



(12) 发明专利申请

(10) 申请公布号 CN 105068763 A

(43) 申请公布日 2015. 11. 18

(21) 申请号 201510496008. 6

(22) 申请日 2015. 08. 13

(71) 申请人 武汉噢易云计算有限公司

地址 430074 湖北省武汉市东湖新技术开发区关山大道 465 号光谷创意大厦 17 层 1701 室

(72) 发明人 李文文

(74) 专利代理机构 武汉智权专利代理事务所

(特殊普通合伙) 42225

代理人 张凯

(51) Int. Cl.

G06F 3/06(2006. 01)

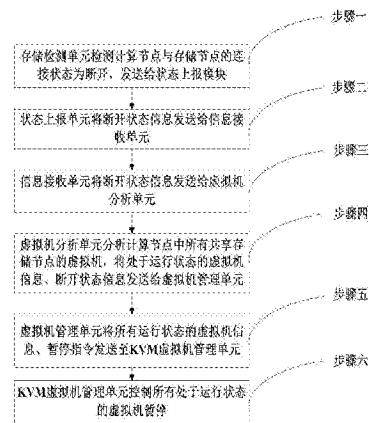
权利要求书2页 说明书5页 附图2页

(54) 发明名称

一种针对存储故障的虚拟机容错系统和方法

(57) 摘要

本发明公开了一种针对存储故障的虚拟机容错系统和方法,涉及云计算领域,系统包括分析管理模块和至少一个检测上报模块;分析管理模块与控制节点相对应,每个检测上报模块均与一计算节点相对应;检测上报模块用于实时检测相应的计算节点与相对应的存储节点的通信状态并在通信状态为断开时上报断开状态至分析管理模块;分析管理模块用于接收检测上报模块上报的计算节点与相对应的存储节点的通信状态,并在通信状态为断开时,使计算节点中与存储节点相应且处于运行状态的虚拟机暂停。本发明不仅能够缩短业务的恢复时间,而且能够避免虚拟机的数据丢失。



1. 一种针对存储故障的虚拟机容错系统,其特征在于:该系统包括控制节点、至少一个计算节点和至少一个存储节点;

每个计算节点与至少一台虚拟机相对应,每个计算节点至少与一存储节点对应,虚拟机将数据发送至与对应计算节点相应的存储节点;

所述系统还包括分析控制平台,该平台包括分析管理模块和至少一个检测上报模块;

所述分析管理模块与控制节点相对应,每个检测上报模块均与一计算节点相对应;

所述检测上报模块用于实时检测相应的计算节点与相对应的存储节点的通信状态并在通信状态为断开时上报断开状态至分析管理模块;

所述分析管理模块用于接收检测上报模块上报的计算节点与相对应的存储节点的通信状态,并在通信状态为断开时,使所述计算节点中与所述存储节点相应且处于运行状态的虚拟机暂停。

2. 如权利要求 1 所述的针对存储故障的虚拟机容错系统,其特征在于:所述分析管理模块还用于在计算节点与相对应的存储节点由断开状态恢复到连通状态时,使所述计算节点中与所述存储节点相应且在断开状态下被暂停的虚拟机恢复运行。

3. 如权利要求 2 所述的针对存储故障的虚拟机容错系统,其特征在于:所述分析管理模块包括虚拟机分析单元、状态信息接收单元和虚拟机管理单元,所述状态信息接收单元用于接收所述检测上报模块上报的通信状态,并将通信状态发送至虚拟机分析单元,所述虚拟机分析单元用于在状态上报单元上报的通信状态为断开时,标记所述计算节点中与所述存储节点相应且处于运行状态的虚拟机,并将被标记虚拟机的信息发送至虚拟机分析单元;

所述虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送暂停指令和被标记的虚拟机信息,所述检测上报模块控制被标记的虚拟机暂停。

