



(12) 发明专利申请

(10) 申请公布号 CN 103389884 A

(43) 申请公布日 2013. 11. 13

(21) 申请号 201310322990. 6

(22) 申请日 2013. 07. 29

(71) 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为
总部办公楼

(72) 发明人 王婷 李亚琼 陈克平 陈为珑

(74) 专利代理机构 北京龙双利达知识产权代理
有限公司 11329

代理人 王君 肖鹂

(51) Int. Cl.

G06F 3/06 (2006. 01)

G06F 13/20 (2006. 01)

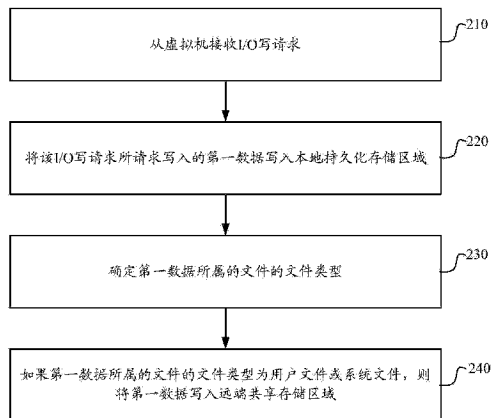
权利要求书4页 说明书16页 附图8页

(54) 发明名称

处理输入 / 输出请求的方法、宿主机、服务器
和虚拟机

(57) 摘要

本发明实施例提供处理 I/O 请求的方法和服务器。该方法包括：从虚拟机接收输入 / 输出 I/O 写请求；将该 I/O 写请求所请求写入的第一数据写入本地持久化存储区域；确定第一数据所属的文件的文件类型；如果第一数据所属的文件的文件类型为用户文件或系统文件，则将第一数据写入远端共享存储区域。本发明实施例中，通过确定本地持久化存储区域存储的第一数据所属的文件的文件类型，在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中，能够降低对远端共享存储区域的访问量。



1. 一种处理输入 / 输出请求的方法,其特征在於,包括:

从虚拟机接收输入 / 输出 I/O 写请求;

将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域;

确定所述第一数据所属的文件的文件类型,其中,所述文件类型包括:临时文件、用户文件和系统文件;

如果所述第一数据所属的文件的文件类型为用户文件或系统文件,则将所述第一数据写入远端共享存储区域。

2. 根据权利要求 1 所述的方法,其特征在於,所述确定所述第一数据所属的文件的文件类型,包括:

从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述所述虚拟机的虚拟磁盘的数据属性信息和文件系统的类型;

根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小,确定所述第一数据所属的文件的文件类型。

3. 根据权利要求 1 所述的方法,其特征在於,所述确定所述第一数据所属的文件的文件类型,包括:

获取所述 I/O 写请求中携带的第一文件信息,其中所述第一文件信息用于指示所述第一数据所属的文件;

根据所述第一文件信息,确定所述第一数据所属的文件的文件类型。

4. 根据权利要求 1 至 3 中任一项所述的方法,其特征在於,所述将所述第一数据写入远端共享存储区域,包括:

将所述第一数据与所述本地持久化存储区域中存储的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件;

将所述聚合后的数据写入所述远端共享存储区域。

5. 根据权利要求 1 至 4 中任一项所述的方法,其特征在於,在将所述第一数据写入远端共享存储区域之后,还包括:

若所述第一数据所属的文件的文件类型为用户文件或系统文件,且所述第一数据的使用频率大于或等于第一阈值,则将所述第一数据写入本地缓存区域。

6. 根据权利要求 1 至 5 中任一项所述的方法,其特征在於,还包括:

从所述虚拟机接收 I/O 读请求;

依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;

向所述虚拟机发送查找到的所述第三数据。

7. 根据权利要求 6 所述的方法,其特征在於,在所述向所述虚拟机发送查找到的所述第三数据之后,还包括:

确定所述第三数据所属的文件的文件类型;

在所述第三数据所属的文件的文件类型为用户文件或系统文件,且所述第三数据的使用频率大于或等于第二阈值时,将所述第三数据写入所述本地缓存区域中。

8. 根据权利要求 7 所述的方法,其特征在於,所述确定所述第三数据所属的文件的文件类型,包括:

从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;

根据所述镜像元数据、在所述虚拟磁盘上所述 I/O 读请求所对应的起始扇区以及所述第三数据的大小,确定所述第三数据所属的文件的文件类型。

9. 根据权利要求 7 所述的方法,其特征在于,所述确定所述第三数据所属的文件的文件类型,包括:

获取所述 I/O 读请求中携带的所述第二文件信息,其中所述第二文件信息是所述虚拟机确定的,所述第二文件信息用于指示所述第三数据所属的文件;

根据所述第二文件信息,确定所述第三数据所属的文件的文件类型。

10. 根据权利要求 2 或 8 所述的方法,其特征在于,在所述从内存虚拟镜像中读取镜像元数据之前,还包括:

生成所述内存虚拟镜像。

11. 一种处理输入 / 输出请求的方法,其特征在于,包括:

确定输入 / 输出 I/O 写请求对应的第一文件信息,所述第一文件信息用于指示所述 I/O 写请求所请求写入的第一数据所属的文件;

向宿主机 host 发送所述 I/O 写请求,所述 I/O 写请求携带所述第一文件信息,以便所述宿主机 host 在将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域之后,根据所述第一文件信息确定所述第一数据所属的文件的文件类型,并在确定所述第一数据所属的文件的文件类型为系统文件时,将所述第一数据写入远端共享存储区域。

12. 根据权利要求 11 所述的方法,其特征在于,还包括:

确定 I/O 读请求对应的第二文件信息,所述第二文件信息用于指示所述 I/O 读请求所请求读取的第三数据所属的文件;

向所述宿主机 host 发送所述 I/O 读请求,所述 I/O 读请求携带所述第二文件信息,以使所述宿主机 host 依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;

接收所述宿主机 host 发送的所述第三数据。

13. 一种宿主机,其特征在于,包括:

接收单元,用于从虚拟机接收输入 / 输出 I/O 写请求;

存储单元,用于将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域;

确定单元,用于确定所述存储单元写入到本地持久化存储区域中的第一数据所属的文件的文件类型,其中,所述文件类型包括:临时文件、用户文件和系统文件;

所述存储单元,还用于在所述第一数据所属的文件的文件类型为系统文件的情况下,将所述第一数据写入远端共享存储区域。

14. 根据权利要求 13 所述的宿主机,其特征在于,所述确定单元具体用于:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小,确定所述第一数据所属的文件的文件类型。

15. 根据权利要求 13 所述的宿主机,其特征在於,所述确定单元具体用于:获取所述 I/O 写请求中携带的所述第一文件信息,其中所述第一文件信息是所述虚拟机确定的,且所述第一文件信息用于指示所述第一数据所属的文件;根据所述第一文件信息,确定所述第一数据所属的文件的文件类型。

16. 根据权利要求 13 至 15 中任一项所述的宿主机,其特征在於,所述存储单元具体用于:将所述第一数据与所述本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件;将所述聚合后的数据写入所述远端共享存储区域。

17. 根据权利要求 13 至 16 中任一项所述的宿主机,其特征在於,所述存储单元还用于在所述将所述第一数据写入远端共享存储区域之后,若所述第一数据所属的文件的文件类型为用户文件或系统文件,且所述第一数据的使用频率大于或等于第一阈值的情况下,则将所述第一数据写入本地缓存区域。

18. 根据权利要求 13 至 17 中任一项所述的宿主机,其特征在於,还包括:
接收单元,还用于从所述虚拟机接收 I/O 读请求;
查找单元,用于依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;
发送单元,用于向所述虚拟机发送查找到的所述第三数据。

19. 根据权利要求 18 所述的宿主机,其特征在於,所述确定单元,还用于在所述向所述虚拟机发送查找到的所述第三数据之后,确定所述第三数据所属的文件的文件类型;
所述存储单元,还用于在所述第三数据所属的文件的文件类型为系统文件,且所述第三数据的使用频率大于或等于第二阈值时,将所述第三数据写入所述本地缓存区域中。

20. 根据权利要求 19 所述的宿主机,其特征在於,所述确定单元具体用于:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、在所述虚拟磁盘上所述 I/O 读请求所对应的起始扇区以及所述第三数据的大小,确定所述第三数据所属的文件的文件类型。

21. 根据权利要求 19 所述的宿主机,其特征在於,所述确定单元具体用于:获取所述 I/O 读请求中携带的所述第二文件信息,其中所述第二文件信息是所述虚拟机确定的,所述第二文件信息用于指示所述第三数据所属的文件;根据所述第二文件信息,确定所述第三数据所属的文件的文件类型。

22. 根据权利要求 14 或 20 所述的宿主机,其特征在於,还包括:
生成单元,用于在从内存虚拟镜像中读取镜像元数据之前,生成所述内存虚拟镜像。

23. 一种服务器,其特征在於,包括:
硬件层、运行在所述硬件层之上的虚拟机监视器,以及运行在所述虚拟机监视器之上的宿主机 Host 和虚拟机 VM,其中,所述硬件层包括输入/输出 I/O 设备、本地持久化存储区域;

其中,所述宿主机 Host,用于从所述虚拟机接收 I/O 写请求,所述 I/O 写请求是客户端

通过 I/O 设备发送给所述虚拟机的；将所述 I/O 写请求所请求写入的第一数据写入所述本地持久化存储区域，确定所述第一数据所属的文件的文件类型，其中，所述文件类型包括：临时文件、用户文件和系统文件；如果所述第一数据所属的文件的文件类型为系统文件，则将所述第一数据写入所述服务器连接的远端共享存储区域。

24. 根据权利要求 23 所述的服务器，其特征在于，所述宿主机 Host 具体用于：从内存虚拟镜像中读取镜像元数据，所述内存虚拟镜像用于存储镜像元数据，所述镜像元数据用于描述所述虚拟机的虚拟磁盘的数据属性信息和文件系统的类型；根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小，确定所述第一数据所属的文件的文件类型。

25. 根据权利要求 23 所述的服务器，其特征在于，所述宿主机 Host 具体用于：获取所述 I/O 写请求中携带的所述第一文件信息，其中所述第一文件信息是所述虚拟机确定的，且所述第一文件信息用于指示所述第一数据所属的文件；根据所述第一文件信息，确定所述第一数据所属的文件的文件类型。

