



(12) 发明专利申请

(10) 申请公布号 CN 117355817 A

(43) 申请公布日 2024. 01. 05

(21) 申请号 202180098502.0

(22) 申请日 2021.07.21

(30) 优先权数据

17/229,709 2021.04.13 US

(85) PCT国际申请进入国家阶段日

2023.11.22

(86) PCT国际申请的申请数据

PCT/US2021/042619 2021.07.21

(87) PCT国际申请的公布数据

W02022/220858 EN 2022.10.20

(71) 申请人 甲骨文国际公司

地址 美国加利福尼亚

(72) 发明人 E·奥托 M·科比 A·库卡尼

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

专利代理师 冯薇

(51) Int.Cl.

G06F 3/06 (2006.01)

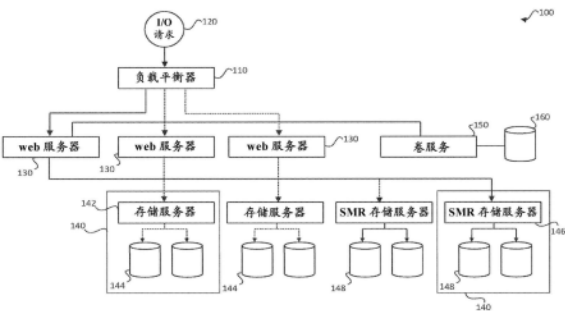
权利要求书3页 说明书23页 附图11页

(54) 发明名称

区片段驱动器管理

(57) 摘要

提供了用于管理分布式存储系统中的数据存储的技术。一种方法可以包括由计算机系统接收将数据写入卷的请求。该方法可以包括由计算机系统识别被映射到卷的区片段。区片段可以包括多个区。该方法可以包括由计算机系统识别指示区片段中的区中的写入位置的片段指针。该方法可以包括由计算机系统从写入位置开始将数据写入区片段中的该多个区中的一个或多个区。该方法还可以包括由计算机系统根据区片段中数据的结束点来更新片段指针。



1. 一种方法, 包括:
 - 由计算机系统接收将数据写入卷的请求;
 - 由计算机系统识别被映射到卷的区片段, 该区片段包括多个区;
 - 由计算机系统识别指示区片段中的区中的写入位置的片段指针;
 - 由计算机系统从写入位置开始将数据写入区片段中的所述多个区中的一个或多个区;以及
 - 由计算机系统根据区片段中数据的结束点来更新片段指针。
2. 如权利要求1所述的方法, 还包括:
 - 确认请求中识别出的数据的尺寸;
 - 确认区的容量, 其中该容量与写入位置与该区的结束点之间的写入路径对应;
 - 将区的容量与数据的尺寸进行比较;
 - 根据数据的尺寸超出区的容量, 将数据分成第一块和第二块, 其中第一块的尺寸与区的容量对应;
 - 从写入位置开始写入第一块; 以及
 - 将第二块写入区片段中的后续区。
3. 如权利要求1或权利要求2所述的方法, 其中更新片段指针包括:
 - 根据区片段中的数据的结束点与该区片段的容量一致, 将该区片段标记为满区片段;
 - 识别被映射到所述卷的新区片段; 以及
 - 将片段指针更新为新区片段的起始点。
4. 如前述权利要求中的任一项所述的方法, 其中识别片段指针包括扫描被映射到所述卷的区片段以查找空的或开放的区。
5. 如权利要求4所述的方法, 其中识别片段指针还包括:
 - 根据被映射到所述卷的区片段不包括空的或开放的区, 向该卷分配新区片段, 其中片段指针是该新区片段中的第一区的起始点。
6. 如前述权利要求中的任一项所述的方法, 还包括:
 - 确认区片段的使用分数;
 - 将使用分数与使用阈值进行比较;
 - 根据使用分数不满足使用阈值:
 - 识别区片段中的所述多个区中的所存储的数据;
 - 将所存储的数据从一个或多个源区映射到所述多个区中的一个或多个空的或开放的区;
 - 将所存储的数据写入空的或开放的区; 以及
 - 将源区重置为开放状态。
7. 如权利要求6所述的方法, 其中识别所述多个区中的所存储的数据包括:
 - 接收描述区片段到卷的映射的映射元数据; 以及
 - 根据映射元数据定位所述多个区中的所存储的数据。
8. 如前述权利要求中的任一项所述的方法, 其中所述多个区在叠瓦式磁记录存储系统中实现。
9. 一种系统, 包括:

存储器,被配置为存储计算机可执行指令;以及

一个或多个处理器,被配置为访问存储器并执行计算机可执行指令以至少:

接收将数据写入卷的请求;

识别被映射到卷的区片段,该区片段包括多个区;

识别指示区片段中的区中的写入位置的片段指针;

从写入位置开始将数据写入区片段中的所述多个区中的一个或多个区;以及

根据区片段中数据的结束点来更新片段指针。

10. 如权利要求9所述的系统,其中所述一个或多个处理器还被配置为执行计算机可执行指令以进一步至少:

确认请求中识别出的数据的尺寸;

确认区的容量,其中该容量与写入位置与该区的结束点之间的写入路径对应;

将区的容量与数据的尺寸进行比较;

根据数据的尺寸超出区的容量,将数据分成第一块和第二块,其中第一块的尺寸与区的容量对应;

从写入位置开始写入第一块;以及

将第二块写入区片段中的后续区。

11. 如权利要求9或权利要求10所述的系统,其中更新片段指针包括:

根据区片段中的数据的结束点与该区片段的容量一致,将该区片段标记为满区片段;

识别被映射到所述卷的新区片段;以及

将片段指针更新为新区片段的起始点。

12. 如权利要求9至11中的任一项所述的系统,其中识别片段指针包括扫描被映射到所述卷的区片段以查找空的或开放的区。

13. 如权利要求12所述的系统,其中识别片段指针还包括:

根据被映射到所述卷的区片段不包括空的或开放的区,向该卷分配新区片段,其中片段指针是该新区片段中的第一区的起始点。

14. 如权利要求9至13中的任一项所述的系统,其中所述一个或多个处理器还被配置为执行计算机可执行指令以进一步至少:

确认区片段的使用分数;

将使用分数与使用阈值进行比较;

根据使用分数不满足使用阈值:

识别区片段中的所述多个区中的所存储的数据;

将所存储的数据从一个或多个源区映射到所述多个区中的一个或多个空的或开放的区;以及

将所存储的数据写入空的或开放的区;以及

将源区重置为开放状态。

15. 如权利要求14所述的系统,其中识别所述多个区中的所存储的数据包括:

接收描述区片段到卷的映射的映射元数据;以及

根据映射元数据定位所述多个区中的所存储的数据。

16. 一种存储可由计算机系统执行的指令的计算机可读存储介质,指令在由计算机系

统的一个或多个处理器执行时使所述一个或多个处理器执行操作,包括:

由计算机系统接收将数据写入卷的请求;

由计算机系统识别被映射到卷的区片段,该区片段包括多个区;

由计算机系统识别指示区片段中的区中的写入位置的片段指针;

由计算机系统从写入位置开始将数据写入区片段中的所述多个区中的一个或多个区;

以及

由计算机系统根据区片段中数据的结束点来更新片段指针。

17. 如权利要求16所述的计算机可读存储介质,其中操作还包括:

确认请求中识别出的数据的尺寸;

确认区的容量,其中该容量与写入位置与该区的结束点之间的写入路径对应;

将区的容量与数据的尺寸进行比较;

根据数据的尺寸超出区的容量,将数据分成第一块和第二块,其中第一块的尺寸与区的容量对应;

从写入位置开始写入第一块;以及

将第二块写入区片段中的后续区。

18. 如权利要求16或权利要求17所述的计算机可读介质,其中更新片段指针包括:

根据区片段中的数据的结束点与该区片段的容量一致,将该区片段标记为满区片段;

识别被映射到所述卷的新区片段;以及

将片段指针更新为新区片段的起始点。

19. 如权利要求16至18中的任一项所述的计算机可读介质,其中识别片段指针包括扫描被映射到所述卷的区片段以查找空的或开放的区。

20. 如权利要求19所述的计算机可读介质,其中识别片段指针还包括:

根据被映射到所述卷的区片段不包括空的或开放的区,向该卷分配新区片段,其中片段指针是该新区片段中的第一区的起始点。

区片段驱动器管理

[0001] 相关申请的交叉引用

[0002] 本申请要求于2021年4月13日提交的编号为17/229,709、标题为“ZONE SEGMENT DRIVE MANAGEMENT”的美国非临时申请的权益和优先权。出于所有目的,美国非临时申请17/229,709的内容通过全文引用并入本文。

背景技术

[0003] 基于云的平台为用户的数据提供了可扩展且灵活的计算资源。此类基于云的平台(也称为基础设施即服务(IaaS))可以围绕客户的数据提供整套云解决方案,例如用于创作变换、加载数据和呈现数据的解决方案。

[0004] 叠瓦式磁记录(SMR)通过部分重叠或搭叠记录轨道来增加硬盘驱动器上的存储容量。重叠的轨道可以被分组为区。通过以保留读取路径的方式搭叠轨道,即使轨道已被部分覆盖,仍可读取轨道。相比之下,由于搭叠,虽然读操作仍可以在区内随机执行,但是写操作将需要顺序地应用于整个区,而不是应用于区内的范围。

发明内容

[0005] 提供了用于管理叠瓦式磁记录驱动器上的云资源数据的对象存储的技术(例如,方法、系统、存储可由一个或多个处理器执行的代码或指令的非暂态计算机可读介质)。

[0006] 在某些实施例中,一种方法可以包括由计算机系统接收将数据写入卷的请求。该方法可以包括由计算机系统识别被映射到卷的区片段(zone segment)。区片段可以包括多个区。该方法可以包括由计算机系统识别指示区片段中的区中的写入位置的片段指针。该方法可以包括由计算机系统从写入位置开始将数据写入区片段中的多个区中的一个或多个区。该方法还可以包括由计算机系统根据区片段中的数据的结束点来更新片段指针。

[0007] 在一些实施例中,该方法还可以包括确认请求中识别出的数据的尺寸。该方法可以包括确认区的容量,其中该容量与写入位置与区的结束点之间的写入路径对应。该方法可以包括将区的容量与数据的尺寸进行比较。该方法可以包括根据数据的尺寸超过区的容量而将数据分成第一块和第二块。第一块的尺寸可以与区的容量对应。该方法可以包括从写入位置开始写入第一块。该方法还可以包括将第二块写入区片段中的后续区。

[0008] 在一些实施例中,更新片段指针可以包括根据区片段中的数据的结束点与区片段的容量一致而将区片段标记为满区片段。更新片段指针还可以包括识别被映射到所述卷的新区片段,以及将片段指针更新为新区片段的起始点。识别片段指针可以包括扫描被映射到所述卷的区片段以查找空的或开放的区。

[0009] 在一些实施例中,一种计算机系统包括一个或多个处理器以及与该一个或多个处理器通信的存储器,存储器被配置为存储计算机可执行指令,其中执行计算机可执行指令使得所述一个或多个处理器执行上述方法或其变体的一个或多个步骤。识别片段指针还可以包括根据被映射到所述卷的区片段不包括空的或开放的区而向该卷分配新区片段。片段指针可以是新区片段中的第一区的起始位置。

[0010] 在一些实施例中,该方法可以包括确认区片段的使用分数。该方法可以包括将使用分数与使用阈值进行比较。该方法可以包括根据使用分数不满足使用阈值而识别区片段中的所述多个区中的所存储的数据,将所存储的数据从一个或多个源区映射到所述多个区中的一个或多个空的或开放的区,将所存储的数据写入空的或开放的区,并将源区重置为开放状态。识别所述多个区中的所存储的数据可以包括接收描述区片段到卷的映射的映射元数据并且根据映射元数据定位所述多个区中的所存储的数据。

[0011] 在一些实施例中,多个区可以在叠瓦式磁记录存储系统中实现。

[0012] 在某些实施例中,一种计算机系统包括一个或多个处理器以及存储指令的非暂态计算机可读存储器,该指令在由一个或多个处理器执行时使该一个或多个处理器执行上述方法或其变体的一个或多个操作。

[0013] 在某些实施例中,计算机可读存储介质存储计算机可执行指令,该指令在被执行时使该计算机系统的一个或多个处理器执行上述方法或其变体的一个或多个操作。

附图说明

[0014] 图1图示了根据一个或多个实施例的用于管理对象存储服务器的示例系统。

[0015] 图2图示了根据一个或多个实施例的示例叠瓦式磁记录驱动器。

[0016] 图3图示了根据一个或多个实施例的用于将对象数据写入区片段的示例流程。

[0017] 图4图示了根据一个或多个实施例的用于从区片段读取数据的示例流程。

[0018] 图5图示了根据一个或多个实施例的用于从区片段删除数据的示例流程。

[0019] 图6图示了根据一个或多个实施例的用于对区片段进行碎片整理(defragment)的示例流程。

[0020] 图7是图示根据至少一个实施例的用于实现云基础设施即服务系统的一种模式的框图。

[0021] 图8是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一种模式的框图。

[0022] 图9是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一种模式的框图。

[0023] 图10是图示根据至少一个实施例的用于实现云基础设施即服务系统的另一种模式的框图。

[0024] 图11是图示根据至少一个实施例的示例计算机系统的框图。

具体实施方式

[0025] 在下面的描述中,将描述各种实施例。为了解释的目的,阐述了具体配置和细节以便提供对实施例的透彻理解。但是,对于本领域技术人员来说也将显而易见的是,可以在没有具体细节的情况下实践这些实施例。此外,可以省略或简化众所周知的特征,以免使所描述的实施例变得模糊。

[0026] 基于云的平台为用户提供了可扩展且灵活的计算资源。此类基于云的平台(也称为基础设施即服务(IaaS))可以围绕客户的数据提供整套云解决方案,例如用于创作变换、加载数据和呈现数据的解决方案。在可以是后端数据存储系统的一部分的数据复制和备份

系统中,叠瓦式磁记录 (SMR) 驱动器可以增加数据存储系统的存储容量。SMR驱动器部分地重叠或搭叠记录轨道,从而可以允许将更多轨道写入给定尺寸的硬盘。重叠的轨道可以被分组为区,这些区可以被组织成区片段。通过以保留读取路径的方式搭叠轨道,即使轨道已被部分覆盖,仍可读取轨道。相比之下,由于搭叠,可以在仅追加过程中应用写操作,通过该过程可以读取部分被重写的数据,但不能重写这些数据。这样,由于写入SMR驱动器时所涉及的附加操作,对SMR驱动器的访问可以由块接口而不是文件系统提供,如下所述。

[0027] 在分布式存储系统中,用户数据可以存储在卷存储装置(例如,块卷存储装置)和对象存储装置中。用户可以直接与卷存储系统交互,而数据库系统的后端存储子系统可以使用对象存储装置,例如用于数据复制、备份或其它数据存储(例如,以补充数据库系统容量)。SMR驱动器在数据复制或备份系统中的实施方式受到向SMR驱动器写入和重写数据的相对复杂性的限制。如参考图2更详细地描述的,当系统尝试向区写入超出区中的可用容量的数据时,在区中组织部分重写的轨道可能造成效率低下。例如,在常规的SMR驱动器中,系统可以拒绝向区写入超出区容量的数据的请求。为了解决这个限制,区可以被组织成区片段,并且管理数据存储的系统可以将写请求寻址到区片段而不是个别的区。以这种方式,可以识别区片段中的写入位置,并且可以跨区片段中的多个区写入由写请求引用的数据。

