,以获取预测信息,其中,所述预测包括对所述待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,所述第三时段的起始时刻晚于所述第一时段的结束时刻;分别对所述第一采样数据、所述第二采样数据和所述预测信息进行特征提取,以分别获取第一关键信息、第二关键信息和第三关键信息,并获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源;根据所述第一关键信息、所述第二关键信息、所述第三关键信息、所述第一可替代资源和所述第二可替代资源对所述待评估云主机进行评估,以获取评估信息;根据所述评估信息生成与所述评估信息对应的优化配置策略,将所述评估信息和所述优化配置策略至显示模块,并通过所述显示模块输出所述评估信息和所述优化配置策略;

[0058] 所述显示模块,用于显示所述评估信息和所述优化配置策略。

[0059] 一种可能的设计中,所述处理模块还用于:

[0060] 创建项目操作需求表,其中,所述项目操作需求表包括项目的预设完成时间、主机资源需求量和所述主机资源需求量对应的最优主机资源分配量;

[0061] 获取用户输入的目标项目信息,并对所述目标项目信息进行数据预处理,其中,所述数据预处理包括缺失值填补处理、去噪处理和数据标准化处理;

[0062] 将经过数据预处理的目标项目信息分成N组,通过反复迭代法对分成N组的目标项目信息进行多次重新分组,以获取最优分组方案;

[0063] 获取所述最优分组方案中各组目标项目信息的项目数据类型;

[0064] 对所述目标项目信息进行分析以获取第四关键信息;

[0065] 根据所述第四关键信息遍历所述项目操作需求表,以获取所述第四关键信息对应的项目操作需求信息。

[0066] 一种可能的设计中,所述处理模块还用于:

[0067] 通过所述资源监控模型,获取所述第一任务中的第一占用率信息和第一优先级,以及获取所述第二任务中的第二占用率信息和第二优先级;

[0068] 根据所述第一优先级,对所述第一任务进行分类,并标识第一类别标签,以及根据所述第二优先级,对所述第二任务进行分类,并标识第二类别标签;

[0069] 根据所述第一占用率信息,对标识所述第一类别标签的第一任务进行分类,并标识第三类别标签,以及所述第二占用率信息,对标识所述第二类别标签的第二任务进行分类,并标识第四类别标签;

[0070] 判断所述第一优先级是否符合第一预设采样条件和/或所述第一占用率信息是否符合第二预设采样条件,以及判断所述第二优先级是否符合第三预设采样条件和/或所述第二占用率信息是否符合第四预设采样条件;

[0071] 若所述第一优先级符合第一预设采样条件和/或所述第一占用率信息符合第二预设采样条件,则按照预设第一采样频率,对符合第一预设采样条件和/或第二预设采样条件的同一类别标签的任务进行采样,以获得第一采样数据;

[0072] 若所述第二优先级符合第三预设采样条件和/或所述第二占用率信息符合第四预

设采样条件,则按照预设第二采样频率,对符合第三预设采样条件和/或第四预设采样条件的同一类别标签的任务进行采样,以获得第二采样数据。

[0073] 一种可能的设计中,所述处理模块在执行所述获得预测信息之后,还用于:

[0074] 获取所述预测信息的第一时间序列数据;

[0075] 对所述第一时间序列数据进行滑动窗口处理,以生成预设数量的具备预设长度的时序子序列;

[0076] 分析所述时序子序列的统计指标以获取统计特征信息,并以所述统计特征信息作为更新后的预测信息,其中,所述统计特征信息包括最大值、最小值、中位数、第一四分位数、第三四分位数、方差和标准差。

[0077] 一种可能的设计中,所述处理模块在执行所述根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测之前,还用于:

[0078] 获取所述第一采样数据的第二时间序列数据,以及获取所述第二采样数据的第三时间序列数据;

[0079] 通过所述数据处理子模型中的指数加权移动平均法EWMA,分别对所述第二时间序列数据和所述第三时间序列数据进行求值,以获得第一平滑处理数据和第一平滑处理数据,对所述第二时间序列数据和所述第三时间序列数据的求值计算如下:

$$[0080] \quad V_t = \frac{\alpha V_{t-1} + (1-\alpha)x_t}{1-\alpha^t},$$

[0081] 其中, x_t 为时刻t的实际第二时间序列数据或实际第三时间序列数据,系数 α 为加权下降的速率, V_t 为t时刻的EWMA值;

[0082] 通过所述数据处理子模型中的马氏距离Mahalanobis算法,分别对所述第一平滑处理数据和所述第二平滑处理数据中的极度异常点进行检测与标识,并删除所标识的极度异常点,检测与标识所述极度异常点的计算如下:

$$[0083] \quad \text{MDist}(\mathbf{b}, \bar{\mathbf{b}}) = \sqrt{(\mathbf{b} - \bar{\mathbf{b}})^T \mathbf{S}^{-1} (\mathbf{b} - \bar{\mathbf{b}})},$$

其中, $\text{MDist}(\mathbf{b}, \bar{\mathbf{b}})$ 是 \mathbf{b} 与 $\bar{\mathbf{b}}$ 的距离, $\bar{\mathbf{b}}$ 是所述平滑处理数据的均值向量, \mathbf{b} 是为所述平滑处理数据中的其他对象, \mathbf{S} 是协方差矩阵。

[0084] 一种可能的设计中,所述处理模块还用于:

[0085] 获取所述第一采样数据中待评估资源的第一特征信息,以及获取所述第二采样数据中待评估资源的第二特征信息,其中,所述第一特征信息和所述第二特征信息均包括数据类型、资源总容量、各数据类型对应的资源的使用占比、资源的性能和特性;

[0086] 根据所述第一特征信息在所述资源数据库中获取与所述第一特征信息对应的第一资源数据,以及根据所述第二特征信息在所述资源数据库中获取与所述第二特征信息对应的第二资源数据;

[0087] 通过预设替换条件对所述第一资源数据进行计算和筛选以获取第一可替代资源,以及通过所述预设替换条件对所述第二资源数据进行计算和筛选以获取第二可替代资源;

[0088] 根据所述第一可替代资源,在所述待评估云主机之外的云主机中进行分析 and 匹配,以获取第一可替代云主机,并获取所述第一可替代云主机的第一资源使用信息,并将所述第一可替代云主机和所述第一资源使用信息标记在所述第一可替代资源上,以获取最终

的第一可替换资源,以及根据所述第二可替换资源,在所述待评估云主机之外的云主机中进行分析 and 匹配,以获取第二可替代云主机,并获取所述第二可替代云主机的第二资源使用信息,并将所述第二可替代云主机和所述第二资源使用信息标记在所述第二可替换资源上,以获取最终的所述第二可替换资源。

[0089] 一种可能的设计中,所述处理模块还用于:

[0090] 获取训练数据,对所述训练数据进行数据预处理;

[0091] 将经过平滑处理和异常数据处理的训练数据存储于训练数据库中,并设置配置文件,其中,所述配置文件包括网络结构、训练时长、训练与测试的比例安排、输出内容、优化学习率的设定、优化参数和存档规则设定;

[0092] 根据所述配置文件,对所述训练数据进行平滑处理,以得到预测信息;

[0093] 根据预设的综合评估规则对所述预测信息进行评估,以获得评估信息,其中,所述评估信息包括评分范围以及分析信息;

[0094] 根据所述评估信息生成与所述评估信息对应的优化配置策略,获得资源监控模型;

[0095] 通过已创建的检测脚本对所述资源监控模型进行准确性检测与性能测试;

[0096] 若准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值,则将所述资源监控模型作为最终的资源监控模型;

[0097] 若准确性检测的结果未达到第一预设阈值和/或性能测试的结果未达到第二预设阈值,则通过不断更新所述训练数据和修改所述预设的综合评估规则,并对所述资源监控模型进行再训练,直至准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值。

[0098] 本申请又一方面提供了一种计算机设备,其包括至少一个连接的处理器、存储器、显示器和输入输出单元,其中,所述存储器用于存储程序代码,所述处理器用于调用所述存储器中的程序代码来执行上述第一方面所述的方法。

[0099] 本申请又一方面提供了一种计算机存储介质,其包括指令,当其在计算机上运行时,使得计算机执行上述第一方面所述的方法。

[0100] 相较于现有技术,本申请提供的方案中,通过训练所得的资源监控模型对目标项目信息、多个云主机的使用资源类型和使用资源情况进行分析,以获取待评估云主机以及所述待评估云主机的待评估资源,获取所述待评估资源的第一采样数据和第二采样数据,根据所述第一采样数据和所述第二采样数据获取预测信息,根据所述第一采样数据、所述第二采样数据和所述预测信息获取评估信息,根据所述评估信息生成与所述评估信息对应的优化配置策略。由于是通过在根据目标项目信息确定待评估云主机和所述待评估云主机的待评估资源,以使在对云主机的资源使用评估的基础上,还能满足对目标项目信息的评估需求,从而提高云主机的资源利用率;通过对云主机的第一采样数据、第二采样数据和预测信息进行评估,所评估的数据信息多方位,以使评估结果更具严谨性和准确性,通过多角度获取预测数据以提高评估的准确性;综上所述,一方面,获取多角度的资源使用情况的数据,以便于用户快速而全面获悉待评估云主机的资源使用情况,另一方面,通过对多角度的数据进行评估以提高评估的严谨性和准确性,因而,本申请能够提高云主机的资源利用率。

附图说明

- [0101] 图1为本申请实施例中评估云主机资源的方法的一种流程示意图；
- [0102] 图2为本申请实施例中用于评估云主机资源的装置的一种结构示意图；
- [0103] 图3为本申请实施例中计算机装置的一种结构示意图。
- [0104] 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

[0105] 应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或模块的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或模块，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或模块，本申请中所出现的模块的划分，仅仅是一种逻辑上的划分，实际应用中实现时可以有另外的划分方式，例如多个模块可以结合成或集成在另一个系统中，或一些特征可以忽略，或不执行。

[0106] 本申请提供一种评估云主机资源的方法、装置、设备及存储介质，可用于云主机的配置使用，为无使用资源的云主机的释放以及由于资源使用量极少而需要执行降低配置操作的云主机提供参考。其中，对于云计算平台厂商，可用于提供资源回收的量化标准，为减少内部计算资源的浪费提供参考支持；对于云计算平台租户，可用于为对业务云主机的资源使用情况和是否使用合理提供参考。

