

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2015 年 2 月 19 日 (19.02.2015)



(10) 国际公布号
WO 2015/021629 A1

(51) 国际专利分类号:
H04L 29/08 (2006.01)

(21) 国际申请号: PCT/CN2013/081514

(22) 国际申请日: 2013 年 8 月 15 日 (15.08.2013)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 顾炯炯 (GU, Jiongjiong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 黄朝意 (HUANG, Zhaoyi); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

(54) Title: RESOURCE ALLOCATION METHOD

(54) 发明名称: 一种资源发放方法

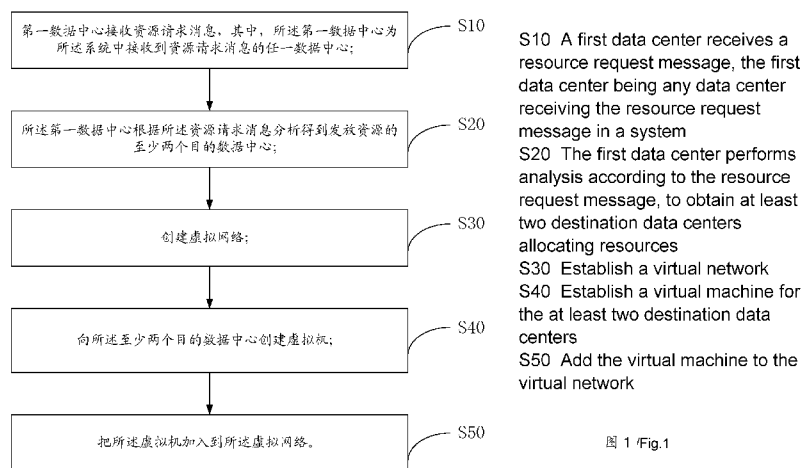


图 1 Fig.1

(57) Abstract: Provided is a resource allocation method, comprising: a first data center receiving a resource request message, the first data center being any data center receiving the resource request message in a system; the first data center performing analysis according to the resource request message, to obtain at least two destination data centers allocating resources; establishing a virtual network; establishing a virtual machine for the at least two destination data centers; and adding the virtual machine to the virtual network. The method improves the resource utilization rate of a data center, simplifies management, maintenance, and operation, and reduces the probability of network connection failure or traffic congestion.

(57) 摘要: 本发明实施例提供一种资源发放方法, 包括: 第一数据中心接收资源请求消息, 其中, 所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心; 所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心; 创建虚拟网络; 向所述至少两个目的数据中心创建虚拟机; 把所述虚拟机加入到所述虚拟网络。通过实施该方法, 提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

WO 2015/021629 A1

一种资源发放方法

技术领域

本发明涉及云计算领域，尤其涉及一种资源发放方法。

5

背景技术

传统的大型企业数据中心是一个物理分层的架构，顶层的数据中心主要负责企业级全球业务的应用集中部署，区域数据中心用来承载全球应用的镜像以及本区域范围内全局共享的应用和数据。在区域范围较大（比如覆盖数
10 个国家）时，在每个区域内还可能进一步建设服务器机房以解决对区域范围全局数据共享没有需求、但对于服务质量有更高要求的应用（如研发仿真、大型软件开发的持续构建、DHCP以及文件服务等）。对于分支机构办公室，则主要通过建设网络机房来解决便携机与个人计算机等日常办公IT设备的纯网络接入问题。

15 传统的数据中心架构面临着一系列的问题：软件应用与物理计算和存储设备耦合绑定，单个应用按照最大预期负荷资源来进行容量规划，资源利用率低，服务器资源利用率更是普遍低于15%；物理资源和应用强耦合，业务部署，扩容以及升级等管理维护运营复杂；跨数据中心的计算、存储、网络资源也不能统一拉通调度与协同管理，多数物理设备在业务量低和空闲时间内
20 处于带电空转的状态；数据中心层次过多导致数据访问延迟加大，尤其是对于处于数据中心分层架构顶端的应用，所必须经过的数据中心站点网络环节和层级过多，加大了网络连接故障或流量拥塞发生的概率。

发明内容

25 本发明实施例提供一种资源发放的方法，用以解决传统数据中心架构面临的资源利用率低、管理维护运营复杂、网络连接故障或流量拥塞发生的概率高的问题。

第一方面，本发明实施例提供一种资源发放的方法，应用于分布式数据中心系统，所述系统包括至少两个数据中心，所述方法包括：

第一数据中心接收资源请求消息，其中，所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心；

5 所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心；

创建虚拟网络；

向所述至少两个目的数据中心创建虚拟机；

把所述虚拟机加入到所述虚拟网络。

10 结合第一方面，在第一方面的第一种可能的实现方式中，所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心，具体包括：

所述第一数据中心的跨数据中心调度器接收到资源请求消息后，将所述资源请求消息转发给所述第一数据中心的智能资源路由模块，其中，所述资源请求消息中包括租户标识和资源模板标识；

15 所述智能资源路由模块根据所述租户标识和资源模板标识，获取对应的资源模板信息，所述资源模板信息至少包括网络资源规格参数和虚拟机描述参数中的一种；

20 所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心。

结合第一方面的第一种可能的实现方式，在第二种可能的实现方式中，所述智能资源路由模块根据所述租户标识和资源模板标识，获取对应的资源模板信息，具体包括：

25 所述智能资源路由模块通过查询所述第一数据中心中的镜像模板库，获取对应于所述租户标识和资源模板标识的资源模板信息。

结合第一方面的第一种可能的实现方式以及第一方面的第二种可能的实现方式，在第三种可能的实现方式中，所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心，具体包括：

所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心中的分布式数据库，获取和所述租户标识以及所述资源模板信息对应的至少两个目的数据中心。

结合第一方面的第一种可能的实现方式，在第四种可能的实现方式中，如果所述资源请求消息中包括跨数据中心资源分配策略，则所述方法进一步包括：

根据所述跨数据中心资源分配策略对所述至少两个目的数据中心进行分析，得到符合所述跨数据中心资源分配策略的目的数据中心。

结合第一方面的第一种可能的实现方式，在第五种可能的实现方式中，所述创建虚拟网络，具体包括：

所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求，所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称；

所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求；

所述第一数据中心的网络控制器根据所述租户标识和网络名称创建虚拟网络，并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟网络的信息包括网络 ID、网络状态、所述网络名称和租户标识；

所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至其他目的数据中心中的分布式数据库。

结合第一方面的第一种可能的实现方式以及第一方面的第五种可能的实现方式，在第六种可能的实现方式中，所述创建虚拟网络，进一步包括：

所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求；

所述第一数据中心的网络控制器根据所述创建虚拟子网的请求创建虚拟子网，并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库；

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目

的数据中心中的分布式数据库。

结合第一方面的第五种可能的实现方式，在第七种可能的实现方式中，所述创建虚拟网络，进一步包括：

所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络 ID、子网名称、子网无类型域间选路 CIDR、子网互联网协议 IP 地址版本号、子网网关 IP 地址、是否启动动态主机设置协议 DHCP 分配 IP 地址、DHCP IP 地址池、待创建虚拟子网的目的数据中心标识、跨数据中心子网切分模式；

所述第一数据中心的网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网，并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、虚拟可扩展本地局域网 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号；

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

结合第一方面的第七种可能的实现方式，在第八种可能的实现方式中，如果所述创建虚拟子网的请求中要求为所述虚拟子网启动 DHCP 分配 IP 地址，且需创建虚拟子网的目的数据中心不少于两个，则所述方法进一步包括：

所述第一数据中心的网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置；

所述第一数据中心的网络控制器向所述待创建虚拟子网的每个数据中心的网络总调度器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池；

所述待创建虚拟子网的每个数据中心的网络总调度器向各自所在的数据

中心中的网络控制器发送所述创建虚拟子网的请求。

结合第一方面的第八种可能的实现方式，在第九种可能的实现方式中，所述方法进一步包括：

收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求中
5 指示的虚拟子网已经存在，则在本数据中心内创建虚拟业务网关。

结合第一方面的第九种可能的实现方式，在第十种可能的实现方式中，所述在本数据中心内创建虚拟业务网关，具体包括：

所述收到所述创建虚拟子网请求的网络控制器向本数据中心的虚拟业务
网关 VSG 管理器发送网关服务请求，所述网关服务请求中携带如下参数：所
10 述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所
述 VxLAN ID 和所述分割后设置的子网网关 IP 地址；

所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

结合第一方面的第十种可能的实现方式，在第十一种可能的实现方式中，所述增加网关服务和 DHCP 服务，具体包括：

15 如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务，则在所述已有的 VSG 虚拟机中增加网卡，设置所述网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池；或者

创建新的 VSG 虚拟机，在所述新创建的 VSG 虚拟机中增加两块网卡，
20 其中一块网卡用于连接外网，从外网获取 IP 地址或者配置固定 IP 地址，另一块网卡用于连接虚拟子网，配置该网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池。

结合第一方面的第十一种可能的实现方式，在第十二种可能的实现方式中，所述方法进一步包括：

所述网络控制器接收所述 VSG 管理器返回的网关服务消息，所述消息包括 VSG 虚拟机标识；

根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网，如果所述 VSG 虚拟机未加入所述虚拟子网，则在所述虚拟子网上创建
30 一个端口，将所述 VSG 虚拟机绑定至该端口。

结合第一方面，在第十三种可能的实现方式中，所述向所述至少两个目的数据中心创建虚拟机，具体包括：

所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

- 5 结合第一方面，在第十四种可能的实现方式中，所述把所述虚拟机加入到所述虚拟网络，具体包括：

所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口；

- 10 所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息，根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口，并将所述虚拟机绑定至该端口。

结合第一方面的第十四种可能的实现方式，在第十五种可能的实现方式中，所述方法进一步包括：

- 15 所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息，确定所述虚拟机成功加入所述虚拟子网后，将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资源调度器所属数据中心中的分布式数据库；

所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

- 20 结合第一方面，在第十六种可能的实现方式中，所述方法进一步包括：

所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求，所述请求中携带如下参数：虚拟机标识、存储大小、磁盘访问服务质量 QoS 等级和虚拟机所在的主机集群；

- 25 所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源，在选定的存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存储卷，生成唯一的卷标识，并将所述卷标识返回给所述计算资源调度器；

所述计算资源调度器将所述存储卷挂载到对应于所述虚拟机标识的虚拟机上，建立所述卷标识和所述虚拟机之间的关联关系，生成唯一的关联关系

标识。

结合第一方面的第十六种可能的实现方式，在第十七种可能的实现方式中，所述方法进一步包括：

所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入所述计算资源调度器所属数据中心的分布式数据库；

所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心中的分布式数据库。

本发明实施例提供的资源发放方法，通过部署在每个数据中心中的分布式云操作系统之间的交互，执行相应的资源发放步骤，使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”，实现了资源的全局拉通，可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

15 附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例。

图 1 为本发明实施例提供的资源发放方法的流程框图；

图 2 为本发明实施例提供的资源发放方法应用于具体场景时的流程示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，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，所描述的实施例是本发明一部分实施例，而不是全部的实施例。

需要说明的是，本文中术语“和/或”，仅仅是一种描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。另外，本文中字符“/”，一般表示前

后关联对象是一种“或”的关系。

进一步需要说明的是，本文中所述的数据中心指的是计算资源、存储资源和网络资源的集合，可以是物理位置上的数据中心，也可以是逻辑概念上的数据中心。相应的，本文所述的方法既可以应用于物理位置上的数据中心，也可以应用于逻辑概念上的数据中心，只要这个数据中心中包含了计算资源、存储资源和网络资源。

本发明实施例提供的技术方案通过在每一个数据中心中，部署同样的分布式云操作系统，将企业分层管理及逐级网络互联的传统数据中心架构，改造为全扁平式管理、点到点全互联的分布式数据中心架构。在该架构中，全网各个物理数据中心节点尽管容纳的资源数量及业务类型各不相同，但从管理调度视角看可以是完全对等的关系，可以不再有主从式管理从属关系，所有基础设施资源作为一个统一的“逻辑资源池”，实现资源全局拉通，进行跨物理数据中心的资源弹性调度。解决了传统数据中心架构面临的资源利用率低、管理维护运营复杂、网络连接故障或流量拥塞发生的概率高的问题。

为了更加清楚的理解本发明实施例提供的技术方案，有必要对部署在每个数据中心中的分布式云操作系统做一简单介绍。该系统主要包括：

跨数据中心调度器：用于进行跨多个数据中心的资源发放；

智能资源路由模块：通过分析资源请求消息中的参数，得到需要申请资源的数据中心列表；

分布式数据库：负责存放分布式云数据中心的全局数据，在每一个数据中心内，都可以从分布式数据库中看到同样的数据；

计算资源调度器：负责本数据中心内计算资源的发放，挂载（attach）存储资源和加入到虚拟网络中；

存储资源调度器：负责本数据中心内存储资源的发放；

网络总调度器：负责建立全局性虚拟网络、虚拟子网、虚拟业务网关，其中所述虚拟网络包括一个或多个虚拟子网，所述虚拟子网可以分布在一个或多个数据中心；

网络控制器：负责数据中心内的软件定义网络的调度，管理虚拟交换机（virtual switch，简称为vSwitch）和端口，以及计算资源加入虚拟网络时端口与虚拟可扩展本地局域网（virtual extensible local area network，简称为

VxLAN) 的映射;

虚拟业务网关 (virtual service gateway, 简称为 VSG) 管理器: 负责创建虚拟网络的虚拟业务网关, 如路由 (Route) 服务、动态主机设置协议 (Dynamic Host Configuration Protocol, 简称为 DHCP) Server 服务、

- 5 VxLAN 网关服务、网络地址转换 (Network Address Translation, 简称为 NAT) 服务、访问控制列表 (Access Control List, 简称为 ACL) 服务、虚拟专用网络 (virtual private network, 简称为 VPN) 服务、软件负载均衡服务等中的一种或者几种服务;

- 10 镜像模板库: 提供虚拟机创建需要的镜像文件和模板服务, 模板可以描述单个虚拟机的规格要求, 也可以包含多个虚拟机甚至组网 (虚拟机之间的网络) 的描述。

基于上述分布式数据中心架构和分布式云操作系统, 本发明实施例提供了一种资源发放方法。图 1 为本发明实施例提供的资源发放方法的流程框图。如图 1 所示, 所述资源发放方法包括以下步骤:

- 15 S10、第一数据中心接收资源请求消息, 其中, 所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心;

- 20 需要说明的是, 本发明实施例提供的资源发放方法应用于分布式数据中心架构, 该架构包括两个或两个以上的物理数据中心, 且该架构中所有的物理数据中心不管容纳的资源数量及业务类型是否相同, 从管理调度视角看可以是完全对等的关系, 各数据中心之间可以不再有主从式管理从属关系。所述第一数据中心为所述分布式数据中心架构中接收到资源请求消息的任一数据中心。

S20、所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心;

- 25 具体的, 所述第一数据中心的跨数据中心调度器接收到资源请求消息后, 将所述资源请求消息转发给所述第一数据中心的智能资源路由模块, 其中, 所述资源请求消息中包括租户标识和资源模板标识; 所述智能资源路由模块根据所述租户标识和资源模板标识, 通过查询所述第一数据中心的镜像模板库, 获取和所述资源模板标识对应的资源模板信息, 所述资源模板信息包括

虚拟机描述参数列表和/或网络资源规格参数；所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心的分布式数据库，获取和所述租户标识以及所述资源模板信息对应的目的数据中心列表。