[0028] 在一些示例中,块存储常常用在期望快速、高效和可靠的数据运输的地方。块存储将数据分解成块,并且可以将这些块存储为单独的片(piece)(例如,每个块都有其自己的标识符)。然后,这些块可以跨不同的系统存储并且每个块可以被配置(例如,分区)为与不同的系统一起工作。可替代地,对象存储将数据文件分成多个片(例如,对象),然后将它们存储在单个储存库中,该储存库可以分布在多个联网系统中。在块存储的上下文中,卷(例如,块卷)可以是可分离(例如,物理上或虚拟上)的块存储设备,其允许开发人员动态扩展实例的存储容量。为了增加或减小块卷存储装置的尺寸,可以分别向上或向下旋转块卷,而不需要太多努力。在数据库系统实现卷存储和对象存储两者的情况下,I/O请求可以通过卷标识符(例如,块卷标识符)来引用数据。通过存储在包括作为数据库系统的一部分的元数据数据库中的卷元数据,卷标识符可以被映射到对象存储系统中的存储位置。在写请求的情况下,数据库系统可以识别SMR系统中存储被映射到写请求中所引用的卷的数据的区片段。一旦被识别出,数据库系统就可以识别区片段指针,该指针描述被包括在区片段中的区中的可以写入新数据的写入位置。SMR系统可以写入该区,而不是重写该区,可以将数据分成要跨多个区写入的块,如果某个区不具有存储数据的能力,那么可以拒绝该请求,或者可以分配(一个或多个)新区片段以存储溢出的数据。以这种方式,SMR系统可以以通过使用SMR存储来增加对象存储系统的有效容量的方式将数据写入硬盘驱动器,同时限制由于数据的尺寸超出区的容量而造成的区写入失败的影响。以类似的方式,还可以参考被映射到存储在区片段中的数据的卷元数据使用SMR系统来实现读取、删除和碎片整理技术。

[0029] 在说明性示例中,数据库系统可以被配置为存储云资源数据作为分布式存储系统的一部分。数据库系统可以包括SMR系统。SMR系统可以包括与一个或多个SMR数据库系统通信的SMR存储服务器,该SMR数据库系统可以是配置用于使用以区片段组织的区的SMR存储的硬盘驱动器或者包括该硬盘驱动器。在这个示例中,系统可以接收输入/输出(I/O)请求,诸如写请求(也称为“put”请求)、读请求(也称为“get”请求)、以及其它I/O过程,包括但不限于删除请求或碎片整理操作。作为SMR配置的一部分,I/O请求可以使数据库系统实现

改进的过程以识别和管理存储在区片段中的SMR系统中的数据。

[0030] 用后端存储子系统实现I/O过程的过程改进了数据复制、存储和备份系统的整体性能,至少因为它允许来自SMR配置的存储系统的高效I/O。例如,SMR系统可以通过至少部分地搭叠先前写入的磁轨来提供高达25%或更多的增加的存储容量,而无需向系统添加硬盘驱动器。通过允许对在SMR驱动器上做出的I/O请求提供服务的改进速度、效率和成功率,将SMR区组织成区片段可以改进I/O过程。此外,通过参考接下来的附图描述的管理区驱动器,数据库系统可以通过允许数据库系统使用主机管理的驱动器来降低系统复杂性,主机管理的驱动器由参考卷元数据以将对象数据映射到卷数据的集成块接口控制。

[0031] 图1图示了根据一个或多个实施例的用于管理对象存储服务器的示例系统100。数据库系统100可以被配置为跨多个存储系统分发和存储用户数据,该多个存储系统可以实现各种磁记录技术(例如,垂直磁记录(PMR)、SMR等),作为后端数据存储子系统的一部分。以这种方式,数据库系统100可以作为数据复制、将卷数据恢复到块卷系统或者对将区组织成区片段的SMR系统进行碎片整理的一部分对存储在对象存储装置中的数据的I/O请求提供服务,作为促进分布式数据存储系统中改进的I/O操作的方法。

[0032] 在一些实施例中,系统100可以包括负载均衡器110,其可以被配置为接收I/O请求120并将I/O请求120发送到web服务器130。负载均衡器110可以将I/O请求分布在作为分布式存储系统的一部分的多个web服务器130上,该分布式存储系统可以位于多个不同的物理位置和/或可以包括在单个物理位置中的多个web服务器130。类似地,web服务器130可以与一个或多个存储系统140进行通信。存储系统140可以被配置为以PMR或SMR配置存储数据。例如,存储系统140可以包括与一个或多个PMR驱动器144进行通信的存储服务器142。此外,web服务器130可以与包括SMR存储服务器146的存储系统进行通信,SMR存储服务器146与一个或多个SMR驱动器148进行通信,如参考图2更详细描述的。系统100可以被配置为例如通过SMR存储服务器146执行软件来对存储在SMR驱动器148中的数据执行I/O请求120。

[0033] 系统100可以通过参考卷服务150将存储在存储系统140中的数据映射到卷数据。以这种方式,I/O请求120可以包括对卷标识符的引用,该卷标识符可以被映射到卷元数据数据库160中的存储系统140的存储装置中的位置。在示例中,读请求可以引用卷标识符,该卷标识符可以被映射到卷元数据数据库160中的条目中的区片段、片段指针和读取尺寸。

[0034] 以这种方式,系统100可以将SMR驱动器148作为分布式数据存储系统的一部分进行管理。有利地,系统100可以允许后端数据存储子系统受益于SMR驱动器148的增加的存储容量,同时维持与前端块卷系统的集成。如参考图3-图6更详细地描述的,数据库系统100可以接收I/O请求120,例如作为数据复制或备份的一部分,并且可以通过对PMR驱动器144和/或SMR驱动器148进行读取、吸入或碎片整理来实现I/O请求120。

[0035] 图2图示了根据一个或多个实施例的示例叠瓦式磁记录SMR驱动器200。SMR驱动器200的物理过程和数据结构基于重写SMR驱动器200的盘上的磁轨210。在PMR驱动器中,磁轨不会被重写。这样,存储在PMR驱动器中的数据可以被随机存取(例如,PMR驱动器可以被用作随机存取硬盘驱动器存储装置)。相比之下,能够通过追加数据而不是随机存取进行写入的SMR驱动器通常被配置用于高效读取不太可能被修改的数据。

[0036] 可以使用块系统代替文件系统来存储SMR驱动器200中的数据,作为仅追加(append only) I/O系统的一部分。这样,数据块可以被写入固定尺寸(例如,256MB)的区

220,这些区可以被组织成区片段230。物理上,每个区可以由具有固定写入宽度240的多个磁轨210来描述,这些轨道210被部分地重写以提供数据轨道250。数据轨道可以由数据轨道间距(pitch)260来描述,数据轨道间距260可以是SMR驱动器200的参数,该参数取决于例如SMR驱动器200的读/写头270的规格。随着数据轨道间距260变窄,SMR驱动器200的总容量增加,但受到包括读/写头270准确读取数据轨道250的能力的因素的限制。

[0037] 磁轨210被示为具有固定宽度,以便展示区220和区段230的分层组织。与在PMR驱动器中所做的那样将数据索引到离散轨道中的单个位置不同,SMR驱动器200可以将数据组织在跨多个磁轨210延伸的区220中,通过指示区片段230的起始点的指针、区220的起始点和/或区220内要找到数据的位置进行索引。例如,指针可以包括写指针,描述区220中的最后写入的位置,其中后续的写操作将追加新数据。

[0038] 在一些情况下,区220可以由可以管控数据是否可以被写入区220的一个或多个状态来描述。例如,区220的状态可以包括但不限于“空(empty)”、“满(full)”、“打开(open)”、“关闭(closed)”或“完成(finished)”。空状态可以指示能够从区220的磁轨210的起始处写入区220(例如,区220的写指针位于区220的起始处)。类似地,满状态可以指示区220不可用于附加的写操作(例如,区220的写指针位于区220的末尾或区220的末尾附近)。打开状态可以指示区220可用于写操作(例如,区220具有可用于将数据写入区220的磁轨210的容量,并且SMR驱动器220具有被指派为实现写操作的资源)。相比之下,关闭状态可以指示区220不可用于写操作,但是可以被打开。完成状态可以指示数据将不会被写入区220(例如,区220的写指针可以被移动到区的末尾以防止进一步的写操作)。可以通过将区220设置为空状态(例如,通过将写指针移动到区220的起始处)来打开完成的区。

[0039] 将区220组织成区片段230可以允许区220以可以定义片段指针的方式一起被管理,描述区片段230的写入位置,而不是每个区的写入位置。在区片段230的区220内,特定于区的写入位置可以遵循上述状态,而片段指针可以描述区片段230的第一可用开放区中的写入位置。例如,区220可以顺序地组织在区片段230中。以这种方式,数据可以跨区片段230中的多个区220写入块中。以顺序方式组织区可以允许SMR驱动器200更好地适应引用其长度超出开放区的容量的数据的写请求。SMR驱动器200可以在区片段中打开多个区,并且可以基于单个写请求来定义多个写操作。如下面参考图3更详细地描述的,可以通过识别指示在区片段230的区220中开始写入数据的位置的片段指针来实现写操作。

[0040] 图3图示了根据一个或多个实施例的用于将对象数据写入区片段的示例流程300。该流程的操作可以被实现为硬件电路系统和/或被存储为计算机系统(诸如图1的系统100)的非暂态计算机可读介质上的计算机可读指令。在实现时,指令表示包括可由计算机系统的(一个或多个)处理器执行的电路系统或代码的模块。此类指令的执行将计算机系统配置为执行本文描述的特定操作。每个电路系统或代码与处理器结合执行相应的(一个或多个)操作。虽然以特定次序图示了操作,但是应当理解的是,特定次序不是必需的,并且一个或多个操作可以被省略、跳过和/或重新排序。

[0041] 在示例中,流程300包括操作310,其中计算机系统接收写请求。可以基于例如将数据存储在对对象存储装置中(而不是在块卷系统中)的用户请求从管理数据复制、存档或其它存储的分布式数据存储系统接收写请求(例如,图1的I/O请求120)。作为接收写请求的一部分,计算机系统可以将请求转发到其它联网系统(例如,作为IaaS基础设施系统的一部分),

如参考图1更详细地描述的。例如,写请求可以被转发到web服务器(例如,图1的web服务器130),该web服务器可以选择存储服务器(例如,图1的SMR存储服务器146)来处置写请求。

[0042] 在示例中,流程300可选地包括操作320,其中计算机系统扫描区片段以查找空的或开放的区。写请求可以包括具有定义长度的数据大块(chunk),其中该大块是要写入区片段(例如,图2的区片段230)的数据块。如参考图2更详细地描述的,区片段可以包括提供等同于区的容量的整数倍的存储容量的多个区(例如,图2的区220)。区片段中的一个或多个区可以是空的或开放的,并且因此可以用于写操作。这些区可以按顺序组织在区片段中,使得开放的区后面可以是空区。应当理解的是,空区不一定指示磁轨为空白的区,而是可以指示该区的写入位置位于该区的起始处的区。

[0043] 在示例中,流程300可选地包括操作330,其中计算机系统向卷分配新区片段。在一些情况下,区片段的扫描可能无法识别区片段的终止区之前的空的或开放的区。在一些实施例中,系统可以向写请求中(例如,通过卷标识符)引用的卷分配新区片段,以增加可用于为所引用的卷写入数据的容量。新区片段可以从为了减少由于缺乏容量而被拒绝的写请求的比例而维护的区片段的空闲池中分配。在一些实施例中,新区片段可以从存储该区片段的同一SMR驱动器分配,或者可以从通过SMR存储服务器(例如,图1的SMR存储服务器146)协调的不同SMR驱动器分配。

[0044] 在示例中,流程300包括操作340,其中计算机系统识别片段指针。在区片段(或新区片段)内,系统可以通过定位由开放或空状态描述的区片段中的第一区来识别片段指针。例如,在区按顺序组织在区片段中的情况下,片段指针可以与第一开放区的写入位置对应。在另一个示例中,片段指针可以与第一空区的起始位置对应,其中该区片段不包括开放区。

[0045] 在示例中,流程300包括操作350,其中计算机系统根据片段指针将数据写入区片段中的区。从片段指针开始,系统可以将写请求引用的数据写入与片段指针对应的区。写操作可以包括系统初始检查与片段指针对应的区是否具有用于由写请求引用的整个大块的容量。在区容量不足的情况下,系统可以定义多个块,使得大块可以跨区片段的多个区被拆分。在说明性示例中,区片段中的区可以被顺序地组织,并且大块可以以使得第一块的尺寸与对应于片段指针的区的写入容量对应的方式被拆分成第一块和第二块。第一块可以从片段指针开始写入,并且第二块可以被写入下一个区。虽然这个示例按顺序描述区,但是方法300可以用其它方式来实现。

[0046] 在示例中,流程300包括操作360,其中计算机系统更新片段指针。在完成写操作(其可以包括将数据写入多个区的多个写操作)之后,系统可以定义指向数据的结束位置的片段指针。以这种方式,相对于与接收到写请求之前的片段指针对应的区,更新后的片段指针可以与区片段中的不同区中的写入位置对应。

[0047] 在示例中,流程300包括操作370,其中计算机系统更新卷元数据。如参考图1更详细地描述的,写入区片段的数据可以通过卷元数据被映射到块卷。卷元数据可以包括片段指针,并且因此,更新卷元数据可以包括更新卷元数据数据库(例如,图1的卷元数据数据库160)中的片段指针,以及在已分配新区片段的情况下更新卷-区片段映射,作为操作320和330的一部分。在一些示例中,卷元数据可以存储数据被写入的起始位置指针,其与在操作360中被更新之前的片段指针值对应。

[0048] 在示例中,流程300包括操作380,其中计算机系统输出写请求的状态。在一些实施

例中,系统可以被配置为输出各种不同类型的状态报告。例如,资源元数据(诸如使用情况和容量数据)可以作为改进IaaS系统操作的一部分被收集和管理。作为另一个示例,可以将写请求的成功或失败返回给系统,使得系统可以重复方法300的一些或全部操作。作为另一个示例,系统可以向IaaS控制台界面的用户输出状态,其中写请求源自人类用户。

[0049] 区片段组织可以为结合SMR驱动器的数据库系统的操作提供显著的技术改进。通过实现写入包括多个区的区片段的技术,可以跨区写入数据,并且可以共享多个区的容量。以这种方式,通过映射到区片段指针的卷元数据促进的数据复制、存储和备份操作可以提高SMR写入和重新写操作的效率。

[0050] 图4图示了根据一个或多个实施例的用于从区片段读取数据的示例流程400。该流程的操作可以被实现为硬件电路系统和/或被存储为计算机系统(诸如图1的系统100)的非暂态计算机可读介质上的计算机可读指令。在实现时,指令表示包括可由计算机系统的(一个或多个)处理器执行的电路系统或代码的模块。此类指令的执行将计算机系统配置为执行本文描述的特定操作。每个电路系统或代码与处理器结合执行相应的(一个或多个)操作。虽然以特定次序图示了操作,但是应当理解的是,特定次序不是必需的,并且一个或多个操作可以被省略、跳过和/或重新排序。

[0051] 在示例中,流程400包括操作410,其中计算机系统接收读请求。如参考图1更详细地描述的,读请求可以引用要读取的卷和数据块。卷可以由卷标识符和块标识符引用,其可以在卷元数据中被映射到SMR驱动器。SMR驱动器可以定义包括区的区片段,区可以按顺序组织在区片段中。以这种方式,系统可以接收读请求,作为数据访问操作的一部分(例如作为数据恢复过程的一部分),并且可以被配置为访问区片段中的多个区以读取数据。

[0052] 在示例中,流程400包括操作420,其中(例如,当在图3的370处更新卷元数据时)计算机系统识别存储在卷元数据中的所存储的指针和读取长度。所存储的指针和读取长度可以与读请求中识别出的卷对应。卷标识符可以通过所存储的指针被映射到区片段以及区片段中的区。与参考图3描述的方法形成对比,流程400中的所存储的指针可以引用读取位置,而不是写入位置。即使如此,也可以类似地理解为参考SMR驱动器读取数据的区片段中的起始位置。

[0053] 读请求还可以包括对要读取的数据的引用,其可以被映射到SMR驱动器上的读取长度。读取长度可以与要读取的轨道(例如,图2的数据轨道250)的维度或长度对应,使得系统可以确定读取长度是否与区片段中的多个区重叠。例如,当区按顺序组织在区片段中时,系统可以指示片段指针处的第一读取位置和区片段中的下一个区的起始位置处的第二读取位置。