[0107] 为解决上述技术问题，本申请主要提供以下技术方案：

[0108] 相较于现有技术，本申请提供的方案中，通过训练所得的资源监控模型对目标项目信息、多个云主机的使用资源类型和使用资源情况进行分析，以获取待评估云主机以及待评估云主机的待评估资源，获取待评估资源的第一采样数据和第二采样数据，根据第一采样数据和第二采样数据获取预测信息，根据第一采样数据、第二采样数据和预测信息获取评估信息，根据评估信息生成与评估信息对应的优化配置策略。由于是通过在根据目标项目信息确定待评估云主机和待评估云主机的待评估资源，以使在对云主机的资源使用评估的基础上，还能满足对目标项目信息的评估需求，从而提高云主机的资源利用率；通过对云主机的第一采样数据、第二采样数据和预测信息进行评估，所评估的数据信息多方位，以使评估结果更具严谨性和准确性，通过多角度获取预测数据以提高评估的准确性；综上，一方面，获取多角度的资源使用情况的数据，以便于用户快速而全面获悉待评估云主机的资源使用情况，另一方面，通过对多角度的数据进行评估以提高评估的严谨性和准确性，因而，本申请能够提高云主机的资源利用率。

[0109] 请参照图1，以下对本申请提供一种评估云主机资源的方法进行举例说明，该方法由计算机设备执行，计算机设备可为服务器或者终端，当图2所示的装置20为应用或者执行程序时，终端为安装图2所示的装置20的终端，本申请不对执行主体的类型作限制，所述方法包括：

[0110] 101、获取训练数据,将训练数据输入至神经网络模型,对神经网络模型进行训练,以获得资源监控模型。

[0111] 其中,训练数据包括多个云主机的多个时段的资源采样数据;训练数据为按照预设采样间隔采样的按照顺向时间顺序排列的多个采样点数据,包括第一时段内的第一数据和第二时段内的第二数据,第一时段的起始时刻晚于第二时段的结束时刻。

[0112] 本实施例中,可通过迁移学习算法对现有的监控模型进行训练,以获取用于对云主机的使用资源进行监控、评估和优化配置的资源监控模型。本实施例的资源监控模型为后向传播神经网络模型。通过结合后向传播神经网络模型的输入输出的非线性映射、进行梯度下降计算、具备一定泛化能力和使用不同传递函数的特点,一方面,使其对云主机的使用资源进行监控、评估和优化配置等的数据挖掘进行多维的特征构造,和新的数据进入神经网络模型的网络进行训练的时候,神经网络能够在调整权值以适应更多的数据;另一方面,使其输出的值可是任意值,和减小资源监控模型的误差,以获取更优的评估信息和优化配置策略。

[0113] 可选的,在本申请的一些实施例中,在获取资源监控模型的情况下,为了进一步加强神经网络模型进行训练以获得资源监控模型,具体操作如下,上述的获取训练数据,将训练数据输入至神经网络模型,对神经网络模型进行训练,以获得资源监控模型,包括:

[0114] 获取训练数据,对训练数据进行数据预处理;

[0115] 将经过平滑处理和异常数据处理的训练数据存储于训练数据库中,并设置配置文件,其中,配置文件包括网络结构、训练时长、训练与测试的比例安排、输出内容、优化学习率的设定、优化参数和存档规则设定;

[0116] 根据配置文件,对训练数据进行平滑处理,以得到预测信息;

[0117] 根据预设的综合评估规则对预测信息进行评估,以获得评估信息,其中,评估信息包括评分范围以及分析信息;

[0118] 根据评估信息生成与评估信息对应的优化配置策略,获得资源监控模型;

[0119] 通过已创建的检测脚本对资源监控模型进行准确性检测与性能测试;

[0120] 若准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值,则将资源监控模型作为最终的资源监控模型;

[0121] 若准确性检测的结果未达到第一预设阈值和/或性能测试的结果未达到第二预设阈值,则通过不断更新训练数据和修改预设的综合评估规则,并对资源监控模型进行再训练,直至准确性检测的结果达到第一预设阈值以及性能测试的结果达到第二预设阈值。

[0122] 对神经网络模型训练以使获取的资源监控模型具备对目标项目信息、多个云主机的使用资源类型和使用资源情况进行分析,以获取待评估云主机以及所述待评估云主机的待评估资源,获取所述待评估资源的第一采样数据和第二采样数据,根据所述第一采样数据和所述第二采样数据获取预测信息,根据所述第一采样数据、所述第二采样数据和所述预测信息获取评估信息,根据所述评估信息生成与所述评估信息对应的优化配置策略的功能,以更好地对云主机的资源利用进行评估,从而提高云主机的资源利用率。

[0123] 可选的,在本申请的一些实施例中,资源监控模型包括数据处理子模型、预测子模型、评估子模型和策略生成子模型,预测子模型与数据处理子模型串联,评估子模型与预测子模型串联,策略生成子模型与评估子模型串联。数据处理子模型用于对资源监控模型获

取的采样数据进行平滑处理和异常数据处理；预测子模型用于对数据处理子模型输出的采样数据进行预测，得到预测信息；评估子模型用于对预测子模型输出的预测信息进行评估，得到评估信息；策略生成子模型用于根据评估子模型输出的评估信息，生成对应的优化配置策略。

[0124] 102、获取输入的目标项目信息，并对目标项目信息进行分析，以获取目标项目信息的项目数据类型和项目操作需求信息。

[0125] 获取用户输入的目标项目信息后，对目标项目信息进行分析以获取完成目标项目所需要操作的数据类型和目标项目的操作需求信息。其中，项目操作需求信息包括但不限于项目完成的计划时间、期限时间和项目完成的功能。例如：目标项目为管理系统，通过分析用户输入的目标项目信息，以获取对管理系统研发完成的计划时间、期限时间、所需的云主机资源量、管理系统具备的性能和管理系统需具备的功能等。

