



# (12)发明专利申请

(10)申请公布号 CN 109063150 A

(43)申请公布日 2018. 12. 21

(21)申请号 201810895300.9

(22)申请日 2018.08.08

(71)申请人 湖南永爱生物科技有限公司

地址 410205 湖南省长沙市长沙高新开发区桐梓坡西路408号麓谷林语小区I区综合体1栋501房

(72)发明人 徐国雄 刘分 刘任仲

(51) Int. Cl.

G06F 17/30(2006.01)

H04L 12/24(2006.01)

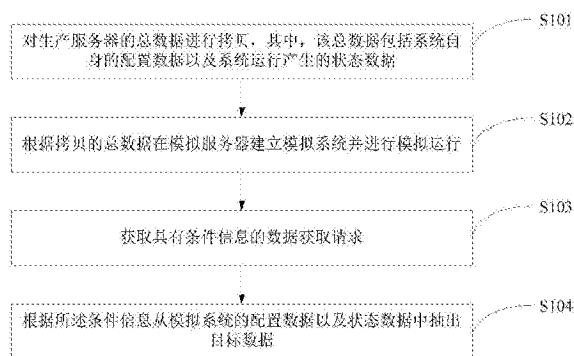
权利要求书2页 说明书5页 附图1页

## (54)发明名称

大数据提取方法、装置、存储介质及服务器

## (57)摘要

本发明实施例提供一种大数据提取方法,用于模拟服务器中,所述方法包括以下步骤:对生产服务器的总数据进行拷贝,其中,该总数据包括系统自身的配置数据以及系统运行产生的状态数据;根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;获取具有条件信息的数据获取请求;根据所述条件信息从模拟系统的配置数据以及状态数据中抽出目标数据。本发明可以降低对生产服务器的压力,提高生产服务器的效率。



1. 一种大数据提取方法,用于模拟服务器中,其特征在于,所述方法包括以下步骤:  
对生产服务器的总数据进行拷贝,其中,该总数据包括系统自身的配置数据以及系统运行产生的状态数据;  
根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;  
获取具有条件信息的数据获取请求;  
根据所述条件信息从模拟系统的配置数据以及状态数据中抽出目标数据。
2. 根据权利要求1所述的大数据提取方法,其特征在于,所述对生产服务器的总数据进行拷贝的步骤包括:  
每隔预设时间段对生产服务器的总数据进行全盘拷贝以替代之前拷贝的总数据。
3. 根据权利要求2所述的大数据提取方法,其特征在于,所述对生产服务器的总数据进行拷贝的步骤包括:  
当获取到拷贝触发信号时,识别拷贝触发信号的种类;  
当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝;  
当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。
4. 根据权利要求1所述的大数据提取方法,其特征在于,所述条件信息包括所述条件信息包括:模拟运行结果特征、数据类型特征、时间特征以及地点特征中的一种或多种。
5. 根据权利要求4所述的大数据提取方法,其特征在于,所述根据所述条件信息从模拟系统中抽出目标数据的步骤:  
根据所述条件信息中的数据类型特征从所述模拟系统中抽出对应数据类型的第二数据;  
根据所述条件信息中的时间特征从所述第二数据中筛选出第三数据;  
根据所述模拟系统的模拟运行结果特征从所述第三数据中抽出目标数据。
6. 一种大数据提取装置,其特征在于,包括:  
拷贝模块,用于对生产服务器的总数据进行拷贝,其中,该总数据包括系统自身的配置数据以及系统运行产生的状态数据;  
模拟运行模块,用于根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;  
获取模块,用于获取具有条件信息的数据获取请求;  
提取模块,用于根据所述条件信息从模拟系统的配置数据以及状态数据中抽出目标数据。
7. 根据权利要求6所述的大数据提取装置,其特征在于,所述拷贝模块用于每隔预设时间段对生产服务器的总数据进行全盘拷贝以替代之前拷贝的总数据。
8. 根据权利要求7所述的大数据提取装置,其特征在于,所述拷贝模块包括:  
识别单元,用于当获取到拷贝触发信号时,识别拷贝触发信号的种类;  
第一拷贝单元,用于当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝;  
第二拷贝单元,用于当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。
9. 一种存储介质,其特征在于,所述存储介质中存储有计算机程序,当所述计算机程序