26. 根据权利要求 23-25 任一项所述的服务器，其特征在于，所述宿主机 Host 具体用于：将所述第一数据与所述本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据，其中第二数据所属的文件的文件类型也为系统文件；将所述聚合后的数据写入所述远端共享存储区域。

27. 根据权利要求 26 所述的服务器，其特征在于，所述硬件层还包括：本地缓存；

所述宿主机 Host，还用于在所述将所述第一数据写入远端共享存储区域之后，若所述第一数据所属的文件的文件类型为系统文件，且所述第一数据的使用频率大于或等于第一阈值的情况下，则将所述第一数据写入本地缓存区域。

28. 一种虚拟机，其特征在于，包括：

确定单元，用于确定输入/输出 I/O 写请求对应的第一文件信息，所述第一文件信息用于指示所述 I/O 写请求所请求写入的第一数据所属的文件；

发送单元，用于向宿主机 host 发送所述 I/O 写请求，所述 I/O 写请求携带所述第一文件信息，以便所述宿主机 host 在将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域之后，根据所述第一文件信息确定所述第一数据所属的文件的文件类型，并在确定所述第一数据所属的文件的文件类型为系统文件时，将所述第一数据写入远端共享存储区域。

29. 根据权利要求 28 所述的虚拟机，其特征在于，所述确定单元还用于，确定 I/O 读请求对应的第二文件信息，所述第二文件信息用于指示所述 I/O 读请求所请求读取的第三数据所属的文件；向所述宿主机 host 发送所述 I/O 读请求，所述 I/O 读请求携带所述第二文件信息，以使所述宿主机 host 依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据，直到查找到所述第三数据为止；

所述虚拟机还包括：接收单元，用于接收所述宿主机 host 发送的所述第三数据。

处理输入 / 输出请求的方法、宿主机、服务器和虚拟机

技术领域

[0001] 本发明涉及信息技术领域,并且具体地,涉及处理输入 / 输出(input/output, I/O)请求的方法、宿主机、服务器和虚拟机。

背景技术

[0002] 用户异地工作时能够快捷地访问平时的桌面环境成为新的需求,因此桌面云解决方案得到了越来越广泛的应用。在桌面云方案中,用户可以通过瘦客户机、平板电脑或者其它与网络连接的设备来访问跨平台的应用程序或者整个客户桌面,能够获得与传统的个人电脑一样的体验。

[0003] 在桌面云方案中,通常采用虚拟桌面基础架构(Virtual Desktop Infrastructure, VDI)形式。这种 VDI 形式采用“集中计算,分布显示”的原则,支持客户端桌面工作负载托管在数据中心,比如操作系统、应用程序和用户数据等。用户可以通过支持远程桌面协议的客户端设备与数据中心的虚拟机进行通信来实现对虚拟桌面环境的访问。由于虚拟磁盘由数据中心的远端共享存储统一管理,因此在用户发起对虚拟桌面环境的写请求时,每个虚拟机需要在远端共享存储中更新虚拟磁盘, I/O 路径较长,从而导致 I/O 写请求处理延迟,影响用户体验。

发明内容

[0004] 本发明实施例提供处理 I/O 请求的方法、宿主机、服务器和虚拟机,能够降低对远端共享存储区域的访问量,并能够减少 I/O 写请求的延迟,提升用户体验。

[0005] 第一方面,提供了一种处理输入 / 输出请求的方法,包括:从虚拟机接收输入 / 输出 I/O 写请求;将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域;确定所述第一数据所属的文件的文件类型,其中,所述文件类型包括:临时文件、用户文件和系统文件;如果所述第一数据所属的文件的文件类型为用户文件或系统文件,则将所述第一数据写入远端共享存储区域。

[0006] 结合第一方面,在第一种可能的实现方式中,所述确定所述第一数据所属的文件的文件类型,包括:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小,确定所述第一数据所属的文件的文件类型。

[0007] 结合第一方面,在第二种可能的实现方式中,所述确定所述第一数据所属的文件的文件类型,包括:获取所述 I/O 写请求中携带的第一文件信息,其中所述第一文件信息是由所述虚拟机确定的,且所述第一文件信息用于指示所述第一数据所属的文件;根据所述第一文件信息,确定所述第一数据所属的文件的文件类型。

[0008] 结合第一方面或第一方面的第一种可能的实现方式或第二种可能的实现方式,在第三种可能的实现方式中,所述将所述第一数据写入远端共享存储区域,包括:将所述第一

数据与所述本地持久化存储区域中存储的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件;将所述聚合后的数据写入所述远端共享存储区域。

[0009] 结合第一方面或第一方面的第一种可能的实现方式至第三种可能的实现方式中任一实现方式,在第四种可能的实现方式中,在所述将所述第一数据写入远端共享存储区域之后,还包括:若所述第一数据所属的文件为用户文件或系统文件,且所述第一数据的使用频率大于或等于第一阈值,则将所述第一数据写入本地缓存区域。

[0010] 结合第一方面或第一方面的第一种可能的实现方式至第三种可能的实现方式中任一实现方式,在第五种可能的实现方式中,还包括:从所述虚拟机接收 I/O 读请求;依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;向所述虚拟机发送查找到的所述第三数据。

[0011] 结合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,在所述向所述虚拟机发送查找到的所述第三数据之后,还包括:确定所述第三数据所属的文件的文件类型;在所述第三数据所属的文件的文件类型为系统文件,且所述第三数据的使用频率大于或等于第二阈值时,将所述第三数据写入所述本地缓存区域中。

[0012] 结合第一方面的第六种可能的实现方式,在第七种可能的实现方式中,所述确定所述第三数据所属的文件的文件类型,包括:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、在所述虚拟磁盘上所述 I/O 读请求所对应的起始扇区以及所述第三数据的大小,确定所述第三数据所属的文件的文件类型。

[0013] 结合第一方面的第六种可能的实现方式,在第八种可能的实现方式中,所述确定所述第三数据所属的文件的文件类型,包括:获取所述 I/O 读请求中携带的所述第二文件信息,其中所述第二文件信息是所述虚拟机确定的,所述第二文件信息用于指示所述第三数据所属的文件;根据所述第二文件信息,确定所述第三数据所属的文件的文件类型。

[0014] 结合第一方面的第一种可能的实现方式或第七种可能的实现方式,在第九种可能的实现方式中,在所述从内存虚拟镜像中读取镜像元数据之前,还包括:生成所述内存虚拟镜像。

[0015] 第二方面,提供了一种处理输入/输出请求的方法,包括:确定输入/输出 I/O 写请求对应的第一文件信息,所述第一文件信息用于指示所述 I/O 写请求所请求写入的第一数据所属的文件;向宿主机 host 发送所述 I/O 写请求,所述 I/O 写请求携带所述第一文件信息,以便所述宿主机 host 在将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域之后,根据所述第一文件信息确定所述第一数据所属的文件的文件类型,并在确定所述第一数据所属的文件的文件类型为系统文件时,将所述第一数据写入远端共享存储区域。

[0016] 结合第二方面,在第一种可能的实现方式中,还包括:确定 I/O 读请求对应的第二文件信息,所述第二文件信息用于指示所述 I/O 读请求所请求读取的第三数据所属的文件;向所述宿主机 host 发送所述 I/O 读请求,所述 I/O 读请求携带所述第二文件信息,以使所述宿主机 host 依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区

域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;接收所述宿主机 host 发送的所述第三数据。

[0017] 第三方面,提供了一种宿主机,包括:接收单元,用于从虚拟机接收输入/输出 I/O 写请求;存储单元,用于将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域;确定单元,用于确定所述存储单元写入到本地持久化存储区域中的第一数据所属的文件的文件类型,其中,所述文件类型包括:临时文件、用户文件和系统文件;所述存储单元,还用于在所述第一数据所属的文件的文件类型为系统文件的情况下,将所述第一数据写入远端共享存储区域。

[0018] 结合第三方面,在第一种可能的实现方式中,所述确定单元具体用于:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小,确定所述第一数据所属的文件的文件类型。

[0019] 结合第三方面,在第二种可能的实现方式中,所述确定单元具体用于:获取所述 I/O 写请求中携带的所述第一文件信息,其中所述第一文件信息是所述虚拟机确定的,且所述第一文件信息用于指示所述第一数据所属的文件;根据所述第一文件信息,确定所述第一数据所属的文件的文件类型。

[0020] 结合第三方面或第三方面的第一种可能的实现方式或第二种可能的实现方式,在第三种可能的实现方式中,所述存储单元具体用于:将所述第一数据与所述本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为系统文件;将所述聚合后的数据写入所述远端共享存储区域。

[0021] 结合第三方面或第三方面的第一种可能的实现方式至第三种可能的实现方式中任一实现方式,在第四种可能的实现方式中,所述存储单元还用于在所述将所述第一数据写入远端共享存储区域之后,若所述第一数据所属的文件为用户文件或系统文件,且所述第一数据的使用频率大于或等于预设的第一阈值,则将所述第一数据写入本地缓存区域。

[0022] 结合第三方面或第三方面的第一种可能的实现方式至第四种可能的实现方式中任一实现方式,在第五种可能的实现方式中,还包括:接收单元,还用于从所述虚拟机接收 I/O 读请求;查找单元,用于依次在本地缓存区域、所述本地持久化存储区域和所述远端共享存储区域查找所述 I/O 读请求所请求读取的第三数据,直到查找到所述第三数据为止;发送单元,用于向所述虚拟机发送查找到的所述第三数据。

[0023] 结合第三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,所述确定单元,还用于在所述向所述虚拟机发送查找到的所述第三数据之后,确定所述第三数据所属的文件的文件类型;所述存储单元,还用于在所述第三数据所属的文件的文件类型为系统文件,且所述第三数据的使用频率大于或等于第二阈值时,将所述第三数据写入所述本地缓存区域中。

[0024] 结合第三方面的第六种可能的实现方式,在第七种可能的实现方式中,所述确定单元具体用于:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、在所述虚拟磁盘上所述 I/O 读请求所对应的起始扇区以及所述第三数据的大小,

