



(12)发明专利

(10)授权公告号 CN 104679595 B

(45)授权公告日 2017. 10. 13

(21)申请号 201510135990.4

(22)申请日 2015.03.26

(65)同一申请的已公布的文献号

申请公布号 CN 104679595 A

(43)申请公布日 2015.06.03

(73)专利权人 南京大学

地址 210008 江苏省南京市鼓楼区汉口路
22号

(72)发明人 曹春 马晓星 朱荣飞

(74)专利代理机构 南京苏高专利商标事务所

(普通合伙) 32204

代理人 柏尚春

(51)Int.Cl.

G06F 9/50(2006.01)

G06F 17/30(2006.01)

(56)对比文件

CN 103458052 A,2013.12.18,

CN 103873587 A,2014.06.18,

CN 104123189 A,2014.10.29,

CN 102427473 A,2012.04.25,

CN 103365713 A,2013.10.23,

CN 103383653 A,2013.11.06,

US 2005081208 A1,2005.04.14,

US 2010070970 A1,2010.03.18,

审查员 邵娜娜

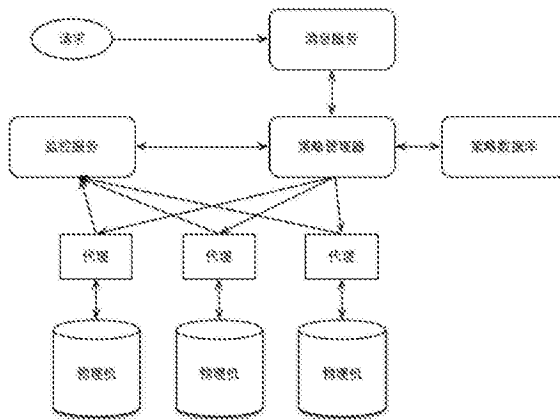
权利要求书1页 说明书5页 附图4页

(54)发明名称

一种面向应用的IaaS层动态资源分配方法

(57)摘要

本发明公开一种面向应用的IaaS层动态资源分配方法,包括以下步骤:系统或用户产生触发资源分配的消息;对产生的消息进行匹配,在策略数据库中搜寻匹配的策略;针对匹配的策略,根据用户自定义的约束条件,与资源监控服务进行通信,判断约束条件是否满足;对满足约束条件的策略,根据用户自定义的行为,与资源监控服务和策略实施服务进行通信,实施具体的资源分配策略。本发明中用户可根据应用的特性定制相应的资源动态分配策略,提高资源利用率并满足用户的多管理目标。



1. 一种面向应用的IaaS层动态资源分配方法,其特征在于:包括数据采集分析、资源调度决策和对云平台的具体组件实施资源分配,具体步骤为:

(1) 系统或用户产生触发资源分配的消息;

(2) 对产生的消息进行匹配,在策略数据库中搜寻匹配的策略;

(3) 针对匹配的策略,根据用户自定义的约束条件,与资源监控服务进行通信,判断约束条件是否满足;

(4) 对满足约束条件的策略,根据用户自定义的行为,与资源监控服务和策略实施服务进行通信,实施具体的资源分配策略;

其中,用户可定义的具体资源调度行为可分为局部调度和全局调度,所述全局调度是指考虑平台的整体资源信息进行的调度,包括用户可自定义全局目标函数及搜索深度,策略语言提供用户自定义目标函数的接口,用户可使用预定义的关键字定义目标函数,策略内置的全局搜索算法将根据目标函数及搜索深度生成调度指令,发送至实施模块。

2. 根据权利要求1所述的面向应用的IaaS层动态资源分配方法,其特征在于:所述步骤(1)中,触发资源分配的消息包括初次分配资源和对已有的资源进行重新调度,可通过用户自主产生或根据环境产生。

3. 根据权利要求1所述的面向应用的IaaS层动态资源分配方法,其特征在于:所述步骤(3)中,用户根据策略语言对资源监控服务定义;资源监控服务采集的数据在纵向上包括有底层的平台信息和高层的应用信息,当数据通过代理进行采集后经过处理;资源监控服务采集的数据载时间维度横向上包括有历史统计数据,当前状态信息以及预测数据。

4. 根据权利要求1所述的面向应用的IaaS层动态资源分配方法,其特征在于:所述步骤(3)中,对匹配的策略的约束条件为用户可定义的,其表达式包含对资源监控服务采集的数据进行等于、大于、小于以及不等于的判断,并可以使用与、或以及非逻辑连接符进行复杂条件的判断。