5 如果所述资源请求消息中还携带参数：指定数据中心、指定数据中心标识列表，则将所述获取的目的数据中心列表与所述指定数据中心标识列表进行取交集运算；

10 如果所述资源请求消息中还包括跨数据中心资源分配策略（如，资源平均分配、按空闲资源比例分配等），则根据所述跨数据中心资源分配策略对所述获取到的目的数据中心列表进行分析，得到符合所述跨数据中心资源分配策略的目的数据中心列表。

S30、创建虚拟网络；

需要说明的是，所述虚拟网络包括一个或多个虚拟子网，所述虚拟子网可以分布在一个或多个数据中心；

15 具体的，所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求，所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称；

所述网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求；

20 所述网络控制器根据所述租户标识和网络名称创建虚拟网络，并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟网络的信息包括网络标识（identity，简称为ID）、网络状态、所述网络名称和租户标识；

所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至所述系统中其他数据中心的分布式数据库。

25 进一步的，所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

通过所述网络总调度器向所述网络控制器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络ID、所述子网名称、子网

无类型域间选路 (Classless Inter-Domain Routing, 简称为 CIDR)、子网互联网协议 (internet protocol, 简称为 IP) 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、DHCP IP 地址池、所述待创建虚拟子网的数据中心标识列表、跨数据中心子网切分模式;

5 所述网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网(包括为虚拟子网分配全局唯一的 VxLAN ID), 并把所述虚拟子网的信息写入所述分布式数据库, 其中, 所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、VxLAN ID、所述子网 CIDR、所述 IP 地址版本号;

10 所述分布式数据库将所述虚拟子网的信息同步至所述系统中其他数据中心的分布式数据库。

进一步的, 如果所述待创建虚拟子网的数据中心标识列表中包括至少两个数据中心且要求为所述虚拟子网启动 DHCP 分配 IP 地址, 则所述网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置;

15 所述网络控制器向所述待创建虚拟子网的数据中心标识列表中的每个数据中心的网络总调度器发送创建虚拟子网的请求, 所述请求中包括下列参数: 所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、
20 分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池;

所述待创建虚拟子网的数据中心标识列表中的每个数据中心的网络总调度器向各自所在的数据中心的网络控制器发送所述创建虚拟子网的请求。

进一步的, 如果收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求的参数中指示的虚拟子网已经存在, 则在本数据中心内创建虚拟业务网关。

25 其中, 在本数据中心内创建虚拟业务网关, 具体包括:

所述收到所述创建虚拟子网请求的网络控制器向本数据中心的 VSG 管理器发送网关服务请求, 所述网关服务请求中携带如下参数: 所述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所述 VxLAN ID

和所述分割后设置的子网网关 IP 地址;

所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

其中, 所述增加网关服务和 DHCP 服务, 具体包括:

- 5 如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务, 则在所述已有的 VSG 虚拟机中增加网卡, 设置所述网卡的 IP 地址为所述创建虚拟子网请求的参数中携带的所述分割后设置的子网网关 IP 地址, 并启动 DHCP 服务, 配置分割后 DHCP IP 地址池; 或者

- 10 创建新的 VSG 虚拟机, 在所述新创建的 VSG 虚拟机中增加两块网卡, 其中一块网卡用于连接外网, 从外网获取 IP 地址或者配置固定 IP 地址, 另一块网卡用于连接虚拟子网, 配置该网卡的 IP 地址为所述创建虚拟子网请求的参数中携带的所述分割后设置的子网网关 IP 地址, 并启动 DHCP 服务, 配置分割后 DHCP IP 地址池。

进一步的, 所述网络控制器接收所述 VSG 管理器返回的网关服务消息, 所述消息包括 VSG 虚拟机标识;

- 15 根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网, 如果所述 VSG 虚拟机未加入所述虚拟子网, 则在所述虚拟子网上创建一个端口, 将所述 VSG 虚拟机绑定至该端口。

S40、向所述至少两个目的数据中心创建虚拟机;

- 20 具体的, 所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

S50、把所述虚拟机加入到所述虚拟网络。

具体的,

- 25 所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口;

所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息, 根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口, 并将所述虚拟机绑定至该端口;

所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息，确定所述虚拟机成功加入所述虚拟子网后，将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资源调度器所属数据中心中的分布式数据库；

所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

进一步的，还可以为所述虚拟机分配存储资源，具体包括：

所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求，所述请求中携带如下参数：虚拟机标识、存储大小、磁盘访问服务质量（Quality of Service，简称为 QoS）等级和虚拟机所在的主机集群；

所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源，在选定的存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存储卷，生成唯一的卷标识，并将所述卷标识返回给所述计算资源调度器；

所述计算资源调度器把所述存储卷挂载到虚拟机标识所代表的虚拟机，建立卷标识和虚拟机之间的关联关系，生成唯一的关联关系标识；

所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入本数据中心的分布式数据库；

所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心的分布式数据库。

本发明实施例提供的资源发放方法，通过部署在每个数据中心中的分布式云操作系统之间的交互，执行相应的资源发放步骤，使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”，实现了资源的全局拉通，可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

为了更清楚地理解本发明实施例的技术方案，下面将结合具体的应用场景对本发明实施例的技术方案做进一步的说明。

图 2 为本发明实施例提供的资源发放方法应用于具体场景时的流程示意图。在图 2 中，有三个数据中心，其中数据中心 1 收到来自租户的资源请求。需要说明的是，在实际应用中，任何一个数据中心都可以接受租户的资源请

求，所有数据中心的功能都是对等的。如图 2 所示，具体的资源发放流程包括以下步骤：

(1) 数据中心 1 的跨数据中心调度器接收资源请求消息，消息参数带有租户标识、资源模板标识、是否指定数据中心、指定数据中心标识列表以及

5 跨数据中心资源分配策略（如，资源平均分配、按空闲资源比例分配等）。

(2) 所述跨数据中心调度器收到请求消息后，把该请求消息转发到数据中心 1 的智能资源路由模块，请求进行参数分析。

(3) 所述智能资源路由模块向数据中心 1 的镜像模板库查询资源模板标识对应的资源模板中需要分配的具体资源要求。

10 (4) 所述镜像模板库向所述智能资源路由模块返回需要分配资源的具体参数要求：

虚拟机描述参数列表：描述参数列表中的每一项都包含一个虚拟机的创建要求，包括虚拟机分配数量、虚拟机规格标识（通过虚拟机规格标识可以获取得到虚拟机的规格信息如虚拟机中央处理器（central processing unit，
15 简称为 CPU）主频、CPU 个数、内存大小、磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级（用于选择磁盘存储类型：固态硬盘、直连硬盘、外置存储等）、每个磁盘备份模式、虚拟网卡个数、每个虚拟网卡的带宽）、虚拟机镜像链接、物理计算资源还是虚拟计算资源、CPU 类型（如果是物理计算资源，指定物理机的 CPU 类型）、HyperVisor 类型 CPU 类型（如果是
20 请求虚拟计算资源，指定 HyperVisor 类型）、是否启动高可用性（Hight Availability，简称为 HA）、是否启动故障冗余（fault tolerance，简称为 FT）、是否需要容灾、容灾类型（同步容灾、异步容灾）、容灾模式（双活、主备模式）、虚拟机启动参数、虚拟机配置参数、安全组、其他参数。

网络资源规格参数：网络名称、子网名称、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、DHCP IP 地址池、
25 外网 IP 地址列表、子网 QoS 参数（包括网络带宽下限、网络带宽上限、网络抖动上限、网络抖动下限、网络时延上限、网络时延下限、网络丢包率上限、网络丢包率下限）、跨数据中心子网切分模式（模式 1：均分 IP 地址，

每个数据中心部署网关；模式 2：单网关单 IP 地址池）。

资源请求消息的参数中可以包含虚拟机描述参数列表和/或网络资源规格参数中的一类参数或者几类参数。当虚拟机描述参数列表中包含容灾相关参数时（容灾类型、容灾模式），一般同时会包含网络资源规格参数中的子

5 网 QoS 参数，以确定容灾的网络要求。

（5）所述智能资源路由模块根据租户标识，查询数据中心 1 的分布式数据库，通过租户标识获得可以申请资源的数据中心列表集合，如果“是否指定数据中心”参数为“是”，则和“指定数据中心标识列表”进行取交集运算，得到初步分析后的数据中心列表集合。

10 智能资源路由模板根据镜像模板库返回的虚拟机描述参数列表、网络资源规格参数等，以及初步分析获得的数据中心列表集合、跨数据中心资源分配策略（如，资源平均分配、按空闲资源比例分配等）进行目的数据中心列表的分析：

首先，智能资源路由模板从数据中心列表集合中取出第一个数据中心列表，根据虚拟机描述参数列表中的虚拟机具体要求和跨数据中心资源分配策略，检查列表中的每个数据中心是否有满足条件的计算和存储资源。如果有足够的计算资源和存储资源，检查是否有网络资源规格参数，如果有网络资源规格参数要求，则分析数据中心间和数据中心内的网络是否满足子网 QoS 参数的要求。根据计算资源、存储资源和网络资源要求，可以分析出虚拟机分配时是否需要在特定的主机集群创建。如果数据中心列表的数据中心满足计算资源、存储资源以及数据中心间的网络参数要求，则该数据中心列表为符合要求的数据中心列表。如果计算资源、存储资源或者网络资源规格参数有任意一个参数不满足要求，则认为该数据中心列表不是符合要求的数据中心列表。

25 重复上述计算资源、存储资源和网络资源的分析，直到数据中心列表集合中的所有数据中心列表都分析完成为止，对符合要求的数据中心列表集合按照满足程度多少从高到低进行排序，得到符合资源发放要求的目的数据中心列表集合。该目的数据中心列表集合中每一个数据中心列表包含如下信息[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]]，其中，数据中

心标识和虚拟机描述参数列表是“和”的关系，[网络资源规格参数]和[数据中心标识,虚拟机描述参数列表]是“和/或”的关系。

5 (6) 智能资源路由模块向跨数据中心调度器返回目的数据中心列表集合，集合中的每一个列表包含[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]]这些信息。

(7) 跨数据中心调度器取出上述目的数据中心列表集合中第一个数据中心列表，启动资源发放事务。以下第(8)步到第(49)步为一个完整的资源发放事务。在整个资源发放事务中，如果有一个资源发放失败，则整个事务进行回滚，事务中已经发放的资源都被完全释放。

10 (8) 数据中心 1 的 跨数据中心调度器检查数据中心列表中的[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]]的信息。如果包含网络资源规格参数，则所述跨数据中心调度器判断在租户标识下网络名称所代表的虚拟网络是否存在，如果不存在，则向其所在数据中心的网络总调度器发出创建虚拟网络的请求，参数带有租户标识、网络名称。

15 (9) 数据中心 1 的网络总调度器判断请求为创建虚拟网络，则向本数据中心的网络控制器发出创建虚拟网络的请求。

(10) 所述网络控制器 根据租户标识、网络名称参数为该租户创建一个虚拟网络，生成唯一网络标识：网络 ID，初始化网络状态为激活状态。把新创建虚拟网络的网络 ID、网络名称、租户标识和网络状态写入分布式数据库，
20 同时要求分布式数据库同步该虚拟网络信息到所有数据中心。

(11) 所述分布式数据库向其他数据中心（图 2 中为数据中心 2 和数据中心 3）的分布式数据库更新新创建的虚拟网络信息,包括租户标识、网络 ID、网络名称、网络状态。

25 (12) 所述分布式数据库向数据中心 1 的网络控制器返回数据库更新结果。

(13) 如果数据库更新成功，则所述网络控制器向所述网络总调度器返回新创建虚拟网络的信息，包括租户标识、网络 ID、网络名称、网络状态。如果数据库更新失败，则向网络总调度器返回创建虚拟网络失败。

(14) 所述网络总调度器向所述跨数据中心调度器返回新创建虚拟网络的信息。

(15) 所述跨数据中心调度器检查虚拟网络的信息, 如果虚拟网络创建失败, 则终止流程。如果虚拟网络信息成功增加到数据库, 则检查[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]], 根据租户标识、网络名称和子网名称判断子网名称所代表的子网是否存在。如果不存在, 需要创建子网, 分析参数[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]]获得待创建虚拟子网的数据中心列表。当数据中心列表中的数据中心个数大于1时, 表明需要在数据中心列表中的所有数据中心创建虚拟子网, 否则表示仅仅在指定的这个数据中心创建虚拟子网。如果仅仅在指定的这个数据中心(图2中为数据中心1)创建虚拟子网, 则所述跨数据中心调度器向数据中心1的网络总调度器发出创建虚拟子网的请求, 请求中携带参数租户标识、网络ID、子网名称、子网CIDR、子网IP地址版本号、子网网关IP地址、是否启动DHCP分配IP地址、数据中心标识列表、跨数据中心子网切分模式(模式1: 均分IP地址, 每个数据中心部署网关; 模式2: 均分IP地址, 单网关模式)。

(16) 所述网络总调度器向其所在数据中心的网络控制器发出创建虚拟子网的请求, 请求中携带参数租户标识、网络ID、子网名称、子网CIDR、子网IP地址版本号、子网网关IP地址、是否启动DHCP分配IP地址、DHCP IP地址池、数据中心标识列表、跨数据中心子网切分模式(模式1: 均分IP地址, 每个数据中心部署网关; 模式2: 单网关单IP地址池)。

(17) 所述网络控制器根据租户标识、网络ID、子网名称参数为该租户创建一个虚拟子网, 生成唯一子网标识: 子网ID, 申请全局唯一的VxLAN ID, 设置子网属性为参数中的子网CIDR、子网IP地址版本号。所述网络控制器把虚拟子网的信息: 租户标识、网络ID、子网名称、子网ID、VxLAN ID、子网CIDR、子网IP地址版本号写入数据中心1的分布式数据库中。同时要求分布式数据库同步虚拟子网信息到所有数据中心。

(18) 数据中心1的分布式数据库向其他数据中心(图2中为数据中心2和数据中心3)的分布式数据库更新新创建的虚拟子网信息, 包括租户标识、

网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号。

(19) 数据中心 1 的分布式数据库向本数据中心的网络控制器返回数据库更新结果。

(20) 数据中心 1 的网络控制器继续分析虚拟子网的数据中心标识列表和跨数据中心子网切分模式(模式 1: 均分 IP 地址, 每个数据中心部署网关; 模式 2: 单网关单 IP 地址池)。如果数据中心标识列表大于 1 个数据中心且参数要求为子网启动 DHCP 分配 IP 地址, 则对待分配的 DHCP IP 地址池进行按照跨数据中心子网切分模式(模式 1: 均分 IP 地址, 每个数据中心部署网关; 模式 2: 单网关单 IP 地址池)进行 IP 地址切分和网关 IP 地址设置。