[0126] 通过获取目标项目信息中的项目数据类型和项目操作需求信息，以便于对云主机的使用资源和资源配置的评估提供更好的依据和分析基础。

[0127] 可选的，在本申请的一些实施例中，在获取目标项目信息的项目数据类型和项目操作需求信息情况下，为了进一步加强信息获取的准确性，具体操作如下，上述的获取输入的目标项目信息，并对目标项目信息进行分析，以获取目标项目信息的项目数据类型和项目操作需求信息，包括：

[0128] 创建项目操作需求表，其中，项目操作需求表包括项目的预设完成时间、主机资源需求量和主机资源需求量对应的最优主机资源分配量；

[0129] 获取用户输入的目标项目信息，并对目标项目信息进行数据预处理，其中，数据预处理包括缺失值填补处理、去噪处理和数据标准化处理；

[0130] 将经过数据预处理的目标项目信息分成N组，通过反复迭代法对分成N组的目标项目信息进行多次重新分组，以获取最优分组方案；

[0131] 获取最优分组方案中各组目标项目信息的项目数据类型；

[0132] 对目标项目信息进行分析以获取第四关键信息；

[0133] 根据第四关键信息遍历项目操作需求表，以获取第四关键信息对应的项目操作需求信息。

[0134] 通过对目标项目信息进行聚类分析，以获得简明结论形式，从而以便于资源监控模型直观而简便地获取目标项目的项目数据类型和项目操作需求信息。

[0135] 通过遍历项目操作需求表以获取关键信息对应的项目操作需求信息，以使获取项目操作需求信息多角度，从而能够更快速而准确地确定待评估云主机以及待评估云主机的待评估资源。

[0136] 103、分析云主机的使用资源类型和使用资源情况，并根据项目数据类型和项目操作需求信息，以确定待评估云主机以及待评估云主机的待评估资源。

[0137] 其中，云主机包括连接使用的所有的云主机；待评估资源包括资源使用情况和剩余资源情况，剩余资源情况为待评估云主机中总资源减去已使用资源后的资源量。

[0138] 通过在根据目标项目信息确定待评估云主机和所述待评估云主机的待评估资源，一方面有目的性、针对性地对云主机资源进行评估，另一方面减少不必要的运行操作以减轻系统的压力和提高运行速度。

[0139] 例如:分析所有的云主机的使用资源类型和使用资源情况,匹配出与项目数据类型对应的使用资源类型,以获取第一云主机,以及匹配出与项目操作需求信息对应的使用资源情况以获取第二云主机,判断第一云主机与第二云主机中既符合项目数据类型又能满足项目操作需求信息的云主机,作为待评估云主机,并获取待评估云主机中的资源使用情况和剩余资源情况。

[0140] 上述举例内容仅作举例说明参考,其内容准确性和实际操作与否不作考虑。

[0141] 104、通过资源监控模型获取待评估资源在第一时段内的第一采样数据,以及获取待评估资源在第二时段内的第二采样数据。

[0142] 其中,第一时段的起始时刻晚于所述第二时段的结束时刻。第一采样数据和第二采样数据均是按照预设采样间隔采样的按照顺向时间顺序排列的多个采样点数据,第一采样数据和第二采样数据均包括中央处理器CPU的占用率信息、内存占用率信息以及输入/输出IO占用率信息,占用率信息包括物理资源与总资源的占比、功能性描述的信息资源与总资源的占比、非功能性描述的信息资源与总资源的占比、网络资源与总资源的占比、已使用的资源与总资源的占比和未使用的资源与总资源的占比。

[0143] 通过采集两个不同时间段的采样数据,增加评估资源的多样性,以提高对云主机资源评估的可参考性和准确性。

[0144] 可选的,在本申请的一些实施例中,在获取目标业务类型和文本预测结果,以及调用目标业务类型词库的情况下,为了进一步加强准确调用目标业务类型相应的目标业务类型词库,具体操作如下,上述的待评估资源包括在第一时段内的第一任务和在第一时段内的第二任务,上述的通过资源监控模型获取待评估资源在第一时段内的第一采样数据,以及获取待评估资源在第二时段内的第二采样数据,包括:

[0145] 通过资源监控模型,获取第一任务中的第一占用率信息和第一优先级,以及获取第二任务中的第二占用率信息和第二优先级;

[0146] 根据第一优先级,对第一任务进行分类,并标识第一类别标签,以及根据第二优先级,对第二任务进行分类,并标识第二类别标签;

[0147] 根据第一占用率信息,对标识第一类别标签的第一任务进行分类,并标识第三类别标签,以及第二占用率信息,对标识第二类别标签的第二任务进行分类,并标识第四类别标签;

[0148] 判断第一优先级是否符合第一预设采样条件和/或第一占用率信息是否符合第二预设采样条件,以及判断第二优先级是否符合第三预设采样条件和/或第二占用率信息是否符合第四预设采样条件;

[0149] 若第一优先级符合第一预设采样条件和/或第一占用率信息符合第二预设采样条件,则按照预设第一采样频率,对符合第一预设采样条件和/或第二预设采样条件的同一类别标签的任务进行采样,以获得第一采样数据;

[0150] 若第二优先级符合第三预设采样条件和/或第二占用率信息符合第四预设采样条件,则按照预设第二采样频率,对符合第三预设采样条件和/或第四预设采样条件的同一类别标签的任务进行采样,以获得第二采样数据。