5. 根据权利要求1所述的面向应用的IaaS层动态资源分配方法,其特征在于:所述局部调度是指考虑平台资源的局部信息进行的调度,具体包括以下三个步骤:源物理机的选取,在源物理机上选择虚拟机以及目的物理机的选取;

其中,对于物理机和虚拟机的选取,采用filter-cost机制,其中filter根据用户自定义的硬约束条件对虚拟机或物理机集合进行过滤,cost根据用户定义的软约束条件进行评价,选择效益最大化的物理机或虚拟机进行调度。

一种面向应用的IaaS层动态资源分配方法

技术领域

[0001] 本发明涉及云平台的资源动态分配技术,具体涉及一种面向应用的IaaS层动态资源分配方法。

背景技术

[0002] 云计算将基础设施、平台和软件作为一种服务以即付即用的方式提供给消费者,在工业界,这些服务被称为基础设施即服务(IaaS),平台即服务(PaaS)和软件即服务(SaaS),对于基础设施即服务,处理cpu,内存,存储,网络等最基本的基础设施资源被作为服务提供给用户。

[0003] 首先,由于云计算按需分配的特性,往往可以以有限的资源提供超过其额定配额的服务,然而用户的需求是多样的,不同的用户与云服务提供商对服务的质量,类型等要求有着不同的协定,我们称为服务等级协议(SLA);其次,由于云服务提供商的日益增加,大型的计算集群及制冷装置往往会消耗大量的电能,同时产生大量的二氧化碳,因此,如何合理地在云平台中分配资源以达到效益最大化是较为关键的问题。

[0004] 现有的资源分配技术在基础设施层将一个虚拟机看作一个资源分配的单位,资源分配即转化成虚拟机在云平台中的调度。资源调度方法包括:1.服务器整合,即将虚拟机尽可能迁入少量的物理机,将多余的物理机关闭以打到节省能耗的需求;2.考虑迁移能耗的调度,即在原有调度方法上考虑迁移虚拟机带来的资源消耗这一因素;3.基于预测的调度,即对平台内部的工作负载进行预测,预先对资源进行预留或迁移,以保障服务质量。

[0005] 然而,上述现有技术局限于平台本身的信息,没有考虑到实际部署的应用的属性,其次不能满足用户多样化的资源分配需求。

发明内容

[0006] 发明目的:本发明的目的在于解决现有技术中存在的问题,提供一种面向应用的IaaS层动态资源分配方法。

[0007] 技术方案:本发明的一种面向应用的IaaS层动态资源分配方法,基于事件驱动的体系架构,即事件-条件-规则(ECA)的模式,包括数据采集分析、资源调度决策和对云平台的具体组件实施资源分配,具体步骤为:

[0008] (1) 系统或用户产生触发资源分配的消息;

[0009] (2) 对产生的消息进行匹配,在策略数据库中搜寻匹配的策略;

[0010] (3) 针对匹配的策略,根据用户自定义的约束条件,与资源监控服务进行通信,判断约束条件是否满足;

[0011] (4) 对满足约束条件的策略,根据用户自定义的行为,与资源监控服务和策略实施服务进行通信,实施具体的资源分配策略。

[0012] 进一步的,所述步骤(1)中,触发资源分配的消息包括初次分配资源和对已有的资源进行重新调度,可通过用户自主产生或根据环境产生。并且,事件决定了触发资源动态分

配的时机,例如用户新建虚拟机时将会发出事件请求资源,系统会为虚拟机初次分配资源;在运行时刻用户可能会发出请求优化当前平台的整体资源利用率使得平台资源进行重分配,当某些虚拟机得不到足够的资源而导致服务质量下降时也会发出请求使得资源进行重分配。策略语言提供给用户event_generator用以产生特定类型的事件,并可以设置消息产生周期,时间等属性。

[0013] 进一步的,所述步骤(3)中,用户根据策略语言对资源监控服务定义;资源监控服务采集的数据在纵向上包括有底层的平台信息和高层的应用信息,当数据通过代理进行采集后经过处理;资源监控服务采集的数据载时间维度横向上包括有历史统计数据,当前状态信息以及预测数据。当资源分配请求到来时,需要分析监控服务所采集的数据以匹配相应的调度策略以实施调度行为。

[0014] 上述平台层数据包括有物理机或虚拟机的cpu使用率,内存和带宽等,应用层数据包括有应用的响应时间等。历史统计数据包括有段时间内的平均cpu使用率,当前状态信息包括有物理机当前负载的虚拟机个数等等,预测数据包括有对虚拟机工作负载的预测等等。