在计算机上运行时,使得所述计算机执行权利要求1至5任一项所述的方法。

10.一种服务器,其特征在于,包括处理器和存储器,所述存储器中存储有计算机程序,所述处理器通过调用所述存储器中存储的所述计算机程序,用于执行权利要求1至5任一项所述的方法。

## 大数据提取方法、装置、存储介质及服务器

### 技术领域

[0001] 本发明涉及大数据处理领域,特别涉及一种大数据提取方法、装置、存储介质及服务器。

### 背景技术

[0002] 现有技术中服务器不仅要接收外部设备的各种数据信息,并根据这些数据信息发出指示信号给各个外部设备,服务器还会对运行过程中满足一定条件的运行结果数据以及原始数据进行提取、分析以及备份等处理。

[0003] 但是,服务器对运行过程中满足一定条件的运行结果数据以及原始数据进行提取、分析以及备份等处理会对占用服务器很大的CPU使用率,影响其对外部设备的正常控制。

[0004] 因此,现有技术存在缺陷,急需改进。

### 发明内容

[0005] 本发明实施例提供一种大数据提取方法、装置、存储介质及服务器,可以减小生产服务器的压力提高生产服务器的运行效率。

[0006] 本发明实施例提供一种大数据提取方法,用于模拟服务器中,所述方法包括以下步骤:

[0007] 对生产服务器的总数据进行拷贝,其中,该总数据包括系统自身的配置数据以及系统运行产生的状态数据;

[0008] 根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;

[0009] 获取具有条件信息的数据获取请求;

[0010] 根据所述条件信息从模拟系统的配置数据以及状态数据中抽出目标数据。

[0011] 在本发明实施例所述的大数据提取方法中,所述对生产服务器的总数据进行拷贝的步骤包括:

[0012] 每隔预设时间段对生产服务器的总数据进行全盘拷贝以替代之前拷贝的总数据。

[0013] 在本发明实施例所述的大数据提取方法中,所述对生产服务器的总数据进行拷贝的步骤包括:

[0014] 当获取到拷贝触发信号时,识别拷贝触发信号的种类;

[0015] 当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝;

[0016] 当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。

[0017] 在本发明实施例所述的大数据提取方法中,所述条件信息包括所述条件信息包括:模拟运行结果特征、数据类型特征、时间特征以及地点特征中的一种或多种。

[0018] 在本发明实施例所述的大数据提取方法中,所述根据所述条件信息从模拟系统中抽出目标数据的步骤:

- [0019] 根据所述条件信息中的数据类型特征从所述模拟系统中抽出对应数据类型的第二数据；
- [0020] 根据所述条件信息中的时间特征从所述第二数据中筛选出第三数据；
- [0021] 根据所述模拟系统的模拟运行结果特征从所述第三数据中抽出目标数据。
- [0022] 一种大数据提取装置,包括:
- [0023] 拷贝模块,用于对生产服务器的总数据进行拷贝,其中,该总数据包括系统自身的配置数据以及系统运行产生的状态数据;
- [0024] 模拟运行模块,用于根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;
- [0025] 获取模块,用于获取具有条件信息的数据获取请求;
- [0026] 提取模块,用于根据所述条件信息从模拟系统的配置数据以及状态数据中抽出目标数据。
- [0027] 在本发明实施例所述的大数据提取装置中,所述拷贝模块用于每隔预设时间段对生产服务器的总数据进行全盘拷贝以替代之前拷贝的总数据。
- [0028] 所述拷贝模块用于每隔预设时间段所述拷贝模块包括:
- [0029] 识别单元,用于当获取到拷贝触发信号时,识别拷贝触发信号的种类;
- [0030] 第一拷贝单元,用于当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝;
- [0031] 第二拷贝单元,用于当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。
- [0032] 由上可知,本发明通过对从生产服务器获取的原始数据进行模拟,并生成模拟的结果数据;获取将要提取的目标数据的条件信息;根据所述条件信息从所述原始数据以及所述结果数据中获取目标数据;对所述目标数据进行分析处理;从而实现对生产服务器中的数据的提取分析处理,可以降低对生产服务器的压力,提高生产服务器的效率。