确定所述第三数据所属的文件的文件类型。

[0025] 结合第三方面的第六种可能的实现方式,在第八种可能的实现方式中,所述确定单元具体用于获取所述 I/O 读请求中携带的所述第二文件信息,其中所述第二文件信息是所述虚拟机确定的,所述第二文件信息用于指示所述第三数据所属的文件;根据所述第二文件信息,确定所述第三数据所属的文件的文件类型。

[0026] 结合第三方面的第一种可能的实现方式或第七种可能的实现方式,在第九种可能的实现方式中,还包括:生成单元,用于在从内存虚拟镜像中读取镜像元数据之前,生成所述内存虚拟镜像。

[0027] 第四方面,提供了一种服务器,包括:硬件层、运行在所述硬件层之上的宿主机 Host、以及运行在所述 Host 之上的虚拟机 VM,其中,所述硬件层包括输入/输出 I/O 设备、本地持久化存储区域;其中,所述宿主机 Host,用于从所述虚拟机接收 I/O 写请求,所述 I/O 写请求是客户端通过 I/O 设备发送给所述虚拟机的;将所述 I/O 写请求所请求写入的第一数据写入所述本地持久化存储区域,确定所述第一数据所属的文件的文件类型,其中,所述文件类型包括:临时文件、用户文件和系统文件;如果所述第一数据所属的文件的文件类型为用户文件或系统文件,则将所述第一数据写入所述服务器连接的远端共享存储区域。

[0028] 结合第四方面,在第一种可能的实现方式中,所述宿主机 Host 具体用于:从内存虚拟镜像中读取镜像元数据,所述内存虚拟镜像用于存储镜像元数据,所述镜像元数据用于描述所述虚拟机的虚拟磁盘的数据属性信息和文件系统的类型;根据所述镜像元数据、所述 I/O 写请求在所述虚拟磁盘上所对应的起始扇区以及所述第一数据的大小,确定所述第一数据所属的文件的文件类型。

[0029] 结合第四方面,或第四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,所述宿主机 Host 具体用于:将所述第一数据与所述本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件;将所述聚合后的数据写入所述远端共享存储区域。

[0030] 第五方面,提供了一种虚拟机,包括:确定单元,用于确定输入/输出 I/O 写请求对应的第一文件信息,所述第一文件信息用于指示所述 I/O 写请求所请求写入的第一数据所属的文件;发送单元,用于向宿主机 host 发送所述 I/O 写请求,所述 I/O 写请求携带所述第一文件信息,以便所述宿主机 host 在将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域之后,根据所述第一文件信息确定所述第一数据所属的文件的文件类型,并在确定所述第一数据所属的文件的文件类型为用户文件或系统文件时,将所述第一数据写入远端共享存储区域。

[0031] 本发明实施例提供的以上技术方案,通过确定本地持久化存储区域存储的第一数据所属的文件的文件类型,在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

附图说明

[0032] 为了更清楚地说明本发明实施例的技术方案,下面将对本发明实施例中所需要使

用的附图作简单地介绍,显而易见地,下面所描述的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

- [0033] 图 1 是可应用本发明实施例的一个场景的示意图;
- [0034] 图 2 是根据本发明实施例的处理输入 / 输出请求的方法的流程图;
- [0035] 图 3 是根据本发明实施例的内存虚拟镜像的结构示意框图;
- [0036] 图 4a 是根据本发明一个实施例的服务器的结构示意图;
- [0037] 图 4b 是根据本发明另一实施例的服务器的结构示意图;
- [0038] 图 5 是根据本发明一个实施例的处理输入 / 输出请求的方法流程图;
- [0039] 图 6 是根据本发明另一实施例的处理输入 / 输出请求的方法流程图;
- [0040] 图 7 是根据本发明另一实施例的处理输入 / 输出请求的方法流程图;
- [0041] 图 8 是根据本发明一个实施例的服务器的示意框图;
- [0042] 图 9 是根据本发明实施例的虚拟机的示意框图;
- [0043] 图 10 是根据本发明另一实施例的服务器的示意框图;
- [0044] 图 11 是根据本发明另一实施例的服务器的示意框图。

具体实施方式

[0045] 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的一部分实施例,而不是全部实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都应属于本发明保护的范围。

[0046] 为了方便理解本发明实施例,首先在此介绍本发明实施例描述中会引入的几个要素;

[0047] 虚拟机 VM:通过虚拟机软件可以在一台物理计算机或服务器上模拟出一台或者多台虚拟的计算机,而这些虚拟机就像真正的计算机那样进行工作,虚拟机上可以安装操作系统和应用程序,虚拟机还可访问网络资源。对于在虚拟机中运行的应用程序而言,虚拟机就像是在真正的计算机中进行工作。

[0048] 硬件层:虚拟化环境运行的硬件平台。其中,硬件层可包括多种硬件,例如某服务器的硬件层可包括 CPU 和内存,还可以包括网卡、存储器、高速 / 低速输入 / 输出(I/O, Input/Output) 设备,及具有特定处理功能的其它设备,如输入输出内存管理单元(IOMMU, Input/Output Memory Management Unit),其中 IOMMU 可用于虚拟机物理地址和 Host 物理地址的转换。

[0049] 宿主机(Host):作为管理层,用以完成硬件资源的管理、分配;为虚拟机呈现虚拟硬件平台;实现虚拟机的调度和隔离。更具体地,在本发明的所有实施例中宿主机(Host)是指服务器中具有虚拟磁盘服务功能的系统组件,比如,Host 可能是一个特权虚拟机,也可以是虚拟机监控器(VMM);此外,有时 VMM 和一个特权虚拟机配合,两者结合组成 Host,在不同的虚拟化架构中,其名称可能有差异,本发明实施例不做特别的限定。其中,虚拟硬件平台对其上运行的各个虚拟机提供各种硬件资源,如提供虚拟 CPU、内存、虚拟磁盘、虚拟网卡等等。其中,该虚拟磁盘可对应 Host 的一个文件或者一个逻辑块设备。

[0050] 图 1 是可应用本发明实施例的一个场景的示意图。

[0051] 如图 1 所示,该场景可以包括计算集群池和存储集群池。计算集群池可以包括一组服务器,如图 1 所示,计算集群池可以包括服务器 110a、110b、110c 和 110d。服务器 110a 至 110d 之间可以通过松散集成的计算机软件和 / 或硬件连接起来协同作业。每个服务器上可以运行一个或多个虚拟机,同时还部署有宿主机 host,用以实现虚拟磁盘服务功能。另外,每个服务器可以具有本地存储,本地存储可以包括本地缓存和本地持久化存储。每个服务器可以在自己的本地存储上进行数据的读写操作。

[0052] 存储集群池可以包括多个存储集群,如图 1 所示的存储集群 120a 和 120b。每个存储集群可以包括多个存储设备,多个存储设备的存储空间可以形成一个共享存储池,能够给计算集群池中的服务器提供统一访问接口和管理界面。因此,相对于服务器的本地存储而言,存储集群池可以称为远端共享存储。

[0053] 如图 1 所示,计算集群池的服务器 110a 至 110d 均可以访问存储集群池中的存储设备来进行数据的读写操作。例如,服务器 110a 至 110d 可以通过标准的网络拓扑结构与存储集群池连接,这种情况下,存储集群池可以称为网络附加存储(Network Attached Storage, NAS)。或者,服务器 110a 至 110d 可以通过光纤通道技术与存储集群池连接,这种情况下,存储集群池可以称为存储区域网络(Storage Area Network, SAN)。

[0054] 因此,本发明实施例中,存储方式可以包括本地持久化存储和远端共享存储等方式。本地持久化存储区域可以实现本地持久化存储的方式。本地持久化存储方式具有访问速度较快和数据非易失等特点。本地持久化存储区域可以包括固态存储产品(Solid Storage Disk, SDD)或硬盘驱动器(Hard Disk Driver, HDD)等。远端共享存储区域可以实现远端共享存储的方式。远端共享存储方式具有访问速度较慢、数据非易失和多服务器共享等特点。远端共享存储区域可以包括 NAS 或网际协议存储区域网络(Internet Protocol Storage Area Network, IPSAN)。

[0055] 图 2 是根据本发明实施例的处理 I/O 请求的方法的示意性流程图。图 2 的方法由服务器执行,例如可以由图 1 所示的服务器 110a 至 110d 之一执行,或者更确切地说,由服务器的宿主机 host,或虚拟磁盘服务组件来执行。

[0056] 210,从虚拟机接收 I/O 写请求;

[0057] 220,将该 I/O 写请求所请求写入的第一数据写入本地持久化存储区域;

[0058] 例如,在客户端设备向虚拟机发起访问虚拟桌面的写请求时,虚拟机可以基于客户端设备的写请求,向服务器(或者更具体地是服务器的虚拟磁盘服务系统组件)发送 I/O 写请求。服务器可以在接收到虚拟机的 I/O 写请求之后,将第一数据写入本地持久化存储区域中。

[0059] 230,确定本地持久化存储区域中存储的第一数据所属的文件的文件类型,其中,文件类型包括:临时文件、用户文件和系统文件。

[0060] 例如,对于新技术文件系统(New Technology File System, NTFS)来说,第一数据所属的文件可以是 \$Boot、\$MFT 或 \xxx\aaa.txt 等。比如,\$Boot 和 \$MFT 文件可以为系统文件。 \xxx\aaa.txt 文件可以为用户文件。

[0061] 240,如果第一数据所属的文件的文件类型为用户文件或系统文件,则将第一数据写入远端共享存储区域。

[0062] 由上述可见,现有技术中,无论何种类型的数据均要写入远端共享存储区域中。本发明实施例中,将第一数据写入本地持久化存储区域中,然后在第一数据所属的文件类型为用户文件或系统文件的情况下,再将本地持久化存储的第一数据写入远端共享存储区域。如果第一数据所属的文件为临时文件,则可以仅存储在本地,而无需写入远端共享存储区域。因此本发明实施例能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0063] 本发明实施例中,通过确定本地持久化存储区域存储的第一数据所属的文件的文件类型,在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0064] 可选地,作为一个实施例,在步骤 230 之前,可以生成内存虚拟镜像,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型。