[0151] 通过上述实施方式,能够实现便于获悉云主机中资源的使用状况和对资源占比的影响程度,从而对后续得到的评估信息的准确性和严谨性起到一定的支持作用的有益效

果。

[0152] 由于用户提交的每一个任务(即为应用请求)在提交时都带有特定的优先级,而任何一个相对高优先级的任务可以抢占低优先级使用的资源,且处于不同档次的高优先级的任务具有对应的特征,例如:正常生产档的高优先级的任务对延迟敏感,一般不会因为过度使用资源而被剔除、挂起或暂停,因而为了能够便于清晰了解各占用率信息的使用状况和对资源占比的影响程度,任务为用户发起的应用请求,可采用上述操作。

[0153] 105、根据第一采样数据和第二采样数据对待评估云主机进行预测,以分别获取第一预测数据、第二预测数据和第三预测数据,并将第三预测数据作为主信息,以及将第一预测数据和第二预测数据作为辅助信息,以获得预测信息。

[0154] 其中,预测包括对待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,第三时段的起始时刻晚于第一时段的结束时刻。第一预测数据包括根据第一采样数据对待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势进行预测而得的预测数据,第二预测数据包括根据第二采样数据对待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势进行预测而得的预测数据,第三预测数据包括根据第一采样数据和第二采样数据对待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势进行预测而得的预测数据,第三时段的起始时刻晚于第一时段的结束时刻。主信息和辅助信息指的是首先对主信息进行处理后,再结合辅助信息进行处理。

[0155] 通过采用多中预测数据和将第三预测数据作为主信息,以及将第一预测数据和第二预测数据作为辅助信息,以提高评估的准确性。

[0156] 可选的,在本申请的一些实施例中,在获得预测信息的情况下,为了进一步加强对预测信息的获取和预测信息的获取内容,可采取如下操作,上述的获得预测信息之后,方法还包括:

[0157] 获取预测信息的第一时间序列数据;

[0158] 对第一时间序列数据进行滑动窗口处理,以生成预设数量的具备预设长度的时序子序列;

[0159] 分析时序子序列的统计指标以获取统计特征信息,并以统计特征信息作为更新后的预测信息,其中,统计特征信息包括最大值、最小值、中位数、第一四分位数、第三四分位数、方差和标准差。

[0160] 通过对预测信息进行统计特征提取,使得预测信息的数据更集中、更系统、更清楚地反映客观实际,以使后续的未来预测值更偏向现实值,促进预测的准确性。

[0161] 可选的,在本申请的一些实施例中,在获取预测信息的情况下,为了进一步加强对预测信息的获取,可采取如下操作,上述的根据第一采样数据和第二采样数据对待评估云主机进行预测之前,方法还包括:

[0162] 获取第一采样数据的第二时间序列数据,以及获取第二采样数据的第三时间序列数据;

[0163] 通过数据处理子模型中的指数加权移动平均法EWMA,分别对第二时间序列数据和第三时间序列数据进行求值,以获得第一平滑处理数据和第一平滑处理数据,对第二时间序列数据和第三时间序列数据的求值计算如下:

$$[0164] \quad V_t = \frac{\alpha V_{t-1} + (1-\alpha)x_t}{1-\alpha^t},$$

[0165] 其中, x_t 为时刻 t 的实际第二时间序列数据或实际第三时间序列数据, 系数 α 为加权下降的速率, V_t 为 t 时刻的 EWMA 值;

[0166] 通过数据处理子模型中的马氏距离 Mahalanobis 算法, 分别对第一平滑处理数据和第二平滑处理数据中的极度异常点进行检测与标识, 并删除所标识的极度异常点, 检测与标识极度异常点的计算如下:

$$[0167] \quad \text{MDist}(b, \bar{b}) = \sqrt{(b, \bar{b})^T S^{-1} (b - \bar{b})},$$

[0168] 其中, $\text{MDist}(b, \bar{b})$ 是 b 与 \bar{b} 的距离, \bar{b} 是平滑处理数据的均值向量, b 是为平滑处理数据中的其他对象, S 是协方差矩阵。

[0169] 通过对第一采样数据和第二采样数据进行平滑处理和异常数据处理, 以获得相对均衡、稳定的时间序列数据, 减少误差, 为后续的未来预测值的准确值提供进一步的支持。

[0170] 106、分别对第一采样数据、第二采样数据和预测信息进行特征提取, 以分别获取第一关键信息、第二关键信息和第三关键信息, 并获取第一采样数据中待评估资源的第一可替代资源, 和获取第二采样数据中待评估资源的第二可替代资源。

[0171] 其中, 分析包括对第一采样数据、第二采样数据和预测信息的各资源的类型、使用时间、操作速率和资源占比的识别和获取。第一关键信息包括对第一采样数据进行分析而获取的各资源的类型、使用时间、操作速率和资源占比, 第二关键信息包括对第二采样数据进行分析而获取的各资源的类型、使用时间、操作速率和资源占比, 第三关键信息包括对预测信息进行分析而获取的各资源的类型、使用时间、操作速率和资源占比。

[0172] 一方面, 通过获取多角度的资源使用情况的数据, 以便于用户快速而全面获悉待评估云主机的资源使用, 另一方面, 通过结合多样化的数据进行评估, 以提高评估的准确性。