### 附图说明

- [0033] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍。显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。
- [0034] 图1是本发明实施例提供的大数据提取方法的流程示意图。
- [0035] 图2是本发明实施例提供的大数据提取装置的结构示意图。

### 具体实施方式

- [0036] 下面将结合本发明实施例中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域技术人员在没有付出创造性劳动前提下所获得的所有其他实施例,都属于本发明的保护范围。
- [0037] 本发明的说明书和权利要求书以及上述附图中的术语“第一”、“第二”、“第三”等

(如果存在)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。应当理解,这样描述的对象在适当情况下可以互换。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排除的包含。例如,包含了一系列步骤的过程、方法或包含了一系列模块或单元的装置、服务器、系统不必限于清楚地列出的那些步骤或模块或单元,还可以包括没有清楚地列出的步骤或模块或单元,也可以包括对于这些过程、方法、装置、服务器或系统固有的其它步骤或模块或单元。

[0038] 参考图1,图1为本发明实施例中的大数据提取方法的流程图。方法可以应用于模拟服务器中。在本实施例中,该方法包括以下步骤:

[0039] S101、对生产服务器的总数据进行拷贝,其中该总数据包括系统自身配置数据以及运行数据。

[0040] 其中,该总数据可以包括该构成该生产服务器的运行系统的系统自身配置数据以及该系统中运行的运行数据。该系统自身配置数据是每隔一断时间更新一次的,就像常用的操作系统一样。

[0041] 在一些实施例中,该步骤S101包括:每隔预设时间段对生产服务器的总数据进行拷贝。例如,可以每隔12小时进行一次复次,后一次拷贝的内容全面覆盖前一次拷贝的内容。

[0042] 在一些实施例中,该步骤S101包括:

[0043] S1011、当获取到拷贝触发信号时,识别拷贝触发信号的种类。

[0044] 其中,该拷贝触发信号的种类根据预设生成。例如,系统设定是每周三对生产服务器的系统进行升级,因此,当周三时,模拟服务器的模拟系统同样需要进行更新,因此其生成的拷贝触发信号为第一类,其他时间生成的拷贝触发信号则为第二类。

[0045] S1012、当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝。

[0046] 当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。

[0047] S102、根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行。

[0048] 该模拟服务器中设置有虚拟运行平台,其从生产服务器获取原始数据,并按照与生产服务器相同的逻辑,对该原始数据进行模拟。其中,该总数据中的运行数据可以包括以下多种数据:影像数据、音频数据、图片数据、数字数据、文字数据、指纹数据等。

[0049] 在一些实施例中,该步骤S102包括:

[0050] S1021、获取所述目标数据块要求满足的数据特征信息;

[0051] S1022、获取所述目标数据块要求满足的结果类型。

[0052] 其中,该数据特征信息包括以下特征中的至少一种:数据类型特征、时间特征、地点特征、关键字特征。该结果类型可以是对原始数据进行识别的结果、加工的结果或者其他处理的结果。

[0053] S103、获取具有条件信息的数据获取请求。

[0054] 优选地,所述条件信息包括所述条件信息包括:模拟运行结果特征、数据类型特征、时间特征以及地点特征中的一种或多种。

[0055] S104、根据所述条件信息从模拟系统中抽出目标数据。

[0056] 其中,该目标数据可以是最开始的总数据中的运行数据,也可以是运行数据经过

模拟运行后产生的运行结果数据。

[0057] 其中,可以根据现有的技术对提取出的目标数据进行备份、分类、关联、人工智能学习等处理。

[0058] 例如,多个目标数据块的数据特征信息可以相同也可以不相同,多个目标数据块的模式不同,例如,可能有一些目标数据块是图片数据,有一些目标数据块是文字数据。而对图片数据的是识别对图片数据进行识别,而对文字数据的模拟是识别其中的关键字。