[0065] 内存虚拟镜像可以包括元数据区,元数据区用于存储镜像元数据。镜像元数据可以用于描述虚拟磁盘的数据属性信息以及文件系统类型。

[0066] 远端共享存储区域可以包括数据区和上述元数据区,远端共享存储区域上的数据区和元数据区可以称为虚拟磁盘。虚拟磁盘的数据的载体可以是普通文件、块设备或其它格式的虚拟机镜像文件等。虚拟磁盘的数据属性信息可以用于指示虚拟磁盘的数据的属性。虚拟磁盘的文件系统类型可以为现有技术中各种文件系统的类型,例如 NTFS。

[0067] 下面以文件系统为 NTFS 为例,说明内存虚拟镜像的结构。图 3 是根据本发明实施例的内存虚拟镜像的结构示意框图。

[0068] 如图 3 所示,在文件系统为 NTFS 时,内存虚拟镜像可以包括多个虚拟分区(Virtual Partition)、虚拟主引导记录(Virtual Master Boot Record, VMBR)和虚拟磁盘分区表(Virtual Disk Partition Table, VDPT)。在各个虚拟分区中,可以包括虚拟卷(Virtual Volume)启动扇区(Boot Sector)、虚拟主文件表(Virtual Master File Table, VMFT)和虚拟数据(Virtual DATA, VDATA)等部分。每个部分可以由一个或多个块区(Block Region, BR)组成。每个 BR 可以表示内存虚拟镜像中某数据段的真实存放路径和文件系统的相关特性。镜像元数据可以存储在各个 BR 中。

[0069] 可选地,作为另一实施例,在步骤 230 中,可以从内存虚拟镜像中读取镜像元数据,根据镜像元数据、I/O 写请求在虚拟磁盘上所对应的起始扇区以及第一数据的大小,确定第一数据所属的文件的文件类型。

[0070] 具体地,服务器通过从内存虚拟镜像读取镜像元数据,可以解析虚拟磁盘的布局。然后根据 I/O 写请求所对应的起始扇区和第一数据的大小,确定在文件系统中第一数据的类型。此处,第一数据的大小是指第一数据所需要占用的存储空间,其具体可以使用字节(byte)或比特(bit)等常用的单位来衡量。例如,对于文件系统类型为 NTFS 的虚拟磁盘来说,第一数据的类型可以为 Boot Sector(启动扇区)、文件元数据或文件数据等。根据第一数据的类型,可以确定第一数据所属的文件。例如,对于文件系统类型为 NTFS 的虚拟磁盘来说,第一数据所属的文件可以为 \$Boot、\$MFT 或 \xxx\aaa.txt 等,从而确定第一数据所属

的文件的文件类型。

[0071] 具体地,在 Xen 架构下,服务器上可以运行多个虚拟域(domain),虚拟域可以包括特权域和非特权域。特权域通常指在服务器上运行的第一个虚拟域,能够对服务器上的其它域进行管理和控制。在本发明实施例中,我们称特权域为宿主机(host),非特权域为虚拟机。

[0072] 例如,可以在服务器的特权域的文件系统上添加驱动层,由驱动层完成上述确定第一文件信息的过程。图 4a 是根据本发明一个实施例的服务器的结构示意图。如图 4a 所示,在特权域的文件系统上增加了驱动层,通过驱动层可以完成上述确定第一文件信息的过程。

[0073] 可选地,作为另一实施例,在步骤 230 中,可以获取 I/O 写请求中携带的第一文件信息,其中第一文件信息用于指示第一数据所属的文件。然后可以根据第一文件信息,确定第一数据所属的文件的文件类型。

[0074] 具体地,虚拟机可以将文件信息和 I/O 请求进行关联。例如,可以在虚拟机的文件系统和块设备层上下分别添加驱动层来实现文件信息和 I/O 请求之间的关联。图 4b 是根据本发明另一实施例的服务器的结构示意图。如图 4b 所示,在非特权域(也就是虚拟机)的文件系统上面添加了驱动层,并在非特权域的块设备层下面添加了驱动层。这样,非特权域,也就是虚拟机,可以通过两个驱动层,完成确定第一文件信息的过程,并将第一文件信息发送给特权域。第一文件信息可以用于指示第一数据所属的文件。服务器的特权域可以直接从 I/O 写请求获取第一文件信息,从而确定第一数据所属的文件的文件类型。

[0075] 可选地,作为另一实施例,可以将本地持久化存储区域中存储的第一数据与本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件,然后将聚合后的数据写入远端共享存储区域。

[0076] 宿主机可以将本地持久化存储区域中的第一数据与其它所属文件为用户文件或系统文件的数据进行聚合,也就是将本地持久化存储区域中的第一数据与第二数据进行聚合。该聚合可以是指对第一数据对应的上述 I/O 写请求与第二数据对应的 I/O 写请求聚合。例如,第二数据和第一数据可以是地址连续的数据,通过聚合得到聚合后的数据,然后通过一次 I/O 写请求将聚合后的数据写入远端共享存储区域中。该操作也可以称为写 I/O 聚合。服务器可以定期或不定期地执行该操作。

[0077] 本发明实施例中,可以将本地持久化存储区域中存储的第一数据与本地持久化存储区域中的多份第二数据进行聚合。例如,在本地持久化存储区域中,第一数据的起始地址为 5,大小为 10,第一份第二数据的起始地址为 20,大小为 5,第二份第二数据的起始地址为 15,大小为 10,可以将第一数据和两份第二数据聚合得到聚合后的数据,聚合后的数据的起始地址为 5,大小为 25,这样可以通过一次 I/O 写请求将聚合后的数据写入远端共享存储区域。

[0078] 本发明实施例中,通过将本地持久化存储区域中的第一数据与属于用户文件或系统文件的第二数据聚合后再写入远端共享存储区域,能够进一步降低远端共享存储区域的访问量,提升用户体验。

[0079] 可选地,作为另一实施例,在步骤 230 之后,若第一数据所属的文件的文件类型为系统文件或用户文件,且第一数据的使用频率大于或等于第一阈值,则可以将第一数据写

入本地缓存区域。

[0080] 上述存储方式还可以包括本地缓存方式。本地缓存方式具有访问速度快和数据易失等特点。本地缓存区域可以用于实现本地缓存方式。本地缓存区域可以使用内存等存储介质。

[0081] 如果第一数据属于用户文件或系统文件，而且是常用的数据，那么可以将第一数据也写入本地缓存区域中，从而能够提升数据访问速度。

[0082] 上述第一阈值可以是宿主机根据第一数据的使用情况以及其它数据的使用情况确定的，因此，第一阈值可以是变化的。例如，服务器经过一段时间的统计发现第一数据的使用频率为 40% 以上，而其它数据的使用频率均小于 40%，那么可以将第一阈值设置为 40%。

[0083] 可选地，作为另一实施例，可以从虚拟机接收 I/O 读请求，依次在本地缓存区域、本地持久化存储区域和远端共享存储区域查找 I/O 读请求所请求读取的第三数据，直到找到第三数据为止，并向虚拟机发送查找到的第三数据。

[0084] 具体地，宿主机可以根据虚拟机的 I/O 读请求，首先在本地缓存区域中查找第三数据。如果在本地缓存区域中查找到第三数据，则向虚拟机返回第三数据，如果未找到，则在本地持久化存储区域中查找第三数据。如果在本地持久化存储区域中找到第三数据，则向虚拟机返回第三数据，如果未找到，则在远端共享存储区域中查找第三数据，然后将查找到的第三数据发送给虚拟机。

[0085] 因此，本发明实施例中，由于依次在本地缓存区域、本地持久化存储区域和远端共享存储区域查找 I/O 读请求所请求读取的第三数据，如果在本地找到第三数据，那么就能够降低对远端共享存储的访问量，从而能够减少 I/O 读请求的延迟，提升用户体验。

[0086] 可选地，作为另一实施例，可以在向虚拟机发送查找到的第三数据之后，确定第三数据所属的文件的文件类型。在第三数据所属的文件的文件类型为系统文件，且第三数据的使用频率大于或等于第二阈值时，可以将第三数据写入本地缓存区域中。

[0087] 具体地，在第三数据是在本地持久化存储区域或远端共享存储区域中查找到的情况下，如果第三数据属于用户文件或系统文件，而且是常用的数据，那么可以将第三数据存储在本地的缓存区域中，从而能够提升数据访问速度。

[0088] 第二阈值也可以是宿主机根据第三数据的使用情况以及其它数据的使用情况确定的。确定方式类似于第一阈值的确定方式。

[0089] 可选地，作为另一实施例，可以从内存虚拟镜像中读取镜像元数据，并根据镜像元数据、在虚拟磁盘上 I/O 读请求所对应的起始扇区以及第三数据的大小，确定第三数据所属的文件的文件类型。

[0090] 该过程类似于确定第一数据所属的文件的文件类型的过程，不再赘述。

[0091] 可选地，作为另一实施例，可以获取 I/O 读请求中携带的第二文件信息，其中第二文件信息是虚拟机根据文件信息与 I/O 读请求之间的对应关系确定的。

[0092] 图 5 是根据本发明一个实施例的处理 I/O 写请求的方法的过程的示意性流程图，图 5 的方法由服务器执行，例如可以由图 1 所示的服务器 110a 至 110d 之一执行，或者更确切地说，由服务器的宿主机 host，或虚拟磁盘服务组件来执行。

[0093] 501，将 I/O 写请求所请求写入的第一数据写入本地持久化存储区域。

[0094] 502，确定本地持久化存储区域存储的第一数据所属的文件的文件类型，文件类型

可以包括：临时文件、用户文件和系统文件。

[0095] 503, 确定第一数据所属的文件的文件类型是否为用户文件或系统文件。

[0096] 如果第一数据所属的文件的文件类型为用户文件或系统文件, 转到步骤 504。

[0097] 如果第一数据所属的文件的文件类型为临时文件, 流程结束。

[0098] 504, 如果第一数据所属文件的文件类型为用户文件或系统文件, 将本地持久化存储区域中的第一数据写入远端共享存储区域。

[0099] 在步骤 503 之后, 可以根据实际情况执行步骤 504, 例如可以立即执行步骤 504, 也可以在一段时间后执行步骤 504。例如, 服务器可以在其系统空闲时, 或者每隔一段时间, 或者本地持久化存储区域中的数据量超过一定阈值时, 将本地化存储区域中的第一数据写入远端共享存储区域。