10 在模式 1 下平均分割 IP 地址并设置不同的子网网关 IP 地址: 假如 CIDR 为 192.168.0.0/24, IP 地址版本号为 IPV4, DHCP IP 地址池为 192.168.0.1 ~ 192.168.0.254, 同时数据中心标识列表包含数据中心 2 和数据中心 3, 则该子网在数据中心 2 的 DHCP IP 地址池是 192.168.0.2 ~ 192.168.0.127, 网关 IP 地址是 192.168.0.1; 该子网在数据中心 3 的 DHCP IP 地址池是 192.168.0.129 ~ 192.168.0.254, 网关 IP 地址是 192.168.0.128。在模式 2 下单网关单 IP 地址池, 不分割 IP 地址, 但是对第一个数据中心设置子网网关 IP 地址: 假如 CIDR 为 192.168.0.0/24, IP 地址版本号为 IPV4, DHCP IP 地址池为 192.168.0.1 ~ 192.168.0.254, 同时数据中心标识列表包含数据中心 2 和数据中心 3, 则该子网在数据中心 2 的 DHCP IP 地址池是 192.168.0.2 ~ 192.168.0.254, 网关 IP 地址是 192.168.0.1; 该子网在数据中心 3 的 DHCP IP 地址池是 192.168.0.2 ~ 192.168.0.254, 是否启动 DHCP 服务设置为否, 网关 IP 为无效 IP 地址。向数据中心标识列表中的每一个数据中心(图 2 中为数据中心 2 和数据中心 3)的网络总调度器发出创建虚拟子网的请求, 请求中携带参数: 租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号、是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址、分割后的 DHCP IP 地址池。

如果数据中心标识列表只包含一个数据中心标识, 则不进行 DHCP IP 地址池的分割和子网网关 IP 地址的重新设置。

(21) 数据中心 2 和数据中心 3 的网络总调度器分别向各自所在数据中

心的网络控制器发出创建虚拟子网的请求，请求中携带参数：租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号、是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址、分割后的 DHCP IP 地址池。

- 5 需要说明的是：由于在数据中心 2 和数据中心 3 中创建虚拟子网的过程相同，为了描述方便，以下的步骤（22）-步骤（28）仅以数据中心 2 为例进行阐述。

10 （22）数据中心 2 的网络控制器判断创建虚拟子网请求参数中虚拟子网已经存在，且无数据中心列表，则在本数据中心内继续虚拟子网虚拟业务网关创建工作。如果子网网关 IP 地址有效，表明需要申请网关服务，所述网络控制器向本数据中心的 VSG 管理器发送网关服务请求，参数中携带子网 ID、是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、VxLAN ID、分割后设置的子网网关 IP 地址。

15 （23）所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。增加网关服务和 DHCP 服务有两种方式：

20 方式一：如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务（如，Route 服务、DHCP Server 服务、VxLAN 网关服务、NAT 服务、ACL 服务、VPN 服务、软件负载均衡服务中的一种服务或者几种服务），则在该已有的 VSG 虚拟机中增加网卡，设置该网卡的 IP 地址为创建虚拟子网请求参数中的网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池；

25 方式二：新创建 VSG 虚拟机，在该虚拟机中增加两块网卡，一块网卡用于连接外网，从外网获取 IP 地址或者配置固定 IP 地址，另一块网卡用于连接虚拟子网，配置该网卡的 IP 地址为创建虚拟子网请求参数中的网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池。

 VSG 管理器建立虚拟子网 ID 和新增加 VSG 虚拟机、网关服务及网卡的映射关系。VSG 管理器向本数据中心的网络控制器返回网关服务请求的结果，包括 VSG 虚拟机标识、新增加网卡的 MAC 地址。

(24) 所述网络控制器收到 VSG 管理器返回的网关服务消息, 判断返回的 VSG 虚拟机是否已经加入虚拟子网。如果 VSG 虚拟机未加入虚拟子网, 则在虚拟子网上创建一个端口: 所述网络控制器向 VSG 虚拟机所在的主机下发 VSG 虚拟机绑定到虚拟子网的命令, 携带参数包括: 租户标识、虚拟网络 ID、虚拟子网 ID、VSG 虚拟机的 MAC 地址、端口状态为激活态、网关 IP 地址、VSG 虚拟机标识。VSG 虚拟机所在的主机部署了网络控制器的代理, 该网络控制器的代理接收到命令后负责在该主机上管理虚拟交换机 (virtual Switch, 简称为 vSwitch), 在 vSwitch 上申请一个端口, 并绑定 VSG 虚拟机的 MAC 地址, 打上虚拟子网 VxLAN 的标识, 并对这个端口分配一个唯一的端口标识: 端口 ID。成功绑定虚拟机后, 所述网络控制器把端口信息写入本地数据库中。

(25) 所述网络控制器判断是否开启 DHCP 服务参数, 如果需要申请 DHCP 服务, 则所述网络控制器向本数据中心的 VSG 管理器发送 DHCP 服务请求, 参数中携带子网 ID、分割后的 DHCP IP 地址池。

(26) 所述 VSG 管理器根据子网 ID 获取该子网对应的 VSG 虚拟机。如果不能够获得 VSG 虚拟机, 表明需要为该虚拟子网分配 VSG 虚拟机提供 DHCP 服务。申请 VSG 虚拟机有如下两种方式:

方式一: 如果已经存在的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务 (如, Route 服务、DHCP Server 服务、VxLAN 网关服务、NAT 服务、ACL 服务、VPN 服务、软件负载均衡服务中的一种服务或者几种服务), 则在该已有的 VSG 虚拟机中增加网卡, 设置该网卡的 IP 地址为 IP 地址池的第一个 IP 地址。

方式二: 新创建 VSG 虚拟机, 在该虚拟机中增加一网卡, 该网卡用于连接虚拟子网, 配置该网卡的 IP 地址为 IP 地址池的第一个 IP 地址。

在该 VSG 虚拟机下启动 DHCP 服务, 设置该 DHCP 服务可分配的 IP 地址为分割后的 DHCP IP 地址池, DHCP 服务绑定虚拟子网的网关 IP 地址, 为该虚拟子网提供 DHCP 服务。所述 VSG 管理器启动 DHCP 服务。所述 VSG 管理器向所述网络控制器返回 DHCP 服务请求的结果, 包括 VSG 虚拟机标识、新增加网卡的 MAC 地址。

(27) 所述网络控制器收到返回结果，并判断 VSG 虚拟机是否已经加入虚拟子网，如果没有加入虚拟子网，则进行第 (24) 步 VSG 虚拟机加入虚拟子网的操作。之后，所述网络控制器向本数据中心的网络总调度器返回虚拟子网信息。

5 (28) 所述网络总调度器向数据中心 1 的网络控制器返回虚拟子网信息。

同理，对于数据中心 3 来说，执行完上述第 (22) 步-第 (27) 步后，数据中心 3 的网络总调度器也向数据中心 1 的网络控制器返回虚拟子网创建结果。

10 (29) 数据中心 1 的网络控制器向数据中心 1 的网络总调度器返回虚拟子网信息。

(30) 数据中心 1 的网络总调度器向数据中心 1 的跨数据中心调度器返回虚拟子网信息，包括：租户标识、网络 ID、子网名称、子网 ID、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、IP 地址池、VxLAN ID 以及跨数据中心子网切分模式。

15 (31) 数据中心 1 的跨数据中心调度器检查数据中心列表中[[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]]中的虚拟机信息，向每一个需要创建虚拟机的数据中心下发虚拟机创建任务，这些任务是并行下发的。当有多个不同规格的虚拟机需要创建的时候，所述跨数据中心调度器分别下发多个创建虚拟机的任务。当需要批量创建相同规格的虚拟机时，调用批量

20 创建的接口，带上需创建虚拟机的数量。图 2 中示例为向数据中心 2 和数据中心 3 中的计算资源调度器下发单个虚拟机创建任务，携带参数包括：虚拟机规格标识（通过虚拟机规格标识可以获得虚拟机的规格信息，如虚拟机 CPU 主频、CPU 个数、内存大小、磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级(用于选择磁盘存储类型：固态硬盘、直连硬盘、外置存储等)、

25 每个磁盘备份模式、虚拟网卡个数、每个虚拟网卡的带宽）、虚拟机镜像链接、物理计算资源还是虚拟计算资源、CPU 类型（如果是物理计算资源，指定物理机的 CPU 类型）、HyperVisor 类型 CPU 类型（如果是请求虚拟计算资源，指定 HyperVisor 类型）、是否启动 HA、是否启动 FT、是否需要容灾、容灾类型（同步容灾、异步容灾）、容灾模式（双活、主备模式）、虚拟机

启动参数、虚拟机配置参数、安全组、指定主机集群、待加入的虚拟子网标识列表、其他参数。

需要说明的是：由于在数据中心 2 和数据中心 3 中创建虚拟机并加入虚拟子网的过程相同，为了描述方便，以下的步骤（32）-步骤（49）仅以数据
5 中心 2 为例进行阐述。

（32）数据中心 2 的计算资源调度器创建虚拟机逻辑对象，为虚拟机创建唯一标识，记录虚拟机的参数信息。

（33）所述计算资源调度器把虚拟机信息写入数据中心 2 的分布式数据库，不需要同步到其他数据中心。

10 （34）所述分布式数据库向所述计算资源调度器返回数据库更新结果。

（35）所述计算资源调度器向数据中心 1 的跨数据中心调度器返回虚拟机标识。

（36）所述数据中心 1 的跨数据中心调度器向数据中心 2 的计算资源调度器轮询虚拟机的创建状态，携带参数为虚拟机标识，直到虚拟机创建成功
15 后结束虚拟机创建任务。

需要说明的是，所述跨数据中心调度器可以周期性的向数据中心 2 和 3 的计算资源调度器轮询虚拟机的创建状态，直至本次资源发放事务结束。

（37）数据中心 2 的计算资源调度器对计算资源进行调度：根据虚拟机规格标识获得虚拟机的规格：虚拟机 CPU 主频、CPU 个数、内存大小、磁盘
20 个数、每个磁盘的存储大小，判断虚拟机建立的资源条件：磁盘访问 QoS 等级，物理计算资源还是虚拟计算资源、CPU 类型（如果是物理计算资源，指定物理机的 CPU 类型）、HyperVisor 类型 CPU 类型（如果是请求虚拟计算资源，指定 HyperVisor 类型），选定满足所有条件的主机。如果参数中指定主机集群，则在指定的主机集群中选择一个主机。在选中的主机上，根据
25 参数的镜像链接，从数据中心 2 的镜像模板库中获得镜像，创建虚拟机。

虚拟机根据虚拟机启动参数、虚拟机配置参数进行虚拟机的启动和配置。

（38）数据中心 2 的计算资源调度器逐一从待加入的虚拟子网标识列表

中取出虚拟子网 ID，向本数据中心的网络总调度器发出创建端口的请求，请求中携带参数：网络 ID、子网 ID、虚拟机标识、加入该虚拟子网的虚拟机网卡的 MAC 地址及端口初始状态（激活态）。

5 (39) 所述网络总调度器把创建端口的请求消息转发给数据中心 2 的网络控制器。

(40) 所述网络控制器根据子网 ID 获得虚拟子网的信息：租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号。然后向虚拟机所在的主机下发虚拟机绑定到虚拟子网的命令，携带参数包括：租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、端口状态。

10 虚拟机所在的主机上部署了网络的代理，该网络的代理接收到命令后负责在该主机上管理 vSwitch，在 vSwitch 上申请一个端口，并绑定虚拟机的 MAC 地址，打上虚拟子网 VxLAN 的标识，虚拟机加入到 VxLAN 二层网络后，从该虚拟子网的 DHCP 服务器中获得 IP 地址。所述网络控制器对这个端口分配一个唯一的端口标识：端口 ID。成功绑定虚拟机后，所述网络控制器把端口信息写入数据库中。

15 所述网络控制器向本数据中心的网络总调度器返回端口信息：端口 ID，租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、虚拟机 IP 地址，端口状态。

(41) 所述网络总调度器向本数据中心的计算资源调度器返回虚拟机加入到虚拟子网的端口信息：端口 ID，租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、虚拟机 IP 地址，端口状态。

20

(42) 所述计算资源调度器获得虚拟机加入到虚拟子网的端口信息，判断虚拟机成功加入到虚拟子网后，把虚拟机 ID、端口 ID 之间的映射关系写入数据中心 2 的分布式数据库，并同步到其他数据中心。

(43) 数据中心 2 的分布式数据库向本数据中心的计算资源调度器返回数据库更新结果。

25

(44) 所述计算资源调度器根据磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级向数据中心 2 的存储资源调度器发出创建存储资源的请求，请求中携带参数：虚拟机标识、存储大小、磁盘访问 QoS 等级、虚拟机所在的

主机集群。

(45) 所述存储资源调度器根据参数：虚拟机标识、存储大小、磁盘访问 QoS 等级及虚拟机所在的主机集群，选择具体的磁盘存储类型：固态硬盘、直连硬盘、外置存储等。在选定的磁盘上创建满足磁盘访问 QoS 等级、存储大小 5 的卷。生成唯一的卷标识：卷 ID。向所述计算资源调度器返回卷标识。

(46) 所述计算资源调度器把卷挂载到虚拟机标识所代表的虚拟机，建立卷标识和虚拟机之间的关联关系，生成唯一的关联关系标识。

(47) 如果成功挂载了卷，所述计算资源调度器把虚拟机 ID、卷标识和关联关系标识写入数据中心 2 的分布式数据库，并同步到其他数据中心。

10 (48) 所述分布式数据库向所述计算资源调度器返回数据库更新结果。如果还需要创建其他卷，则重复第 (45) - (47) 步，直到所有卷创建挂载成功。

(49) 当数据中心 1 的跨数据中心调度器通过轮询虚拟机创建任务执行结果，获知数据中心 2 和 3 分别完成上述资源发放事务后，数据中心 1 的跨 15 数据中心调度器结束第一个数据中心列表的资源发放事务，如果在事务中任何一步资源申请失败，则释放本事务中所有分配的资源。然后选取数据中心列表集合中的次优数据中心列表，重复资源发放事务，重复次数以系统配置为上限。

(50) 数据中心 1 的跨数据中心调度器返回资源发放结果，如果成功， 20 返回发放的计算资源、存储资源和网络资源。

本发明实施例提供的资源发放方法，通过部署在每个数据中心中的分布式云操作系统之间的交互，执行相应的资源发放步骤，使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”，实现了资源的全局拉通，可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、 25 简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

应理解，在本发明实施例中，“与 A 相应的 B”表示 B 与 A 相关联，根据 A 可以确定 B。但还应理解，根据 A 确定 B 并不意味着仅仅根据 A 确定 B，还可以根据 A 和/或其它信息确定 B。

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

5 示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。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。

10 在本申请所提供的实施例中，应该理解到，所揭露的系统，可以通过其它的方式实现。例如，以上所描述的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。另外，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，也可以是电的，机械的或其它的形式连接。

15 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发明实施例方案的目的。

20 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，也可以是两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

25 所述集成的单元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，或者该技术方案的全部或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、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

30 以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以权利要求的保护范围为准。

权利要求书

1、一种资源发放方法，应用于分布式数据中心系统，所述系统包括至少两个数据中心，其特征在于，所述方法包括：

5 第一数据中心接收资源请求消息，其中，所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心；

所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心；

创建虚拟网络；

向所述至少两个目的数据中心创建虚拟机；

10 把所述虚拟机加入到所述虚拟网络。

2、根据权利要求1所述的资源发放方法，其特征在于，所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心，具体包括：

15 所述第一数据中心的跨数据中心调度器接收到资源请求消息后，将所述资源请求消息转发给所述第一数据中心的智能资源路由模块，其中，所述资源请求消息中包括租户标识和资源模板标识；

所述智能资源路由模块根据所述租户标识和资源模板标识，获取对应的资源模板信息，所述资源模板信息至少包括网络资源规格参数和虚拟机描述参数中的一种；

20 所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心。

3、根据权利要求2所述的资源发放方法，其特征在于，所述智能资源路由模块根据所述租户标识和资源模板标识，获取对应的资源模板信息，具体包括：

25 所述智能资源路由模块通过查询所述第一数据中心中的镜像模板库，获取对应于所述租户标识和资源模板标识的资源模板信息。

4、根据权利要求2或3所述的资源发放方法，其特征在于，所述智能资

源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心，具体包括：

所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心中的分布式数据库，获取和所述租户标识以及所述资源模板信息对应的至少两个目的数据中心。