[0059] 在一些实施例中,该步骤S104包括:

[0060] S1041、根据所述条件信息中的数据类型特征从所述模拟系统中抽出对应数据类型的第二数据;

[0061] S1042、根据所述条件信息中的时间特征从所述第二数据中筛选出第三数据;

[0062] S1043、根据所述模拟系统的模拟运行结果特征从所述第三数据中抽出目标数据。例如,该原始数据中包括有图片数据、文字数据以及视频数据。数据特征信息包括时间特征、数据类型特征、地点特征。在这里,该数据特征信息的筛选条件为XX年YY日ZZ时、图片数据以及地点位于天安门广场。而在本实施例中,模拟的类型为图片识别,运行结果类型的条件为:识别出待筛选图片中的黑色人种。

[0063] 根据所述第二数据块的数据类型特征、时间特征以及地点特征以及所述第三数据块对所述多个目标数据块进行分组存储。

[0064] 由上可知,本发明通过对从生产服务器获取的原始数据进行模拟,并生成模拟的结果数据;获取将要提取的目标数据的条件信息;根据所述条件信息从所述原始数据以及所述结果数据中获取目标数据;对所述目标数据进行分析处理;从而实现对生产服务器中的数据的提取分析处理,可以降低对生产服务器的压力,提高生产服务器的效率。

[0065] 请参照图2,图2是本发明实施例中的一种大数据提取装置的结构图。该装置包括:

[0066] 拷贝模块201,用于对生产服务器的总数据进行拷贝,其中该总数据包括系统自身配置数据以及运行数据;

[0067] 模拟运行模块202,用于根据拷贝的总数据在模拟服务器建立模拟系统并进行模拟运行;

[0068] 获取模块203,用于获取具有条件信息的数据获取请求;

[0069] 提取模块204,用于根据所述条件信息从模拟系统中抽出目标数据。

[0070] 优选地,所述拷贝模块201用于每隔预设时间段对生产服务器的总数据进行拷贝。

[0071] 优选地,拷贝模块201包括:识别单元、第一拷贝单元以及第二拷贝单元。

[0072] 其中,识别单元用于当获取到拷贝触发信号时,识别拷贝触发信号的种类。

[0073] 其中,第一拷贝单元用于当所述拷贝触发信号的为第一类时,对所述总数据中的系统自身配置数据以及运行数据进行拷贝。

[0074] 其中,第二拷贝单元用于当所述拷贝触发信号的为第二类时,对所述总数据中的运行数据进行拷贝。其中,该拷贝触发信号的种类根据预设生成。例如,系统设定是每周三对生产服务器的系统进行升级,因此,当周三时,模拟服务器的模拟系统同样需要进行更新,因此其生成的拷贝触发信号为第一类,其他时间生成的拷贝触发信号则为第二类。

[0075] 本发明实施例还提供一种存储介质,所述存储介质中存储有计算机程序,当所述计算机程序在计算机上运行时,使得所述计算机执行上述实施例中的大数据提取方法。

[0076] 本发明实施例还提供一种服务器,包括处理器和存储器,所述存储器中存储有计算机程序,所述处理器通过调用所述存储器中存储的所述计算机程序,用于执行上述实施例中的大数据提取方法。

[0077] 需要说明的是,本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成,该程序可以存储于计算机可读存储介质中,该存储介质可以包括但不限于:只读存储器(ROM,Read Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁盘或光盘等。

[0078] 以上对本发明实施例所提供的大数据提取方法、装置、存储介质及服务器进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。