[0100] 可选地, 可以将本地持久化存储区域中的第一数据与本地持久化存储区域中的第二数据聚合得到聚合后的数据, 然后将聚合后的数据写入远端共享存储区域。第二数据所属的文件的文件类型也为用户文件或系统文件。

[0101] 505, 如果第一数据所属文件的文件类型为用户文件或系统文件, 确定第一数据的使用频率是否大于或等于第一阈值。

[0102] 如果第一数据的使用频率大于或等于第一阈值, 转到步骤 506。

[0103] 如果第一数据的使用频率小于第一阈值, 流程结束。

[0104] 506, 如果第一数据的使用频率大于或等于第一阈值, 则将本地持久化存储区域中的第一数据写入本地缓存区域。

[0105] 在将本地持久化存储区域中的第一数据写入本地缓存区域之后, 流程结束。

[0106] 本发明实施例中, 通过确定本地持久化存储区域存储的第一数据所属的文件的文件类型, 在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中, 而非将所有类型的数据均写入远端共享存储区域中, 能够降低对远端共享存储区域的访问量, 从而能够减少 I/O 写请求的延迟, 提升用户体验。

[0107] 图 6 是根据本发明另一实施例的处理 I/O 读请求的方法的过程的示意性流程图, 其中, 图 6 的方法由服务器执行, 例如可以由图 1 所示的服务器 110a 至 110d 之一执行, 或者更确切地说, 由服务器的宿主机 host, 或虚拟磁盘服务组件来执行。

[0108] 601, 根据从虚拟机接收到的 I/O 读请求, 在本地缓存区域中查找 I/O 读请求所请求读取的第三数据。

[0109] 602, 判断是否在本地缓存区域中查找到第三数据。

[0110] 603, 如果在本地缓存区域中查找到第三数据, 向虚拟机发送第三数据。

[0111] 604, 如果在本地缓存区域中没有查找到第三数据, 则在本地持久化存储区域中查找第三数据。

[0112] 605, 判断是否在本地持久化存储区域中查找到第三数据。

[0113] 606, 如果在本地持久化存储区域中查找到第三数据, 则向虚拟机发送第三数据。

[0114] 607, 确定第三数据所属的文件的文件类型。

[0115] 例如, 第三数据所属的文件的文件类型可以为临时文件、用户文件或系统文件。

[0116] 608, 判断第三数据所属的文件的文件类型是否为用户文件或系统文件, 且第三数

据的使用频率是否大于或等于第二阈值。

[0117] 如果第三数据所属的文件的文件类型不为用户文件或系统文件,或者第三数据的使用频率小于第二阈值,则流程结束。

[0118] 609,如果第三数据所属的文件的文件类型为用户文件或系统文件,且第三数据的使用频率大于或等于第二阈值,则将第三数据写入本地缓存区域。

[0119] 610,如果在本地持久化存储区域中没有查找到第三数据,则在远端共享存储区域中查找第三数据。

[0120] 611,判断是否在远端共享存储区域中查找到第三数据。

[0121] 612,如果在远端共享存储区域中查找到第三数据,则向虚拟机发送第三数据。

[0122] 步骤 613 至 615 与步骤 607 至 609 类似,为了避免重复,不再赘述。

[0123] 本发明实施例中,由于依次在本地缓存区域、本地持久化存储区域和远端共享存储区域查找 I/O 读请求所请求读取的第三数据,如果在本地找到第三数据,那么就能够降低对远端共享存储的访问量,从而能够减少 I/O 读请求的延迟,提升用户体验。

[0124] 图 7 是根据本发明另一实施例的处理 I/O 请求的方法的示意性流程图。图 7 的方法由虚拟机执行,例如可以由图 1 所示的服务器 110a 至 110b 之一上运行的虚拟机执行。

[0125] 710,确定输入 / 输出 I/O 写请求对应的第一文件信息,第一文件信息用于指示 I/O 写请求所请求写入的第一数据所属的文件。

[0126] 例如,如图 4b 所示,在虚拟机的文件系统上面添加了驱动层,并在虚拟机的块设备层下面添加了驱动层。这样,虚拟机可以通过两个驱动层,完成确定第一文件信息的过程。

[0127] 720,向宿主机 host 发送 I/O 写请求, I/O 写请求携带第一文件信息,以便宿主机 host 在将所述 I/O 写请求所请求写入的第一数据写入本地持久化存储区域之后,根据第一文件信息确定第一数据所属的文件的文件类型,并在第一数据所属的文件的文件类型为用户文件或系统文件的情况下,将第一数据写入远端共享存储区域。

[0128] 虚拟机可以向宿主机 host 发送 I/O 写请求,以便宿主机 host 对该 I/O 写请求进行处理。例如,宿主机 host 可以将 I/O 写请求所请求写入的第一数据写入本地持久化存储区域中。并且宿主机 host 可以根据该 I/O 写请求携带的第一文件信息确定第一数据所属的文件的文件类型。这样在第一数据所属的文件为用户文件或系统文件的情况下,将本地持久化存储区域中存储的第一数据写入远端共享存储区域。

[0129] 本发明实施例中,通过确定用于指示 I/O 写请求所请求写入的第一数据的文件的第一文件信息,并向宿主机 host 发送携带第一文件信息的 I/O 写请求,使得宿主机 host 在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0130] 可选地,作为一个实施例,虚拟机可以确定 I/O 读请求对应的第二文件信息,第二文件信息用于指示 I/O 读请求所请求读取的第三数据所属的文件,并可以向宿主机 host 发送 I/O 读请求, I/O 读请求携带第二文件信息。

[0131] 类似与确定第一文件信息的过程,虚拟机可以通过文件系统上面所添加的驱动层以及块设备层下面所添加的驱动层,完成确定第二文件信息的过程。宿主机 host 可以根据

第二文件信息确定第三数据所属的文件的文件类型。在第三数据是在本地持久化存储区域或远端共享存储区域中查找到的情况下,如果第三数据属于用户文件或系统文件,而且是常用的数据,那么可以将第三数据存储在本地的缓存区域中,从而能够提升数据访问速度。

[0132] 图 8 是根据本发明一个实施例的宿主机的示意框图。例如,图 8 的宿主机 800 可以为图 1 所示的服务器 110a、110b、110c 或 110d 上的特权虚拟机或虚拟磁盘服务组件。宿主机 800 包括接收单元 810、确定单元 820 和存储单元 830,其中:

[0133] 接收单元 810 从虚拟机接收 I/O 写请求。存储单元 830 将该 I/O 写请求所请求写入的第一数据写入宿主机 800 的本地持久化存储区域。确定单元 820 确定本地持久化存储区域中存储的第一数据所属的文件的文件类型,其中,文件类型包括:临时文件、用户文件和系统文件。存储单元 830 在第一数据所属的文件的文件类型为系统文件的情况下,将第一数据写入远端共享存储区域。

[0134] 可见,本发明实施例中,通过确定本地持久化存储区域中存储的第一数据所属的文件的文件类型,在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0135] 可选地,作为一个实施例,确定单元 820 可以用于从内存虚拟镜像中读取镜像元数据,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型。然后确定单元 820 可以根据镜像元数据、I/O 写请求在虚拟磁盘上所对应的起始扇区以及第一数据的大小,确定第一数据的所属的文件的文件类型。

[0136] 可选地,作为另一实施例,确定单元 820 可以用于获取 I/O 写请求中携带的第一文件信息,其中第一文件信息是虚拟机确定的,第一文件信息可以用于指示第一数据所属的文件。确定单元 820 可以根据第一文件信息,确定第一数据所属的文件的文件类型。

[0137] 可选地,作为另一实施例,存储单元 830 可以用于将本地持久化存储区域中存储的第一数据与本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为系统文件,并将聚合后的数据写入远端共享存储区域。

[0138] 可选地,作为另一实施例,存储单元 830 还可以用于将第一数据写入远端共享存储区域之后,若第一数据所属的文件的文件类型为系统文件,且第一数据的使用频率大于或等于第一阈值,则将第一数据写入本地缓存区域。

[0139] 可选地,作为另一实施例,宿主机 800 还可以包括接收单元 830、查找单元 840 和发送单元 850,其中:

[0140] 接收单元 830 还可以用于从虚拟机接收 I/O 读请求。查找单元 840 可以依次在本地的缓存区域、本地持久化存储区域和远端共享存储区域查找 I/O 读请求所请求读取的第三数据,直到查找到第三数据为止。发送单元 850 可以向虚拟机发送查找到的第三数据。

[0141] 可选地,作为另一实施例,确定单元 820 还可以用于确定第三数据所属的文件的文件类型。

[0142] 存储单元 830 还可以用于在第三数据所属的文件的文件类型为系统文件,且第三数据的使用频率大于或等于第二阈值时,将第三数据写入本地缓存区域中。

[0143] 可选地,作为另一实施例,确定单元 820 可以用于从内存虚拟镜像中读取镜像元数据,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型。然后确定单元 820 可以根据镜像元数据、在虚拟磁盘上 I/O 读请求所对应的起始扇区以及第三数据的大小,确定第三数据所属的文件的文件类型。

[0144] 可选地,作为另一实施例,确定单元 820 可以用于获取 I/O 读请求中携带的第二文件信息,其中第二文件信息是虚拟机确定的,第二文件信息可以用于指示第三数据所属的文件。确定单元 820 可以根据第二文件信息,确定第三数据所属的文件的文件类型。

[0145] 可选地,作为另一实施例,宿主机 800 还可以包括生成单元 860。生成单元 860 可以用于在确定单元 820 从内存虚拟镜像中读取镜像元数据之前,生成内存虚拟镜像。

[0146] 宿主机 800 的其它功能和操作可以参照上面图 1 至图 6 的方法实施例的过程,为了避免重复,此处不再赘述。

[0147] 图 9 是根据本发明一个实施例的虚拟机的示意框图。图 9 的虚拟机 900 为部署在物理主机上的虚拟机,即部署在上述服务器上的虚拟机。例如,虚拟机 900 可以为图 1 所示的服务器 110a 至 110d 上部署的虚拟机。虚拟机 900 包括确定单元 910 和发送单元 920。