[0173] 可选的, 在本申请的一些实施例中, 在获取第一可替代资源和第二可替代资源以提供评估的参考数据支撑的情况下, 为了进一步加强对评估的参考数据支撑, 可采取如下操作, 本方法包括资源数据库, 资源数据库包括收集的多种数据类型对应的资源数据, 上述的获取第一采样数据中待评估资源的第一可替代资源, 和获取第二采样数据中待评估资源的第二可替代资源, 包括:

[0174] 获取第一采样数据中待评估资源的第一特征信息, 以及获取第二采样数据中待评估资源的第二特征信息, 其中, 第一特征信息和第二特征信息均包括数据类型、资源总容量、各数据类型对应的资源的使用占比、资源的性能和特性;

[0175] 根据第一特征信息在资源数据库中获取与第一特征信息对应的第一资源数据, 以及根据第二特征信息在资源数据库中获取与第二特征信息对应的第二资源数据;

[0176] 通过预设替换条件对第一资源数据进行计算和筛选以获取第一可替代资源, 以及通过预设替换条件对第二资源数据进行计算和筛选以获取第二可替代资源;

[0177] 根据第一可替代资源, 在待评估云主机之外的云主机中进行分析 and 匹配, 以获取第一可替代云主机, 并获取第一可替代云主机的第一资源使用信息, 并将第一可替代云主机和第一资源使用信息标记在第一可替代资源上, 以获取最终的第一可替代资源, 以及根

据第二可替换资源,在待评估云主机之外的云主机中进行分析 and 匹配,以获取第二可替代云主机,并获取第二可替代云主机的第二资源使用信息,并将第二可替代云主机和第二资源使用信息标记在第二可替换资源上,以获取最终的第二可替换资源。

[0178] 通过标记可替代云主机和可替代云主机的资源使用信息,一方面,以便于用户获知可替代使用的云主机,以达到节省资源的效果,另一方面,以为评估提供多角度多方面的数据,从而提高评估的准确性和可行性。

[0179] 107、根据第一关键信息、第二关键信息、第三关键信息、第一可替代资源和第二可替代资源对待评估云主机进行评估,以获取评估信息。

[0180] 根据第一关键信息、第二关键信息和第三关键信息对待评估云主机的当前的使用资源和未来某一段时间的预估使用资源进行评估,并根据第一可替代资源和第二可替代资源评估可替换待评估云主机的其他云主机,以合理使用云主机和提高云主机的资源利用率。

[0181] 例如:甲云主机为待评估云主机,乙云主机为待评估云主机之外的云主机。根据第一关键信息、第二关键信息和第三关键信息对甲云主机进行评估,所得甲云主机的可使用资源已剩10%,且各项指标不满足继续使用要求和满足项目的要求,而乙云主机的可使用资源还有90%,和已空置较长一段时间,根据第一关键信息、第二关键信息、第三关键信息、第一可替代资源和第二可替代资源对甲云主机和乙云主机分析,所得乙云主机的资源类型与甲云主机相同和/或相似,可进行替换使用,且各项指标均满足继续使用要求和满足项目的要求,则对于输入的目标项目信息,乙云主机可作为甲云主机的可替换云主机。其中,分析过程所产生的过程信息和结果信息,作为评估信息。

[0182] 上述举例内容仅作举例说明参考,其内容准确性和实际操作与否不作考虑。

[0183] 108、根据评估信息,生成与评估信息对应的优化配置策略,输出评估信息和优化配置策略。

[0184] 其中,优化配置策略包括云主机的资源使用状态、云主机的各物理资源占用率信息、是否需要云主机进行配置提升或配置降低或资源释放,以及云主机的预设未来各时间段的资源使用情况、用户行为匹配和语义匹配。

[0185] 与现有机制相比,本申请实施例中,通过训练所得的资源监控模型对目标项目信息、多个云主机的使用资源类型和使用资源情况进行分析,以获取待评估云主机以及待评估云主机的待评估资源,获取待评估资源的第一采样数据和第二采样数据,根据第一采样数据和第二采样数据获取预测信息,根据第一采样数据、第二采样数据和预测信息获取评估评估信息,根据评估信息,生成与评估信息对应的优化配置策略。由于是通过在根据目标项目信息确定待评估云主机和待评估云主机的待评估资源,以使在对云主机的资源使用评估的基础上,还能满足对目标项目信息的评估需求,从而提高云主机的资源利用率;通过对云主机的第一采样数据、第二采样数据和预测信息进行评估,所评估的数据信息多方位,以使评估结果更具严谨性和准确性,通过多角度获取预测数据以提高评估的准确性;综上,一方面,获取多角度的资源使用情况的数据,以便于用户快速而全面获悉待评估云主机的资源使用情况,另一方面,通过对多角度的数据进行评估以提高评估的严谨性和准确性,因而,本申请能够提高云主机的资源利用率。