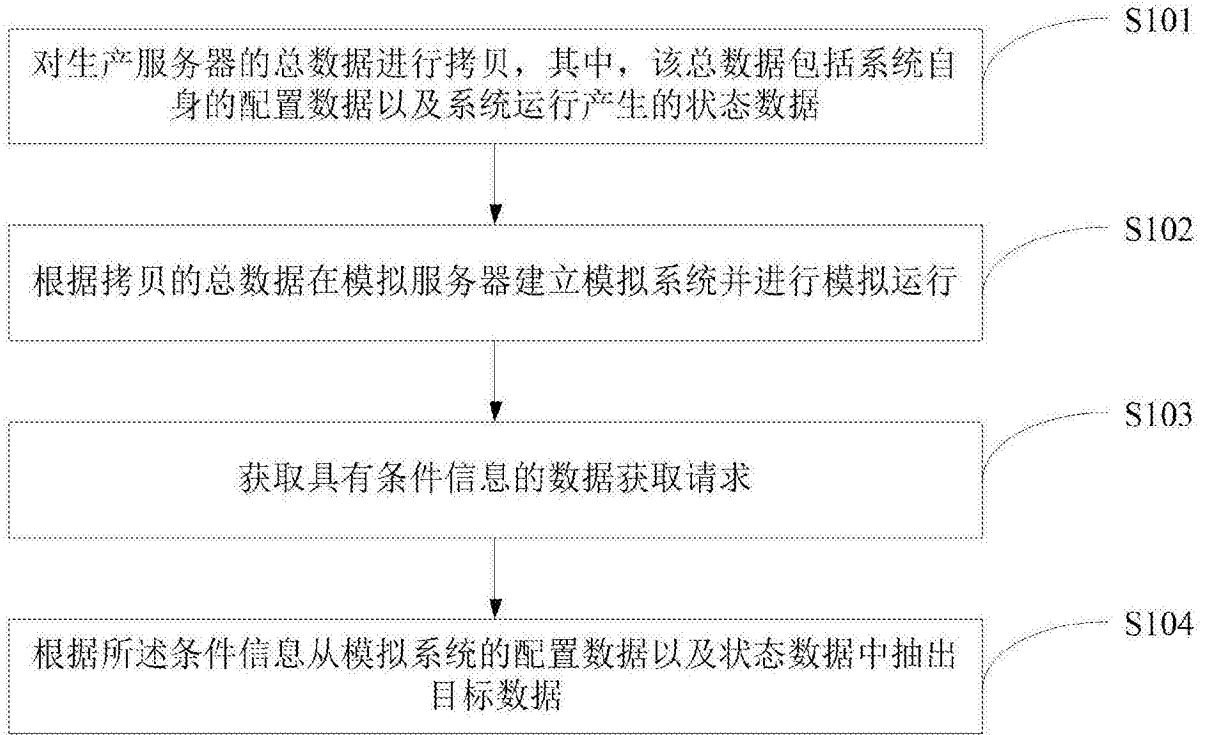


图1

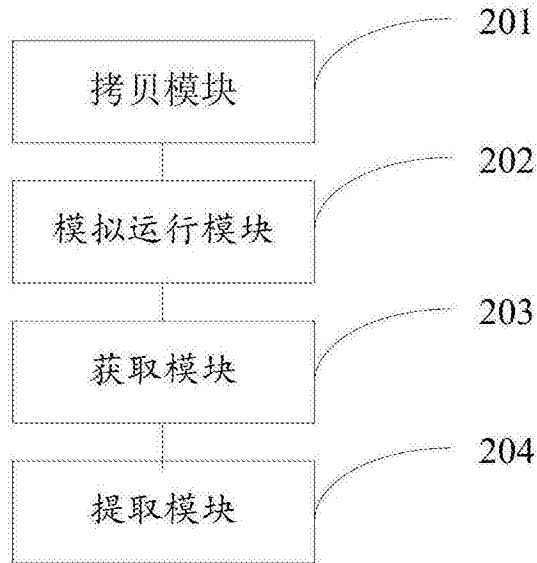


图2