[0148] 确定单元 910 用于确定 I/O 写请求对应的第一文件信息,第一文件信息用于指示 I/O 写请求所请求写入的第一数据所属的文件。

[0149] 发送单元 920 用于向宿主机 host 发送 I/O 写请求,I/O 写请求携带第一文件信息,以便宿主机 host 根据第一文件信息确定第一数据所属的文件的文件类型,并在第一数据所属的文件的文件类型为系统文件的情况下,将第一数据写入远端共享存储区域。

[0150] 可见,本发明实施例中,通过确定用于指示 I/O 写请求所请求写入的第一数据的文件的第一文件信息,并向宿主机 host 发送携带第一文件信息的 I/O 写请求,使得宿主机 host 在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0151] 可选地,作为一个实施例,确定单元 910 还可以确定 I/O 读请求对应的第二文件信息,第二文件信息用于指示 I/O 读请求所请求读取的第三数据所属的文件。发送单元 920 还可以向宿主机 host 发送 I/O 读请求,I/O 读请求携带第二文件信息,以使宿主机 host 依次在虚拟机 900 所在服务器的本地缓存区域、本地持久化存储区域和远端共享存储区域查找该 I/O 读请求所请求读取的第三数据,直到查找到第三数据为止;

[0152] 相应地,虚拟机 900 还包括:接收单元 930,用于接收宿主机 host 发送的查找到的第三数据。

[0153] 虚拟机 900 的其它功能和操作可以参照上面图 7 的方法实施例的过程,为了避免重复,此处不再赘述。

[0154] 图 10 是根据本发明另一实施例的服务器的示意框图。图 10 的服务器 1000 也可以称为物理主机。图 10 的服务器 1000 包括存储器 1010 和处理器 1020。

[0155] 存储器 1010 可以包括随机存储器、闪存、只读存储器、可编程只读存储器、非易失性存储器或寄存器等。处理器 1020 可以是中央处理器(Central Processing Unit, CPU)。

[0156] 存储器 1010 用于存储可执行指令。处理器 1020 可以执行存储器 410 中存储的可执行指令,用于:接收输入/输出 I/O 写请求,将该 I/O 写请求所请求写入的第一数据写入本地持久化存储区域,确定第一数据所属的文件的文件类型,其中,文件类型包括:临时文件、用户文件和系统文件;在第一数据所属的文件的文件类型为系统文件的情况下,将第一数据写入远端共享存储区域。

[0157] 可见,本发明实施例中,通过确定本地持久化存储区域中存储的第一数据所属的文件的文件类型,在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0158] 可选地,作为另一实施例,处理器 1020 可以用于从内存虚拟镜像中读取镜像元数据,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型,并可以根据镜像元数据、I/O 写请求在虚拟磁盘上所对应的起始扇区以及第一数据的大小,确定第一数据所属的文件的文件类型。

[0159] 可选地,作为另一实施例,处理器 1020 可以用于获取 I/O 写请求中携带的第一文件信息,其中第一文件信息是虚拟机确定的,第一文件信息可以用于指示第一数据所属的文件。处理器 1020 可以根据第一文件信息,确定第一数据所属的文件的文件类型。

[0160] 可选地,作为另一实施例,处理器 1020 可以用于将本地持久化存储区域中存储的第一数据与本地持久化存储区域中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为系统文件,并将聚合后的数据写入远端共享存储区域。

[0161] 可选地,作为另一实施例,处理器 1020 还可以用于在将第一数据写入远端共享存储区域中之后,若第一数据所属的文件为用户文件或系统文件,且第一数据的使用频率大于或等于预设的第一阈值,则将第一数据写入本地缓存区域。

[0162] 可选地,作为另一实施例,服务器 1000 还可以包括接收器 1030 和发送器 1040。

[0163] 接收器 1030 还可以用于接收 I/O 读请求。处理器 1020 可以依次在本地缓存区域、本地持久化存储区域和远端共享存储区域查找 I/O 读请求所请求读取的第三数据,直到查找到第三数据为止。发送器 1040 可以向虚拟机发送查找到的第三数据。

[0164] 可选地,作为另一实施例,处理器 1020 还可以用于确定第三数据所属的文件的文件类型。

[0165] 处理器 1020 还可以用于在第三数据所属的文件的文件类型为系统文件,且第三数据的使用频率大于或等于第二阈值时,将第三数据写入本地缓存区域中。

[0166] 可选地,作为另一实施例,处理器 1020 可以用于从内存虚拟镜像中读取镜像元数据,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟磁盘的数据属性信息和文件系统的类型,并可以根据镜像元数据、在虚拟磁盘上 I/O 读请求所对应的起始扇区以及第三数据的大小,确定第三数据所属的文件的文件类型。

[0167] 可选地,作为另一实施例,处理器 1020 可以用于获取 I/O 读请求中携带的第二文件信息,其中第二文件信息是虚拟机确定的,第二文件信息可以用于指示第三数据所属的文件。处理器 1020 可以用于根据第二文件信息,确定第三数据所属的文件的文件类型。

[0168] 可选地,作为一个实施例,处理器 1020 可以在从内存虚拟镜像中读取镜像元数据之前,生成内存虚拟镜像。

[0169] 服务器 1000 的其它功能和操作可以参照上面图 1 至图 6 的方法实施例的过程,为了避免重复,此处不再赘述。

[0170] 图 11 是根据本发明另一实施例的服务器的示意框图。

[0171] 如图 11 所示,本发明实施例还提供一种服务器,根据图 11,服务器 1200 包括:

[0172] 硬件层、运行在所述硬件层之上的 VMM (Virtual Machine Monitor,虚拟机监视器),以及运行在 VMM 之上的宿主机 Host1201 和虚拟机 (VM, Virtual Machine) 1202,其中,硬件层包括 I/O 设备、本地持久化存储区域;

[0173] 其中,宿主机 Host1201,用于从虚拟机 1202 接收 I/O 写请求,该 I/O 写请求是客户端通过 I/O 设备发送给虚拟机 1202 的;将该 I/O 写请求所请求写入的第一数据写入本地持久化存储区域 1203,确定第一数据所属的文件的文件类型,其中,文件类型包括:临时文件、用户文件和系统文件;如果第一数据所属的文件的文件类型为用户文件或系统文件,则将第一数据写入所述服务器连接的远端共享存储区域 1204。

[0174] 在一个实施例中,宿主机 Host1201 从内存虚拟镜像中读取镜像元数据,;根据该镜像元数据、该 I/O 写请求在该虚拟磁盘上所对应的起始扇区以及第一数据的大小,确定第一数据所属的文件的文件类型;其中,内存虚拟镜像用于存储镜像元数据,镜像元数据用于描述虚拟机的虚拟磁盘的数据属性信息和文件系统的类型。

[0175] 在另一个实施例中,宿主机 Host1201 获取 I/O 写请求中携带的第一文件信息,其中第一文件信息是虚拟机确定的,且第一文件信息用于指示第一数据所属的文件;根据第一文件信息,确定第一数据所属的文件的文件类型。优选地,宿主机 Host 可以采用如下方式将第一数据写入远端共享存储区域 1204:将第一数据与本地持久化存储区域 1203 中的第二数据进行聚合以得到聚合后的数据,其中第二数据所属的文件的文件类型也为用户文件或系统文件;将聚合后的数据写入远端共享存储区域 1204。

[0176] 优选地,服务器 1200 的硬件层还包括有本地缓存区域(cache),宿主机 Host1201 在所述将所述第一数据写入远端共享存储区域之后,若所述第一数据所属的文件的文件类型为用户文件或系统文件,且所述第一数据的使用频率大于或等于第一阈值的情况下,还将所述第一数据写入本地缓存区域。

[0177] 上述第一阈值可以是宿主机根据第一数据的使用情况以及其它数据的使用情况确定的,因此,第一阈值可以是变化的。例如,宿主机经过一段时间的统计发现第一数据的使用频率为 40% 以上,而其它数据的使用频率均小于 40%,那么可以将第一阈值设置为 40%。

[0178] 需要说明的是,服务器 1200 的其它功能和操作可以参照上面图 1 至图 6 的方法实施例的过程,为了避免重复,此处不再赘述。

[0179] 可见,本发明实施例中,通过确定本地持久化存储区域中存储的第一数据所属的文件的文件类型,在第一数据所属的文件为用户文件或系统文件的情况下将本地持久化存储区域中存储的第一数据写入远端共享存储区域中,而非将所有类型的数据均写入远端共享存储区域中,能够降低对远端共享存储区域的访问量,从而能够减少 I/O 写请求的延迟,提升用户体验。

[0180] 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、或者计算机软件和电子硬件的结合来实现。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。

[0181] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0182] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置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。另一点,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,可以是电性,机械或其它的形式。

[0183] 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实施例方案的目的。

[0184] 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,也可以两个或两个以上单元集成在一个单元中。

[0185] 所述功能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)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。

[0186] 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,可轻易想到变化或替换,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所述权利要求的保护范围为准。