[0015] 进一步的,所述步骤(3)中,对匹配的策略的约束条件为用户可定义的,其表达式包含对资源监控服务采集的数据进行等于、大于、小于以及不等于的判断,并可以使用与、或以及非等逻辑连接符进行复杂条件的判断。

[0016] 所述步骤(4)中,在云平台的基础设施层,虚拟机作为资源分配的基本单位,对平台资源的动态分配可以转化为对虚拟机的创建,调度,关闭或对物理机开启,关闭等管理方法。决策服务中便是产生一系列对物理机或虚拟机管理的指令。用户可定义的具体资源调度行为可分为局部调度和全局调度。

[0017] 进一步的,所述局部调度是指考虑平台资源的局部信息进行的调度,例如一台物理机的cpu负载低于我们设定的阈值,需要对其运行的虚拟机进行调度以实现服务器整合,具体包括以下三个步骤:源物理机的选取,在源物理机上选择虚拟机以及目的物理机的选取;

[0018] 其中,对于物理机和虚拟机的选取,采用filter-cost机制,其中filter根据用户自定义的硬约束条件对虚拟机或物理机集合进行过滤,cost根据用户定义的软约束条件进行评价,选择效益最大化的虚拟机或虚拟机进行调度。

[0019] 进一步的,所述全局调度是指考虑平台的整体资源信息进行的调度,包括用户可自定义全局目标函数及搜索深度,策略语言提供了用户自定义目标函数的接口,用户可使用一些预定义的关键字如MIGRATION_COST(迁移开销)、ENERGY_COST(能耗)和UTILIZATION(资源利用率)等定义目标函数,策略内置的全局搜索算法将根据目标函数及搜索深度生成调度指令,发送至实施模块。

[0020] 有益效果:与现有技术相比,本发明具有以下优点:

[0021] (1) 本发明将基础设施层云平台的资源调度过程进行抽象,提供了一种可编程的策略语言以及基于ECA模式的策略框架,由此可供用户根据平台及平台上部署应用的信息进行策略的定制,实现面向应用及多目标的动态资源分配。

[0022] (2) 用户可根据应用的特性定制相应的资源动态分配策略,提高资源利用率并满足用户的多管理目标。

附图说明

- [0023] 图1为本发明的系统架构示意图；
- [0024] 图2为本发明中的监控服务数据处理过程图；
- [0025] 图3为本发明中的资源调度决策过程图；
- [0026] 图4为本发明中的物理机/虚拟机选择的filter-cost机制图；
- [0027] 图5为本发明中的全局搜索算法示意图；
- [0028] 图6为实施例的流程图。

具体实施方式

[0029] 下面对本发明技术方案进行详细说明,但是本发明的保护范围不局限于所述实施例。

[0030] 如图1所示,本发明的整个系统基于ECA(event-condition-action)模式,所以必须定义事件,条件,动作等各个模块的管理方式以供用户进行策略的自定义,面向应用的IaaS层动态资源分配方法的具体步骤如下:

[0031] (1)事件服务发出事件来触发云平台的资源调度动作,产生的事件可以来自用户输入,系统环境或者策略管理器。用户输入包括向云平台请求新的资源分配(新建虚拟机等)或者对云平台的资源进行重分配(服务器整合等);系统环境包括产生的时钟信号,例如我们定义每2小时对云平台进行一次服务器整合,那么系统环境便会以2小时的频率产生事件;策略管理器在匹配到相应策略后执行相应操作也有可能产生新的事件触发资源分配行为。

[0032] (2)策略管理器收到事件服务发出的事件后在策略数据库中进行查找,匹配到具有相同事件的策略。策略管理器由策略引擎和一个策略数据库组成,用户将自定义的策略存入数据库中,策略引擎会根据相应的事件对策略进行匹配。策略即一系列事件,条件,规则的集合,当事件匹配到相应的策略后,将会判断策略定义的条件,条件包含了对监控服务提供的数据的判断,例如物理机可用内存是否大于1G,虚拟机的cpu使用率是否小于20%等。策略语言提供了&,|,!等逻辑连接符允许用户定义复杂的条件。

[0033] (3)当策略被匹配且条件满足后,策略引擎将会自动化地执行用户自定义的资源分配行为。其中,策略引擎的实现可以采用一些开源的规则引擎如drools,clips等。