5 5、根据权利要求 1 所述的资源发放方法，其特征在于，如果所述资源请求消息中包括跨数据中心资源分配策略，则所述方法进一步包括：

根据所述跨数据中心资源分配策略对所述至少两个目的数据中心进行分析，得到符合所述跨数据中心资源分配策略的目的数据中心。

10 6、根据权利要求 2 所述的资源发放方法，其特征在于，所述创建虚拟网络，具体包括：

所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求，所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称；

15 所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求；

所述第一数据中心的网络控制器根据所述租户标识和网络名称创建虚拟网络，并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟网络的信息包括网络 ID、网络状态、所述网络名称和租户标识；

20 所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至其他目的数据中心中的分布式数据库。

7、根据权利要求 2 或 6 所述的资源发放方法，所述创建虚拟网络进一步包括：

25 所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求；

所述第一数据中心的网络控制器根据所述创建虚拟子网的请求创建虚拟子网，并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库；

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

8、根据权利要求 6 所述的资源发放方法，其特征在于，所述创建虚拟网络，进一步包括：

- 5 所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络 ID、子网名称、子网无类型域间选路 CIDR、子网互联网协议 IP 地址版本号、子网网关 IP 地址、是否启动动态主机设置协议 DHCP 分配 IP 地址、DHCP IP 地址池、待创建虚拟子网的目的数据中心标识、跨数据中心子网切分模式；

10

所述第一数据中心的网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网，并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、虚拟可扩展本地局域网 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号；

15

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

- 20 9、根据权利要求 8 所述的资源发放方法，其特征在于，如果所述创建虚拟子网的请求中要求为所述虚拟子网启动 DHCP 分配 IP 地址，且需创建虚拟子网的目的数据中心不少于两个，则所述方法进一步包括：

所述第一数据中心的网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置；

- 25 所述第一数据中心的网络控制器向所述待创建虚拟子网的每个数据中心的网络总调度器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池；

- 30 所述待创建虚拟子网的每个数据中心的网络总调度器向各自所在的数据

中心中的网络控制器发送所述创建虚拟子网的请求。

10、根据权利要求 9 所述的资源发放方法，其特征在于，所述方法进一步包括：

收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求中
5 指示的虚拟子网已经存在，则在本数据中心内创建虚拟业务网关。

11、根据权利要求 10 所述的资源发放方法，其特征在于，所述在本数据中心内创建虚拟业务网关，具体包括：

所述收到所述创建虚拟子网请求的网络控制器向本数据中心的虚拟业务
网关 VSG 管理器发送网关服务请求，所述网关服务请求中携带如下参数：所
10 述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所
述 VxLAN ID 和所述分割后设置的子网网关 IP 地址；

所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

12、根据权利要求 11 所述的资源发放方法，其特征在于，所述增加网
关服务和 DHCP 服务，具体包括：

15 如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务，则在所述已有的 VSG 虚拟机中增加网卡，设置所述网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池；或者

创建新的 VSG 虚拟机，在所述新创建的 VSG 虚拟机中增加两块网卡，
20 其中一块网卡用于连接外网，从外网获取 IP 地址或者配置固定 IP 地址，另一块网卡用于连接虚拟子网，配置该网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址，并启动 DHCP 服务，配置分割后 DHCP IP 地址池。

13、根据权利要求 12 所述的资源发放方法，其特征在于，所述方法进
25 一步包括：

所述网络控制器接收所述 VSG 管理器返回的网关服务消息，所述消息包括 VSG 虚拟机标识；

根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网，如果所述 VSG 虚拟机未加入所述虚拟子网，则在所述虚拟子网上创建

一个端口，将所述 VSG 虚拟机绑定至该端口。

14、根据权利要求 1 所述的资源发放方法，其特征在于，所述向所述至少两个目的数据中心创建虚拟机，具体包括：

5 所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

15、根据权利要求 1 所述的资源发放方法，其特征在于，所述把所述虚拟机加入到所述虚拟网络，具体包括：

10 所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口；

所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息，根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口，并将所述虚拟机绑定至该端口。

15 16、根据权利要求 15 所述的资源发放方法，其特征在于，所述方法进一步包括：

所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息，确定所述虚拟机成功加入所述虚拟子网后，将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资源调度器所属数据中心中的分布式数据库；

20 所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

17、根据权利要求 1 所述的资源发放方法，其特征在于，所述方法进一步包括：

25 所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求，所述请求中携带如下参数：虚拟机标识、存储大小、磁盘访问服务质量 QoS 等级和虚拟机所在的主机集群；

所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源，在选定的存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存

储卷，生成唯一的卷标识，并将所述卷标识返回给所述计算资源调度器；

所述计算资源调度器将所述存储卷挂载到对应于所述虚拟机标识的虚拟机上，建立所述卷标识和所述虚拟机之间的关联关系，生成唯一的关联关系标识。

- 5 18、根据权利要求 17 所述的资源发放方法，其特征在于，所述方法进一步包括：

所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入所述计算资源调度器所属数据中心的分布式数据库；

- 10 所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心中的分布式数据库。

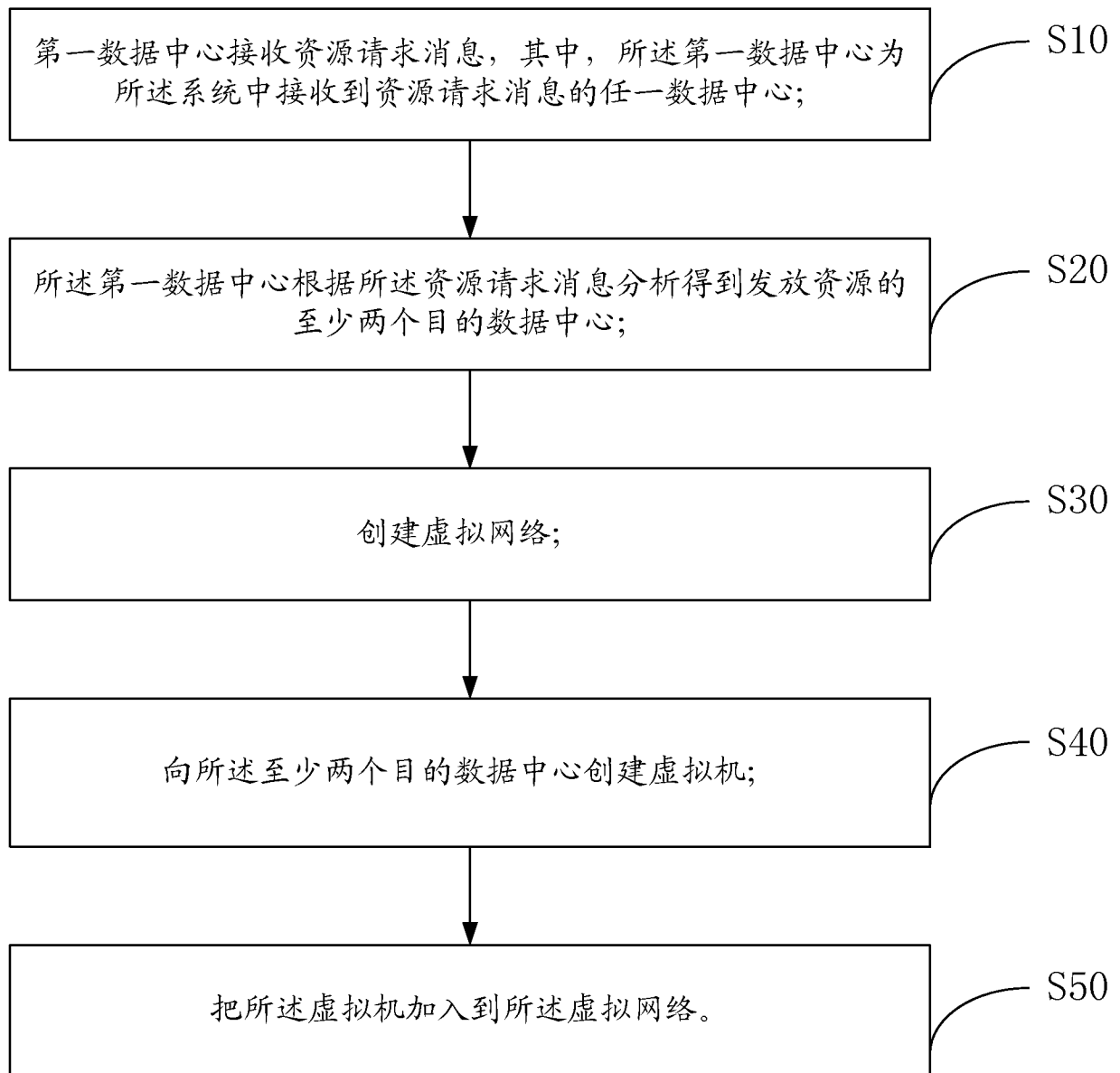


图 1

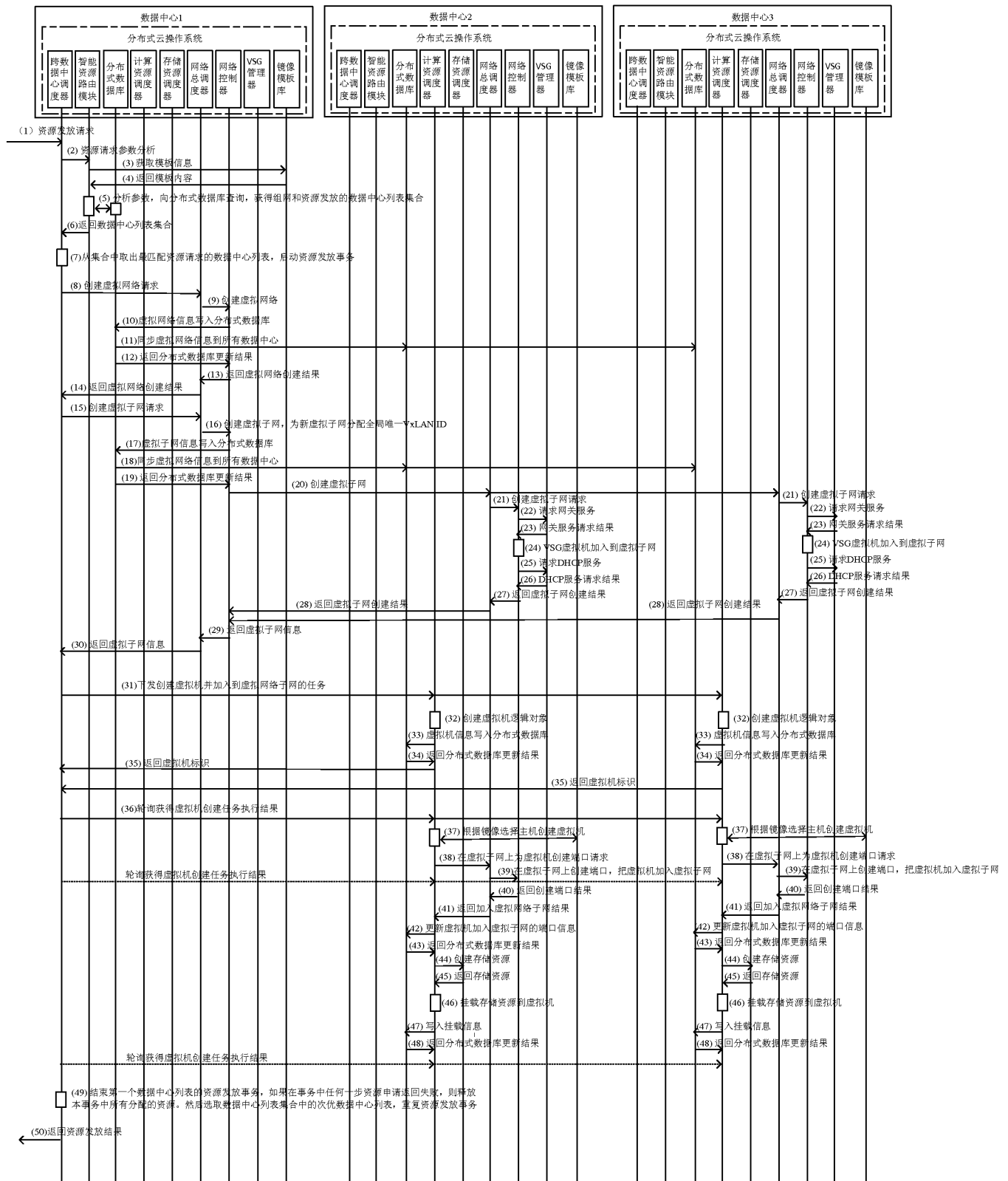


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/081514

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS; CNKI; VEN; GOOGLE: academic, produce, resource, data, information, issue, distribute, assign, data center, request, virtual, network, machine, create, add

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YU, Jun, "Application of Virtual Cluster In Network Management Centre", COMPUTER KNOWLEDGE AND TECHNOLOGY, no. 27, vol. 8, 30 September 2012 (30.09.2012), pages 6606-6612	1
A	YU, Jun, "Application of Virtual Cluster In Network Management Centre", COMPUTER KNOWLEDGE AND TECHNOLOGY, no. 27, vol. 8, 30 September 2012 (30.09.2012), pages 6606-6612	2-18
Y	CN 102185928 A (GUANGZHOU GCI SCIENCE & TECHNOLOGY CO., LTD.), 14 September 2011 (14.09.2011), abstract, and description, paragraphs 5-18	1
A	CN 102185928 A (GUANGZHOU GCI SCIENCE & TECHNOLOGY CO., LTD.), 14 September 2011 (14.09.2011), abstract, and description, paragraphs 5-18	2-18
Y	CN 102404385 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY), 04 April 2012 (04.04.2012), abstract, and description, paragraphs 4-17	1
A	CN 102404385 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY), 04 April 2012 (04.04.2012), abstract, and description, paragraphs 4-17	2-18

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 10 May 2014 (10.05.2014)	Date of mailing of the international search report 14 July 2014 (14.07.2014)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer WANG, Xuelian Telephone No.: (86-10) 62411747

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/081514

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	CN 103703724 A (HUAWEI TECHNOLOGIES CO., LTD.), 02 April 2014 (02.04.2014), claims 1-18	1-18

International application No.
PCT/CN2013/081514

Form PCT/ISA/210 (patent family annex) (July 2009)

A. 主题的分类		
H04L 29/08(2006.01) i		
按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
H04L		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CPRSABS; CNKI; VEN; GOOGLE 学术: 资源, 数据, 信息, 发放, 发布, 分配, 数据中心, 请求, 虚拟, 网, 机, 创建, 产生, 加入, resource, data, information, issue, distribute, assign, data center, request, virtual, network, machine, create, add		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	干俊. “虚拟化集群在网络管理中心的应用” 电脑知识与技术, 第27期卷, 第8卷期, 2012年 9月 30日 (2012 - 09 - 30), 第6606 - 6612页	1
A	干俊. “虚拟化集群在网络管理中心的应用” 电脑知识与技术, 第27期卷, 第8卷期, 2012年 9月 30日 (2012 - 09 - 30), 第6606-6612页	2-18
Y	CN 102185928 A (广州杰赛科技股份有限公司) 2011年 9月 14日 (2011 - 09 - 14) 摘要, 说明书第5-18段	1
A	CN 102185928 A (广州杰赛科技股份有限公司) 2011年 9月 14日 (2011 - 09 - 14) 摘要, 说明书第5-18段	2-18
Y	CN 102404385 A (华中科技大学) 2012年 4月 04日 (2012 - 04 - 04) 摘要, 说明书第4-17段	1
A	CN 102404385 A (华中科技大学) 2012年 4月 04日 (2012 - 04 - 04) 摘要, 说明书第4-17段	2-18
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2014年 5月 10日		2014年 7月 14日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国		王雪莲
传真号 (86-10)62019451		电话号码 (86-10)62411747