[0054] 在示例中,流程400可选地包括操作430,其中计算机系统确定用于区片段中的每个区的单个读命令。在从操作420中识别出的片段指针起始的读取长度延伸到多个区的情况下,系统可以确定用于区片段中的不同区的多个读命令。多个单个读命令可以被交叉引用,使得可以根据在操作420中确定的读取长度来重构由读命令返回的数据。以这种方式,跨多个区写入的数据可以由引用区片段的单个读请求来读取,而不是通过在各个区寻址的多个读请求来读取。

[0055] 在示例中,流程400包括操作440,其中计算机系统向SMR驱动器发出读命令。如参考图1更详细地描述的,系统可以经由存储服务器(例如,图1的存储服务器142或SMR服务器

146) 管理数据存储。读命令可以包括在操作420中确定的片段指针和读取长度,并且系统可以向通过映射元数据(诸如参考图1更详细地描述的卷元数据)引用的SMR存储服务器发出读命令。

[0056] 在示例中,流程400包括操作450,其中计算机系统输出读取数据和读请求的状态。数据的输出可以包括将数据从SMR驱动器传送到读请求中引用的目的地。例如,系统可以将数据输出到卷存储系统,或者作为从一个对象存储驱动器到另一个对象存储驱动器的数据复制操作的一部分。与流程300类似,状态可以包括请求的成功/失败信息,以及其它信息,包括但不限于描述读命令中所引用的数据和区的元数据。

[0057] 使用流程400的操作促进读命令可以为结合SMR驱动器的数据库系统的操作提供显著的技术改进。通过实现为在组织成区片段的多个区处寻址的读请求提供服务的技术,可以更高效地从SMR驱动器读取数据,而无需为区片段中的每个区生成多个读请求。

[0058] 图5图示了根据一个或多个实施例的用于从区片段删除数据的示例流程500。该流程的操作可以被实现为硬件电路系统和/或被存储为计算机系统(诸如图1的系统100)的非暂态计算机可读介质上的计算机可读指令。在实现时,指令表示包括可由计算机系统的(一个或多个)处理器执行的电路系统或代码的模块。此类指令的执行将计算机系统配置为执行本文描述的特定操作。每个电路系统或代码与处理器结合执行相应的(一个或多个)操作。虽然以特定次序图示了操作,但是应当理解的是,特定次序不是必需的,并且一个或多个操作可以被省略、跳过和/或重新排序。

[0059] 在示例中,流程500包括操作510,其中计算机系统接收删除请求。与读操作或写操作相反,删除操作可以在不参考片段指针作为起始位置的情况下进行。代替地,删除请求可以包括卷标识符信息,该卷标识符信息可以被映射到一个或多个驱动器(诸如参考图2更详细地描述的SMR驱动器(例如,图2的SMR驱动器200))中的对象存储装置。

[0060] 在示例中,流程500包括操作520,其中计算机系统读取卷对象的映射元数据。如参考图1更详细地描述的,映射元数据可以引用存储受删除请求影响的块卷的数据的区片段。以这种方式,系统可以通过参考映射元数据来识别分配给删除请求中识别出的卷的一个或多个区片段。

[0061] 在示例中,流程500包括操作530,其中计算机系统更新卷对象的状态。更新卷对象的状态可以包括但不限于向存储服务器(例如,图1的SMR存储服务器146)发送指令以将分配的区片段中包括的区的状态从“满”或“开放”修改为“空”,使得对于后续的写请求,片段指针与区片段中的第一区的起始位置对应。附加地或可替代地,系统可以修改卷元数据数据库(例如,图1的卷元数据数据库160)中的卷元数据以移除映射信息,使得包括在后续I/O请求(例如,图1的I/O请求120)中的卷标识符不与SMR驱动器中的区片段相关联。例如,在参考图3描述的流程300的操作中,识别片段指针可以返回没有区片段被分配给所引用的卷标识符。以这种方式,删除请求可以被寻址到区片段的区的一部分。代替擦除整个区,这可以允许SMR驱动器以更类似于随机存取驱动器而不是仅追加驱动器的方式进行操作,除了写请求不能写入部分重写的轨道之外。

[0062] 在示例中,流程500包括操作540,其中计算机系统输出删除请求的状态。与图3-图4的前述流程一样,系统可以向例如中央元数据管理系统返回删除请求的状态以指示删除请求所寻址的资源已经被释放。附加地或可替代地,系统可以向删除请求的发起者返回状

态成功/失败消息,以指示是否应当重新发送该请求或者将重新尝试该请求。

[0063] 图6图示了根据一个或多个实施例的用于对区片段进行碎片整理的示例流程600。该流程的操作可以被实现为硬件电路系统和/或被存储为计算机系统(诸如图1的系统100)的非暂态计算机可读介质上的计算机可读指令。在实现时,指令表示包括可由计算机系统的(一个或多个)处理器执行的电路系统或代码的模块。此类指令的执行将计算机系统配置为执行本文描述的特定操作。每个电路系统或代码与处理器结合执行相应的(一个或多个)操作。虽然以特定次序图示了操作,但是应当理解的是,特定次序不是必需的,并且一个或多个操作可以被省略、跳过和/或重新排序。

[0064] 除了改进SMR驱动器上的读/写操作之外,在一些实施例中,还可以通过将SMR驱动器的区组织成区片段并且重新映射存储在区中的数据以减少存储装置的低效使用来改进碎片整理操作。有利地,对区片段进行碎片整理可以允许SMR驱动器的功能更像随机存取驱动器,而不是严格地作为仅追加驱动器,但仍将限于仅追加写操作。在说明性示例中,如参考图5更详细地描述的,删除请求可以修改映射元数据而不是直接写入驱动器的数据。在操作过程中,如当起始于区片段中的一个区并结束于后续区的数据已被删除时,SMR驱动器(例如,图2的SMR驱动器200)可以使用区片段的容量的一小部分。由于部分重叠的轨道中的数据无法在SMR驱动器中被重写,因此剩余的未使用容量可能无法用于写操作,这可能引入SMR驱动器上的存储容量效率低下且利用率低下。对SMR驱动器进行碎片整理可以允许系统通过将数据重新写入新区并重新映射卷元数据来释放未使用的存储容量。

[0065] 在示例中,流程600包括操作610,其中计算机系统确认区片段的使用分数(fraction)。作为碎片整理的一部分,系统可以基于超出阈值的使用分数来实现自动(例如,无需人工干预)过程以触发碎片整理操作(例如,周期性检查使用分数、在删除请求之后刷新使用分数,等等)。为此,系统可以访问卷元数据,以确认使用分数。使用分数可以描述在卷元数据中映射的区片段的区中的存储容量的部分。例如,卷元数据可以引用存储在区片段中的数据的起始点、结束点、大块长度等,其仅相当于区片段的总容量的一部分,例如基于具有固定容量的区片段。

[0066] 在示例中,流程600包括操作620,其中计算机系统将使用分数与使用阈值进行比较。使用阈值可以关于使用分数来表达。例如,在系统被配置为当使用分数下降到50%以下时实现碎片整理的情况下,使用阈值可以被表达为0.5。

[0067] 在示例中,流程600包括操作630,其中计算机系统识别存储在区片段的区中的数据。当碎片整理过程由于操作620的比较而被触发时,系统可以使用卷元数据和/或通过扫描区片段来识别要在区片段的碎片整理期间被移动的数据。虽然随机存取驱动器的碎片整理可以允许移动一些数据,同时将其它数据原位留在驱动器上,但SMR驱动器的仅追加写入可以包括识别数据并将其重新映射到新区,以及擦除从其复制数据的区。

[0068] 在示例中,流程600包括操作640,其中计算机系统将存储的数据映射到空的或开放的区。在一些情况下,这是两部分操作,其中第一部分发生在操作640处(例如,这类似于如何在图3的操作340处确定写入目的地位置)并且第二部分发生在操作650之后(后面描述)。映射在操作630中识别出的数据可以包括修改卷元数据,使得与卷相关联的数据被映射到区片段的新区、新分配的区、或者缓冲存储装置,其可以在维持空闲以供SMR驱动器上的碎片整理使用的区片段中。类似地,随机存取存储器可以用作缓冲存储器。重新映射可以

包括确定用于随后将存储的数据写入区片段的存储的数据的配置。

[0069] 在示例中,流程600包括操作650,其中计算机系统将存储的数据写入空的或开放的区。写入所存储的数据可以包括流程300的操作,如参考图3更详细地描述的。例如,系统可以向SMR存储服务器(例如,图1的SMR存储服务器146)发送写命令,以使所存储的数据被写入一个或多个空的或开放的区。在对未保留用于碎片整理的区执行写操作的情况下,数据可以留在被写入的位置,并且可以如操作640所提供的那样维持映射。

[0070] 在示例中,在操作650处复制数据之后,在操作640处上述两部分操作的第二部分发生在流程600的操作655处。例如,一旦数据在新位置是安全的,就更新卷元数据(例如,类似于图3的操作370)。

[0071] 在示例中,流程600包括操作660,其中计算机系统重置所存储的数据的源区。此外,系统可以重置区片段的源区以用于后续I/O操作(例如,图1的I/O操作120)。重置源区可以包括将区的状态从“满”更新为“空”或其它修改以反映区片段中的数据的新配置。

[0072] 在示例中,流程600可选地包括操作670,其中计算机系统释放多余的区。如参考图3针对流程300的操作330所描述的,要写入的数据超出所引用的区片段的容量的写操作可以包括分配新区片段以提供附加的存储容量。以这种方式,碎片整理可以通过减少区片段中数据容量的未使用部分来使得数据可以存储在单个区片段中。这样,可以释放先前被用于存储用于该区片段的数据的区片段。在一些实施例中,多余的区可以被识别为640的映射操作的一部分,并且在根据新映射写入所存储的数据之后被释放,如参考操作650所描述的。

[0073] 在示例中,流程600可选地包括操作680,其中计算机系统输出碎片整理过程的状态。与前述图3-图5的输出操作一样,操作680可以包括向资源元数据的中央数据库报告资源使用状态,和/或指示碎片整理过程的成功/失败。在一些实施例中,操作680还可以包括确认区的新使用分数或者重置定时器或其它触发条件。

[0074] 基于区-片段的碎片整理可以为结合SMR驱动器的数据库系统的操作提供显著的技术改进。通过实现区-片段碎片整理技术,可以至少通过提高区片段的使用分数来提高SMR驱动器容量的有效使用,其中数据可以跨区片段中的区写入。以这种方式,可以减少由仅追加写操作(其是SMR驱动器的特点)引入的限制。

[0075] 如上所述,基础设施即服务(IaaS)是一种特定类型的云计算。IaaS可以被配置为通过公共网络(例如,互联网)提供虚拟化计算资源。在IaaS模型中,云计算提供商可以托管基础设施组件(例如,服务器、存储设备、网络节点(例如,硬件)、部署软件、平台虚拟化(例如,管理程序层)等)。在一些情况下,IaaS提供商还可以提供各种服务来伴随这些基础设施组件(例如,计费、监视、日志记录、安全性、负载平衡和聚类等)。因此,由于这些服务可能是策略驱动的,因此IaaS用户可以能够实现策略来驱动负载平衡,以维持应用的可用性和性能。

[0076] 在一些情况下,IaaS客户可以通过诸如互联网之类的广域网(WAN)访问资源和服务,并且可以使用云提供商的服务来安装应用堆栈的剩余元素。例如,用户可以登录到IaaS平台以创建虚拟机(VM)、在每个VM上安装操作系统(OS)、部署诸如数据库之类的中间件、为工作负载和备份创建存储桶、甚至将企业软件安装到该VM中。然后,客户可以使用提供商的服务来执行各种功能,包括平衡网络流量、解决应用问题、监视性能、管理灾难恢复等。

[0077] 在大多数情况下,云计算模型将需要云提供商的参与。云提供商可以但不一定是专门提供(例如,供应、出租、销售)IaaS的第三方服务。实体也可能选择部署私有云,从而成为其自己的基础设施服务提供商。

[0078] 在一些示例中,IaaS部署是将新应用或应用的新版本放置到准备好的应用服务器等上的处理。它还可以包括准备服务器(例如,安装库、守护进程等)的处理。这通常由云提供商管理,位于管理程序层(例如,服务器、存储装置、网络硬件和虚拟化)之下。因此,客户可以负责处理(OS)、中间件和/或(例如,在(例如可以按需启动的)自助服务虚拟机上的)应用部署等。

[0079] 在一些示例中,IaaS供给可以指获取计算机或虚拟主机以供使用,甚至在它们上安装所需的库或服务。大多数情况下,部署不包括供给,并且供给可能需要被首先执行。

[0080] 在一些情况下,IaaS供给存在两个不同的问题。首先,在任何东西运行之前供给初始基础设施集合存在最初的挑战。其次,一旦所有东西已被供给,就存在演进现有基础设施(例如,添加新服务、更改服务、移除服务等)的挑战。在一些情况下,可以通过启用以声明方式定义基础设施的配置来解决这两个挑战。换句话说,基础设施(例如,需要哪些组件以及它们如何交互)可以由一个或多个配置文件来定义。因此,基础设施的总体拓扑(例如,哪些资源依赖于哪些资源,以及它们如何协同工作)可以以声明的方式描述。在一些情况下,一旦定义了拓扑,就可以生成创建和/或管理配置文件中描述的不同组件的工作流。

[0081] 在一些示例中,基础设施可以具有许多互连的元素。例如,可能存在一个或多个虚拟私有云(VPC)(例如,可配置和/或共享计算资源的潜在按需池),也称为核心网络。在一些示例中,还可以供给一个或多个安全性组规则以定义将如何设置网络的安全性以及一个或多个虚拟机(VM)。也可以供给其它基础设施元素,诸如负载均衡器、数据库等。随着期望和/或添加越来越多的基础设施元素,基础设施可以逐步演进。

[0082] 在一些情况下,可以采用连续部署技术来使得能够跨各种虚拟计算环境来部署基础设施代码。此外,所描述的技术可以使得能够在这些环境内进行基础设施管理。在一些示例中,服务团队可以编写期望部署到一个或多个但通常是许多不同的生产环境(例如,跨各种不同的地理位置,有时跨越整个世界)的代码。但是,在一些示例中,必须首先设置将在其上部署代码的基础设施。在一些情况下,供给可以手动完成,可以利用供给工具来供给资源,和/或一旦供给基础设施就可以利用部署工具来部署代码。

[0083] 图7是图示根据至少一个实施例的IaaS体系架构的示例模式的框图700。服务运营商702可以通信地耦合到可以包括虚拟云网络(VCN)706和安全主机子网708的安全主机租赁704。在一些示例中,服务运营商702可以使用一个或多个客户端计算设备(客户端计算设备可以是便携式手持设备(例如,**iPhone®**、蜂窝电话、**iPad®**、计算平板、个人数字助理(PDA))或可穿戴设备(例如,**GoogleGlass®**头戴式显示器)),运行软件(诸如Microsoft **Windows Mobile®**)和/或各种移动操作系统(诸如iOS、Windows Phone、Android、BlackBerry 7、Palm OS等),并且支持互联网、电子邮件、短消息服务(SMS)、**Blackberry®**或其它通信协议。可替代地,客户端计算设备可以是通用个人计算机,包括例如运行各种版本的Microsoft **Windows®**、Apple **Macintosh®**和/或Linux操作系统的个人计算机和/或膝上型计算机。客户端计算设备可以是运行各种商业上可获得的

UNIX®或类UNIX操作系统(包括但不限于各种GNU/Linux操作系统(诸如例如Google Chrome OS))中的任何操作系统的工作站计算机。替代地或附加地,客户端计算设备可以是任何其它电子设备,诸如瘦客户端计算机、支持互联网的游戏系统(例如,具有或不具有**Kinect®**手势输入设备的Microsoft Xbox游戏控制台)、和/或能够通过可以访问VCN 706和/或互联网的网络进行通信的个人消息传递设备。