4. 如权利要求 3 所述的用于针对存储故障的虚拟机容错系统,其特征在于:所述虚拟机分析单元还用于在状态上报单元上报的通信状态由断开恢复为连通时,并将在断开时被标记虚拟机的信息发送至虚拟机分析单元;

所述虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送恢复运行指令和被标记的虚拟机信息,所述检测上报模块控制被标记的虚拟机恢复运行。

5. 如权利要求 4 所述的针对存储故障的虚拟机容错系统,其特征在于:所述检测上报模块包括存储状态监测单元、状态上报单元和 KVM 虚拟机管理单元;

所述存储状态监测单元用于实时监测相应计算节点与对应存储节点的通信状态,并将通信状态发送至状态上报单元;

状态上报单元用于接收存储状态监测单元上报的通信状态,并将通信状态发送至信息接收单元;

KVM 虚拟机管理单元用于接收并执行虚拟机管理单元发送的指令,根据指令控制相应虚拟机的暂停或恢复运行。

6. 一种针对存储故障的虚拟机容错的方法,该方法用于对与存储节点对应的计算节点相应的虚拟机容错,其特征在于:该方法包括以下步骤:

a、检测计算节点与存储节点之间的通信状态,并在通信状态为断开时上报断开状态;

b、在通信状态为断开时控制当前计算节点中与所述存储节点相应的处于运行的虚拟机暂停。

7. 如权利要求6所述的针对存储故障的虚拟机容错的方法,其特征在于:在所述步骤b具体包括以下步骤:

通信状态为断开时,对所述计算节点中与所述存储节点相应的处于运行状态的虚拟机进行标记,控制所有被标记的虚拟机暂停。

8. 如权利要求7所述的针对存储故障的虚拟机容错的方法,其特征在于:在步骤b之后还包括以下步骤:

检测计算节点与相对应的存储节点的断开状态是否恢复到连通,并在恢复连通时控制所有被标记的虚拟机恢复运行。

## 一种针对存储故障的虚拟机容错系统和方法

### 技术领域

[0001] 本发明涉及云计算领域,具体涉及一种针对存储故障的虚拟机容错系统和方法。

### 背景技术

[0002] 目前虚拟化的技术发展已日渐成熟,基于开关机的虚拟机高可用技术在目前的生产环境中也已经部署较多,但是虚拟机高在使用过程中,当虚拟机的存储系统出现故障时,高可用技术难以感知,需要等待虚拟机系统 IO(输入输出)完成读写,IO读写超时时才会将超时信息反馈给虚拟机系统,使得 windows 虚拟机系统蓝屏或 Linux 虚拟机内核 Panic(报错),此时,管理平台感仍然难以知晓虚拟机已出现异常,造成虚拟机内核崩溃,虚拟机内核崩溃后管理平台上显示的该虚拟机状态仍为正常通信状态,排错难度较大,延长了了业务恢复的时间。

[0003] 当主备存储切换的过程中,计算节点与存储节点断开,虚拟机的数据无法正常写入和读出,虚拟机非正常写入和读出的时间超过 10s 后(不同操作系统发生异常的时间不一样),会发生异常,当主备切换操作完成以后,虚拟机也不能恢复正常,需要在管理平台上对虚拟机进行重置,重置不仅会延长虚拟机业务的恢复时间,而且会导致虚拟机在重置过程中由于电源关闭引起的数据丢失。

### 发明内容

[0004] 针对现有技术中存在的缺陷,本发明的目的在于提供一种针对存储故障的虚拟机容错系统和方法,不仅能够缩短业务的恢复时间,而且能够避免虚拟机的数据丢失。

[0005] 为达到以上目的,本发明采取的技术方案是:一种针对存储故障的虚拟机容错系统,其特征在于:该系统包括控制节点、至少一个计算节点和至少一个存储节点;

[0006] 每个计算节点与至少一台虚拟机相对应,每个计算节点至少与一存储节点对应,虚拟机将数据发送至与对应计算节点相应的存储节点;