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
E	CN 103703724 A（华为技术有限公司）2014年 4月 02日（2014 - 04 - 02） 权利要求1-18	1-18

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2013/081514

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 102185928 A	2011年 9月 14日	CN 102185928 B	2013年 7月 17日
CN 102404385 A	2012年 4月 04日	无	
CN 103703724 A	2014年 4月 02日	无	



(12) 发明专利申请

(10) 申请公布号 CN 103703724 A

(43) 申请公布日 2014. 04. 02

(21) 申请号 201380000795. X

(22) 申请日 2013. 08. 15

(85) PCT国际申请进入国家阶段日
2013. 08. 30

(86) PCT国际申请的申请数据
PCT/CN2013/081514 2013. 08. 15

(71) 申请人 华为技术有限公司
地址 518129 中国广东省深圳市龙岗区坂田
华为总部办公楼

(72) 发明人 顾炯炯 黄朝意

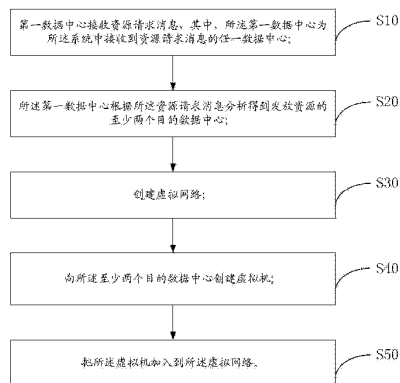
(51) Int. Cl.
H04L 12/24 (2006. 01)
H04L 29/08 (2006. 01)

权利要求书4页 说明书17页 附图2页

(54) 发明名称
一种资源发放方法

(57) 摘要

本发明实施例提供一种资源发放方法,包括:第一数据中心接收资源请求消息,其中,所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心;所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心;创建虚拟网络;向所述至少两个目的数据中心创建虚拟机;把所述虚拟机加入到所述虚拟网络。通过实施该方法,提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。



1. 一种资源发放方法,应用于分布式数据中心系统,所述系统包括至少两个数据中心,其特征在于,所述方法包括:

第一数据中心接收资源请求消息,其中,所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心;

所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心;

创建虚拟网络;

向所述至少两个目的数据中心创建虚拟机;

把所述虚拟机加入到所述虚拟网络。

2. 根据权利要求1所述的资源发放方法,其特征在于,所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心,具体包括:

所述第一数据中心的跨数据中心调度器接收到资源请求消息后,将所述资源请求消息转发给所述第一数据中心的智能资源路由模块,其中,所述资源请求消息中包括租户标识和资源模板标识;

所述智能资源路由模块根据所述租户标识和资源模板标识,获取对应的资源模板信息,所述资源模板信息至少包括网络资源规格参数和虚拟机描述参数中的一种;

所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心。

3. 根据权利要求2所述的资源发放方法,其特征在于,所述智能资源路由模块根据所述租户标识和资源模板标识,获取对应的资源模板信息,具体包括:

所述智能资源路由模块通过查询所述第一数据中心中的镜像模板库,获取对应于所述租户标识和资源模板标识的资源模板信息。

4. 根据权利要求2或3所述的资源发放方法,其特征在于,所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心,具体包括:

所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心中的分布式数据库,获取和所述租户标识以及所述资源模板信息对应的至少两个目的数据中心。

5. 根据权利要求1所述的资源发放方法,其特征在于,如果所述资源请求消息中包括跨数据中心资源分配策略,则所述方法进一步包括:

根据所述跨数据中心资源分配策略对所述至少两个目的数据中心进行分析,得到符合所述跨数据中心资源分配策略的目的数据中心。

6. 根据权利要求2所述的资源发放方法,其特征在于,所述创建虚拟网络,具体包括:

所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求,所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称;

所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求;

所述第一数据中心的网络控制器根据所述租户标识和网络名称创建虚拟网络,并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库,其中,所述虚拟网络的信息包

括网络 ID、网络状态、所述网络名称和租户标识；

所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至其他目的数据中心中的分布式数据库。

7. 根据权利要求 2 或 6 所述的资源发放方法,所述创建虚拟网络进一步包括:

所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息,确定需创建虚拟子网的目的数据中心;

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求;

所述第一数据中心的网络控制器根据所述创建虚拟子网的请求创建虚拟子网,并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库;

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

8. 根据权利要求 6 所述的资源发放方法,其特征在于,所述创建虚拟网络,进一步包括:

所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息,确定需创建虚拟子网的目的数据中心;

通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求,所述请求中包括下列参数:所述租户标识、所述网络 ID、子网名称、子网无类型域间选路 CIDR、子网互联网协议 IP 地址版本号、子网网关 IP 地址、是否启动动态主机设置协议 DHCP 分配 IP 地址、DHCP IP 地址池、待创建虚拟子网的目的数据中心标识、跨数据中心子网切分模式;

所述第一数据中心的网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网,并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库,其中,所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、虚拟可扩展本地局域网 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号;

所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

9. 根据权利要求 8 所述的资源发放方法,其特征在于,如果所述创建虚拟子网的请求中要求为所述虚拟子网启动 DHCP 分配 IP 地址,且需创建虚拟子网的目的数据中心不少于两个,则所述方法进一步包括:

所述第一数据中心的网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置;

所述第一数据中心的网络控制器向所述待创建虚拟子网的每个数据中心的网络总调度器发送创建虚拟子网的请求,所述请求中包括下列参数:所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池;

所述待创建虚拟子网的每个数据中心的网络总调度器向各自所在的数据中心中的网络控制器发送所述创建虚拟子网的请求。

10. 根据权利要求 9 所述的资源发放方法,其特征在于,所述方法进一步包括:

收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求中指示的虚拟子网已经存在,则在本数据中心内创建虚拟业务网关。

11. 根据权利要求 10 所述的资源发放方法,其特征在于,所述在本数据中心内创建虚拟业务网关,具体包括:

所述收到所述创建虚拟子网请求的网络控制器向本数据中心的虚拟业务网关 VSG 管理器发送网关服务请求,所述网关服务请求中携带如下参数:所述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所述 VxLAN ID 和所述分割后设置的子网网关 IP 地址;

所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

12. 根据权利要求 11 所述的资源发放方法,其特征在于,所述增加网关服务和 DHCP 服务,具体包括:

如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务,则在所述已有的 VSG 虚拟机中增加网卡,设置所述网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池;或者

创建新的 VSG 虚拟机,在所述新创建的 VSG 虚拟机中增加两块网卡,其中一块网卡用于连接外网,从外网获取 IP 地址或者配置固定 IP 地址,另一块网卡用于连接虚拟子网,配置该网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池。

13. 根据权利要求 12 所述的资源发放方法,其特征在于,所述方法进一步包括:

所述网络控制器接收所述 VSG 管理器返回的网关服务消息,所述消息包括 VSG 虚拟机标识;

根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网,如果所述 VSG 虚拟机未加入所述虚拟子网,则在所述虚拟子网上创建一个端口,将所述 VSG 虚拟机绑定至该端口。

14. 根据权利要求 1 所述的资源发放方法,其特征在于,所述向所述至少两个目的数据中心创建虚拟机,具体包括:

所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

15. 根据权利要求 1 所述的资源发放方法,其特征在于,所述把所述虚拟机加入到所述虚拟网络,具体包括:

所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口;

所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息,根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口,并将所述虚拟机绑定至该端口。

16. 根据权利要求 15 所述的资源发放方法,其特征在于,所述方法进一步包括:

所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息,确定所述虚拟机成功加入所述虚拟子网后,将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资源调度器所属数据中心中的分布式数据库;

所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

17. 根据权利要求 1 所述的资源发放方法,其特征在于,所述方法进一步包括:

所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求,所述请求中携带如下参数:虚拟机标识、存储大小、磁盘访问服务质量 QoS 等级和虚拟机所在的主机集群;

所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源,在选定的存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存储卷,生成唯一的卷标识,并将所述卷标识返回给所述计算资源调度器;

所述计算资源调度器将所述存储卷挂载到对应于所述虚拟机标识的虚拟机上,建立所述卷标识和所述虚拟机之间的关联关系,生成唯一的关联关系标识。

18. 根据权利要求 17 所述的资源发放方法,其特征在于,所述方法进一步包括:

所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入所述计算资源调度器所属数据中心的分布式数据库;

所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心中的分布式数据库。

一种资源发放方法

技术领域

[0001] 本发明涉及云计算领域,尤其涉及一种资源发放方法。

背景技术

[0002] 传统的大型企业数据中心是一个物理分层的架构,顶层的数据中心主要负责企业级全球业务的应用集中部署,区域数据中心用来承载全球应用的镜像以及本区域范围内全局共享的应用和数据。在区域范围较大(比如覆盖数个国家)时,在每个区域内还可能进一步建设服务器机房以解决对区域范围全局数据共享没有需求、但对于服务质量有更高要求的应用(如研发仿真、大型软件开发的持续构建、DHCP 以及文件服务等)。对于分支机构办公室,则主要通过建设网络机房来解决便携机与个人计算机等日常办公 IT 设备的纯网络接入问题。

[0003] 传统的数据中心架构面临着一系列的问题:软件应用与物理计算和存储设备耦合绑定,单个应用按照最大预期负荷资源来进行容量规划,资源利用率低,服务器资源利用率更是普遍低于 15%;物理资源和应用强耦合,业务部署,扩容以及升级等管理维护运营复杂;跨数据中心的计算、存储、网络资源也不能统一拉通调度与协同管理,多数物理设备在业务量低和空闲时间内处于带电空转的状态;数据中心层次过多导致数据访问延迟加大,尤其是对于处于数据中心分层架构顶端的应用,所必须经过的数据中心站点网络环节和层级过多,加大了网络连接故障或流量拥塞发生的概率。

发明内容

[0004] 本发明实施例提供一种资源发放的方法,用以解决传统数据中心架构面临的资源利用率低、管理维护运营复杂、网络连接故障或流量拥塞发生的概率高的问题。

[0005] 第一方面,本发明实施例提供一种资源发放的方法,应用于分布式数据中心系统,所述系统包括至少两个数据中心,所述方法包括:

[0006] 第一数据中心接收资源请求消息,其中,所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心;

[0007] 所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心;

[0008] 创建虚拟网络;

[0009] 向所述至少两个目的数据中心创建虚拟机;

[0010] 把所述虚拟机加入到所述虚拟网络。

[0011] 结合第一方面,在第一方面的第一种可能的实现方式中,所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心,具体包括:

[0012] 所述第一数据中心的跨数据中心调度器接收到资源请求消息后,将所述资源请求消息转发给所述第一数据中心的智能资源路由模块,其中,所述资源请求消息中包括租户标识和资源模板标识;

[0013] 所述智能资源路由模块根据所述租户标识和资源模板标识,获取对应的资源模板信息,所述资源模板信息至少包括网络资源规格参数和虚拟机描述参数中的一种;

[0014] 所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心。

[0015] 结合第一方面的第一种可能的实现方式,在第二种可能的实现方式中,所述智能资源路由模块根据所述租户标识和资源模板标识,获取对应的资源模板信息,具体包括:

[0016] 所述智能资源路由模块通过查询所述第一数据中心中的镜像模板库,获取对应于所述租户标识和资源模板标识的资源模板信息。

[0017] 结合第一方面的第一种可能的实现方式以及第一方面的第二种可能的实现方式,在第三种可能的实现方式中,所述智能资源路由模块通过分析所述资源模板信息得到发放资源的至少两个目的数据中心,具体包括:

[0018] 所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心中的分布式数据库,获取和所述租户标识以及所述资源模板信息对应的至少两个目的数据中心。

[0019] 结合第一方面的第一种可能的实现方式,在第四种可能的实现方式中,如果所述资源请求消息中包括跨数据中心资源分配策略,则所述方法进一步包括:

[0020] 根据所述跨数据中心资源分配策略对所述至少两个目的数据中心进行分析,得到符合所述跨数据中心资源分配策略的目的数据中心。

[0021] 结合第一方面的第一种可能的实现方式,在第五种可能的实现方式中,所述创建虚拟网络,具体包括:

[0022] 所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求,所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称;

[0023] 所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求;

[0024] 所述第一数据中心的网络控制器根据所述租户标识和网络名称创建虚拟网络,并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库,其中,所述虚拟网络的信息包括网络 ID、网络状态、所述网络名称和租户标识;

[0025] 所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至其他目的数据中心中的分布式数据库。

[0026] 结合第一方面的第一种可能的实现方式以及第一方面的第五种可能的实现方式,在第六种可能的实现方式中,所述创建虚拟网络,进一步包括:

[0027] 所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息,确定需创建虚拟子网的目的数据中心;

[0028] 通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求;

[0029] 所述第一数据中心的网络控制器根据所述创建虚拟子网的请求创建虚拟子网,并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库;

[0030] 所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据

中心中的分布式数据库。

[0031] 结合第一方面的第五种可能的实现方式,在第七种可能的实现方式中,所述创建虚拟网络,进一步包括:

[0032] 所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息,确定需创建虚拟子网的目的数据中心;

[0033] 通过所述第一数据中心的网络总调度器向所述第一数据中心的网络控制器发送创建虚拟子网的请求,所述请求中包括下列参数:所述租户标识、所述网络 ID、子网名称、子网无类型域间选路 CIDR、子网互联网协议 IP 地址版本号、子网网关 IP 地址、是否启动动态主机设置协议 DHCP 分配 IP 地址、DHCP IP 地址池、待创建虚拟子网的目的数据中心标识、跨数据中心子网切分模式;

[0034] 所述第一数据中心的网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网,并把所述虚拟子网的信息写入所述第一数据中心的分布式数据库,其中,所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、虚拟可扩展本地局域网 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号;

[0035] 所述第一数据中心的分布式数据库将所述虚拟子网的信息同步至其他目的数据中心中的分布式数据库。

[0036] 结合第一方面的第七种可能的实现方式,在第八种可能的实现方式中,如果所述创建虚拟子网的请求中要求为所述虚拟子网启动 DHCP 分配 IP 地址,且需创建虚拟子网的目的数据中心不少于两个,则所述方法进一步包括:

[0037] 所述第一数据中心的网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置;

[0038] 所述第一数据中心的网络控制器向所述待创建虚拟子网的每个数据中心的网络总调度器发送创建虚拟子网的请求,所述请求中包括下列参数:所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池;

[0039] 所述待创建虚拟子网的每个数据中心的网络总调度器向各自所在的数据中心中的网络控制器发送所述创建虚拟子网的请求。

[0040] 结合第一方面的第八种可能的实现方式,在第九种可能的实现方式中,所述方法进一步包括:

[0041] 收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求中指示的虚拟子网已经存在,则在本数据中心内创建虚拟业务网关。

[0042] 结合第一方面的第九种可能的实现方式,在第十种可能的实现方式中,所述在本数据中心内创建虚拟业务网关,具体包括:

[0043] 所述收到所述创建虚拟子网请求的网络控制器向本数据中心的虚拟业务网关 VSG 管理器发送网关服务请求,所述网关服务请求中携带如下参数:所述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所述 VxLAN ID 和所述分割后设置的子网网关 IP 地址;

[0044] 所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

[0045] 结合第一方面的第十种可能的实现方式,在第十一种可能的实现方式中,所述增

加网关服务和 DHCP 服务,具体包括:

[0046] 如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务,则在所述已有的 VSG 虚拟机中增加网卡,设置所述网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池;或者

[0047] 创建新的 VSG 虚拟机,在所述新创建的 VSG 虚拟机中增加两块网卡,其中一块网卡用于连接外网,从外网获取 IP 地址或者配置固定 IP 地址,另一块网卡用于连接虚拟子网,配置该网卡的 IP 地址为所述创建虚拟子网请求中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池。

[0048] 结合第一方面的第十一种可能的实现方式,在第十二种可能的实现方式中,所述方法进一步包括:

[0049] 所述网络控制器接收所述 VSG 管理器返回的网关服务消息,所述消息包括 VSG 虚拟机标识;

[0050] 根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网,如果所述 VSG 虚拟机未加入所述虚拟子网,则在所述虚拟子网上创建一个端口,将所述 VSG 虚拟机绑定至该端口。

[0051] 结合第一方面,在第十三种可能的实现方式中,所述向所述至少两个目的数据中心创建虚拟机,具体包括:

[0052] 所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

[0053] 结合第一方面,在第十四种可能的实现方式中,所述把所述虚拟机加入到所述虚拟网络,具体包括:

[0054] 所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口;

[0055] 所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息,根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口,并将所述虚拟机绑定至该端口。

[0056] 结合第一方面的第十四种可能的实现方式,在第十五种可能的实现方式中,所述方法进一步包括:

[0057] 所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息,确定所述虚拟机成功加入所述虚拟子网后,将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资源调度器所属数据中心中的分布式数据库;

[0058] 所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

[0059] 结合第一方面,在第十六种可能的实现方式中,所述方法进一步包括:

[0060] 所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求,所述请求中携带如下参数:虚拟机标识、存储大小、磁盘访问服务质量 QoS 等级和虚拟机所在的主机集群;

[0061] 所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源,在选定的

存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存储卷,生成唯一的卷标识,并将所述卷标识返回给所述计算资源调度器;

[0062] 所述计算资源调度器将所述存储卷挂载到对应于所述虚拟机标识的虚拟机上,建立所述卷标识和所述虚拟机之间的关联关系,生成唯一的关联关系标识。

[0063] 结合第一方面的第十六种可能的实现方式,在第十七种可能的实现方式中,所述方法进一步包括:

[0064] 所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入所述计算资源调度器所属数据中心的分布式数据库;

[0065] 所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心中的分布式数据库。

[0066] 本发明实施例提供的资源发放方法,通过部署在每个数据中心中的分布式云操作系统之间的交互,执行相应的资源发放步骤,使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”,实现了资源的全局拉通,可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

附图说明

[0067] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例。

[0068] 图 1 为本发明实施例提供的资源发放方法的流程框图;

[0069] 图 2 为本发明实施例提供的资源发放方法应用于具体场景时的流程示意图。

具体实施方式

[0070] 为使本发明实施例的目的、技术方案和优点更加清楚,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,所描述的实施例是本发明一部分实施例,而不是全部的实施例。

[0071] 需要说明的是,本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A 和/或 B,可以表示:单独存在 A,同时存在 A 和 B,单独存在 B 这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。

[0072] 进一步需要说明的是,本文中所述的数据中心指的是计算资源、存储资源和网络资源的集合,可以是物理位置上的数据中心,也可以是逻辑概念上的数据中心。相应的,本文所述的方法既可以应用于物理位置上的数据中心,也可以应用于逻辑概念上的数据中心,只要这个数据中心中包含了计算资源、存储资源和网络资源。

[0073] 本发明实施例提供的技术方案通过在每一个数据中心中,部署同样的分布式云操作系统,将企业分层管理及逐级网络互联的传统数据中心架构,改造为全扁平式管理、点到点全互联的分布式数据中心架构。在该架构中,全网各个物理数据中心节点尽管容纳的资源数量及业务类型各不相同,但从管理调度视角看可以是完全对等的关系,可以不再有主从式管理从属关系,所有基础设施资源作为一个统一的“逻辑资源池”,实现资源全局拉通,

进行跨物理数据中心的资源弹性调度。解决了传统数据中心架构面临的资源利用率低、管理维护运营复杂、网络连接故障或流量拥塞发生的概率高的问题。

[0074] 为了更加清楚的理解本发明实施例提供的技术方案,有必要对部署在每个数据中心中的分布式云操作系统做一简单介绍。该系统主要包括:

[0075] 跨数据中心调度器:用于进行跨多个数据中心的资源发放;

[0076] 智能资源路由模块:通过分析资源请求消息中的参数,得到需要申请资源的数据中心列表;

[0077] 分布式数据库:负责存放分布式云数据中心的全局数据,在每一个数据中心内,都可以从分布式数据库中看到同样的数据;

[0078] 计算资源调度器:负责本数据中心内计算资源的发放,挂载(attach)存储资源和加入到虚拟网络中;

[0079] 存储资源调度器:负责本数据中心内存储资源的发放;

[0080] 网络总调度器:负责建立全局性虚拟网络、虚拟子网、虚拟业务网关,其中所述虚拟网络包括一个或多个虚拟子网,所述虚拟子网可以分布在一个或多个数据中心;

[0081] 网络控制器:负责数据中心内的软件定义网络的调度,管理虚拟交换机(virtual switch,简称为 vSwitch)和端口,以及计算资源加入虚拟网络时端口与虚拟可扩展本地局域网(virtual extensible local area network,简称为 VxLAN)的映射;

[0082] 虚拟业务网关(virtual service gateway,简称为 VSG)管理器:负责创建虚拟网络的虚拟业务网关,如路由(Route)服务、动态主机设置协议(Dynamic Host Configuration Protocol,简称为 DHCP) Server 服务、VxLAN 网关服务、网络地址转换(Network Address Translation,简称为 NAT)服务、访问控制列表(Access Control List,简称为 ACL)服务、虚拟专用网络(virtual private network,简称为 VPN)服务、软件负载均衡服务等中的一种或者几种服务;

[0083] 镜像模板库:提供虚拟机创建需要的镜像文件和模板服务,模板可以描述单个虚拟机的规格要求,也可以包含多个虚拟机甚至组网(虚拟机之间的网络)的描述。

[0084] 基于上述分布式数据中心架构和分布式云操作系统,本发明实施例提供了一种资源发放方法。图 1 为本发明实施例提供的资源发放方法的流程框图。如图 1 所示,所述资源发放方法包括以下步骤:

[0085] S10、第一数据中心接收资源请求消息,其中,所述第一数据中心为所述系统中接收到资源请求消息的任一数据中心;

[0086] 需要说明的是,本发明实施例提供的资源发放方法应用于分布式数据中心架构,该架构包括两个或两个以上的物理数据中心,且该架构中所有的物理数据中心不管容纳的资源数量及业务类型是否相同,从管理调度视角看可以是完全对等的关系,各数据中心之间可以不再有主从式管理从属关系。所述第一数据中心为所述分布式数据中心架构中接收到资源请求消息的任一数据中心。

[0087] S20、所述第一数据中心根据所述资源请求消息分析得到发放资源的至少两个目的数据中心;

[0088] 具体的,所述第一数据中心的跨数据中心调度器接收到资源请求消息后,将所述资源请求消息转发给所述第一数据中心的智能资源路由模块,其中,所述资源请求消息中

包括租户标识和资源模板标识；所述智能资源路由模块根据所述租户标识和资源模板标识，通过查询所述第一数据中心的镜像模板库，获取和所述资源模板标识对应的资源模板信息，所述资源模板信息包括虚拟机描述参数列表和 / 或网络资源规格参数；所述智能资源路由模块根据所述租户标识和所述资源模板信息查询所述第一数据中心的分布式数据库，获取和所述租户标识以及所述资源模板信息对应的目的数据中心列表。

[0089] 如果所述资源请求消息中还携带参数：指定数据中心、指定数据中心标识列表，则将所述获取的目的数据中心列表与所述指定数据中心标识列表进行取交集运算；

[0090] 如果所述资源请求消息中还包括跨数据中心资源分配策略（如，资源平均分配、按空闲资源比例分配等），则根据所述跨数据中心资源分配策略对所述获取到的目的数据中心列表进行分析，得到符合所述跨数据中心资源分配策略的目的数据中心列表。

[0091] S30、创建虚拟网络；

[0092] 需要说明的是，所述虚拟网络包括一个或多个虚拟子网，所述虚拟子网可以分布在一个或多个数据中心；

[0093] 具体的，所述第一数据中心的跨数据中心调度器向所述第一数据中心的网络总调度器发送创建虚拟网络的请求，所述创建虚拟网络的请求携带所述租户标识和所述网络资源规格参数中的网络名称；

[0094] 所述网络总调度器向所述第一数据中心的网络控制器发送所述创建虚拟网络的请求；

[0095] 所述网络控制器根据所述租户标识和网络名称创建虚拟网络，并把所述虚拟网络的信息写入所述第一数据中心的分布式数据库，其中，所述虚拟网络的信息包括网络标识（identity，简称为 ID）、网络状态、所述网络名称和租户标识；

[0096] 所述第一数据中心的分布式数据库将所述虚拟网络的信息同步至所述系统中其他数据中心的分布式数据库。

[0097] 进一步的，所述第一数据中心的跨数据中心调度器通过分析所述资源模板信息，确定需创建虚拟子网的目的数据中心；

[0098] 通过所述网络总调度器向所述网络控制器发送创建虚拟子网的请求，所述请求中包括下列参数：所述租户标识、所述网络 ID、所述子网名称、子网无类型域间选路（Classless Inter-Domain Routing，简称为 CIDR）、子网互联网协议（internet protocol，简称为 IP）地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、DHCP IP 地址池、所述待创建虚拟子网的数据中心标识列表、跨数据中心子网切分模式；

[0099] 所述网络控制器根据所述租户标识、网络 ID 和子网名称创建虚拟子网（包括为虚拟子网分配全局唯一的 VxLAN ID），并把所述虚拟子网的信息写入所述分布式数据库，其中，所述虚拟子网的信息包括所述租户标识、所述网络 ID、所述子网名称、子网 ID、VxLAN ID、所述子网 CIDR、所述 IP 地址版本号；

[0100] 所述分布式数据库将所述虚拟子网的信息同步至所述系统中其他数据中心的分布式数据库。

[0101] 进一步的，如果所述待创建虚拟子网的数据中心标识列表中包括至少两个数据中心且要求为所述虚拟子网启动 DHCP 分配 IP 地址，则所述网络控制器对待分配的所述 DHCP IP 地址池按照所述跨数据中心子网切分模式进行 IP 地址切分和网关设置；

[0102] 所述网络控制器向所述待创建虚拟子网的数据中心标识列表中的每个数据中心的网络总调度器发送创建虚拟子网的请求,所述请求中包括下列参数:所述租户标识、所述网络 ID、所述子网名称、所述子网 ID、所述 VxLAN ID、所述子网 CIDR、所述 IP 地址版本号、所述是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址和分割后 DHCP IP 地址池;

[0103] 所述待创建虚拟子网的数据中心标识列表中的每个数据中心的网络总调度器向各自所在的数据中心的网络控制器发送所述创建虚拟子网的请求。

[0104] 进一步的,如果收到所述创建虚拟子网请求的网络控制器判断所述创建虚拟子网请求的参数中指示的虚拟子网已经存在,则在本数据中心内创建虚拟业务网关。

[0105] 其中,在本数据中心内创建虚拟业务网关,具体包括:

[0106] 所述收到所述创建虚拟子网请求的网络控制器向本数据中心的 VSG 管理器发送网关服务请求,所述网关服务请求中携带如下参数:所述子网 ID、所述是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、所述 VxLAN ID 和所述分割后设置的子网网关 IP 地址;

[0107] 所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。

[0108] 其中,所述增加网关服务和 DHCP 服务,具体包括:

[0109] 如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务,则在所述已有的 VSG 虚拟机中增加网卡,设置所述网卡的 IP 地址为所述创建虚拟子网请求的参数中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池;或者

[0110] 创建新的 VSG 虚拟机,在所述新创建的 VSG 虚拟机中增加两块网卡,其中一块网卡用于连接外网,从外网获取 IP 地址或者配置固定 IP 地址,另一块网卡用于连接虚拟子网,配置该网卡的 IP 地址为所述创建虚拟子网请求的参数中携带的所述分割后设置的子网网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池。

[0111] 进一步的,所述网络控制器接收所述 VSG 管理器返回的网关服务消息,所述消息包括 VSG 虚拟机标识;

[0112] 根据所述 VSG 虚拟机标识判断所述 VSG 虚拟机是否已经加入所述虚拟子网,如果所述 VSG 虚拟机未加入所述虚拟子网,则在所述虚拟子网上创建一个端口,将所述 VSG 虚拟机绑定至该端口。

[0113] S40、向所述至少两个目的数据中心创建虚拟机;

[0114] 具体的,所述第一数据中心的跨数据中心调度器指示所述至少两个目的数据中心中的计算资源调度器在各自所属的数据中心中创建虚拟机。

[0115] S50、把所述虚拟机加入到所述虚拟网络。

[0116] 具体的,

[0117] 所述至少两个目的数据中心中的计算资源调度器通过各自所属数据中心中的网络总调度器请求本数据中心中的网络控制器在虚拟子网上分别为所述虚拟机创建端口;

[0118] 所述网络控制器根据端口创建请求中携带的子网 ID 获得虚拟子网的信息,根据所述虚拟子网的信息在虚拟子网上为所述虚拟机创建端口,并将所述虚拟机绑定至该端口;

[0119] 所述计算资源调度器获得所述虚拟机加入到所述虚拟子网的端口信息,确定所述虚拟机成功加入所述虚拟子网后,将虚拟机 ID 和端口 ID 之间的映射关系写入所述计算资

源调度器所属数据中心中的分布式数据库；

[0120] 所述分布式数据库将所述虚拟机 ID 和端口 ID 之间的映射关系同步至其他目的数据中心中的分布式数据库。

[0121] 进一步的,还可以为所述虚拟机分配存储资源,具体包括:

[0122] 所述至少两个目的数据中心中的计算资源调度器向各自所属数据中心中的存储资源调度器分别发送创建存储资源的请求,所述请求中携带如下参数:虚拟机标识、存储大小、磁盘访问服务质量(Quality of Service,简称为 QoS)等级和虚拟机所在的主机集群;

[0123] 所述存储资源调度器根据所述请求中携带的参数选择具体的存储资源,在选定的存储资源上创建满足所述磁盘访问 QoS 等级和所述存储大小的存储卷,生成唯一的卷标识,并将所述卷标识返回给所述计算资源调度器;

[0124] 所述计算资源调度器把所述存储卷挂载到虚拟机标识所代表的虚拟机,建立卷标识和虚拟机之间的关联关系,生成唯一的关联关系标识;

[0125] 所述计算资源调度器将所述虚拟机标识、所述卷标识和所述关联关系标识写入本数据中心的分布式数据库;

[0126] 所述分布式数据库将所述虚拟机标识、所述卷标识和所述关联关系标识同步至其他目的数据中心的分布式数据库。

[0127] 本发明实施例提供的资源发放方法,通过部署在每个数据中心中的分布式云操作系统之间的交互,执行相应的资源发放步骤,使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”,实现了资源的全局拉通,可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