[0084] VCN 706可以包括本地对等网关(LPG) 710,该VCN 706可以经由包含在安全壳(SSH)VCN 712中的LPG 710通信地耦合到SSH VCN 712。SSH VCN 712可以包括SSH子网714,并且SSH VCN 712可以经由包含在控制平面VCN 716中的LPG 710通信地耦合到控制平面VCN 716。此外,SSH VCN 712可以经由LPG 710通信地耦合到数据平面VCN 718。控制平面VCN 716和数据平面VCN 718可以包含在可以由IaaS提供商拥有和/或操作的服务租赁719中。

[0085] 控制平面VCN 716可以包括充当外围网络(例如,公司内部网络和外部网络之间的公司网络的部分)的控制平面非军事区(DMZ)层720。基于DMZ的服务器可以承担有限责任并有助于控制安全性违规。此外,DMZ层720可以包括一个或多个负载平衡器(LB)子网722、可以包括(一个或多个)应用(app)子网726的控制平面应用层724、可以包括(一个或多个)数据库(DB)子网730(例如,(一个或多个)前端DB子网和/或(一个或多个)后端DB子网)的控制平面数据层728。包含在控制平面DMZ层720中的(一个或多个)LB子网722可以通信地耦合到包含在控制平面应用层724中的(一个或多个)应用子网726和可以包含在控制平面VCN 716中的互联网网关734,并且(一个或多个)应用子网726可以通信地耦合到包含在控制平面数据层728中的(一个或多个)DB子网730以及服务网关736和网络地址转换(NAT)网关738。控制平面VCN 716可以包括服务网关736和NAT网关738。

[0086] 控制平面VCN 716可以包括数据平面镜像应用层740,其可以包括(一个或多个)应用子网726。包含在数据平面镜像应用层740中的(一个或多个)应用子网726可以包括可以执行计算实例744的虚拟网络接口控制器(VNIC) 742。计算实例744可以将数据平面镜像应用层740的(一个或多个)应用子网726通信地耦合到可以包含在数据平面应用层746中的(一个或多个)应用子网726。

[0087] 数据平面VCN 718可以包括数据平面应用层746、数据平面DMZ层748和数据平面数据层750。数据平面DMZ层748可以包括(一个或多个)LB子网722,其可以通信地耦合到数据平面应用层746的(一个或多个)应用子网726和数据平面VCN 718的互联网网关734。(一个或多个)应用子网726可以通信地耦合到数据平面VCN 718的服务网关736和数据平面VCN 718的NAT网关738。数据平面数据层750还可以包括可以通信地耦合到数据平面应用层746的(一个或多个)应用子网726的(一个或多个)DB子网730。

[0088] 控制平面VCN 716和数据平面VCN 718的互联网网关734可以通信地耦合到元数据管理服务752,该元数据管理服务752可以通信地耦合到公共互联网754。公共互联网754可以通信地耦合到控制平面VCN 716和数据平面VCN 718的NAT网关738。控制平面VCN 716和数据平面VCN 718的服务网关736可以通信地耦合到云服务756。

[0089] 在一些示例中,控制平面VCN 716或数据平面VCN 718的服务网关736可以对云服务756进行应用编程接口(API)调用,而无需通过公共互联网754。从服务网关736到云服务756的API调用可以是单向的:服务网关736可以对云服务756进行API调用,并且云服务756

可以将请求的数据发送到服务网关736。但是,云服务756可以不发起对服务网关736的API调用。

[0090] 在一些示例中,安全主机租赁704可以直接连接到服务租赁719,服务租赁719否则可以被隔离。安全主机子网708可以通过LPG 710与SSH子网714通信,LPG 710可以使得能够在否则隔离的系统上进行双向通信。将安全主机子网708连接到SSH子网714可以使安全主机子网708访问服务租赁719内的其它实体。

[0091] 控制平面VCN 716可以允许服务租赁719的用户设置或以其它方式供给期望资源。在控制平面VCN 716中供给的期望资源可以在数据平面VCN 718中部署或以其它方式使用。在一些示例中,控制平面VCN 716可以与数据平面VCN 718隔离,并且控制平面VCN 716的数据平面镜像应用层740可以经由VNIC 742与数据平面VCN 718的数据平面应用层746通信,VNIC 742可以包含在数据平面镜像应用层740和数据平面应用层746中。

[0092] 在一些示例中,系统的用户或客户可以通过可以将请求传送到元数据管理服务752的公共互联网754来做出请求,例如创建、读取、更新或删除(CRUD)操作。元数据管理服务752可以通过互联网网关734将请求传送到控制平面VCN 716。请求可以由包含在控制平面DMZ层720中的(一个或多个)LB子网722接收。(一个或多个)LB子网722可以确定请求是有效的,并且响应于该确定,(一个或多个)LB子网722可以将请求传输到包含在控制平面应用层724中的(一个或多个)应用子网726。如果请求被验证并且使用对公共互联网754的调用,那么对公共互联网754的调用可以被传输到可以对公共互联网754进行调用的NAT网关738。请求可能期望存储的存储器可以存储在(一个或多个)DB子网730中。

[0093] 在一些示例中,数据平面镜像应用层740可以促进控制平面VCN 716和数据平面VCN 718之间的直接通信。例如,可能期望对包含在数据平面VCN 718中的资源应用对配置的更改、更新或其它适当的修改。经由VNIC 742,控制平面VCN 716可以直接与包含在数据平面VCN 718中的资源通信,并且从而可以执行对这些资源的配置的更改、更新或其它适当的修改。

[0094] 在一些实施例中,控制平面VCN 716和数据平面VCN 718可以包含在服务租赁719中。在这种情况下,系统的用户或客户可能不拥有或操作控制平面VCN 716或数据平面VCN 718。替代地,IaaS提供商可以拥有或操作控制平面VCN 716和数据平面VCN 718,这两种平面都可以包含在服务租赁719中。该实施例可以使得能够隔离可能阻止用户或客户与其它用户或其它客户的资源进行交互的网络。此外,该实施例可以允许系统的用户或客户私自存储数据库,而无需依赖可能不具有期望安全级别的公共互联网754以进行存储。

[0095] 在其它实施例中,包含在控制平面VCN 716中的(一个或多个)LB子网722可以被配置为从服务网关736接收信号。在这个实施例中,控制平面VCN 716和数据平面VCN 718可以被配置为由IaaS提供商的客户调用而无需调用公共互联网754。IaaS提供商的客户可能期望这个实施例,因为客户使用的(一个或多个)数据库可以由IaaS提供商控制并且可以存储在服务租赁719上,服务租赁719可能与公共互联网754隔离。

[0096] 图8是图示根据至少一个实施例的IaaS体系架构的另一个示例模式的框图800。服务运营商802(例如,图7的服务运营商702)可以通信地耦合到安全主机租赁804(例如,图7的安全主机租赁704),该安全主机租赁804可以包括虚拟云网络(VCN)806(例如,图7的VCN 706)和安全主机子网808(例如,图7的安全主机子网708)。VCN 806可以包括本地对等网关

(LPG) 810 (例如, 图7的LPG 710), 该VCN 806可以经由包含在安全壳 (SSH) VCN 812 (例如, 图7的SSH VCN 712) 中的LPG 710通信地耦合到SSH VCN 812。SSH VCN 812可以包括SSH子网814 (例如, 图7的SSH子网714), 并且SSH VCN 812可以经由包含在控制平面VCN 816 (例如, 图7的控制平面VCN 716) 中的LPG 810通信地耦合到控制平面VCN 816。控制平面VCN 816可以包含在服务租赁819 (例如, 图7的服务租赁719) 中, 并且数据平面VCN 818 (例如, 图7的数据平面VCN 718) 可以包含在可能由系统的用户或客户拥有或操作的客户租赁821中。

[0097] 控制平面VCN 816可以包括控制平面DMZ层820 (例如, 图7的控制平面DMZ层720), 其可以包括 (一个或多个) LB子网822 (例如, 图7的 (一个或多个) LB子网722)、可以包括 (一个或多个) 应用子网826 (例如, 图7的 (一个或多个) 应用子网726) 的控制平面应用层824 (例如, 图7的控制平面应用层724)、可以包括 (一个或多个) 数据库 (DB) 子网830 (例如, 类似于图7的 (一个或多个) DB子网730) 的控制平面数据层828 (例如, 图7的控制平面数据层728)。包含在控制平面DMZ层820中的 (一个或多个) LB子网822可以通信地耦合到包含在控制平面应用层824中的 (一个或多个) 应用子网826和可以包含在控制平面VCN 816中的互联网网关834 (例如, 图7的互联网网关734), 并且 (一个或多个) 应用子网826可以通信地耦合到包含在控制平面数据层828中的 (一个或多个) DB子网830以及服务网关836 (例如, 图7的服务网关) 和网络地址转换 (NAT) 网关838 (例如, 图7的NAT网关738)。控制平面VCN 816可以包括服务网关836和NAT网关838。

[0098] 控制平面VCN 816可以包括可以包括 (一个或多个) 应用子网826的数据平面镜像应用层840 (例如, 图7的数据平面镜像应用层740)。包含在数据平面镜像应用层840中的 (一个或多个) 应用子网826可以包括可以执行计算实例844 (例如, 类似于图7的计算实例744) 的虚拟网络接口控制器 (VNIC) 842 (例如, 图7的VNIC 742)。计算实例844可以促进数据平面镜像应用层840中的 (一个或多个) 应用子网826和可以包含在数据平面应用层846 (例如, 图7的数据平面应用层746) 中的 (一个或多个) 应用子网826之间的经由包含在数据平面镜像应用层840中的VNIC 842以及包含在数据平面应用层846中的VNIC 842的通信。

[0099] 包含在控制平面VCN 816中的互联网网关834可以通信地耦合到元数据管理服务852 (例如, 图7的元数据管理服务752), 该元数据管理服务852可以通信地耦合到公共互联网854 (例如, 图7的公共互联网754)。公共互联网854可以通信地耦合到包含在控制平面VCN 816中的NAT网关838。包含在控制平面VCN 816中的服务网关836可以通信地耦合到云服务856 (例如, 图7的云服务756)。

[0100] 在一些示例中, 数据平面VCN 818可以包含在客户租赁821中。在这种情况下, IaaS提供商可以为每个客户提供控制平面VCN 816, 并且IaaS提供商可以为每个客户设置包含在服务租赁819中的唯一计算实例844。每个计算实例844可以允许包含在服务租赁819中的控制平面VCN 816和包含在客户租赁821中的数据平面VCN 818之间的通信。计算实例844可以允许在包含在服务租赁819中的控制平面VCN 816中供给的资源被部署或以其它方式用于包含在客户租赁821中的数据平面VCN 818中。

[0101] 在其它示例中, IaaS提供商的客户可以具有存在于客户租赁821中的数据库。在这个示例中, 控制平面VCN 816可以包括数据平面镜像应用层840, 该数据平面镜像应用层840可以包括 (一个或多个) 应用子网826。数据平面镜像应用层840可以驻留在数据平面VCN 818中, 但数据平面镜像应用层840可能不在数据平面VCN 818中。换句话说, 数据平面镜像

应用层840可以访问客户租赁821,但是数据平面镜像应用层840可能不存在于数据平面VCN 818中或者由IaaS提供商的客户拥有或操作。数据平面镜像应用层840可以被配置为对数据平面VCN 818进行调用,但可以被配置为对包含在控制平面VCN 816中的任何实体进行调用。客户可能期望在数据平面VCN 818中部署或以其它方式使用在控制平面VCN 816中供给的资源,并且数据平面镜像应用层840可以促进客户的期望部署或资源的其它使用。

[0102] 在一些实施例中,IaaS提供商的客户可以将过滤器应用到数据平面VCN 818。在这个实施例中,客户可以确定数据平面VCN 818可以访问什么,并且客户可以限制从数据平面VCN 818对公共互联网854的访问。IaaS提供商可能无法应用过滤器或以其它方式控制数据平面VCN 818对任何外部网络或数据库的访问。客户将过滤器和控制应用到包含在客户租赁821中的数据平面VCN 818上可以帮助将数据平面VCN 818与其它客户和公共互联网854隔离开。

[0103] 在一些实施例中,云服务856可以由服务网关836调用以访问公共互联网854、控制平面VCN 816或数据平面VCN 818上可能不存在的服务。云服务856与控制平面VCN 816或数据平面VCN 818之间的连接可以不是实时的或连续的。云服务856可以存在于由IaaS提供商拥有或操作的不同网络上。云服务856可以被配置为接收来自服务网关836的调用并且可以被配置为不接收来自公共互联网854的调用。一些云服务856可以与其它云服务856隔离,并且控制平面VCN 816可以与可能与控制平面VCN 816不在同一区域的云服务856隔离。例如,控制平面VCN 816可能位于“区域1”,并且云服务“部署6”可能位于区域1和“区域2”。如果包含在位于区域1中的控制平面VCN 816中的服务网关836对部署6进行调用,那么该调用可以被传输到区域1中的部署6。在这个示例中,控制平面VCN 816或区域1中的部署6可能不与区域2中的部署6通信地耦合或以其它方式通信。

[0104] 图9是图示根据至少一个实施例的IaaS体系架构的另一个示例模式的框图900。服务运营商902(例如,图7的服务运营商702)可以通信地耦合到安全主机租赁904(例如,图7的安全主机租赁704),该安全主机租赁904可以包括虚拟云网络(VCN) 906(例如,图7的VCN 706)和安全主机子网908(例如,图7的安全主机子网708)。VCN 906可以包括LPG 910(例如,图7的LPG 710),该VCN 906可以经由包含在SSH VCN 912(例如,图7的SSH VCN 712)中的LPG 910通信地耦合到SSH VCN 912。SSH VCN 912可以包括SSH子网914(例如,图7的SSH子网714),并且SSH VCN 912可以经由包含在控制平面VCN 916(例如,图7的控制平面VCN 716)中的LPG 910通信地耦合到控制平面VCN 916并且经由包含在数据平面VCN 918(例如,图7的数据平面718)中的LPG 910通信地耦合到数据平面VCN 918。控制平面VCN 916和数据平面VCN 918可以包含在服务租赁919(例如,图7的服务租赁719)中。

[0105] 控制平面VCN 916可以包括可以包括(一个或多个)负载平衡器(LB)子网922(例如,图7的(一个或多个)LB子网722)的控制平面DMZ层920(例如,图7的控制平面DMZ层720)、可以包括(一个或多个)应用子网926(例如,类似于图7的(一个或多个)应用子网726)的控制平面应用层924(例如,图7的控制平面应用层724)、可以包括(一个或多个)DB子网930的控制平面数据层928(例如,图7的控制平面数据层728)。包含在控制平面DMZ层920中的(一个或多个)LB子网922可以通信地耦合到包含在控制平面应用层924中的(一个或多个)应用子网926和可以包含在控制平面VCN 916中的互联网网关934(例如,图7的互联网网关734),并且(一个或多个)应用子网926可以通信地耦合到包含在控制平面数据层928中的(一个或

多个) DB子网930以及服务网关936 (例如,图7的服务网关) 和网络地址转换 (NAT) 网关938 (例如,图7的NAT网关738)。控制平面VCN 916可以包括服务网关936和NAT网关938。

[0106] 数据平面VCN 918可以包括数据平面应用层946 (例如,图7的数据平面应用层746)、数据平面DMZ层948 (例如,图7的数据平面DMZ层748),以及数据平面数据层950 (例如,图7的数据平面数据层750)。数据平面DMZ层948可以包括可以通信地耦合到数据平面应用层946的 (一个或多个) 可信应用子网960和 (一个或多个) 不可信应用子网962以及包含在数据平面VCN 918中的互联网网关934的 (一个或多个) LB子网922。(一个或多个) 可信应用子网960可以通信地耦合到包含在数据平面VCN 918中的服务网关936、包含在数据平面VCN 918中的NAT网关938、以及包含在数据平面数据层950中的 (一个或多个) DB子网930。(一个或多个) 不可信应用子网962可以通信地耦合到包含在数据平面VCN 918中的服务网关936和包含在数据平面数据层950中的 (一个或多个) DB子网930。数据平面数据层950可以包括可以通信地耦合到包含在数据平面VCN 918中的服务网关936的 (一个或多个) DB子网930。