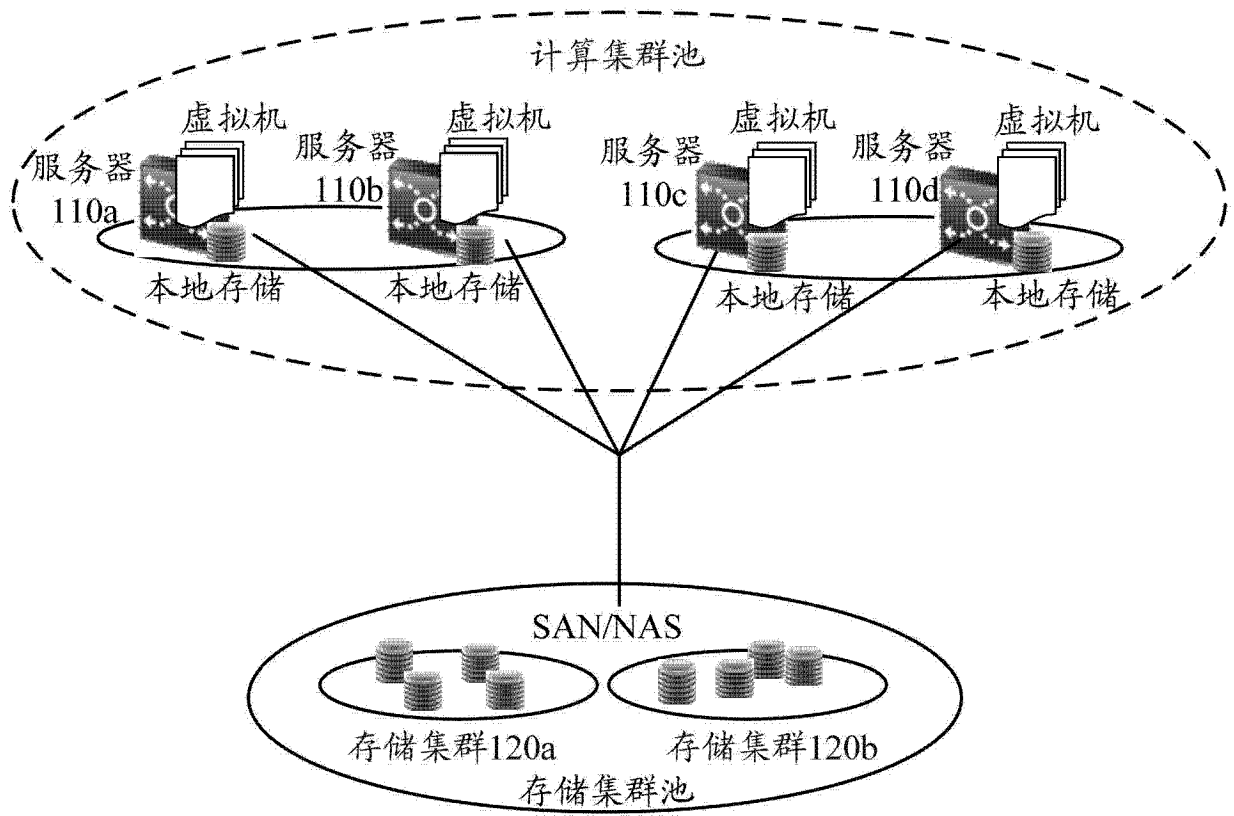


图 1

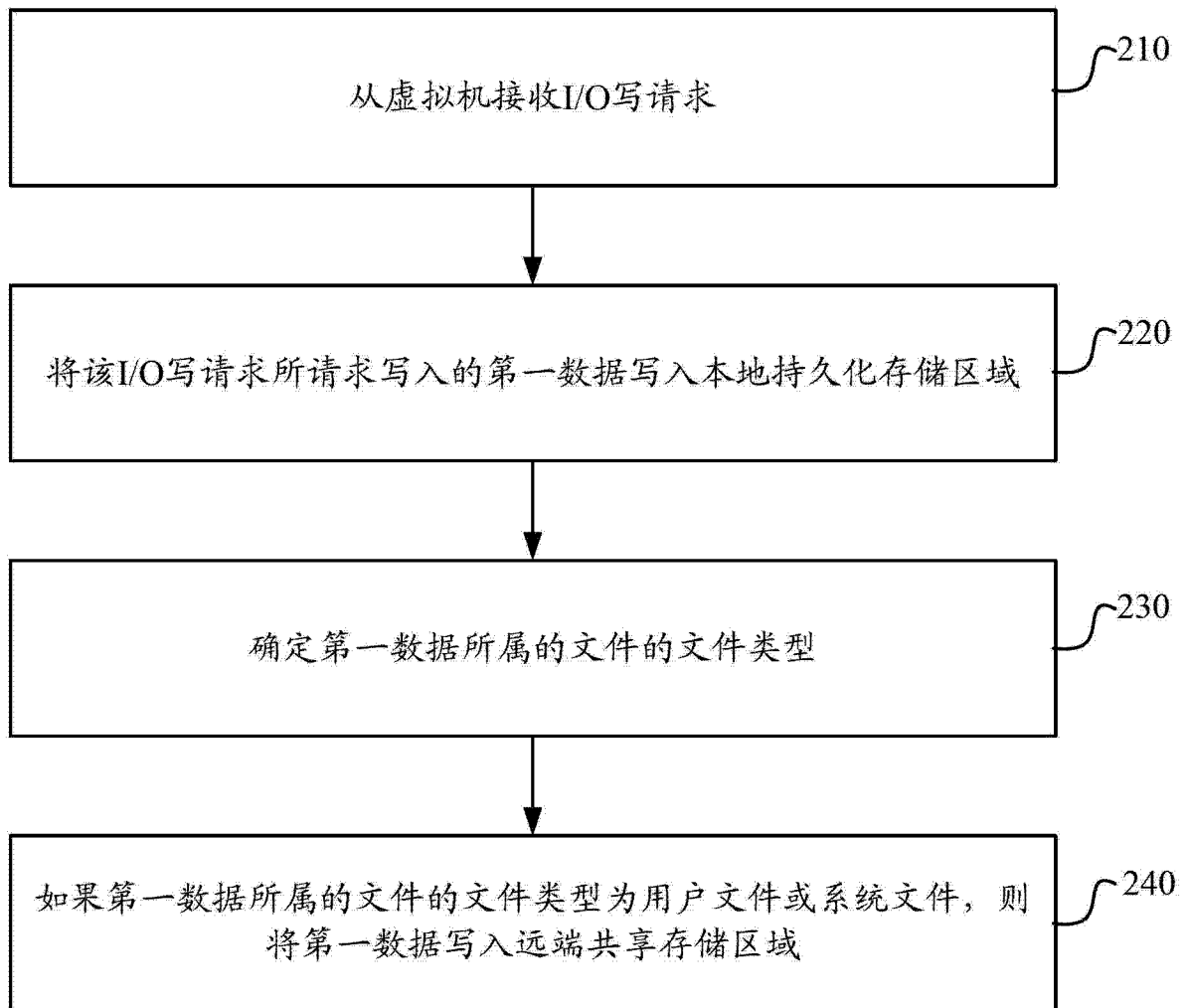


图 2

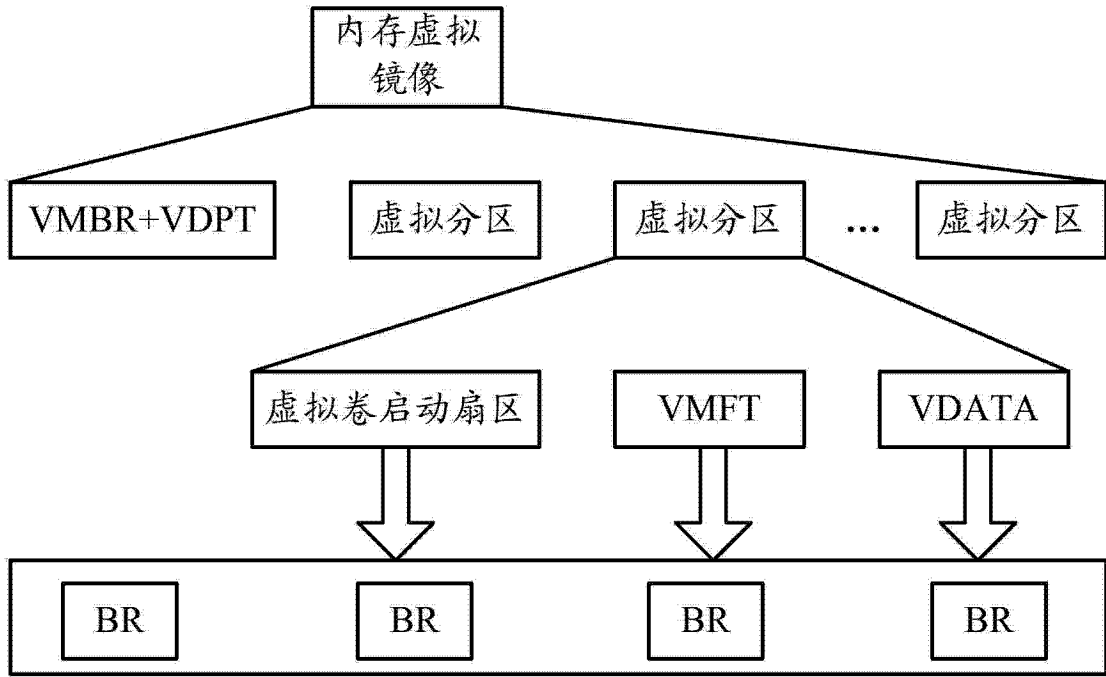


图 3

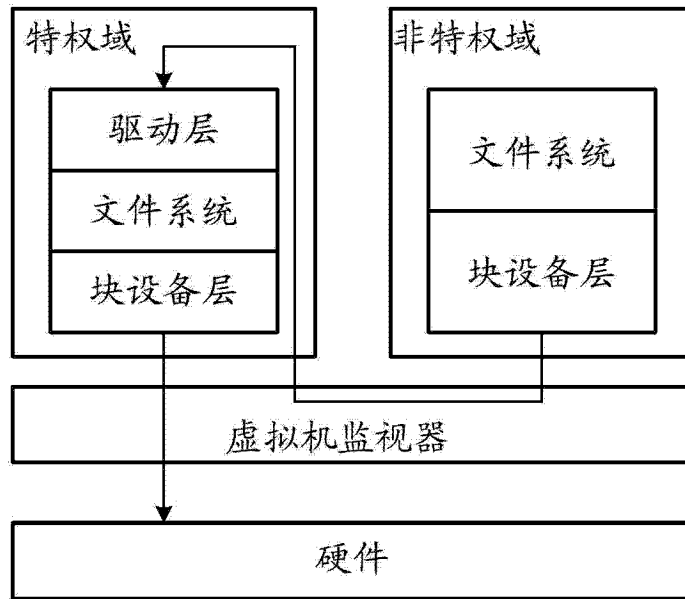


图 4a

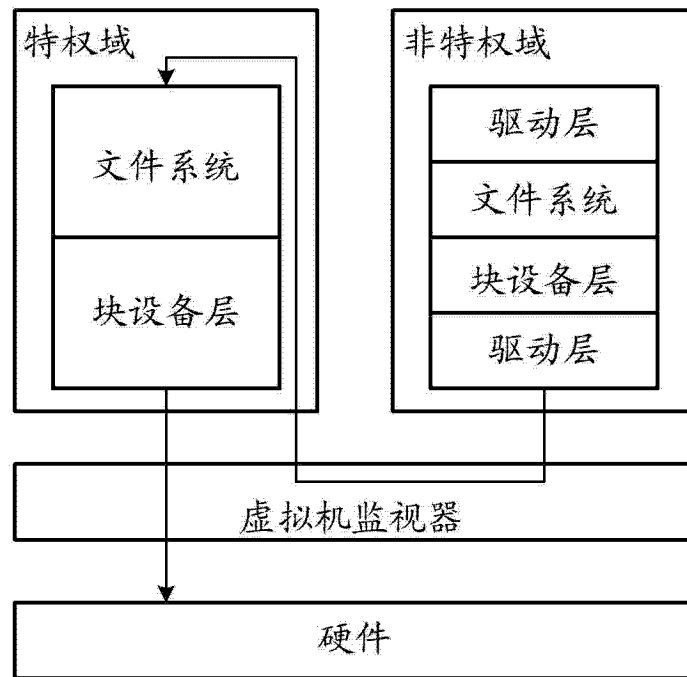