[0034] 如图2所示,在纵向上将可采集的数据分为底层的平台信息和高层的应用信息,平台信息例如物理机的cpu使用信息等,应用信息可包括应用的相应事件等。当数据通过代理被采集后,将会有有一个信息处理模块对其进行处理,从时间维度上将数据横向分为历史统计数据,当前状态数据和预测数据。历史统计数据包括物理机在一段时间内的平均cpu使用率等;当前状态数据描述了当前与时间无关的状态信息,例如物理机上运行的虚拟机的个数;预测数据包括物理机的未来工作负载等。

[0035] 如图3所示,策略语言提供了虚拟机/物理机的filter-cost,随机等选择机制和可供用户定义的目标函数以及根据目标函数进行选择的全局搜索算法,资源管理器将根据用户定义的策略选择出源物理机,需要迁移的虚拟机,目的物理机供执行模块进行资源的迁移,将这样的三元组表示为(srcHost,instance,destHost)。资源调度决策的具体步骤如

下:

[0036] 步骤一,判断调度类型,分为局部调度和全局调度。其中局部调度即根据局部的物理机或虚拟机的信息进行调度,如选择cpu只用率低于某个阈值的物理机进行调度;全局调度则是根据用户的全局目标函数使用全局搜索算法进行调度。

[0037] 步骤二,如果是局部调度,则通常可分为三部分,源物理机的选取,需要调度虚拟机的选取以及目的物理机的选取,选取方式采用filter-cost以及其他内置方法。

[0038] 如图4所示,将目标备选虚拟机/物理机集合作为输入,其中filter根据硬约束条件进行过滤,cost根据软约束条件进行评价。

[0039] 硬约束即虚拟机/物理机必须满足的条件,例如选取空闲内存大于1G的物理机,使用json数据格式来描述filter的约束条件如: {freeMem: {value:1,op:lt,unit:G}},其中value代表空闲内存的值,op:lt代表大于,unit为数值单位。

[0040] 软约束是非硬性约束条件,用于对满足条件的机器按某种条件进行评价。策略语言提供了供用户定义价值函数的接口,例如需要尽可能选取cpu负载较低的物理机,本发明可以定义价值函数为cost function=-CPU_UTIL(startTime,endTime,'average'),其中cpu_util是通过监控服务采集的信息,通过上述价值函数选取一组物理机中价值函数值最高的物理机,即在给定事件段内平均cpu负载最低的物理机。

[0041] 如果是全局调度,则首先用户自定义全局目标函数,策略语言提供了用户自定义目标函数的接口,用户可使用一些预定义的关键字如MIGRATION_COST(迁移开销),ENERGY_COST(能耗),UTILIZATION(资源利用率)等定义目标函数,例如cost=MIGRATION(\$instance)+ENERGY_COST(\$pm)+UTILIZATION(\$pm)。其次,需要定义搜索的深度,资源管理器将使用全局搜索算法生成调度的三元组。

[0042] 如图5所示的全局搜索算法示意图中,每个点代表了当前的资源分配状态,每条边代表一个虚拟机的一次调度,标记为L0的点表示起始状态,标记为L1的点表示从其标记为L0的父节点经过一次调度到达的状态,标记为L2的点表示从其标记为L1的父节点经过一次调度到达的状态,以此类推。用户需要定义的搜索深度即调度的次数,标记为L1的点代表深度1,标记为L2的点代表深度2,以此类推。图5中每个点对应的状态都需要使用用户自定义的目标函数进行度量,从而从同一深度的状态中选取使得目标函数度量值最大或最小的状态点。搜索算法将根据自定义的目标函数和搜索深度进行最佳状态的搜索,并生成调度三元组(srcHost,instance,destHost)序列返回给决策服务。

[0043] 步骤三,根据局部调度或全局调度生成的调度三元组,发送信号给云平台相应的组件,实施虚拟机的迁移。

[0044] 实施例

[0045] 假设应用场景如下:两组运行MATLAB应用的虚拟机部署在多台物理机中,由于相同集群内的虚拟机互相通信需要占用大量带宽并产生通信开销,管理目标需要尽可能减少运行MATLAB虚拟机的物理机以降低能耗同时减少通信开销。

[0046] 如图6所示,具体步骤包括:

[0047] 1)、系统以两小时的频率产生调度消息,对现有平台的资源进行优化;

[0048] 2)、过滤掉运行MATLAB主节点的机器,并在运行MATLAB从节点的物理机中选取cpu负载较低的物理机作为源物理机;