[0107] (一个或多个) 不可信应用子网962可以包括可以通信地耦合到租户虚拟机 (VM) 966(1) - (N) 的一个或多个主VNIC 964(1) - (N)。每个租户VM 966(1) - (N) 可以通信地耦合到可以包含在相应容器出口VCN 968(1) - (N) 中的相应应用子网967(1) - (N),该相应容器出口VCN 968(1) - (N) 可以包含在相应客户租赁970(1) - (N) 中。相应辅VNIC 972(1) - (N) 可以促进包含在数据平面VCN 918中的 (一个或多个) 不可信应用子网962与包含在容器出口VCN 968(1) - (N) 中的应用子网之间的通信。每个容器出口VCN 968(1) - (N) 可以包括NAT网关938,该NAT网关938可以通信地耦合到公共互联网954 (例如,图7的公共互联网754)。

[0108] 包含在控制平面VCN 916中以及包含在数据平面VCN 918中的互联网网关934可以通信地耦合到元数据管理服务952 (例如,图7的元数据管理系统752),该元数据管理服务952可以通信地耦合到公共互联网954。公共互联网954可以通信地耦合到包含在控制平面VCN 916中以及包含在数据平面VCN 918中的NAT网关938。包含在控制平面VCN 916中以及包含在数据平面VCN 918中的服务网关936可以通信地耦合到云服务956。

[0109] 在一些实施例中,数据平面VCN 918可以与客户租赁970集成。在一些情况下,诸如在执行代码时可能期望支持的情况下,这种集成对于IaaS提供商的客户可能是有用的或期望的。客户可能提供可能具有破坏性、可能与其它客户资源通信或可能以其它方式导致非期望效果的代码来运行。作为对此的响应,IaaS提供商可以确定是否运行由客户给予IaaS提供商的代码。

[0110] 在一些示例中,IaaS提供商的客户可以向IaaS提供商授予临时网络访问,并请求附加到数据平面应用层946的功能。运行该功能的代码可以在VM 966(1) - (N) 中执行,并且该代码可以不被配置为在数据平面VCN 918上的其它任何地方运行。每个VM 966(1) - (N) 可以连接到一个客户租赁970。包含在VM 966(1) - (N) 中的相应容器971(1) - (N) 可以被配置为运行代码。在这种情况下,可以存在双重隔离 (例如,容器971(1) - (N) 运行代码,其中容器971(1) - (N) 可能至少包含在 (一个或多个) 不可信应用子网962中所包含的VM 966(1) - (N) 中),这可以帮助防止不正确的或以其它方式非期望的代码损坏IaaS提供商的网络或损坏不同客户的网络。容器971(1) - (N) 可以通信地耦合到客户租赁970并且可以被配置为传输或接收来自客户租赁970的数据。容器971(1) - (N) 可以不被配置为从数据平面VCN 918中的任何其它实体传输或接收数据。在运行代码完成后,IaaS提供商可以终止或以其它方式处

置容器971(1) - (N)。

[0111] 在一些实施例中, (一个或多个) 可信应用子网960可以运行可以由IaaS提供商拥有或操作的代码。在这个实施例中, (一个或多个) 可信应用子网960可以通信地耦合到 (一个或多个) DB子网930并且被配置为在 (一个或多个) DB子网930中执行CRUD操作。

[0112] (一个或多个) 不可信应用子网962可以通信地耦合到 (一个或多个) DB子网930, 但是在这个实施例中, (一个或多个) 不可信应用子网可以被配置为在 (一个或多个) DB子网930中执行读取操作。可以包含在每个客户的VM 966(1) - (N) 中并且可以运行来自客户的代码的容器971(1) - (N) 可以不与 (一个或多个) DB子网930通信地耦合。

[0113] 在其它实施例中, 控制平面VCN 916和数据平面VCN 918可以不直接通信地耦合。在这个实施例中, 控制平面VCN 916和数据平面VCN 918之间可能不存在直接通信。但是, 通信可以通过至少一个方法而间接地发生。LPG 910可以由IaaS提供商建立, 其可以促进控制平面VCN 916和数据平面VCN 918之间的通信。在另一个示例中, 控制平面VCN 916或数据平面VCN 918可以经由服务网关936对云服务956进行调用。例如, 从控制平面VCN 916对云服务956的调用可以包括对可以与数据平面VCN 918通信的服务的请求。

[0114] 图10是图示根据至少一个实施例的IaaS体系架构的另一个示例模式的框图1000。服务运营商1002(例如, 图7的服务运营商702) 可以通信地耦合到安全主机租赁1004(例如, 图7的安全主机租赁704), 该安全主机租赁1004可以包括虚拟云网络 (VCN) 1006(例如, 图7的VCN 706) 和安全主机子网1008(例如, 图7的安全主机子网708)。VCN 1006可以包括LPG 1010(例如, 图7的LPG 710), 该VCN 1006可以经由包含在SSH VCN 1012(例如, 图7的SSH VCN 712) 中的LPG 1010通信地耦合到SSH VCN 1012。SSH VCN 1012可以包括SSH子网1014(例如, 图7的SSH子网714), 并且SSH VCN 1012可以经由包含在控制平面VCN 1016(例如, 图7的控制平面VCN 716) 中的LPG 1010通信地耦合到控制平面VCN 1016并且经由包含在数据平面VCN 1018(例如, 图7的数据平面718) 中的LPG 1010通信地耦合到数据平面VCN 1018。控制平面VCN 1016和数据平面VCN 1018可以包含在服务租赁1019(例如, 图7的服务租赁719) 中。

[0115] 控制平面VCN 1016可以包括可以包括 (一个或多个) LB子网1022(例如, 图7的 (一个或多个) LB子网722) 的控制平面DMZ层1020(例如, 图7的控制平面DMZ层720)、可以包括 (一个或多个) 应用子网1026(例如, 图7的 (一个或多个) 应用子网726) 的控制平面应用层1024(例如, 图7的控制平面应用层724)、可以包括 (一个或多个) DB子网1030(例如, 图9的 (一个或多个) DB子网930) 的控制平面数据层1028(例如, 图7的控制平面数据层728)。包含在控制平面DMZ层1020中的 (一个或多个) LB子网1022可以通信地耦合到包含在控制平面应用层1024中的 (一个或多个) 应用子网1026和可以包含在控制平面VCN 1016中的互联网网关1034(例如, 图7的互联网网关734), 并且 (一个或多个) 应用子网1026可以通信地耦合到包含在控制平面数据层1028中的 (一个或多个) DB子网1030以及服务网关1036(例如, 图7的服务网关) 和网络地址转换 (NAT) 网关1038(例如, 图7的NAT网关738)。控制平面VCN 1016可以包括服务网关1036和NAT网关1038。

[0116] 数据平面VCN 1018可以包括数据平面应用层1046(例如, 图7的数据平面应用层746)、数据平面DMZ层1048(例如, 图7的数据平面DMZ层748)、以及数据平面数据层1050(例如, 图7的数据平面数据层750)。数据平面DMZ层1048可以包括可以通信地耦合到数据平面

应用层1046的(一个或多个)可信应用子网1060(例如,图9的(一个或多个)可信应用子网960)和(一个或多个)不可信应用子网1062(例如,图9的(一个或多个)不可信应用子网962)以及包含在数据平面VCN 1018中的互联网网关1034的(一个或多个)LB子网1022。(一个或多个)可信应用子网1060可以通信地耦合到包含在数据平面VCN 1018中的服务网关1036、包含在数据平面VCN 1018中的NAT网关1038以及包含在数据平面数据层1050中的(一个或多个)DB子网1030。(一个或多个)不可信应用子网1062可以通信地耦合到包含在数据平面VCN 1018中的服务网关1036和包含在数据平面数据层1050中的(一个或多个)DB子网1030。数据平面数据层1050可以包括可以通信地耦合到包含在数据平面VCN 1018中的服务网关1036的(一个或多个)DB子网1030。

[0117] (一个或多个)不可信应用子网1062可以包括可以通信地耦合到驻留在(一个或多个)不可信应用子网1062内的租户虚拟机 (VM) 1066 (1) - (N) 的主VNIC 1064 (1) - (N) 。每个租户VM 1066 (1) - (N) 可以运行相应容器1067 (1) - (N) 中的代码,并且可通信地耦合到可以包含在数据平面应用层1046中的应用子网1026,该数据平面应用层1046可以包含在容器出口VCN 1068中。相应辅VNIC 1072 (1) - (N) 可以促进包含在数据平面VCN 1018中的(一个或多个)不可信应用子网1062和包含在容器出口VCN 1068中的应用子网之间的通信。容器出口VCN可以包括可以通信地耦合到公共互联网1054(例如,图7的公共互联网754)的NAT网关1038。

[0118] 包含在控制平面VCN 1016中以及包含在数据平面VCN 1018中的互联网网关1034可以通信地耦合到元数据管理服务1052(例如,图7的元数据管理系统752),该元数据管理服务1052可以通信地耦合到公共互联网1054。公共互联网1054可以通信地耦合到包含在控制平面VCN 1016中以及包含在数据平面VCN 1018中的NAT网关1038。包含在控制平面VCN 1016中以及包含在数据平面VCN 1018中的服务网关1036可以通信地耦合到云服务1056。

[0119] 在一些示例中,图10的框图1000的体系架构所示的模式可以被认为是图9的框图900的体系架构所示的模式的例外,并且在IaaS提供商不能直接与客户通信(例如,断开连接的区域)的情况下,这种模式可能是IaaS提供商的客户所期望的。客户可以实时访问每个客户的包含在VM 1066 (1) - (N) 中的相应容器1067 (1) - (N) 。容器1067 (1) - (N) 可以被配置为对包含在数据平面应用层1046的(一个或多个)应用子网1026中的相应辅VNIC 1072 (1) - (N) 进行调用,该数据平面应用层1046可以包含在容器出口VCN 1068中。辅VNIC 1072 (1) - (N) 可以将调用传输到NAT网关1038,该NAT网关1038可以将调用传输到公共互联网1054。在这个示例中,可以由客户实时访问的容器1067 (1) - (N) 可以与控制平面VCN 1016隔离,并且可以与包含在数据平面VCN 1018中的其它实体隔离。容器1067 (1) - (N) 也可以与来自其它客户的资源隔离。

[0120] 在其它示例中,客户可以使用容器1067 (1) - (N) 来调用云服务1056。在这个示例中,客户可以运行容器1067 (1) - (N) 中的从云服务1056请求服务的代码。容器1067 (1) - (N) 可以将该请求传输到辅VNIC 1072 (1) - (N) ,该辅VNIC 1072 (1) - (N) 可以将请求传输到NAT网关,该NAT网关可以将请求传输到公共互联网1054。公共互联网1054可以经由互联网网关1034将请求传输到包含在控制平面VCN 1016中的(一个或多个)LB子网1022。响应于确定请求有效,(一个或多个)LB子网可以将请求传输到(一个或多个)应用子网1026,该(一个或多个)应用子网1026可以经由服务网关1036将请求传输到云服务1056。

[0121] 应当认识到的是,各图中描绘的IaaS体系架构700、800、900、1000可以具有除所描绘的组件之外的其它组件。另外,各图中所示的实施例仅仅是可以结合本公开的实施例的云基础设施系统的一些示例。在一些其它实施例中,IaaS系统可以具有比各图中所示更多或更少的组件、可以组合两个或更多个组件,或者可以具有不同的组件布置或配置。

[0122] 在某些实施例中,本文描述的IaaS系统可以包括以自助服务、基于订阅、弹性可伸缩、可靠、高度可用和安全的方式交付给客户的应用套件、中间件和数据库服务产品。此类IaaS系统的示例是本受让方提供的Oracle云基础设施(OCI)。

[0123] 图11图示了其中可以实现各种实施例的示例计算机系统1100。系统1100可以用于实现上述任何计算机系统。如图所示,计算机系统1100包括经由总线子系统1102与多个外围子系统通信的处理单元1104。这些外围子系统可以包括处理加速单元1106、I/O子系统1108、存储子系统1118和通信子系统1124。存储子系统1118包括有形计算机可读存储介质1122和系统存储器1110。

[0124] 总线子系统1102提供用于让计算机系统1100的各种部件和子系统按意图彼此通信的机制。虽然总线子系统1102被示意性地示出为单条总线,但是总线子系统的替代实施例可以利用多条总线。总线子系统1102可以是若干种类型的总线结构中的任何一种,包括存储器总线或存储器控制器、外围总线、以及使用任何各种总线体系架构的局部总线。例如,这种体系架构可以包括工业标准体系架构(ISA)总线、微通道体系架构(MCA)总线、增强型ISA(EISA)总线、视频电子标准协会(VESA)局部总线和外围部件互连(PCI)总线,其可以被实现为按IEEE P1386.1标准制造的Mezzanine总线。

[0125] 可以被实现为一个或多个集成电路(例如,常规微处理器或微控制器)的处理单元1104控制计算机系统1100的操作。一个或多个处理器可以被包括在处理单元1104中。这些处理器可以包括单核或多核处理器。在某些实施例中,处理单元1104可以被实现为一个或多个独立的处理单元1132和/或1134,其中在每个处理单元中包括单核或多核处理器。在其它实施例中,处理单元1104也可以被实现为通过将两个双核处理器集成到单个芯片中形成的四核处理单元。

[0126] 在各种实施例中,处理单元1104可以响应于程序代码执行各种程序并且可以维护多个并发执行的程序或进程。在任何给定的时间,要被执行的程序代码中的一些或全部代码可以驻留在(一个或多个)处理器1104中和/或存储子系统1118中。通过适当的编程,(一个或多个)处理器1104可以提供上述各种功能。计算机系统1100可以附加地包括处理加速单元1106,其可以包括数字信号处理器(DSP)、专用处理器,等等。

[0127] I/O子系统1108可以包括用户接口输入设备和用户接口输出设备。用户接口输入设备可以包括键盘、诸如鼠标或轨迹球的定点设备、结合到显示器中的触摸板或触摸屏、滚动轮、点击轮、拨盘、按钮、开关、键盘、具有语音命令识别系统的音频输入设备、麦克风以及其它类型的输入设备。用户接口输入设备可以包括例如运动感测和/或手势识别设备,诸如Microsoft**Kinect**®运动传感器,其使得用户能够使用手势和语音命令通过自然用户接口来控制诸如Microsoft**Xbox**® 360游戏控制器的输入设备并与之交互。用户接口输入设备也可以包括眼睛姿势识别设备,诸如从用户检测眼睛活动(例如,当拍摄照片和/或做出菜单选择时的“眨眼”)并且将眼睛姿势转换为去往输入设备(例如,Google**Glass**®)的输

入的Google **Glass®**眨眼检测器。此外,用户接口输入设备可以包括使用户能够通过语音命令与语音识别系统(例如,**Siri®**导航器)交互的语音识别感测设备。

[0128] 用户接口输入设备也可以包括但不限于三维(3D)鼠标、操纵杆或指向棒、游戏面板和绘图板、以及音频/视频设备,诸如扬声器、数码相机、数码摄录机、便携式媒体播放器、网络摄像头、图像扫描仪、指纹扫描仪、条形码阅读器3D扫描仪、3D打印机、激光测距仪和视线跟踪设备。此外,用户接口输入设备可以包括例如医学成像输入设备,诸如计算机断层扫描、磁共振成像、正电子发射断层摄影术、医疗超声设备。用户接口输入设备也可以包括例如音频输入设备,诸如MIDI键盘、数字乐器等。

[0129] 用户接口输出设备可以包括显示子系统、指示灯,或者诸如音频输出设备的非可视显示器等。显示子系统可以是阴极射线管(CRT)、诸如使用液晶显示器(LCD)或等离子显示器的平板设备、投影设备、触摸屏等。一般而言,术语“输出设备”的使用意在包括用于从计算机系统1100向用户或其它计算机输出信息的所有可能类型的设备和机制。例如,用户接口输出设备可以包括但不限于可视地传达文本、图形和音频/视频信息的各种显示设备,诸如监视器、打印机、扬声器、耳机、汽车导航系统、绘图仪、语音输出设备、以及调制解调器。