[0007] 所述系统还包括分析控制平台,该平台包括分析管理模块和至少一个检测上报模块;

[0008] 所述分析管理模块与控制节点相对应,每个检测上报模块均与一计算节点相对应;

[0009] 所述检测上报模块用于实时检测相应的计算节点与相对应的存储节点的通信状态并在通信状态为断开时上报断开状态至分析管理模块;

[0010] 所述分析管理模块用于接收检测上报模块上报的计算节点与相对应的存储节点的通信状态,并在通信状态为断开时,使所述计算节点中与所述存储节点相应且处于运行状态的虚拟机暂停。

[0011] 在上述技术方案的基础上,所述分析管理模块还用于在计算节点与相对应的存储节点由断开状态恢复到连通状态时,使所述计算节点中与所述存储节点相应且在断开状态下被暂停的虚拟机恢复运行。

[0012] 在上述技术方案的基础上,所述分析管理模块包括虚拟机分析单元、状态信息接收单元和虚拟机管理单元,所述状态信息接收单元用于接收所述检测上报模块上报的通信状态,并将通信状态发送至虚拟机分析单元,所述虚拟机分析单元用于在状态上报单元上报的通信状态为断开时,标记所述计算节点中与所述存储节点相应且处于运行状态的虚拟机,并将被标记虚拟机的信息发送至虚拟机分析单元;

[0013] 所述虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送暂停指令和被标记的虚拟机信息,所述检测上报模块控制被标记的虚拟机暂停。

[0014] 在上述技术方案的基础上,所述虚拟机分析单元还用于在状态上报单元上报的通信状态由断开恢复为连通时,并将在断开时被标记虚拟机的信息发送至虚拟机分析单元;

[0015] 所述虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送恢复运行指令和被标记的虚拟机信息,所述检测上报模块控制被标记的虚拟机恢复运行。

[0016] 在上述技术方案的基础上,所述检测上报模块包括存储状态监测单元、状态上报单元和 KVM 虚拟机管理单元;

[0017] 所述存储状态监测单元用于实时监测相应计算节点与对应存储节点的通信状态,并将通信状态发送至状态上报单元;

[0018] 状态上报单元用于接收存储状态监测单元上报的通信状态,并将通信状态发送至信息接收单元;

[0019] KVM 虚拟机管理单元用于接收并执行虚拟机管理单元发送的指令,根据指令控制相应虚拟机的暂停或恢复运行。

[0020] 一种针对存储故障的虚拟机容错的方法,该方法用于对与存储节点对应的计算节点相应的虚拟机容错,该方法包括以下步骤:

[0021] a、检测计算节点与存储节点之间的通信状态,并在通信状态为断开时上报断开状态;

[0022] b、在通信状态为断开时控制当前计算节点中与所述存储节点相应的处于运行的虚拟机暂停。

[0023] 在上述技术方案的基础上,在所述步骤 b 具体包括以下步骤:

[0024] 通信状态为断开时,对所述计算节点中与所述存储节点相应的处于运行状态的虚拟机进行标记,控制所有被标记的虚拟机暂停。

[0025] 在上述技术方案的基础上,在步骤 b 之后还包括以下步骤:

[0026] 检测计算节点与相对应的存储节点的断开状态是否恢复到连通,并在恢复连通时控制所有被标记的虚拟机恢复运行。

[0027] 在上述技术方案的基础上,

[0028] 与现有技术相比,本发明的优点在于:

[0029] (1) 本发明中针对存储故障的虚拟机容错系统,包括分析管理模块和若干检测上报模块,能够实时检测计算节点与存储节点的通信状态,并根据通信状态控制虚拟机的运行,与现有技术中计算节点与存储节点断开时,虚拟机仍然向存储节点发送数据,容易造成虚拟机内核崩溃、数据丢失相比,本发明的系统能够在计算节点与存储节点断开时,及时使