[0186] 上述图1对应的实施例或图1对应的实施例中的任一可选实施例或可选实施方式

中所提及的技术特征也同样适用于本申请中的图2和图3所对应的实施例,后续类似之处不再赘述。

[0187] 以上对本申请中一种评估云主机资源的方法进行说明,以下对执行上述评估云主机资源的方法的装置进行描述。

[0188] 如图2所示的一种用于评估云主机资源的装置20的结构示意图,其可应用于云主机的配置使用,为无使用资源的云主机的释放以及由于资源使用量极少而需要执行降低配置操作的云主机提供参考。其中,对于云计算平台厂商,可用于提供资源回收的量化标准,为减少内部计算资源的浪费提供参照支持;对于云计算平台租户,可用于为对业务云主机的资源使用情况和是否使用合理提供参考。本申请实施例中的装置20能够实现对应于上述图1对应的实施例或图1对应的实施例中的任一可选实施例或可选实施方式中所执行的评估云主机资源的方法的步骤。装置20实现的功能可以通过硬件实现,也可以通过硬件执行相应的软件实现。硬件或软件包括一个或多个与上述功能相对应的模块,所述模块可以是软件和/或硬件。所述装置20可包括输入输出模块201、处理模块202和显示模块203,所述输入输出模块201、处理模块202和显示模块203的功能实现可参考图1对应的实施例或图1对应的实施例中的任一可选实施例或可选实施方式中所执行的操作,此处不作赘述。所述处理模块202可用于控制所述输入输出模块201的收发操作,显示模块203可用于显示处理模块202的处理操作。

[0189] 一些实施方式中,输入输出模块201,用于获取训练数据,用于获取输入的目标项目信息;

[0190] 处理模块202,用于将输入输出模块201获取的训练数据输入至神经网络模型,对神经网络模型进行训练,以获得资源监控模型;对输入输出模块201获取的目标项目信息进行分析,以获取目标项目信息的项目数据类型和项目操作需求信息;获取云主机的使用资源类型和使用资源情况,并根据所述使用资源类型、所述使用资源情况、所述项目数据类型和所述项目操作需求信息,确定待评估云主机以及所述待评估云主机的待评估资源;通过所述资源监控模型获取所述待评估资源在第一时段内的第一采样数据,以及获取所述待评估资源在第二时段内的第二采样数据;根据所述第一采样数据和所述第二采样数据对所述待评估云主机进行预测,以分别获取第一预测数据、第二预测数据和第三预测数据,并将所述第三预测数据作为主信息,以及将所述第一预测数据和所述第二预测数据作为辅助信息,以获取预测信息;分别对所述第一采样数据、所述第二采样数据和所述预测信息进行特征提取,以分别获取第一关键信息、第二关键信息和第三关键信息,并获取所述第一采样数据中待评估资源的第一可替代资源,和获取所述第二采样数据中待评估资源的第二可替代资源;根据所述第一关键信息、所述第二关键信息、所述第三关键信息、所述第一可替代资源和所述第二可替代资源对所述待评估云主机进行评估,以获取评估信息;根据评估信息生成与评估信息对应的优化配置策略,将评估信息和优化配置策略至显示模块203,并通过显示模块203输出评估信息和优化配置策略;

[0191] 显示模块203,用于显示评估信息和优化配置策略。

[0192] 其中,训练数据包括多个云主机的多个时段的资源采样数据;第一时段的起始时刻晚于第二时段的结束时刻;预测包括对待评估云主机在第三时段的各资源使用状态和各资源使用量的趋势的预测,第三时段的起始时刻晚于所述第一时段的结束时刻。

[0193] 本申请实施例中,处理模块202通过训练所得的资源监控模型对目标项目信息、多个云主机的使用资源类型和使用资源情况进行分析,以获取待评估云主机以及待评估云主机的待评估资源,获取待评估资源的第一采样数据和第二采样数据,根据第一采样数据和第二采样数据获取预测信息,根据第一采样数据、第二采样数据和预测信息获取评估评估信息,根据评估信息生成与评估信息对应的优化配置策略。由于是通过在根据目标项目信息确定待评估云主机和待评估云主机的待评估资源,以使在对云主机的资源使用评估的基础上,还能满足对目标项目信息的评估需求,从而提高云主机的资源利用率;通过对云主机的第一采样数据、第二采样数据和预测信息进行评估,所评估的数据信息多方位,以使评估结果更具严谨性和准确性,通过多角度获取预测数据以提高评估的准确性;综上,一方面,获取多角度的资源使用情况的数据,以便于用户快速而全面获悉待评估云主机的资源使用情况,另一方面,通过对多角度的数据进行评估以提高评估的严谨性和准确性,因而,本申请能够提高云主机的资源利用率。

[0194] 可选的,在本申请的一些实施方式中,上述评估云主机资源的方法的任一实施例或实施方式中所提及的技术特征也同样适用于本申请中的对执行上述评估云主机资源的方法的装置20,后续类似之处不再赘述。

[0195] 上面从模块化功能实体的角度分别介绍了本申请实施例中的装置20,以下从硬件角度介绍一种计算机装置,如图3所示,其包括:处理器、存储器、显示器、输入输出单元(也可以是收发器,图3中未标识出)以及存储在所述存储器中并可在所述处理器上运行的计算机程序。例如,该计算机程序可以为图1对应的实施例或图1对应的实施例中的任一可选实施例或可选实施方式中评估云主机资源的方法对应的程序。例如,当计算机装置实现如图2所示的装置20的功能时,所述处理器执行所述计算机程序时实现上述图2所对应的实施例中由装置20执行的评估云主机资源的方法中的各步骤;或者,所述处理器执行所述计算机程序时实现上述图2所对应的实施例的装置20中各模块的功能。又例如,该计算机程序可以为图1对应的实施例或图1对应的实施例中的任一可选实施例或可选实施方式的方法对应的程序。

[0196] 所称处理器可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述计算机装置的控制中心,利用各种接口和线路连接整个计算机装置的各个部分。

[0197] 所述存储器可用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述计算机装置的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、视频数据等)等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存

卡 (Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0198] 所述输入输出单元也可以用接收器和发送器代替,可以为相同或者不同的物理实体。为相同的物理实体时,可以统称为输入输出单元。该输入输出单元可以为收发器。