[0130] 计算机系统1100可以包括包含软件元件、被示为当前位于系统存储器1110内的存储子系统1118。系统存储器1110可以存储在处理单元1104上可执行且可加载的程序指令,以及在这些程序的执行期间所产生的数据。

[0131] 取决于计算机系统1100的配置和类型,系统存储器1110可以是易失性的(诸如随机存取存储器(RAM))和/或非易失性的(诸如只读存储器(ROM)、闪存存储器等)。RAM通常包含可被处理单元1104立即访问和/或目前正被处理单元1104操作和执行的数据和/或程序模块。在一些实施方案中,系统存储器1110可以包括多种不同类型的存储器,诸如静态随机存取存储器(SRAM)或动态随机存取存储器(DRAM)。在一些实施方案中,包含有助于在诸如启动期间在计算机系统1100内的元件之间传送信息的基本例程的基本输入/输出系统(BIOS)通常可以被存储在ROM中。作为示例而非限制,系统存储器1110也示出了可以包括客户端应用、Web浏览器、中间层应用、关系数据库管理系统(RDBMS)等的应用程序1112、程序数据1114、以及操作系统1116。作为示例,操作系统1116可以包括各种版本的Microsoft **Windows®**、Apple **Macintosh®**和/或Linux操作系统、各种可商业获得的 **UNIX®**或类UNIX操作系统(包括但不限于各种GNU/Linux操作系统、Google **Chrome®** OS等)和/或诸如iOS、**Windows®** Phone、**Android®** OS、**BlackBerry®** 110S和**Palm®** OS操作系统的移动操作系统。

[0132] 存储子系统1118也可以提供用于存储提供一些实施例的功能的基本编程和数据构造的有形计算机可读存储介质。当被处理器执行时提供上述功能的软件(程序、代码模块、指令)可以被存储在存储子系统1118中。这些软件模块或指令可以被处理单元1104执行。存储子系统1118也可以提供用于存储根据本公开使用的数据的储存库。

[0133] 存储子系统1100也可以包括可被进一步连接到计算机可读存储介质1122的计算机可读存储介质读取器1120。与系统存储器1110一起,并且可选地与其相结合,计算机可读存储介质1122可以全面地表示用于临时和/或更持久地包含、存储、传输和检索计算机可读

信息的远程、本地、固定和/或可移除的存储设备加存储介质。

[0134] 包含代码或代码的部分的计算机可读存储介质1122也可以包括本领域已知或使用的任何适当的介质,包括存储介质和通信介质,诸如但不限于:以用于信息的存储和/或传输的任何方法或技术实现的易失性和非易失性、可移除和不可移除的介质。这可以包括有形的计算机可读存储介质,诸如RAM、ROM、电可擦除可编程ROM(EEPROM)、闪存存储器或其它存储器技术、CD-ROM、数字多功能盘(DVD)或其它光学存储装置、磁带盒、磁带、磁盘存储装置或其它磁存储设备,或者其它有形的计算机可读介质。这也可以包括非有形的计算机可读介质,诸如数据信号、数据传输、或者可以用于传输期望信息并且可以被计算系统1100访问的任何其它介质。

[0135] 举例来说,计算机可读存储介质1122可以包括从不可移除的非易失性磁介质读取或写到其的硬盘驱动器、从可移除的非易失性磁盘读取或写到其的磁盘驱动器、以及从可移除的非易失性光盘(诸如CD ROM、DVD和Blu-Ray®盘或其它光学介质)读取或写到其的光盘驱动器。计算机可读存储介质1122可以包括但不限于:Zip®驱动器、闪存卡、通用串行总线(USB)闪存驱动器、安全数字(SD)卡、DVD盘、数字视频带等。计算机可读存储介质1122也可以包括基于非易失性存储器的固态驱动器(SSD)(诸如基于闪存存储器的SSD、企业闪存驱动器、固态ROM等)、基于易失性存储器的SSD(诸如基于固态RAM、动态RAM、静态RAM、DRAM的SSD)、磁阻RAM(MRAM) SSD、以及使用基于DRAM的SSD和基于闪存存储器的SSD的组合的混合SSD。盘驱动器及其关联的计算机可读介质可以为计算机系统1100提供计算机可读指令、数据结构、程序模块及其它数据的非易失性存储。

[0136] 通信子系统1124提供到其它计算机系统和网络的接口。通信子系统1124用作用于从其它系统接收数据和从计算机系统1100向其它系统传输数据的接口。例如,通信子系统1124可以使计算机系统1100能够经由互联网连接到一个或多个设备。在一些实施例中,通信子系统1124可以包括用于访问无线语音和/或数据网络的射频(RF)收发器部件(例如,使用蜂窝电话技术、诸如3G、4G或EDGE(用于全球演进的增强型数据速率)的先进数据网络技术、WiFi(IEEE 802.11系列标准)、或其它移动通信技术、或其任何组合)、全球定位系统(GPS)接收器部件和/或其它部件。在一些实施例中,通信子系统1124可以提供有线网络连接(例如,以太网),作为无线接口的附加或者替代。

[0137] 在一些实施例中,通信子系统1124也可以代表可以使用计算机系统1100的一个或多个用户接收结构化和/或非结构化的数据馈送1126、事件流1128、事件更新1130等形式的输入通信。

[0138] 举例来说,通信子系统1124可以被配置为实时地从社交网络和/或其它通信服务的用户接收数据馈送1126,诸如Twitter®馈送、Facebook®更新、诸如丰富站点摘要(RSS)馈送的web馈送和/或来自一个或多个第三方信息源的实时更新。

[0139] 此外,通信子系统1124也可以被配置为接收连续数据流形式的数据,这可以包括本质上可以是连续的或无界的没有明确终止的实时事件的事件流1128和/或事件更新1130。生成连续数据的应用的示例可以包括例如传感器数据应用、金融报价机、网络性能测量工具(例如,网络监视和流量管理应用)、点击流分析工具、汽车流量监视等。

[0140] 通信子系统1124也可以被配置为向一个或多个数据库输出结构化和/或非结构化

的数据馈送1126、事件流1128、事件更新1130等,该一个或多个数据库可以与耦合到计算机系统1100的一个或多个流式传输数据源计算机进行通信。

[0141] 计算机系统1100可以是各种类型之一,包括手持便携式设备(例如,**iPhone®**蜂窝电话、**iPad®**计算平板电脑、PDA)、可穿戴设备(例如,Google**Glass®**头戴式显示器)、PC、工作站、大型机、信息站、服务器机架、或任何其它数据处理系统。

[0142] 由于计算机和网络的不断变化的本质,在图中绘出的计算机系统1100的描述仅仅要作为具体的示例。具有比图中绘出的系统更多或更少部件的许多其它配置是可能的。例如,定制的硬件也可以被使用和/或特定的元素可以用硬件、固件、软件(包括applets)或其组合来实现。另外,也可以采用到诸如网络输入/输出设备之类的其它计算设备的连接。基于本文提供的公开内容和示教,本领域普通技术人员将认识到实现各种实施例的其它方式和/或方法。

[0143] 虽然已经描述了具体实施例,但是各种修改、变更、替代构造和等效形式也包含在本公开的范围内。实施例不限于在某些特定数据处理环境内操作,而是可以在多个数据处理环境内自由操作。此外,虽然已经使用特定系列的事务和步骤描述了实施例,但是本领域技术人员应该清楚本公开的范围不限于所描述系列的事务和步骤。上述实施例的各种特征和方面可以单独或联合使用。

[0144] 另外,虽然已经使用硬件和软件的特定组合描述了实施例,但是应当认识到硬件和软件的其它组合也在本公开的范围。实施例可以仅用硬件、或仅用软件、或使用它们的组合来实现。本文描述的各种处理可以以任何组合在相同的处理器或在不同的处理器上实现。因而,在组件或模块被描述为被配置为执行某些操作的情况下,可以通过例如设计电子电路来执行操作、通过对可编程电子电路(诸如微处理器)进行编程来执行操作、或通过其任何组合来完成这样的配置。处理可以使用多种技术进行通信,包括但不限于用于进程间通信的常规技术,并且不同的进程对可以使用不同的技术,或者同一进程对可以在不同时间使用不同的技术。

[0145] 因而,说明书和附图被认为是说明性的而不是限制性的意义。但是,在不脱离权利要求中阐述的更广泛的精神和范围的情况下,显然可以对其进行添加、减少、删除和其它修改和改变。因此,虽然已经描述了具体的公开实施例,但这些并不旨在进行限制。各种修改和等效形式都在以下权利要求的范围内。

[0146] 在描述所公开的实施例的上下文中(尤其在以下权利要求的上下文中)使用术语“一”和“一个”和“该”以及类似的指称应被解释为涵盖单数和复数两者,除非本文另有指示或与上下文明显矛盾。除非另有说明,否则术语“包括”、“具有”、“包含”和“含有”应被解释为开放式术语(即,意思是“包括但不限于”)。术语“连接”应被解释为部分或全部包含在内、附接到或连接在一起,即使中间存在一些东西。除非本文另有指示,否则本文中的值范围的描述仅旨在用作单独引用落入该范围内的每个单独值的简略方法,并且每个单独值被并入说明书中,就好像它在本文中单独描述一样。除非本文另有指示或与上下文明显矛盾,否则本文描述的所有方法都可以以任何合适的顺序执行。除非另有声明,否则本文提供的任何和全部示例或示例性语言(例如,“诸如”)的使用仅旨在更好地阐明实施例并且不对本公开的范围构成限制。说明书中的任何语言都不应被解释为指示任何未要求保护的元素对于本公开的实践是必不可少的。

[0147] 除非另有明确说明,否则析取(disjunctive)语言(诸如短语“X、Y或Z中的至少一个”)旨在在上下文中被理解为用作一般地表示项目、术语等可以是X、Y或Z中的一者或者是它们的任何组合(例如,X、Y和/或Z)。因此,这种析取语言通常不旨在也不应暗示某些实施例需要X中的至少一个、Y中的至少一个或Z中的至少一个各自存在。

[0148] 本文描述了本公开的优选实施例,包括已知用于实施本公开的最佳模式。这些优选实施例的变型对于本领域普通技术人员在阅读前述描述后会变得显而易见。普通技术人员应该能够适当地采用这样的变型并且可以以不同于本文具体描述的方式来实践本公开。因而,本公开包括在适用法律允许的情况下对所附权利要求中记载的主题的所有修改和等效形式。此外,除非在本文中另有指示,否则本公开包括上述元素在其所有可能的变型中的任何组合。

[0149] 本文引用的所有参考文献(包括出版物、专利申请和专利)均以相同的程度通过引用并入本文,就好像指示每个参考文献单独且具体地通过引用并入并且在本文中全文阐述一样。

[0150] 在前述的说明书中,本公开的各方面参考其具体实施例进行了描述,但本领域技术人员将认识到的是,本公开不限于此。上述公开的各个特征和方面可以被单独或联合使用。此外,在不脱离本说明书的更广泛精神和范围的情况下,实施例可以在除本文所述的环境和应用之外的任何数量的环境和应用中被使用。因而,本说明书和附图应当被认为是说明性而不是限制性的。