得虚拟机暂停,并在计算节点与存储节点连通后,恢复虚拟机的运行,避免了数据的丢失和虚拟机内核的崩溃。

[0030] (2) 本发明中针对存储故障的虚拟机容错系统,能够实时获取虚拟机的通信状态,与现有技术中虚拟机内核崩溃后管理平台上显示的该虚拟机状态仍为正常通信状态,排错难度较大,延长了业务恢复的时间相比,本发明的虚拟机发生故障时,能够在较短的时间内确定发生故障的虚拟机,能够有效缩短业务恢复的时间。

[0031] (3) 本发明中针对存储故障的虚拟机容错的方法,检测计算节点与存储节点是否为断开状态,并在断开时确定当前计算节点中共享所述存储节点的所有虚拟机,分析所有虚拟机的通信状态,标记所有处于运行的虚拟机,使所有处于运行的虚拟机暂停,暂停需要的时间在 3s 以内,远远小于现有技术中启动和关闭虚拟机所需要的时间,大大缩减了业务的中断时间。

[0032] (4) 本发明中针对存储故障的虚拟机容错的方法,在虚拟机暂停之后,检测计算节点与存储节点是否为连通,并在连通时确定当前计算节点中共享所述存储节点的所有虚拟机,将所有虚拟机中被标记的虚拟机通信状态修改为运行,本发明通过发送指令使得虚拟机暂停和运行,不需要进行人工干预,提升了系统的稳定性,且虚拟机暂停和运行前后,虚拟机内存数据前后一致,在整个存储故障发生到最后恢复,虚拟机能够接着上一条运行的指令继续往下运行,能够避免内存数据的丢失。

## 附图说明

[0033] 图 1 为本发明实施例中针对存储故障的虚拟机容错系统的结构示意图。

[0034] 图 2 为本发明中的分析平台的结构框图;

[0035] 图 3 为本发明实施例中计算节点与存储节点断开时针对存储故障的虚拟机容错的方法的流程图。

[0036]

## 具体实施方式

[0037] 以下结合附图及实施例对本发明作进一步详细说明。

[0038] 参见图 1 所示,本发明实施例提供一种针对存储故障的虚拟机容错系统,该系统包括控制节点、至少一个计算节点和至少一个存储节点,所有计算节点均与控制节点通信,所有存储节点均与控制节点通信,所有计算节点与相应的存储节点通信。

[0039] 每个计算节点与至少一台虚拟机相对应,每个计算节点均与至少一存储节点对应,计算节点的虚拟机能够将数据发送至相应的存储节点,同一计算节点中的不同虚拟机可对应不同的存储节点,且所有存储节点均与所述同一计算节点相应。

[0040] 本实施例中,所有计算节点均通过网络与控制节点通信,所有存储节点均通过存储交换机与控制节点和计算节点进行通信。

[0041] 参见图 2 所示,本发明的系统还包括分析控制平台,该平台包括分析管理模块和至少一检测上报模块,分析管理模块与控制节点对应,每个检测上报模块均与一计算节点对应。

[0042] 检测上报模块用于检测对应计算节点与相应存储节点的通信状态:断开或连通,

将通信状态上报至分析管理模块,接收并执行分析管理模块的指令。

[0043] 分析管理模块用于接收检测上报模块的通信状态,并在计算节点与存储节点断开时,分析计算节点中与存储节点相应的所有虚拟机的通信状态:运行或暂停,并对所有处于运行的虚拟机进行标记,使被标记的虚拟机暂停。

[0044] 分析管理模块还用于通信状态由断开恢复连通时,使所述计算节点中与所述存储节点相应,且由于断开被暂停的虚拟机恢复运行。

[0045] 分析管理模块包括虚拟机分析单元、状态信息接收单元和虚拟机管理单元,状态信息接收单元用于接收检测上报模块上报的通信状态,并将通信状态发送至虚拟机分析单元,虚拟机分析单元用于在状态上报单元上报的通信状态为断开时,标记计算节点中与存储节点相应且处于运行状态的虚拟机,并将被标记虚拟机的信息发送至虚拟机分析单元。