[0128] 为了更清楚地理解本发明实施例的技术方案,下面将结合具体的应用场景对本发明实施例的技术方案做进一步的说明。

[0129] 图 2 为本发明实施例提供的资源发放方法应用于具体场景时的流程示意图。在图 2 中,有三个数据中心,其中数据中心 1 收到来自租户的资源请求。需要说明的是,在实际应用中,任何一个数据中心都可以接受租户的资源请求,所有数据中心的功能都是对等的。如图 2 所示,具体的资源发放流程包括以下步骤:

[0130] (1) 数据中心 1 的跨数据中心调度器接收资源请求消息,消息参数带有租户标识、资源模板标识、是否指定数据中心、指定数据中心标识列表以及跨数据中心资源分配策略(如,资源平均分配、按空闲资源比例分配等)。

[0131] (2) 所述跨数据中心调度器收到请求消息后,把该请求消息转发到数据中心 1 的智能资源路由模块,请求进行参数分析。

[0132] (3) 所述智能资源路由模块向数据中心 1 的镜像模板库查询资源模板标识对应的资源模板中需要分配的具体资源要求。

[0133] (4) 所述镜像模板库向所述智能资源路由模块返回需要分配资源的具体参数要求:

[0134] 虚拟机描述参数列表:描述参数列表中的每一项都包含一个虚拟机的创建要求,包括虚拟机分配数量、虚拟机规格标识(通过虚拟机规格标识可以获取得到虚拟机的规格信息如虚拟机中央处理器(central processing unit,简称为 CPU)主频、CPU 个数、内存

大小、磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级（用于选择磁盘存储类型：固态硬盘、直连硬盘、外置存储等）、每个磁盘备份模式、虚拟网卡个数、每个虚拟网卡的带宽、虚拟机镜像链接、物理计算资源还是虚拟计算资源、CPU 类型（如果是物理计算资源，指定物理机的 CPU 类型）、HyperVisor 类型 CPU 类型（如果是请求虚拟计算资源，指定 HyperVisor 类型）、是否启动高可用性（Hight Availability，简称为 HA）、是否启动故障冗余（fault tolerance，简称为 FT）、是否需要容灾、容灾类型（同步容灾、异步容灾）、容灾模式（双活、主备模式）、虚拟机启动参数、虚拟机配置参数、安全组、其他参数。

[0135] 网络资源规格参数：网络名称、子网名称、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、DHCP IP 地址池、外网 IP 地址列表、子网 QoS 参数（包括网络带宽下限、网络带宽上限、网络抖动上限、网络抖动下限、网络时延上限、网络时延下限、网络丢包率上限、网络丢包率下限）、跨数据中心子网切分模式（模式 1：均分 IP 地址，每个数据中心部署网关；模式 2：单网关单 IP 地址池）。

[0136] 资源请求消息的参数中可以包含虚拟机描述参数列表和 / 或网络资源规格参数中的一类参数或者几类参数。当虚拟机描述参数列表中包含容灾相关参数时（容灾类型、容灾模式），一般同时会包含网络资源规格参数中的子网 QoS 参数，以确定容灾的网络要求。

[0137] （5）所述智能资源路由模块根据租户标识，查询数据中心 1 的分布式数据库，通过租户标识获得可以申请资源的数据中心列表集合，如果“是否指定数据中心”参数为“是”，则和“指定数据中心标识列表”进行取交集运算，得到初步分析后的数据中心列表集合。

[0138] 智能资源路由模板根据镜像模板库返回的虚拟机描述参数列表、网络资源规格参数等，以及初步分析获得的数据中心列表集合、跨数据中心资源分配策略（如，资源平均分配、按空闲资源比例分配等）进行目的数据中心列表的分析：

[0139] 首先，智能资源路由模板从数据中心列表集合中取出第一个数据中心列表，根据虚拟机描述参数列表中的虚拟机具体要求和跨数据中心资源分配策略，检查列表中的每个数据中心是否有满足条件的计算和存储资源。如果有足够的计算资源和存储资源，检查是否有网络资源规格参数，如果有网络资源规格参数要求，则分析数据中心间和数据中心内的网络是否满足子网 QoS 参数的要求。根据计算资源、存储资源和网络资源要求，可以分析出虚拟机分配时是否需要在特定的主机集群创建。如果数据中心列表的数据中心满足计算资源、存储资源以及数据中心间的网络参数要求，则该数据中心列表为符合要求的数据中心列表。如果计算资源、存储资源或者网络资源规格参数有任意一个参数不满足要求，则认为该数据中心列表不是符合要求的数据中心列表。

[0140] 重复上述计算资源、存储资源和网络资源的分析，直到数据中心列表集合中的所有数据中心列表都分析完成为止，对符合要求的数据中心列表集合按照满足程度多少从高到低进行排序，得到符合资源发放要求的目数据的数据中心列表集合。该目的数据中心列表集合中每一个数据中心列表包含如下信息 {[网络资源规格参数]、[数据中心标识，虚拟机描述参数列表]}，其中，数据中心标识和虚拟机描述参数列表是“和”的关系，[网络资源规格参数] 和 [数据中心标识，虚拟机描述参数列表] 是“和 / 或”的关系。

[0141] （6）智能资源路由模块向跨数据中心调度器返回目的数据中心列表集合，集合中的每一个列表包含 {[网络资源规格参数]、[数据中心标识，虚拟机描述参数列表]} 这些信息。

[0142] (7) 跨数据中心调度器取出上述目的数据中心列表集合中第一个数据中心列表, 启动资源发放事务。以下第(8)步到第(49)步为一个完整的资源发放事务。在整个资源发放事务中, 如果有一个资源发放失败, 则整个事务进行回滚, 事务中已经发放的资源都被完全释放。

[0143] (8) 数据中心 1 的跨数据中心调度器检查数据中心列表中的 {[网络资源规格参数]、[数据中心标识, 虚拟机描述参数列表]} 的信息。如果包含网络资源规格参数, 则所述跨数据中心调度器判断在租户标识下网络名称所代表的虚拟网络是否存在, 如果不存在, 则向其所在数据中心的网络总调度器发出创建虚拟网络的请求, 参数带有租户标识、网络名称。

[0144] (9) 数据中心 1 的网络总调度器判断请求为创建虚拟网络, 则向本数据中心的网络控制器发出创建虚拟网络的请求。

[0145] (10) 所述网络控制器根据租户标识、网络名称参数为该租户创建一个虚拟网络, 生成唯一网络标识: 网络 ID, 初始化网络状态为激活状态。把新创建虚拟网络的网络 ID、网络名称、租户标识和网络状态写入分布式数据库, 同时要求分布式数据库同步该虚拟网络信息到所有数据中心。

[0146] (11) 所述分布式数据库向其他数据中心(图 2 中为数据中心 2 和数据中心 3)的分布式数据库更新新创建的虚拟网络信息, 包括租户标识、网络 ID、网络名称、网络状态。

[0147] (12) 所述分布式数据库向数据中心 1 的网络控制器返回数据库更新结果。

[0148] (13) 如果数据库更新成功, 则所述网络控制器向所述网络总调度器返回新创建虚拟网络的信息, 包括租户标识、网络 ID、网络名称、网络状态。如果数据库更新失败, 则向网络总调度器返回创建虚拟网络失败。

[0149] (14) 所述网络总调度器向所述跨数据中心调度器返回新创建虚拟网络的信息。

[0150] (15) 所述跨数据中心调度器检查虚拟网络的信息, 如果虚拟网络创建失败, 则终止流程。如果虚拟网络信息成功增加到数据库, 则检查 {[网络资源规格参数]、[数据中心标识, 虚拟机描述参数列表]}, 根据租户标识、网络名称和子网名称判断子网名称所代表的子网是否存在。如果不存在, 需要创建子网, 分析参数 {[网络资源规格参数]、[数据中心标识, 虚拟机描述参数列表]} 获得待创建虚拟子网的数据中心列表。当数据中心列表中的数据中心个数大于 1 时, 表明需要在数据中心列表中的所有数据中心创建虚拟子网, 否则表示仅仅在指定的这个数据中心创建虚拟子网。如果仅仅在指定的这个数据中心(图 2 中为数据中心 1)创建虚拟子网, 则所述跨数据中心调度器向数据中心 1 的网络总调度器发出创建虚拟子网的请求, 请求中携带参数租户标识、网络 ID、子网名称、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、数据中心标识列表、跨数据中心子网切分模式(模式 1: 均分 IP 地址, 每个数据中心部署网关; 模式 2: 均分 IP 地址, 单网关模式)。

[0151] (16) 所述网络总调度器向其所在数据中心的网络控制器发出创建虚拟子网的请求, 请求中携带参数租户标识、网络 ID、子网名称、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、DHCP IP 地址池、数据中心标识列表、跨数据中心子网切分模式(模式 1: 均分 IP 地址, 每个数据中心部署网关; 模式 2: 单网关单 IP 地址池)。

[0152] (17) 所述网络控制器根据租户标识、网络 ID、子网名称参数为该租户创建一个虚

拟子网,生成唯一子网标识:子网 ID,申请全局唯一的 VxLANID,设置子网属性为参数中的子网 CIDR、子网 IP 地址版本号。所述网络控制器把虚拟子网的信息:租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号写入数据中心 1 的分布式数据库中。同时要求分布式数据库同步虚拟子网信息到所有数据中心。

[0153] (18) 数据中心 1 的分布式数据库向其他数据中心(图 2 中为数据中心 2 和数据中心 3)的分布式数据库更新新创建的虚拟子网信息,包括租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号。

[0154] (19) 数据中心 1 的分布式数据库向本数据中心的网络控制器返回数据库更新结果。

[0155] (20) 数据中心 1 的网络控制器继续分析虚拟子网的数据中心标识列表和跨数据中心子网切分模式(模式 1:均分 IP 地址,每个数据中心部署网关;模式 2:单网关单 IP 地址池)。如果数据中心标识列表大于 1 个数据中心且参数要求为子网启动 DHCP 分配 IP 地址,则对待分配的 DHCP IP 地址池进行按照跨数据中心子网切分模式(模式 1:均分 IP 地址,每个数据中心部署网关;模式 2:单网关单 IP 地址池)进行 IP 地址切分和网关 IP 地址设置。

[0156] 在模式 1 下平均分割 IP 地址并设置不同的子网网关 IP 地址:假如 CIDR 为 192.168.0.0/24,IP 地址版本号为 IPV4,DHCP IP 地址池为 192.168.0.1~192.168.0.254,同时数据中心标识列表包含数据中心 2 和数据中心 3,则该子网在数据中心 2 的 DHCP IP 地址池是 192.168.0.2~192.168.0.127,网关 IP 地址是 192.168.0.1;该子网在数据中心 3 的 DHCP IP 地址池是 192.168.0.129~192.168.0.254,网关 IP 地址是 192.168.0.128。在模式 2 下单网关单 IP 地址池,不分割 IP 地址,但是对第一个数据中心设置子网网关 IP 地址:假如 CIDR 为 192.168.0.0/24,IP 地址版本号为 IPV4,DHCP IP 地址池为 192.168.0.1~192.168.0.254,同时数据中心标识列表包含数据中心 2 和数据中心 3,则该子网在数据中心 2 的 DHCP IP 地址池是 192.168.0.2~192.168.0.254,网关 IP 地址是 192.168.0.1;该子网在数据中心 3 的 DHCP IP 地址池是 192.168.0.2~192.168.0.254,是否启动 DHCP 服务设置为否,网关 IP 为无效 IP 地址。向数据中心标识列表中的每一个数据中心(图 2 中为数据中心 2 和数据中心 3)的网络总调度器发出创建虚拟子网的请求,请求中携带参数:租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号、是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址、分割后的 DHCP IP 地址池。

[0157] 如果数据中心标识列表只包含一个数据中心标识,则不进行 DHCP IP 地址池的分割和子网网关 IP 地址的重新设置。

[0158] (21) 数据中心 2 和数据中心 3 的网络总调度器分别向各自所在数据中心的网络控制器发出创建虚拟子网的请求,请求中携带参数:租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号、是否启动 DHCP 分配 IP 地址、分割后设置的子网网关 IP 地址、分割后的 DHCP IP 地址池。

[0159] 需要说明的是:由于在数据中心 2 和数据中心 3 中创建虚拟子网的过程相同,为了描述方便,以下的步骤(22)~步骤(28)仅以数据中心 2 为例进行阐述。

[0160] (22) 数据中心 2 的网络控制器判断创建虚拟子网请求参数中虚拟子网已经存在,且无数据中心列表,则在本数据中心内继续虚拟子网虚拟业务网关创建工作。如果子网网关 IP 地址有效,表明需要申请网关服务,所述网络控制器向本数据中心的 VSG 管理器发送

网关服务请求,参数中携带子网 ID、是否启动 DHCP 分配 IP 地址、分割后 DHCP IP 地址池、VxLAN ID、分割后设置的子网网关 IP 地址。

[0161] (23) 所述 VSG 管理器向其管理的软件业务网关增加网关服务和 DHCP 服务。增加网关服务和 DHCP 服务有两种方式:

[0162] 方式一:如果已经具备网关服务能力的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务(如,Route 服务、DHCP Server 服务、VxLAN 网关服务、NAT 服务、ACL 服务、VPN 服务、软件负载均衡服务中的一种服务或者几种服务),则在该已有的 VSG 虚拟机中增加网卡,设置该网卡的 IP 地址为创建虚拟子网请求参数中的网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池;

[0163] 方式二:新创建 VSG 虚拟机,在该虚拟机中增加两块网卡,一块网卡用于连接外网,从外网获取 IP 地址或者配置固定 IP 地址,另一块网卡用于连接虚拟子网,配置该网卡的 IP 地址为创建虚拟子网请求参数中的网关 IP 地址,并启动 DHCP 服务,配置分割后 DHCP IP 地址池。

[0164] VSG 管理器建立虚拟子网 ID 和新增加 VSG 虚拟机、网关服务及网卡的映射关系。VSG 管理器向本数据中心的网络控制器返回网关服务请求的结果,包括 VSG 虚拟机标识、新增加网卡的 MAC 地址。

[0165] (24) 所述网络控制器收到 VSG 管理器返回的网关服务消息,判断返回的 VSG 虚拟机是否已经加入虚拟子网。如果 VSG 虚拟机未加入虚拟子网,则在虚拟子网上创建一个端口:所述网络控制器向 VSG 虚拟机所在的主机下发 VSG 虚拟机绑定到虚拟子网的命令,携带参数包括:租户标识、虚拟网络 ID、虚拟子网 ID、VSG 虚拟机的 MAC 地址、端口状态为激活态、网关 IP 地址、VSG 虚拟机标识。VSG 虚拟机所在的主机部署了网络控制器的代理,该网络控制器的代理接收到命令后负责在该主机上管理虚拟交换机(virtual Switch,简称为 vSwitch),在 vSwitch 上申请一个端口,并绑定 VSG 虚拟机的 MAC 地址,打上虚拟子网 VxLAN 的标识,并对这个端口分配一个唯一的端口标识:端口 ID。成功绑定虚拟机后,所述网络控制器把端口信息写入本地数据库中。

[0166] (25) 所述网络控制器判断是否开启 DHCP 服务参数,如果需要申请 DHCP 服务,则所述网络控制器向本数据中心的 VSG 管理器发送 DHCP 服务请求,参数中携带子网 ID、分割后的 DHCP IP 地址池。