[0199] 所述存储器可以集成在所述处理器中,也可以与所述处理器分开设置。

[0200] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献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(如ROM/RAM)中,包括若干指令用以使得一台终端(可以是手机,计算机,服务器或者网络设备)执行本申请各个实施例所述的方法。

[0201] 上面结合附图对本申请的实施例进行了描述,但是本申请并不局限于上述的具体实施方式,上述的具体实施方式仅仅是示意性的,而不是限制性的,本领域的普通技术人员在本申请的启示下,在不脱离本申请宗旨和权利要求所保护的范围情况下,还可做出很多形式,凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,这些均属于本申请的保护之内。

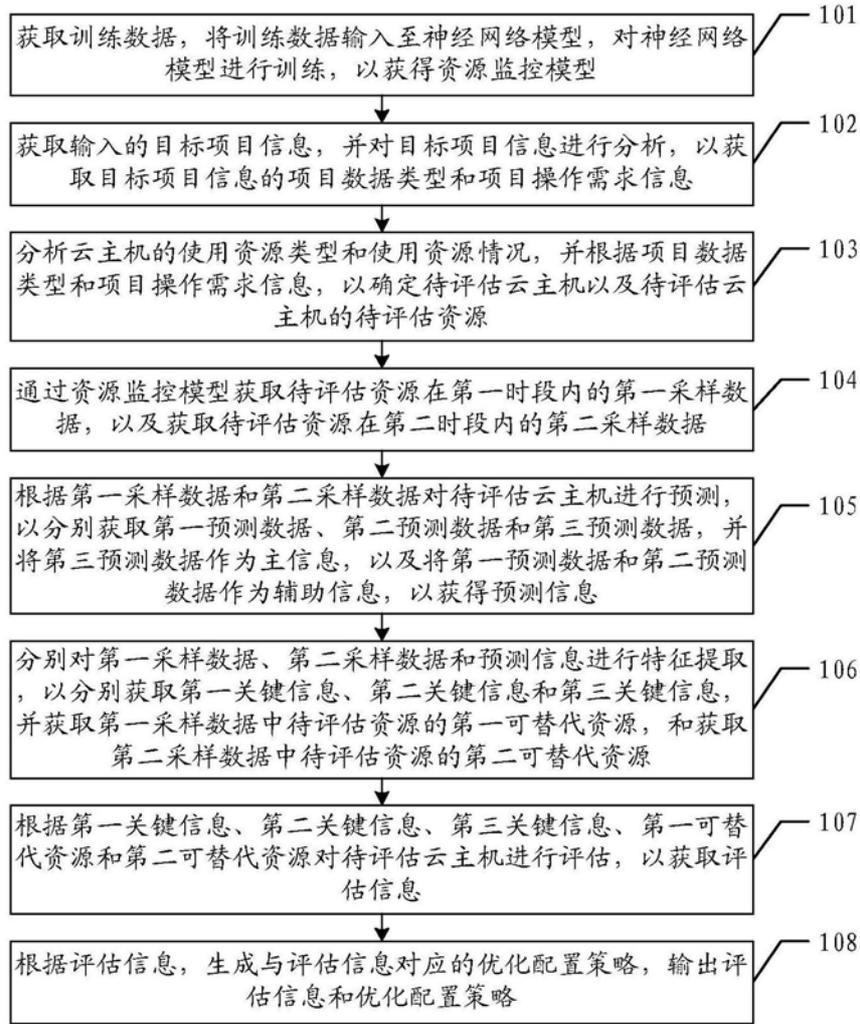


图1

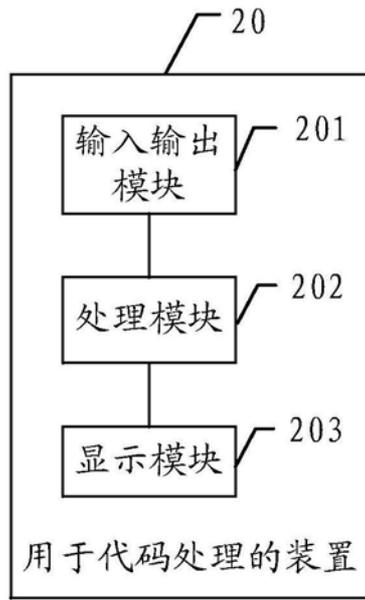


图2

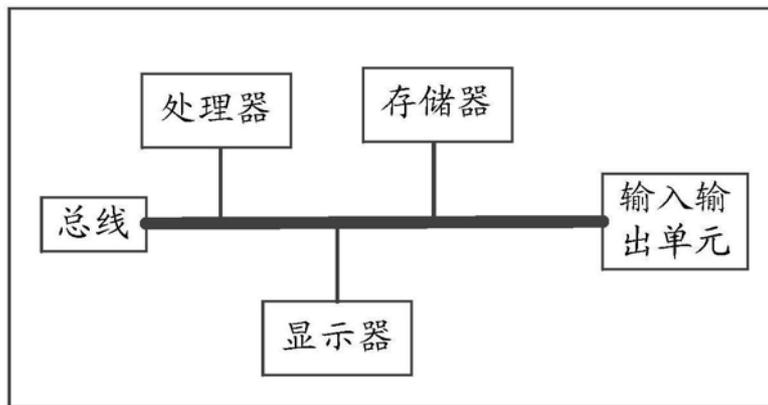


图3