[0046] 虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送暂停指令和被标记的虚拟机信息,检测上报模块控制被标记的虚拟机暂停。

[0047] 虚拟机分析单元还用于在状态上报单元上报的通信状态由断开恢复为连通时,并将在断开时被标记虚拟机的信息发送至虚拟机分析单元。虚拟机分析单元用于接收虚拟机分析上报的被标记虚拟机的信息,并向检测上报模块发送恢复运行指令和被标记的虚拟机信息,检测上报模块控制被标记的虚拟机恢复运行。

[0048] 检测上报模块包括存储状态监测单元、状态上报单元和 KVM 虚拟机管理单元。

[0049] 存储状态监测单元用于实时监测所有计算节点与存储节点的通信状态,并将通信状态发送至状态上报单元。

[0050] 状态上报单元用于接收存储状态监测单元的通信状态,并将通信状态发送至信息接收单元。

[0051] KVM 虚拟机管理单元用于接收并执行虚拟机管理单元发送的指令,根据指令控制相应虚拟机的暂停或恢复运行。

[0052] 一种针对存储故障的虚拟机容错系统的方法,该方法用于对与存储节点对应的计算节点相应的虚拟机容错,包括以下步骤:

[0053] 1、检测计算节点与存储节点之间的通信状态,并在通信状态为断开时上报断开状态;

[0054] 2、对所述计算节点中与所述存储节点相应的处于运行状态的虚拟机进行标记,在通信状态为断开时控制当前计算节点中与所述存储节点相应的处于运行的虚拟机暂停。

[0055] 3、检测计算节点与相对应的存储节点的断开状态是否恢复到连通,并在恢复连通时控制所有被标记的虚拟机恢复运行。

[0056] 参见图 3 所示,当前通信状态为断开时,针对存储故障的虚拟机容错系统的方法包括以下步骤。

[0057] 步骤一、存储状态监测单元监测所有计算节点与相应存储节点的通信状态,得到断开状态信息,并将断开状态信息发送至状态上报单元,转入步骤二。

[0058] 步骤二、状态上报单元接收存储状态监测单元发送的断开状态信息并发送至信息接收单元,转入步骤三。

[0059] 步骤三、信息接收单元接收状态上报单元的断开状态信息并发送至虚拟机分析单元,转入步骤四。

[0060] 步骤四、虚拟机分析单元接收断开状态信息,分析当前状态为断开状态下的计算节点中共享所述存储节点的所有虚拟机,确定所有虚拟机中处于通信状态虚拟机的信息、计算节点与相应存储节点的通信状态发送至虚拟机管理单元,转入步骤五。

[0061] 步骤五、虚拟机管理单元将相应处于通信状态虚拟机的信息和暂停指令发送至 KVM 虚拟机管理单元,转入步骤六。

[0062] 步骤六、KVM 虚拟机管理单元控制相应处于通信状态的虚拟机暂停。

[0063] 计算节点与相应存储节点的通信状态由断开恢复连通时,针对存储故障的虚拟机容错系统的方法包括以下步骤。

[0064] S1:存储状态监测单元监测所有计算节点与相应存储节点的通信状态,得到连通状态信息,并将连通状态信息发送至状态上报单元,转入 S2。

[0065] S2:状态上报单元接收存储状态监测单元发送的连通状态信息并发送至信息接收单元,转入 S3。

[0066] S3:信息接收单元接收状态上报单元的连通状态信息并发送至虚拟机分析单元,转入 S4。

[0067] S4:虚拟机分析单元接收连通状态信息,分析当前状态为连通状态下的计算节点中共享所述存储节点的所有虚拟机,确定在断开状态下被暂停的所有虚拟机,将连通状态信息、所有被暂停的虚拟机信息发送至至虚拟机管理单元,转入 S5。