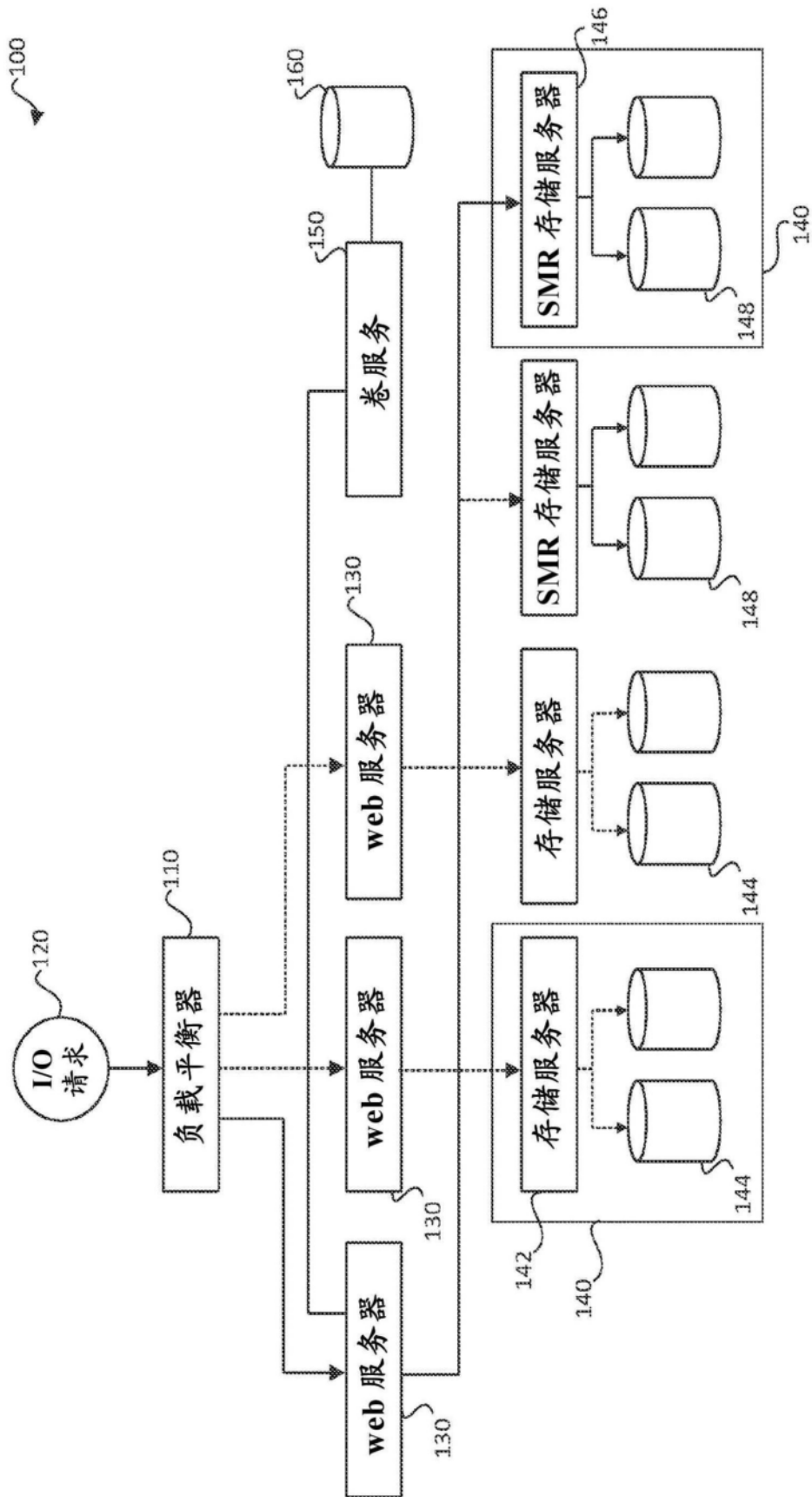


图1

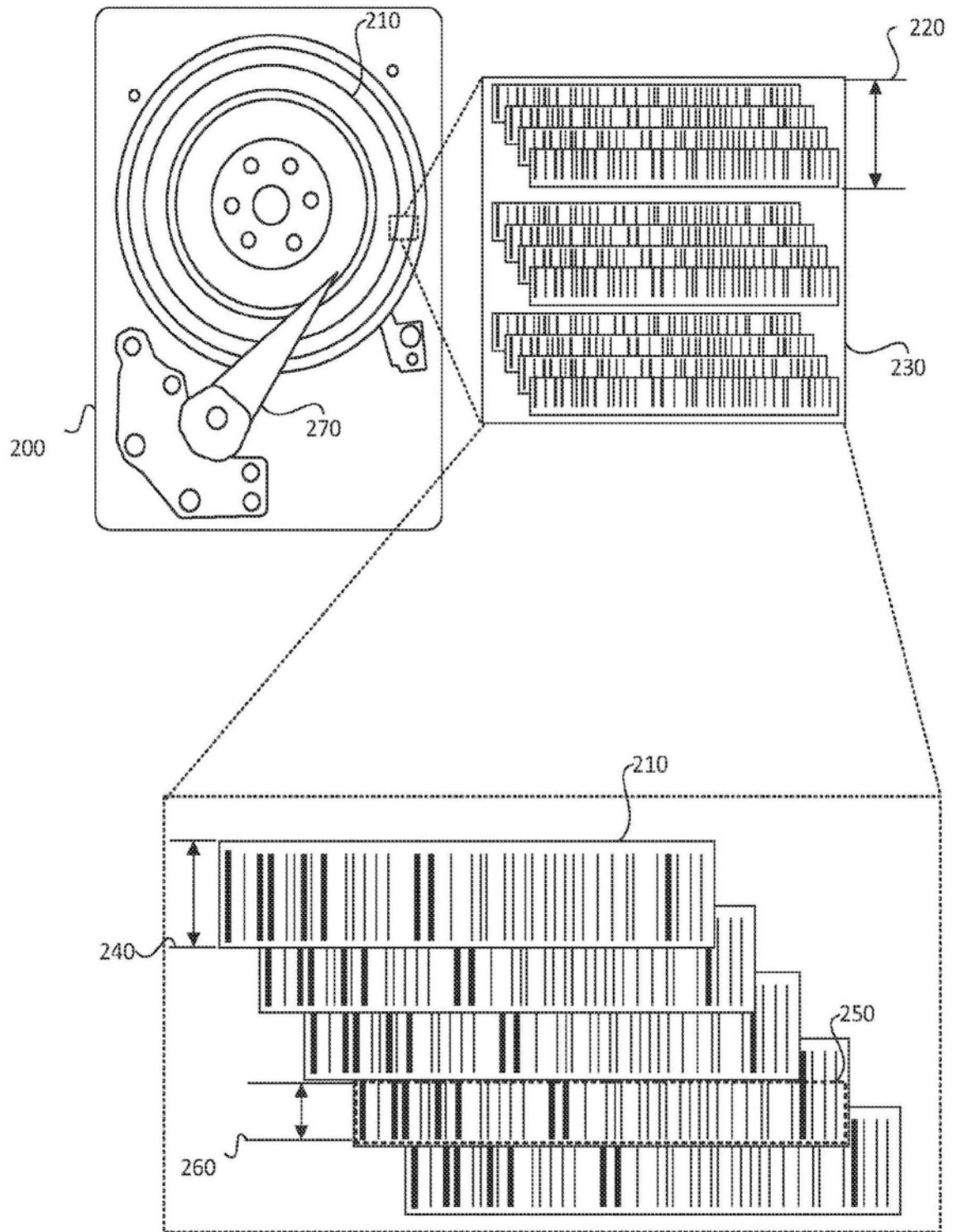


图2

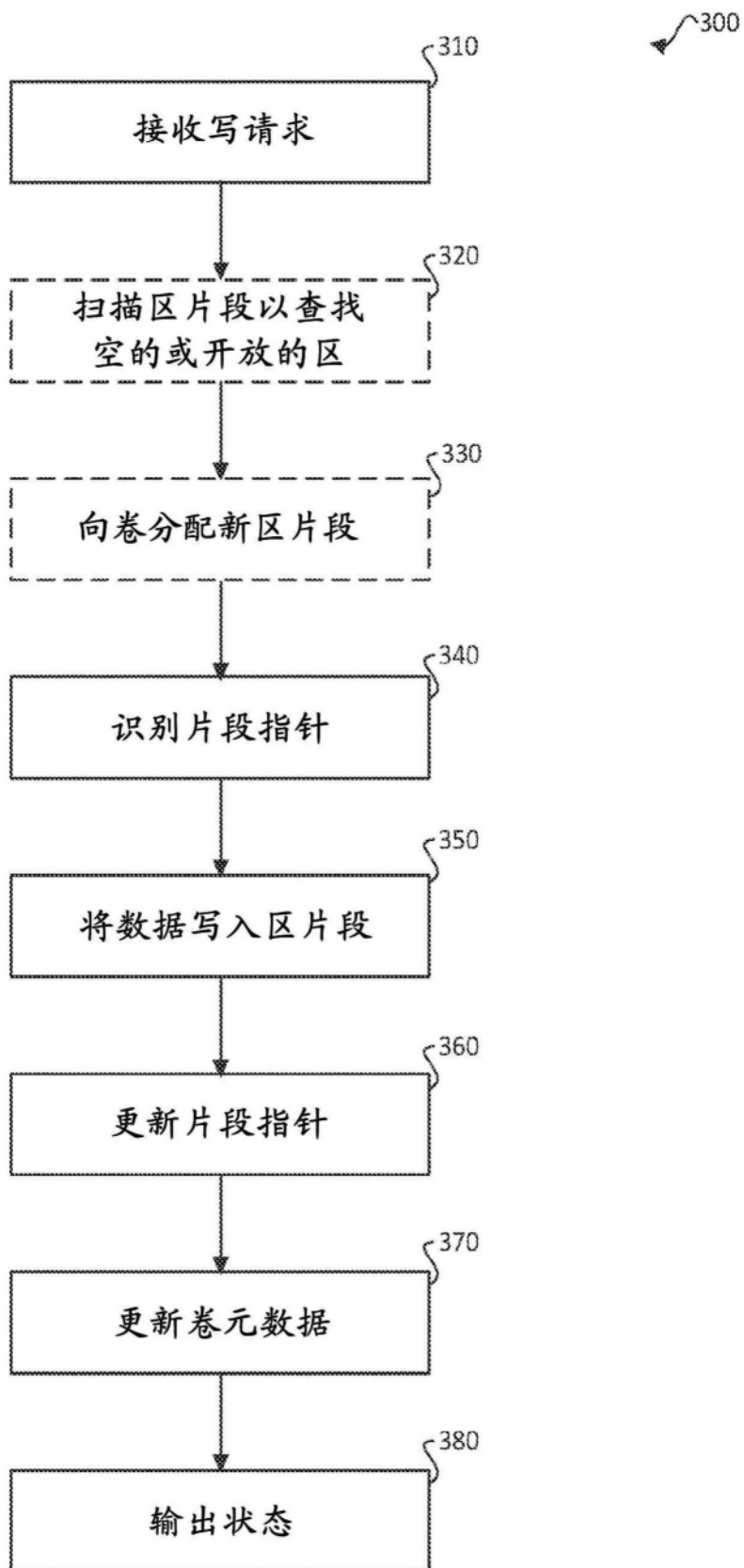


图3

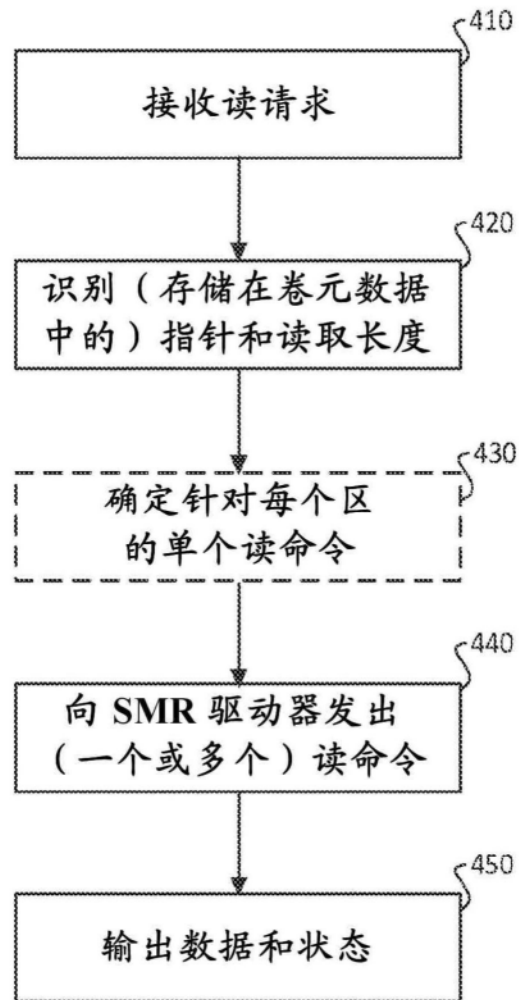


图4

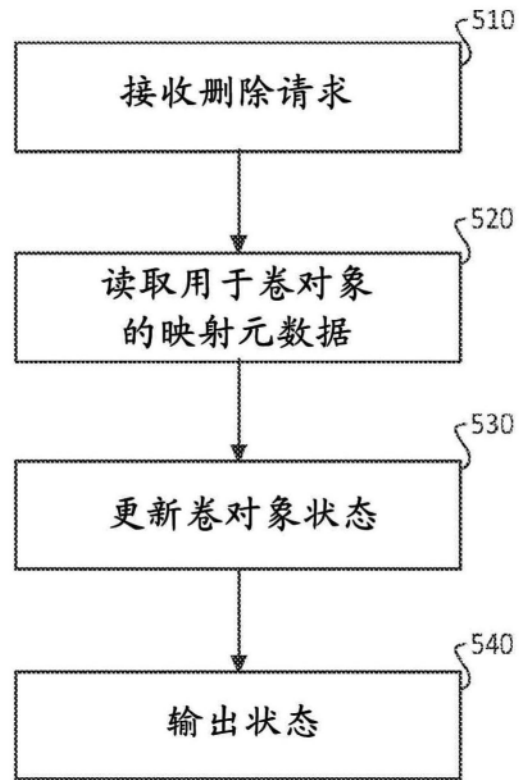


图5