[0167] (26) 所述 VSG 管理器根据子网 ID 获取该子网对应的 VSG 虚拟机。如果不能够获得 VSG 虚拟机,表明需要为该虚拟子网分配 VSG 虚拟机提供 DHCP 服务。申请 VSG 虚拟机有如下两种方式:

[0168] 方式一:如果已经存在的 VSG 虚拟机还有能力为新的虚拟子网提供 VSG 服务(如,Route 服务、DHCP Server 服务、VxLAN 网关服务、NAT 服务、ACL 服务、VPN 服务、软件负载均衡服务中的一种服务或者几种服务),则在该已有的 VSG 虚拟机中增加网卡,设置该网卡的 IP 地址为 IP 地址池的第一个 IP 地址。

[0169] 方式二:新创建 VSG 虚拟机,在该虚拟机中增加一网卡,该网卡用于连接虚拟子网,配置该网卡的 IP 地址为 IP 地址池的第一个 IP 地址。

[0170] 在该 VSG 虚拟机下启动 DHCP 服务,设置该 DHCP 服务可分配的 IP 地址为分割后的 DHCP IP 地址池, DHCP 服务绑定虚拟子网的网关 IP 地址,为该虚拟子网提供 DHCP 服务。所

述 VSG 管理器启动 DHCP 服务。所述 VSG 管理器向所述网络控制器返回 DHCP 服务请求的结果,包括 VSG 虚拟机标识、新增加网卡的 MAC 地址。

[0171] (27) 所述网络控制器收到返回结果,并判断 VSG 虚拟机是否已经加入虚拟子网,如果没有加入虚拟子网,则进行第(24)步 VSG 虚拟机加入虚拟子网的操作。之后,所述网络控制器向本数据中心的网络总调度器返回虚拟子网信息。

[0172] (28) 所述网络总调度器向数据中心 1 的网络控制器返回虚拟子网信息。

[0173] 同理,对于数据中心 3 来说,执行完上述第(22)步-第(27)步后,数据中心 3 的网络总调度器也向数据中心 1 的网络控制器返回虚拟子网创建结果。

[0174] (29) 数据中心 1 的网络控制器向数据中心 1 的网络总调度器返回虚拟子网信息。

[0175] (30) 数据中心 1 的网络总调度器向数据中心 1 的跨数据中心调度器返回虚拟子网信息,包括:租户标识、网络 ID、子网名称、子网 ID、子网 CIDR、子网 IP 地址版本号、子网网关 IP 地址、是否启动 DHCP 分配 IP 地址、IP 地址池、VxLAN ID 以及跨数据中心子网切分模式。

[0176] (31) 数据中心 1 的跨数据中心调度器检查数据中心列表中 {[网络资源规格参数]、[数据中心标识,虚拟机描述参数列表]} 中的虚拟机信息,向每一个需要创建虚拟机的数据中心下发虚拟机创建任务,这些任务是并行下发的。当有多个不同规格的虚拟机需要创建的时候,所述跨数据中心调度器分别下发多个创建虚拟机的任务。当需要批量创建相同规格的虚拟机时,调用批量创建的接口,带上需创建虚拟机的数量。图 2 中示例为向数据中心 2 和数据中心 3 中的计算资源调度器下发单个虚拟机创建任务,携带参数包括:虚拟机规格标识(通过虚拟机规格标识可以获得虚拟机的规格信息,如虚拟机 CPU 主频、CPU 个数、内存大小、磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级(用于选择磁盘存储类型:固态硬盘、直连硬盘、外置存储等)、每个磁盘备份模式、虚拟网卡个数、每个虚拟网卡的带宽)、虚拟机镜像链接、物理计算资源还是虚拟计算资源、CPU 类型(如果是物理计算资源,指定物理机的 CPU 类型)、HyperVisor 类型 CPU 类型(如果是请求虚拟计算资源,指定 HyperVisor 类型)、是否启动 HA、是否启动 FT、是否需要容灾、容灾类型(同步容灾、异步容灾)、容灾模式(双活、主备模式)、虚拟机启动参数、虚拟机配置参数、安全组、指定主机集群、待加入的虚拟子网标识列表、其他参数。

[0177] 需要说明的是:由于在数据中心 2 和数据中心 3 中创建虚拟机并加入虚拟子网的过程相同,为了描述方便,以下的步骤(32)-步骤(49)仅以数据中心 2 为例进行阐述。

[0178] (32) 数据中心 2 的计算资源调度器创建虚拟机逻辑对象,为虚拟机创建唯一标识,记录虚拟机的参数信息。

[0179] (33) 所述计算资源调度器把虚拟机信息写入数据中心 2 的分布式数据库,不需要同步到其他数据中心。

[0180] (34) 所述分布式数据库向所述计算资源调度器返回数据库更新结果。

[0181] (35) 所述计算资源调度器向数据中心 1 的跨数据中心调度器返回虚拟机标识。

[0182] (36) 所述数据中心 1 的跨数据中心调度器向数据中心 2 的计算资源调度器轮询虚拟机的创建状态,携带参数为虚拟机标识,直到虚拟机创建成功后结束虚拟机创建任务。

[0183] 需要说明的是,所述跨数据中心调度器可以周期性的向数据中心 2 和 3 的计算资源调度器轮询虚拟机的创建状态,直至本次资源发放事务结束。

[0184] (37)数据中心 2 的计算资源调度器对计算资源进行调度:根据虚拟机规格标识获得虚拟机的规格:虚拟机 CPU 主频、CPU 个数、内存大小、磁盘个数、每个磁盘的存储大小,判断虚拟机建立的资源条件:磁盘访问 QoS 等级,物理计算资源还是虚拟计算资源、CPU 类型(如果是物理计算资源,指定物理机的 CPU 类型)、HyperVisor 类型 CPU 类型(如果是请求虚拟计算资源,指定 HyperVisor 类型),选定满足所有条件的主机。如果参数中指定主机集群,则在指定的主机集群中选择一个主机。在选中的主机上,根据参数的镜像链接,从数据中心 2 的镜像模板库中获得镜像,创建虚拟机。

[0185] 虚拟机根据虚拟机启动参数、虚拟机配置参数进行虚拟机的启动和配置。

[0186] (38)数据中心 2 的计算资源调度器逐一从待加入的虚拟子网标识列表中取出虚拟子网 ID,向本数据中心的网络总调度器发出创建端口的请求,请求中携带参数:网络 ID、子网 ID、虚拟机标识、加入该虚拟子网的虚拟机网卡的 MAC 地址及端口初始状态(激活态)。

[0187] (39)所述网络总调度器把创建端口的请求消息转发给数据中心 2 的网络控制器。

[0188] (40)所述网络控制器根据子网 ID 获得虚拟子网的信息:租户标识、网络 ID、子网名称、子网 ID、VxLAN ID、子网 CIDR、子网 IP 地址版本号。然后向虚拟机所在的主机下发虚拟机绑定到虚拟子网的命令,携带参数包括:租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、端口状态。虚拟机所在的主机上部署了网络的代理,该网络的代理接收到命令后负责在该主机上管理 vSwitch,在 vSwitch 上申请一个端口,并绑定虚拟机的 MAC 地址,打上虚拟子网 VxLAN 的标识,虚拟机加入到 VxLAN 二层网络后,从该虚拟子网的 DHCP 服务器中获得 IP 地址。所述网络控制器对这个端口分配一个唯一的端口标识:端口 ID。成功绑定虚拟机后,所述网络控制器把端口信息写入数据库中。所述网络控制器向本数据中心的网络总调度器返回端口信息:端口 ID,租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、虚拟机 IP 地址,端口状态。

[0189] (41)所述网络总调度器向本数据中心的计算资源调度器返回虚拟机加入到虚拟子网的端口信息:端口 ID,租户标识、网络 ID、子网 ID、虚拟机标识、虚拟机的 MAC 地址、虚拟机 IP 地址,端口状态。

[0190] (42)所述计算资源调度器获得虚拟机加入到虚拟子网的端口信息,判断虚拟机成功加入到虚拟子网后,把虚拟机 ID、端口 ID 之间的映射关系写入数据中心 2 的分布式数据库,并同步到其他数据中心。

[0191] (43)数据中心 2 的分布式数据库向本数据中心的计算资源调度器返回数据库更新结果。

[0192] (44)所述计算资源调度器根据磁盘个数、每个磁盘的存储大小、磁盘访问 QoS 等级向数据中心 2 的存储资源调度器发出创建存储资源的请求,请求中携带参数:虚拟机标识、存储大小、磁盘访问 QoS 等级、虚拟机所在的主机集群。

[0193] (45)所述存储资源调度器根据参数:虚拟机标识、存储大小、磁盘访问 QoS 等级及虚拟机所在的主机集群,选择具体的磁盘存储类型:固态硬盘、直连硬盘、外置存储等。在选定的磁盘上创建满足磁盘访问 QoS 等级、存储大小的卷。生成唯一的卷标识:卷 ID。向所述计算资源调度器返回卷标识。

[0194] (46)所述计算资源调度器把卷挂载到虚拟机标识所代表的虚拟机,建立卷标识和虚拟机之间的关联关系,生成唯一的关联关系标识。

[0195] (47) 如果成功挂载了卷,所述计算资源调度器把虚拟机 ID、卷标识和关联关系标识写入数据中心 2 的分布式数据库,并同步到其他数据中心。

[0196] (48) 所述分布式数据库向所述计算资源调度器返回数据库更新结果。如果还需要创建其他卷,则重复第(45)–(47)步,直到所有卷创建挂载成功。

[0197] (49) 当数据中心 1 的跨数据中心调度器通过轮询虚拟机创建任务执行结果,获知数据中心 2 和 3 分别完成上述资源发放事务后,数据中心 1 的跨数据中心调度器结束第一个数据中心列表的资源发放事务,如果在事务中任何一步资源申请失败,则释放本事务中所有分配的资源。然后选取数据中心列表集中的次优数据中心列表,重复资源发放事务,重复次数以系统配置为上限。

[0198] (50) 数据中心 1 的跨数据中心调度器返回资源发放结果,如果成功,返回发放的计算资源、存储资源和网络资源。

[0199] 本发明实施例提供的资源发放方法,通过部署在每个数据中心中的分布式云操作系统之间的交互,执行相应的资源发放步骤,使所有数据中心的基础设施资源成为一个统一的“逻辑资源池”,实现了资源的全局拉通,可以进行跨物理数据中心的资源弹性调度。从而提高了数据中心的资源利用率、简化了管理维护运营、降低了网络连接故障或流量拥塞发生的概率。

[0200] 应理解,在本发明实施例中,“与 A 相应的 B”表示 B 与 A 相关联,根据 A 可以确定 B。但还应理解,根据 A 确定 B 并不意味着仅仅根据 A 确定 B,还可以根据 A 和 / 或其它信息确定 B。

[0201] 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、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。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。

[0202] 在本申请所提供的实施例中,应该理解到,所揭露的系统,可以通过其它的方式实现。例如,以上所描述的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。另外,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,也可以是电的,机械的或其它的形式连接。

[0203] 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发明实施例方案的目的。

[0204] 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,也可以是两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0205] 所述集成的单元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

时,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,或者该技术方案的全部或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、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0206] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以权利要求的保护范围为准。

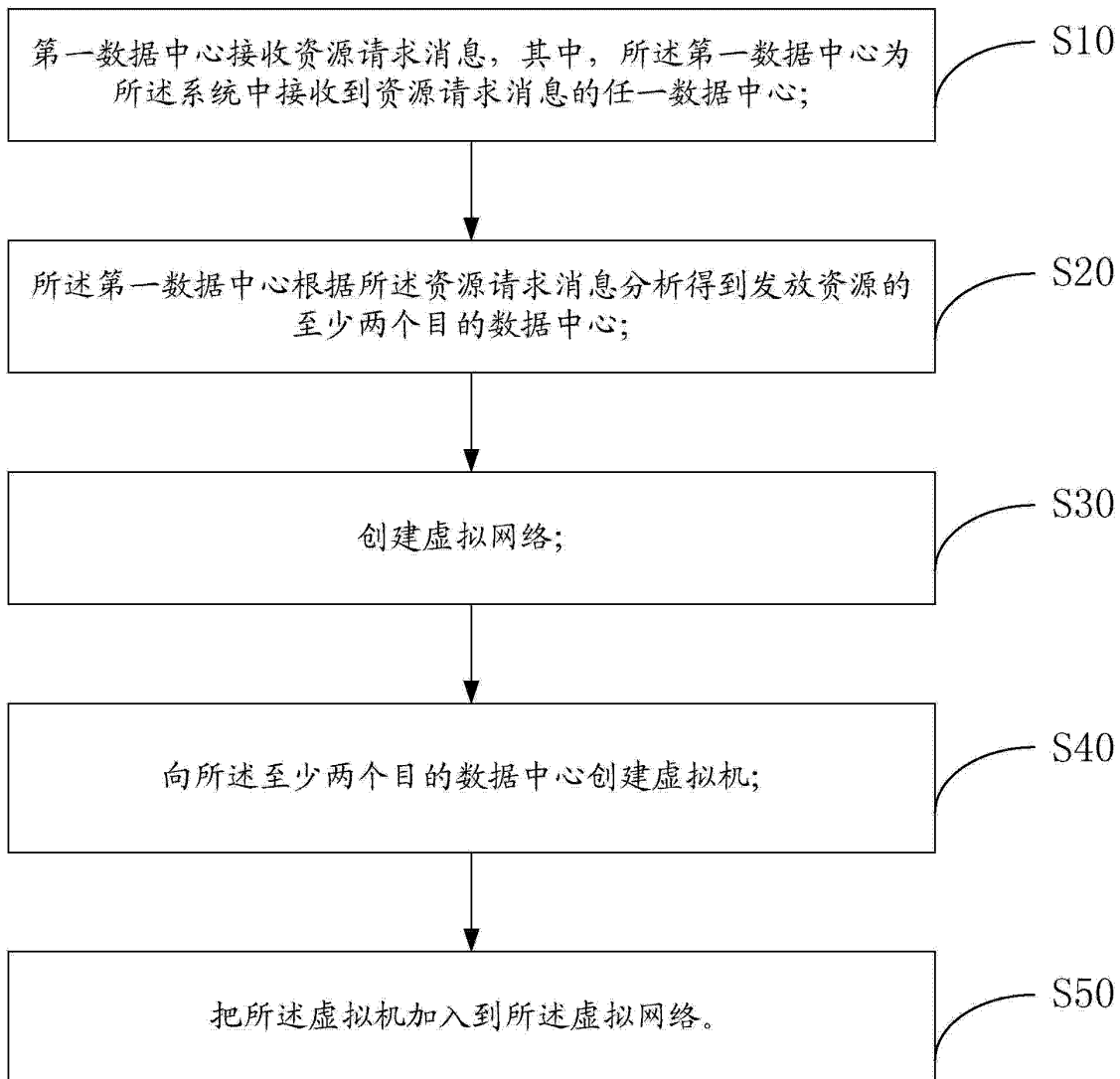


图 1

Abstract:

Embodiments of the present invention provides a resource provisioning method, including: receiving, by a first data center, a resource request message, where the first data center is any data center that receives the resource request message in a system; analyzing, by the first data center according to the resource request message, and obtaining at least two destination data centers that provision resources; creating a virtual network; creating a virtual machine for the at least two destination data centers; and adding the virtual machine to the virtual network. By implementing this method, the resource utilization of the data center is improved, administration, maintenance, and operation are simplified, and an occurrence probability of network connection fault or traffic congestion is reduced.