- [0049] 3)、若源物理机存在,在源物理机中随机选择一台运行MATLAB的虚拟机进行调度;
- [0050] 4)、过滤在过去两小时内平均cpu使用率高于80%的物理机,在候选目的主机组中,以到MATLAB主节点所在物理机的网络传输时延为标准,选取时延最低的物理机作为目的物理机,过滤cpu是为了使该物理机的计算资源没有超过负荷以保障MATLAB的正常运行;
- [0051] 5)、将虚拟机从源物理机热迁移到目的物理机。
- [0052] 从本实施例可以看出,应用本发明的动态资源分配方法能够使得同一集群中的MATLAB从节点尽可能与主节点的通信开销最小,同时,尽可能将低MATLAB负载的物理机中的虚拟机迁入高MATLAB负载的物理机,以减少物理机的使用,降低能耗,相比现有技术考虑到了部署应用本身的特性。

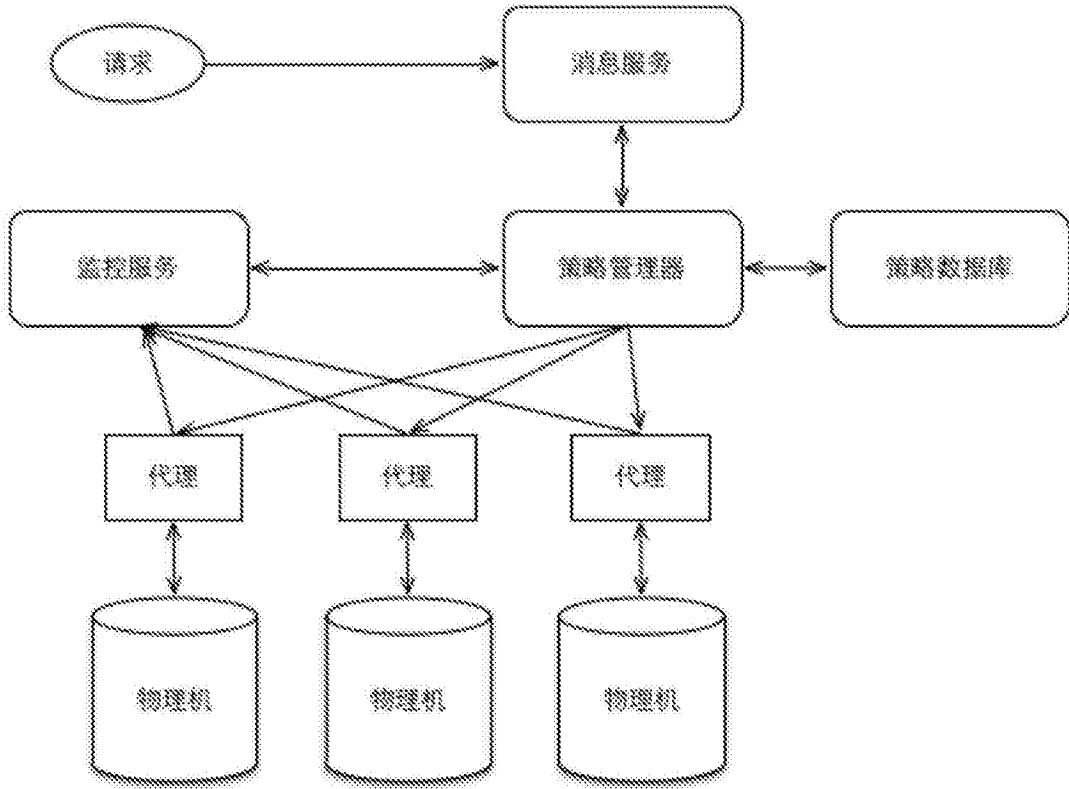


图1

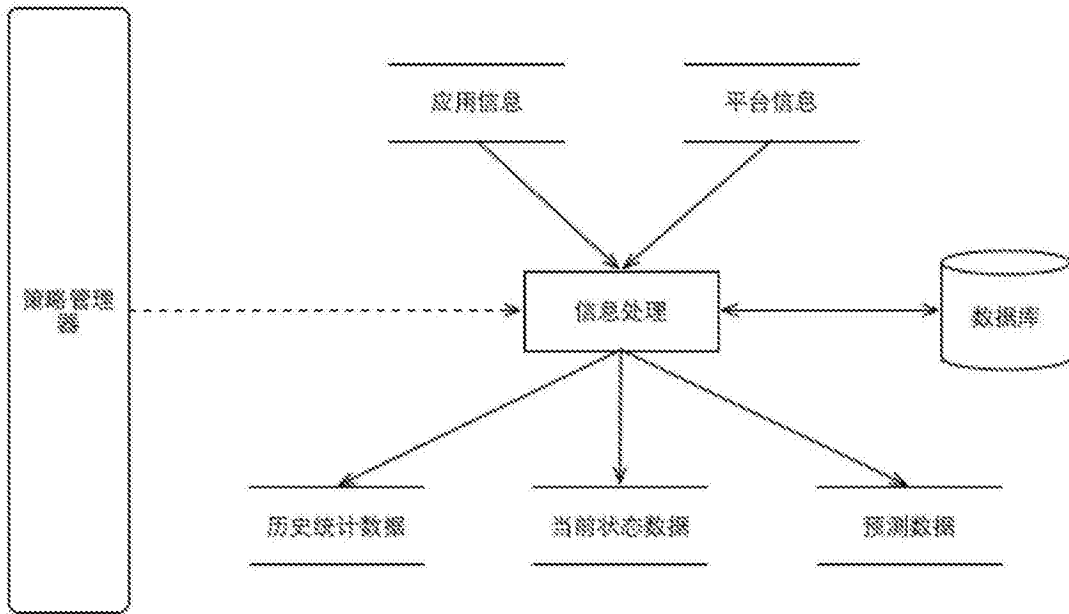


图2

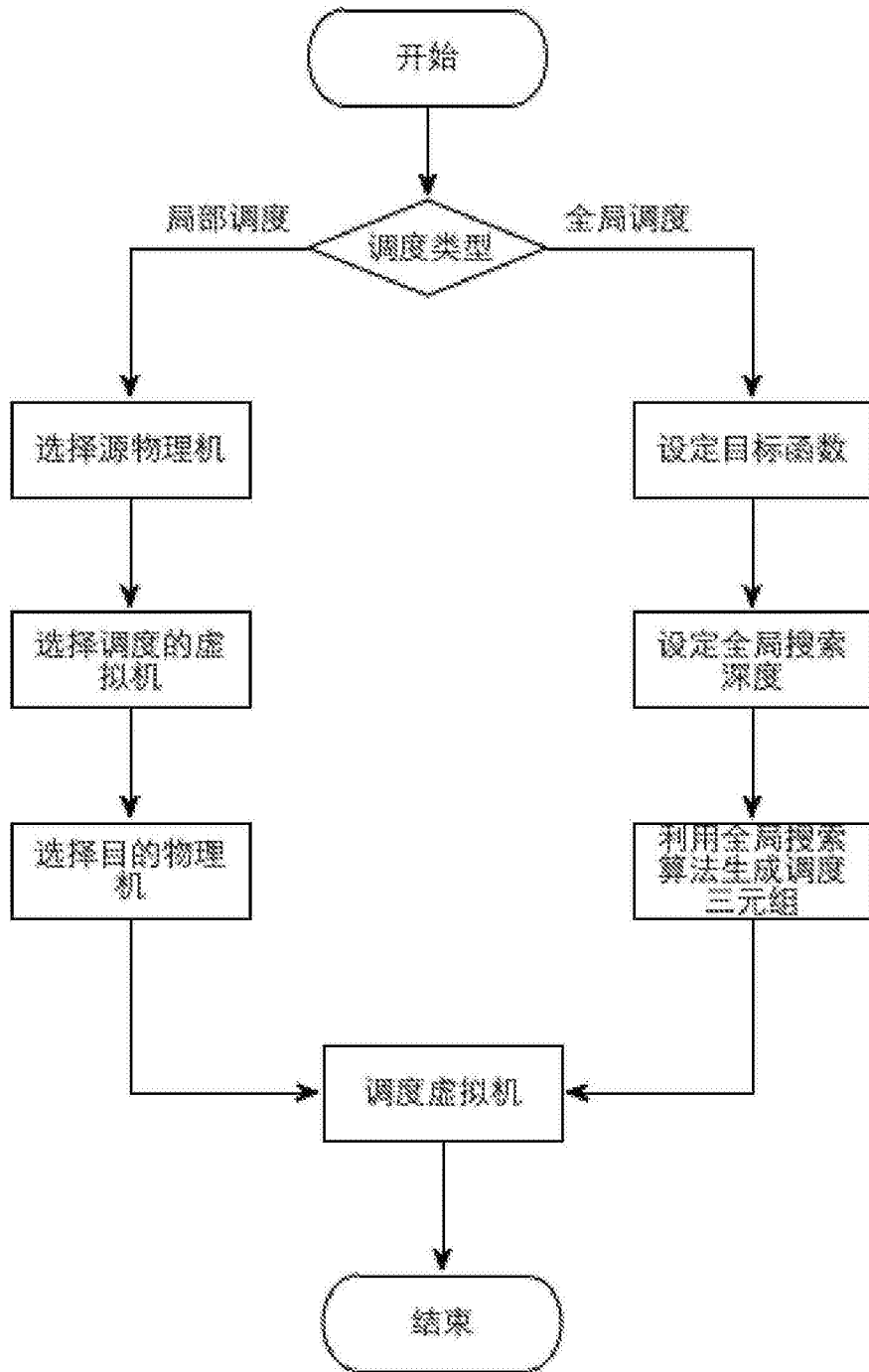