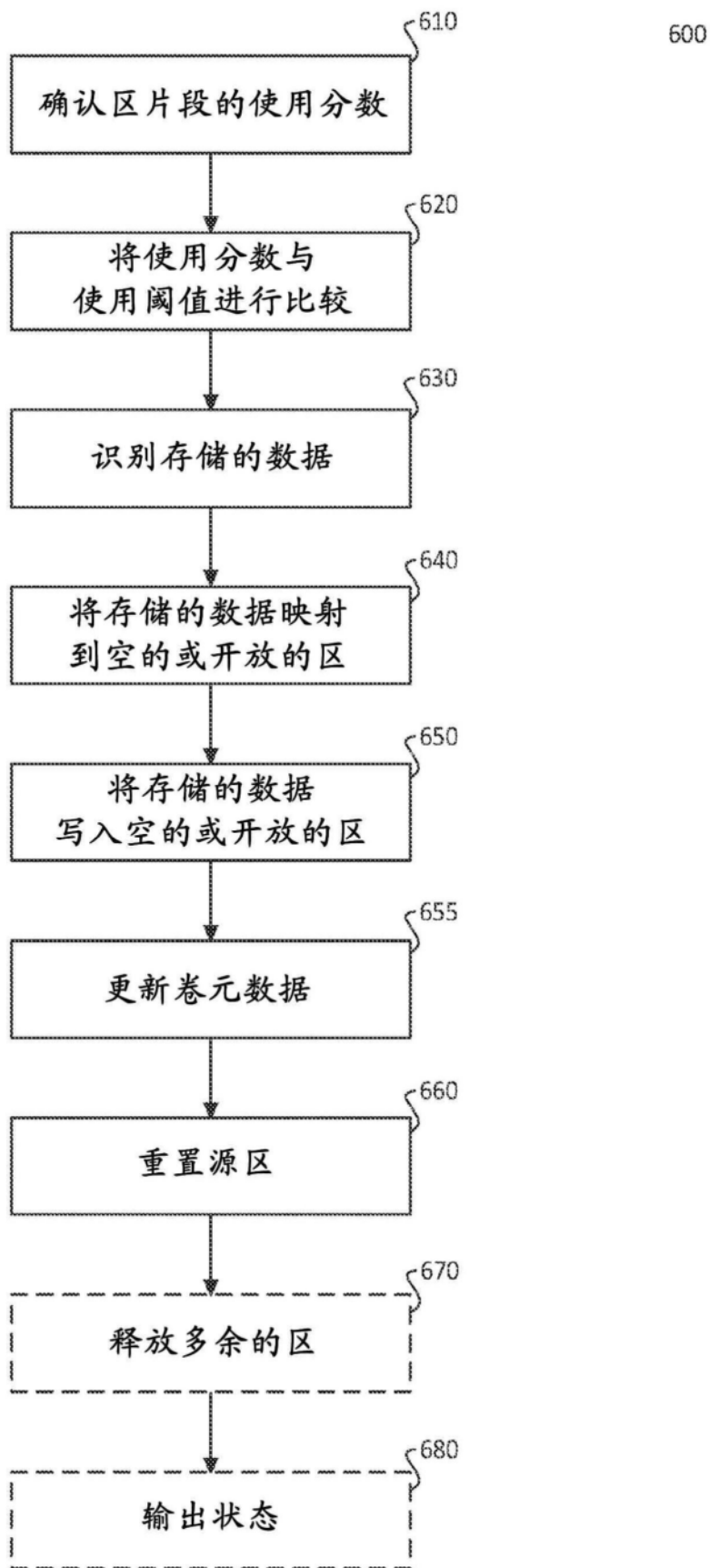


图6

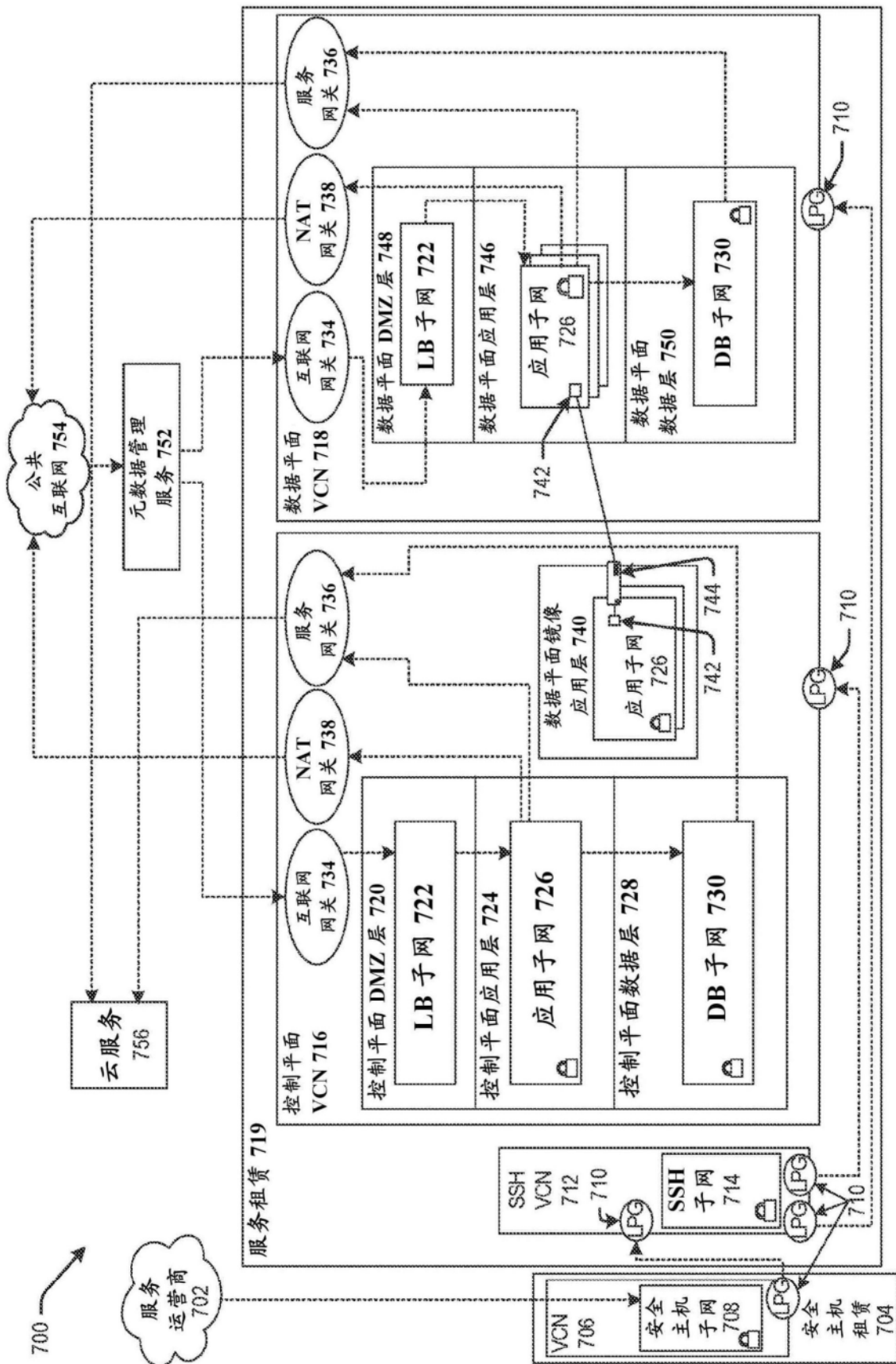


图7

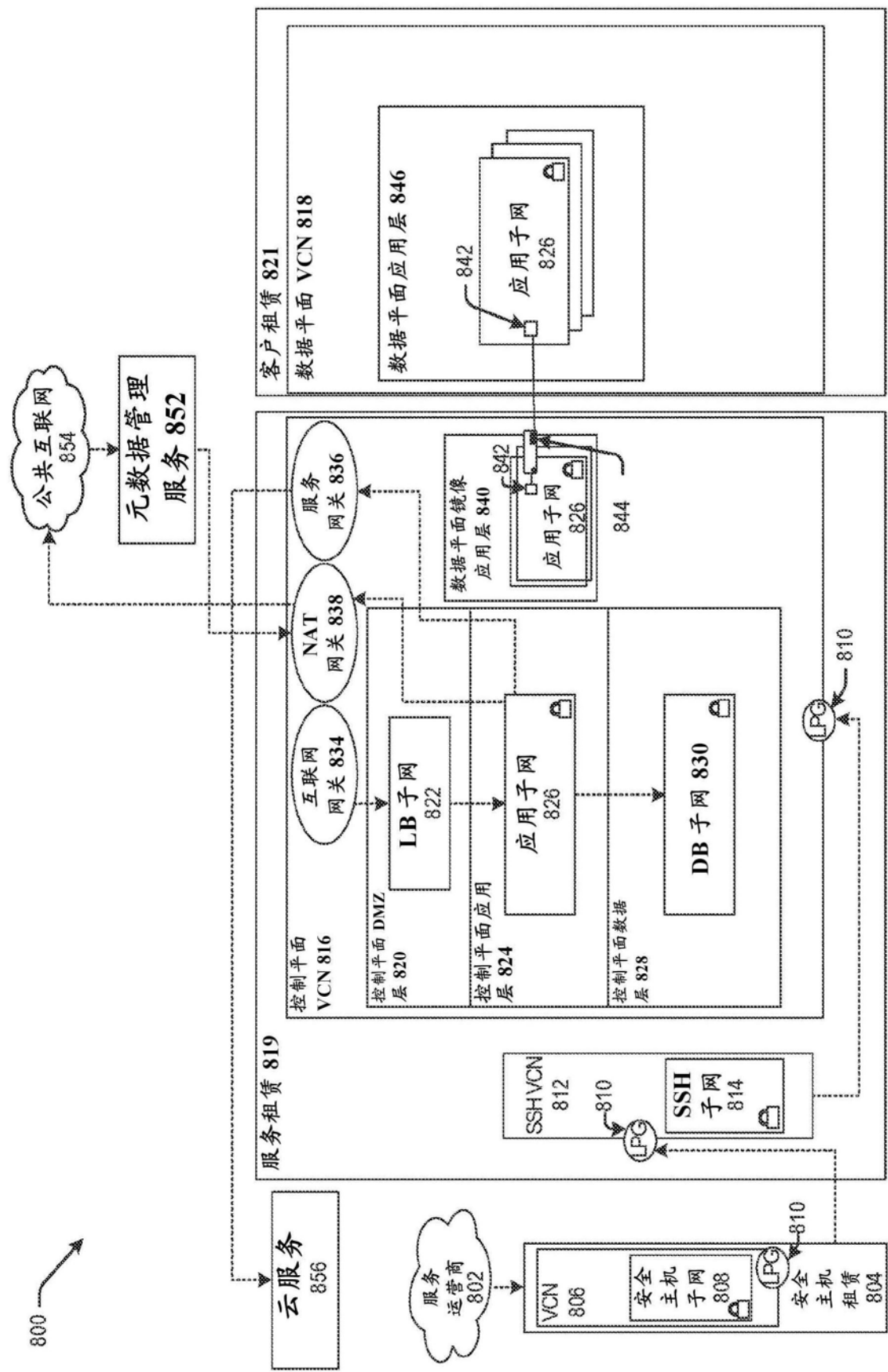


图8

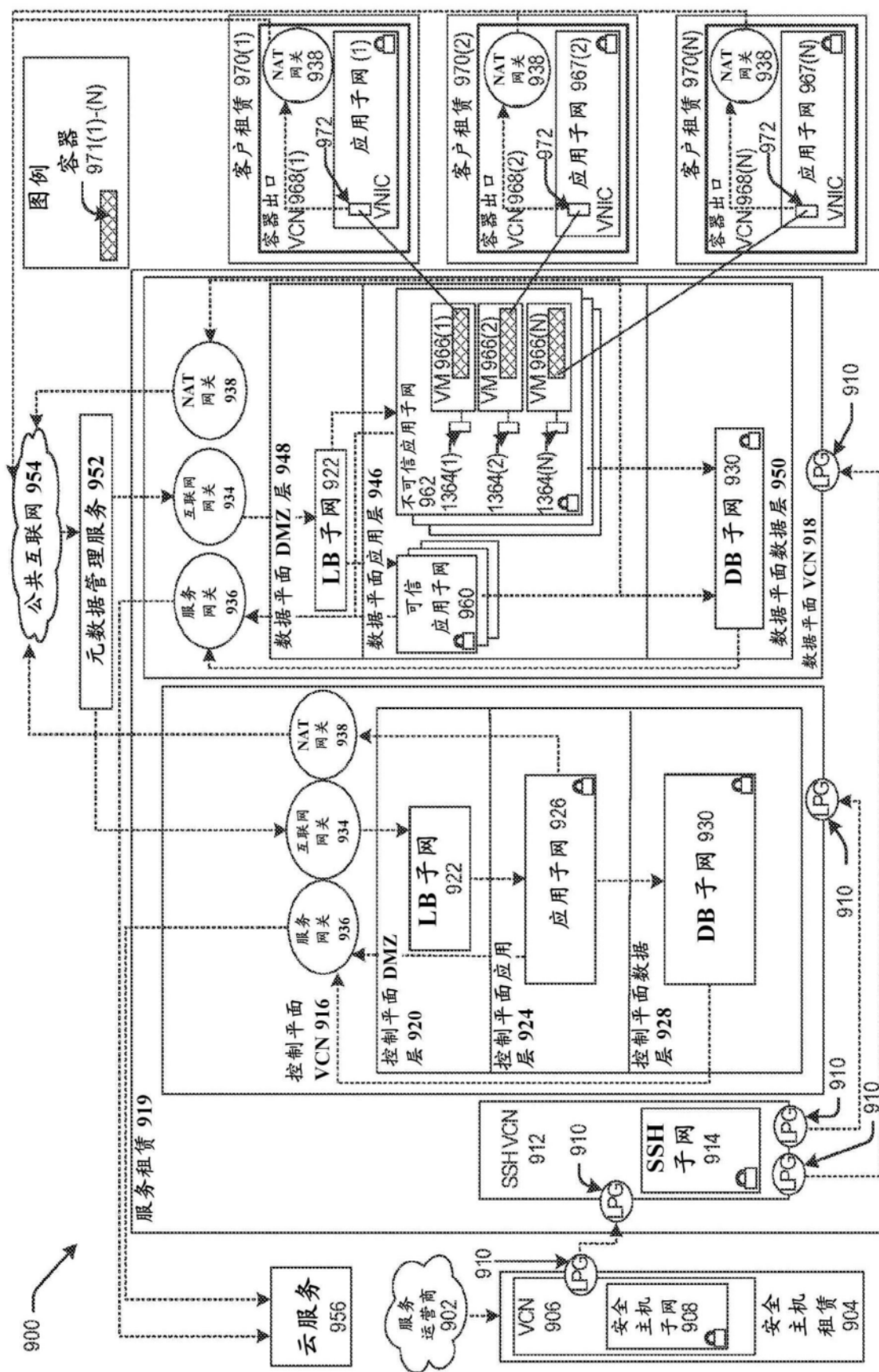
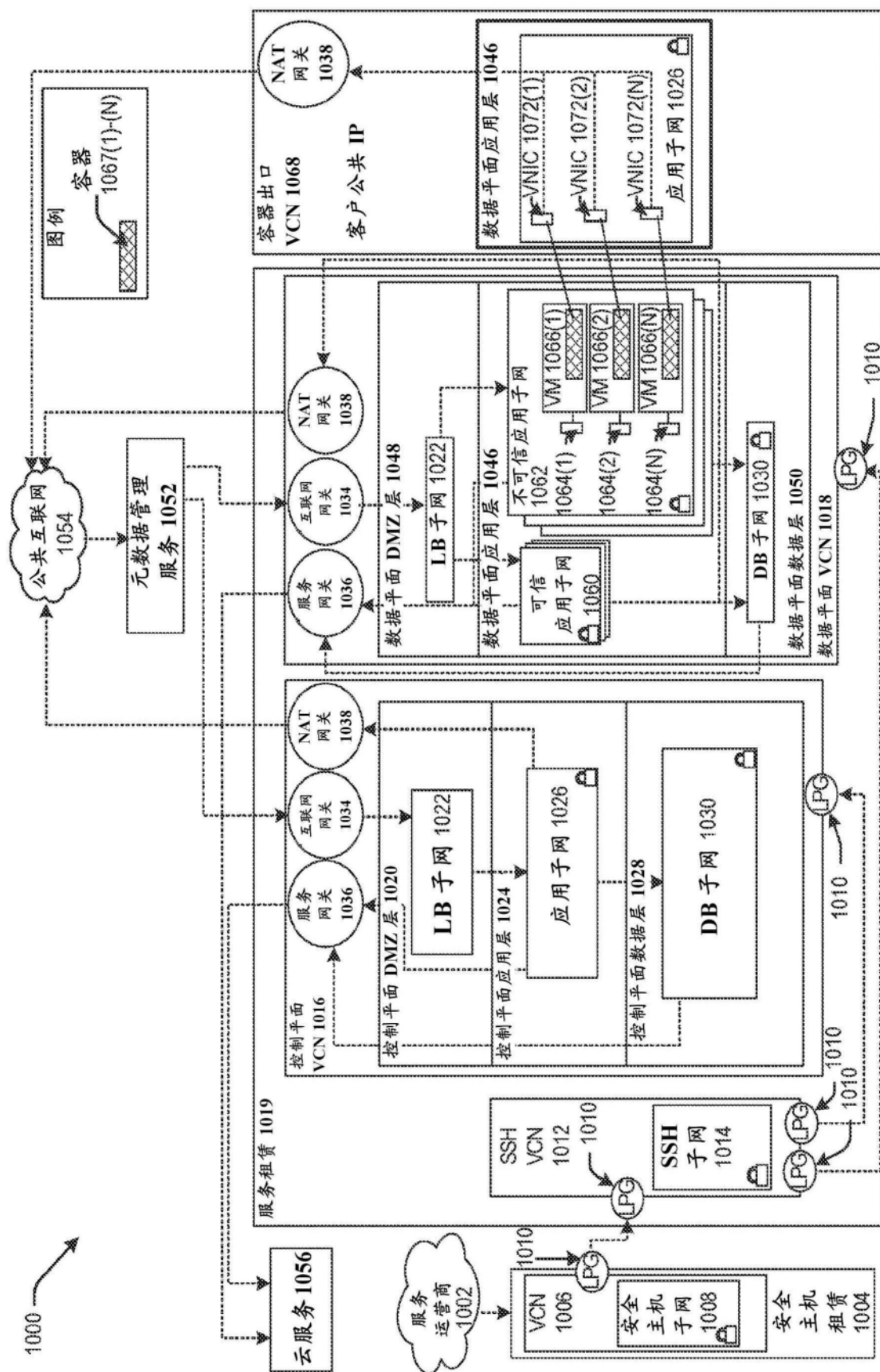


图9



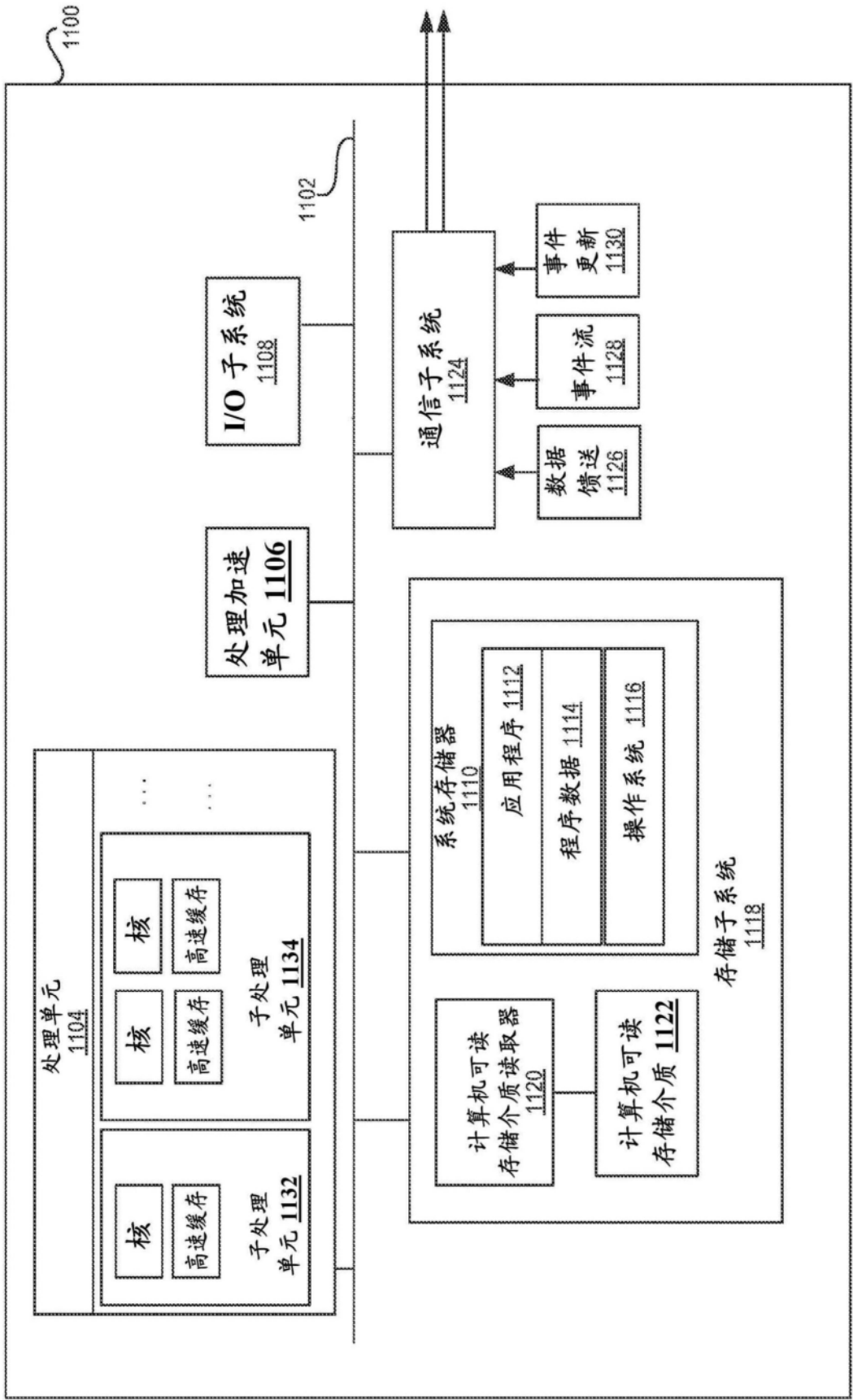


图11