[0068] S5:虚拟机管理单元将所有被暂停的虚拟机信息和运行指令发送至 KVM 虚拟机管理单元,转入 S6。

[0069] S6:KVM 虚拟机管理单元控制所有被暂停的虚拟机重新运行。

[0070] 本发明不局限于上述实施方式,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也视为本发明的保护范围之内。本说明书中未作详细描述的内容属于本领域专业技术人员公知的现有技术。



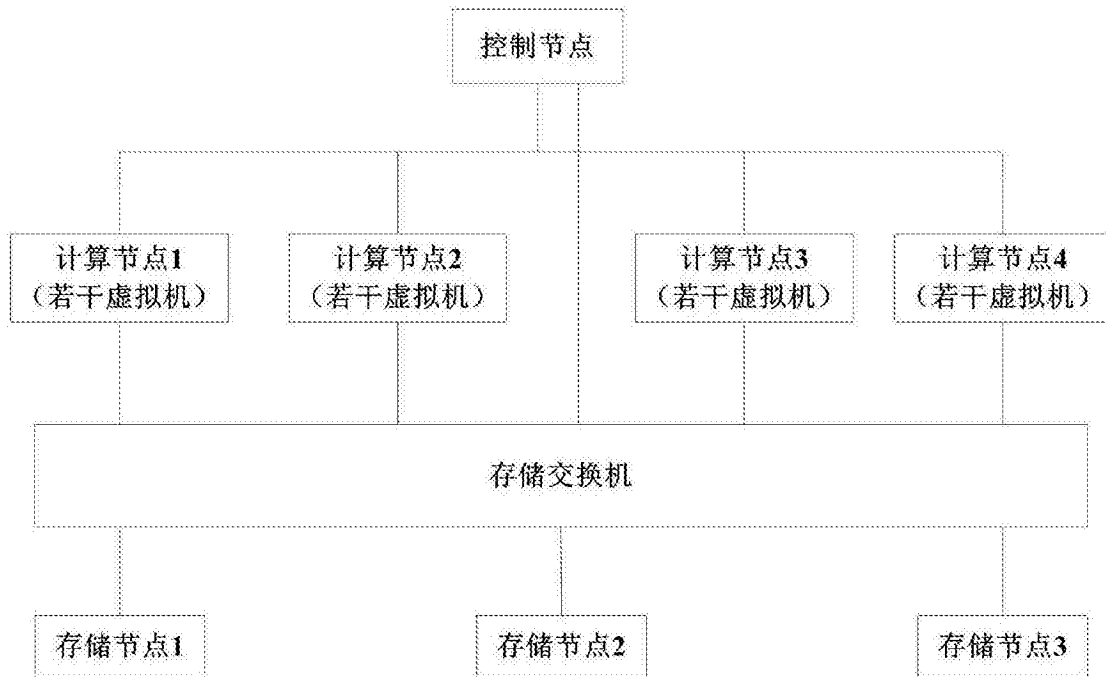


图 1

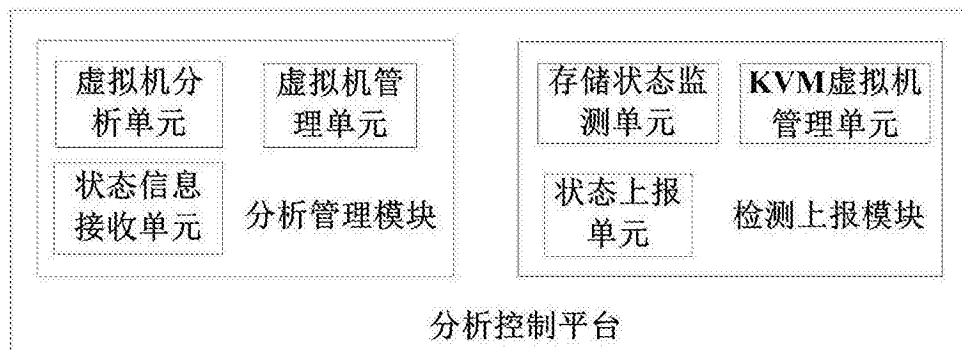


图 2

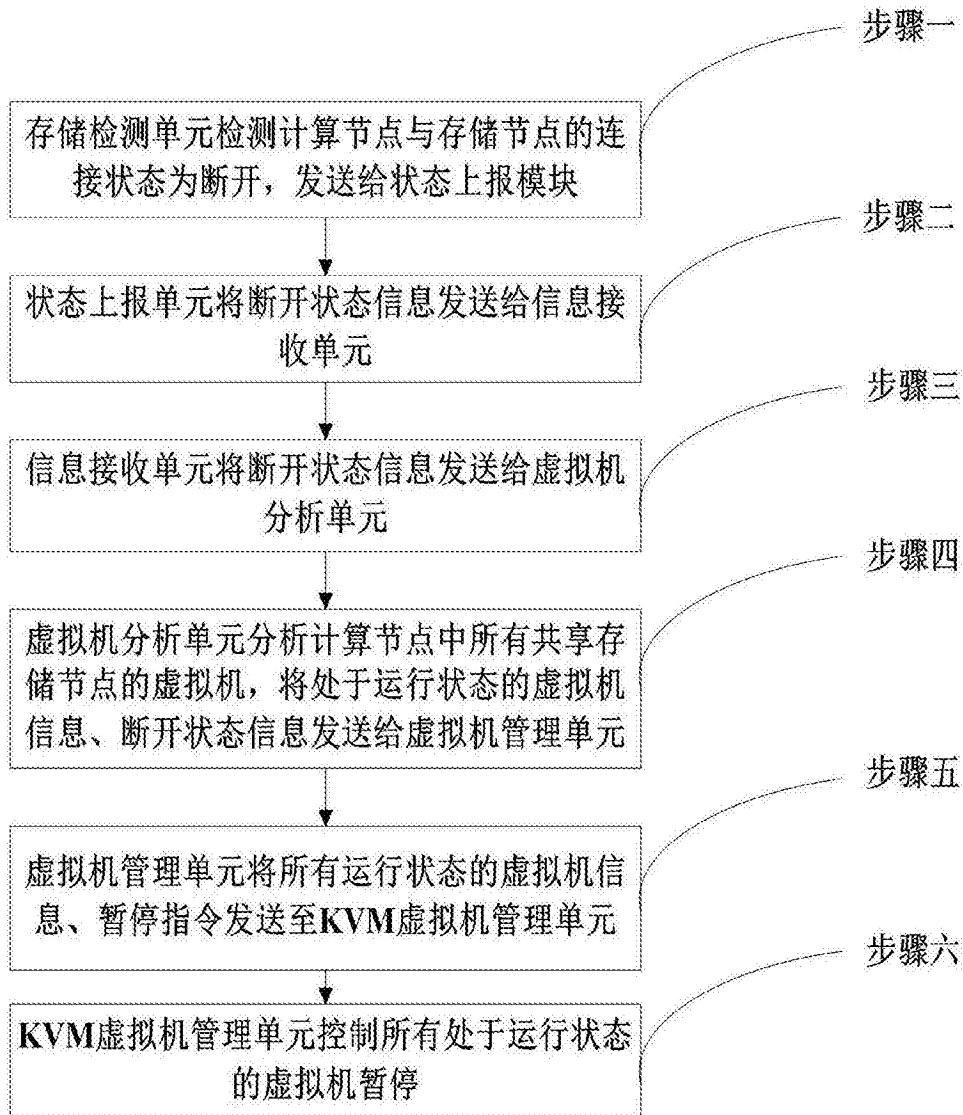


图 3