图3

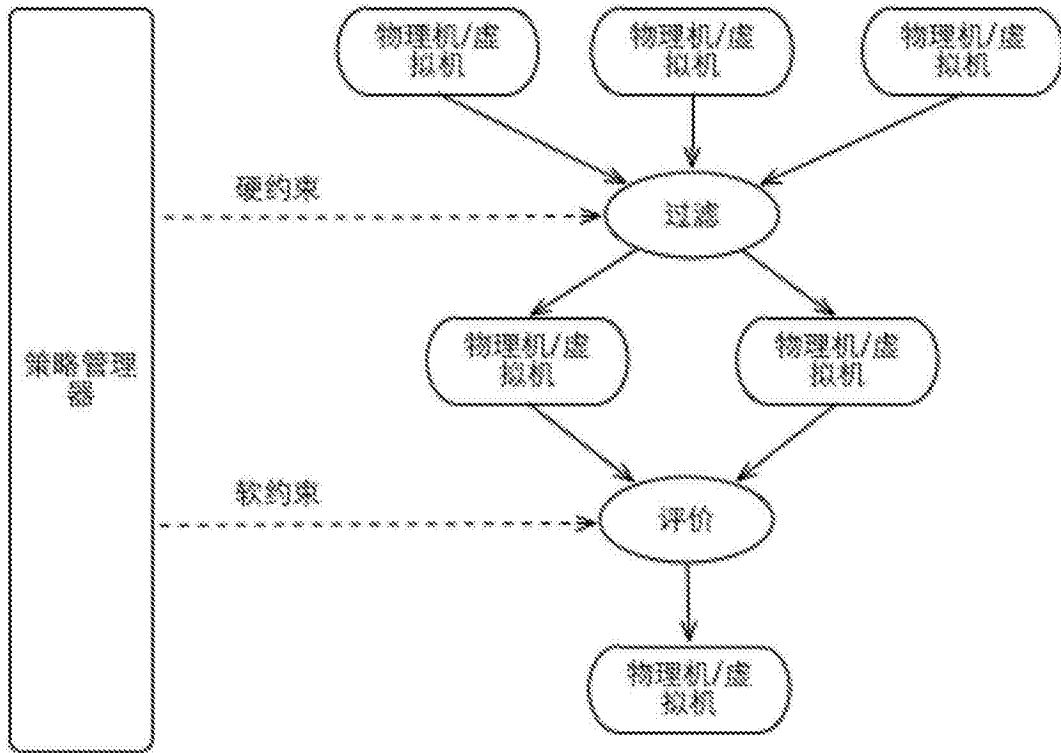


图4

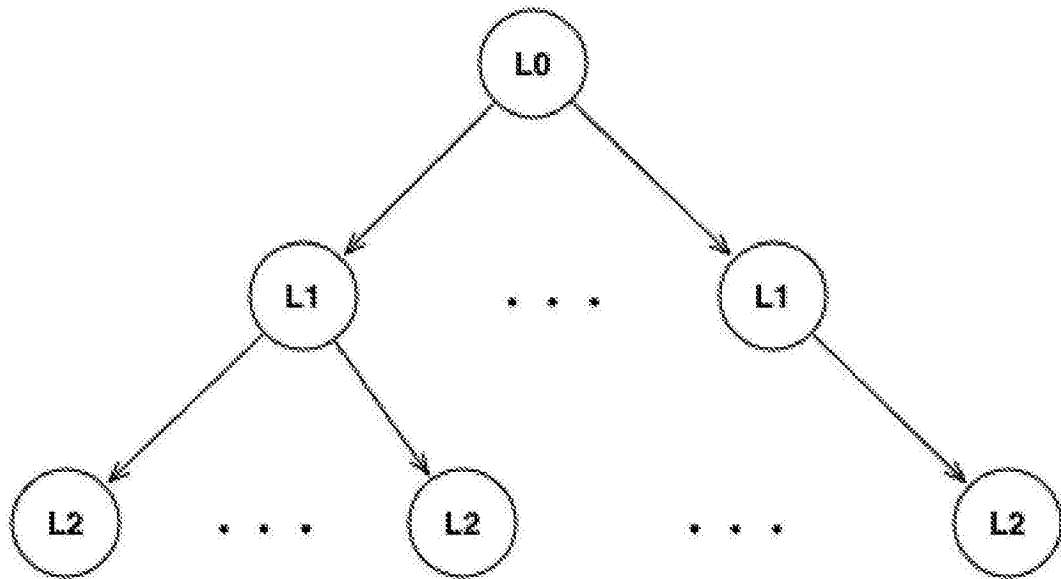


图5

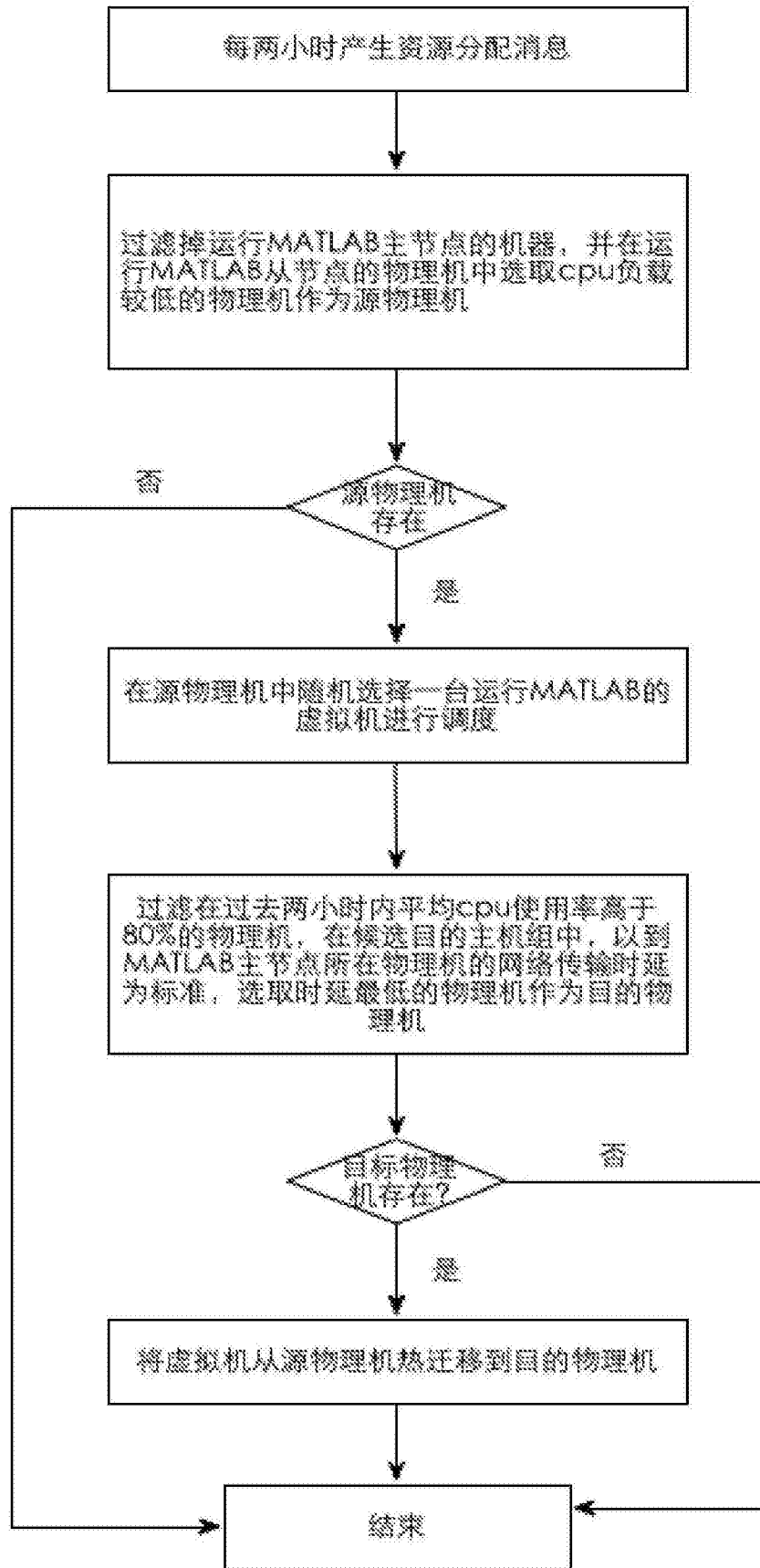


图6