图 4b

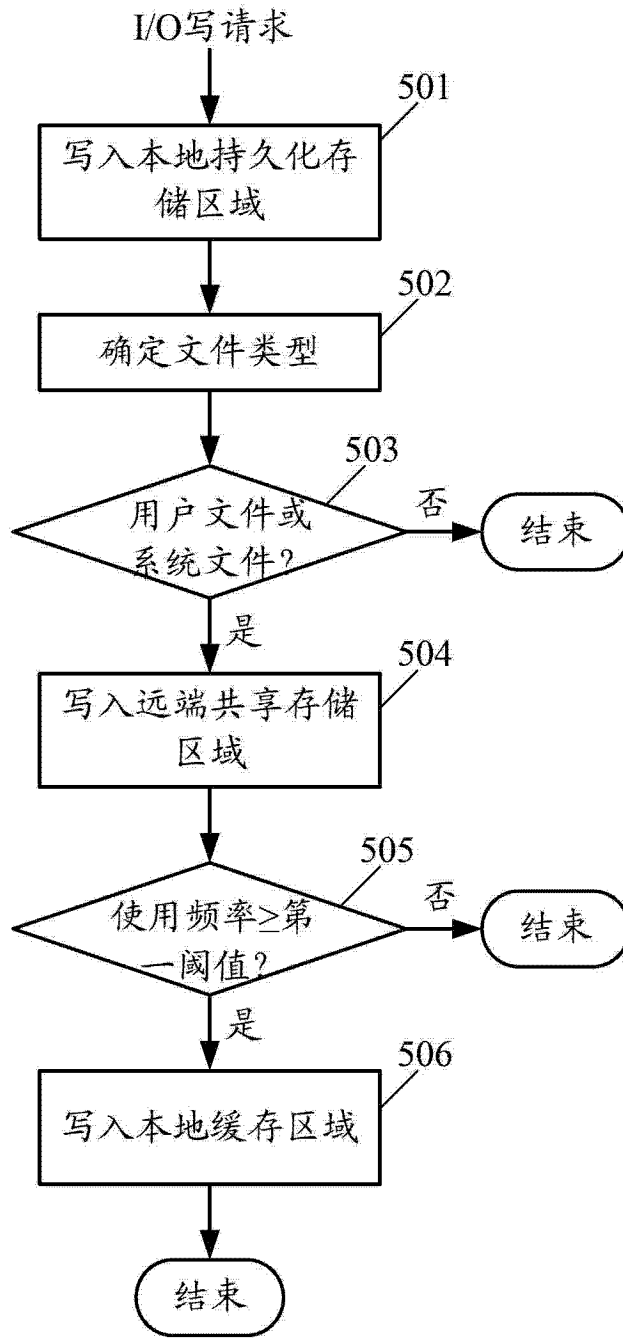


图 5

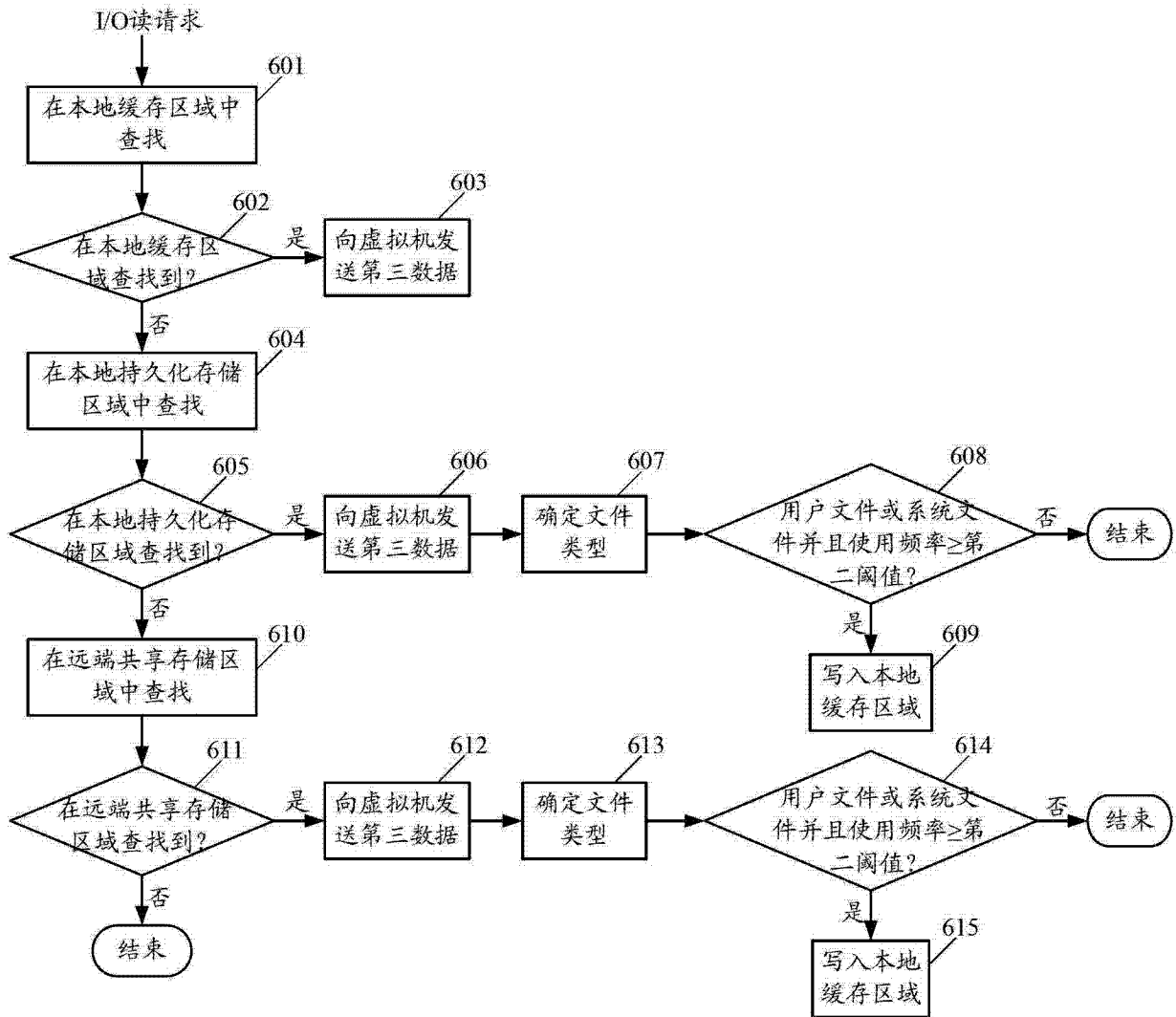


图 6

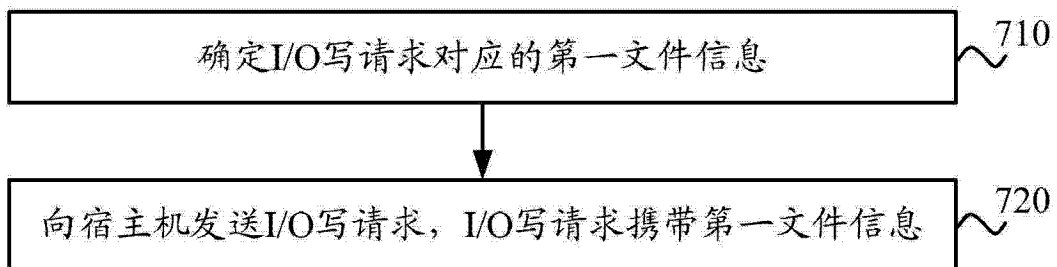


图 7

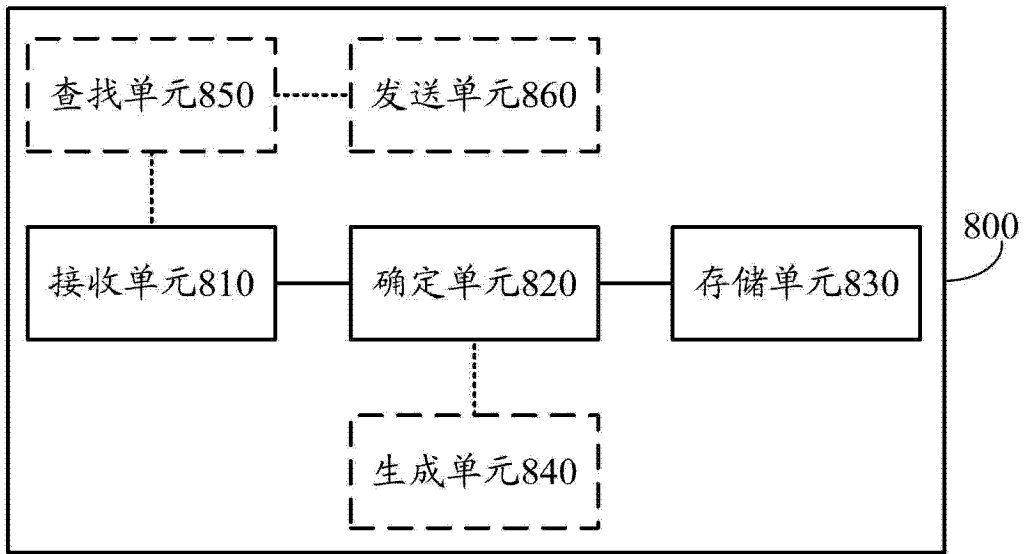


图 8

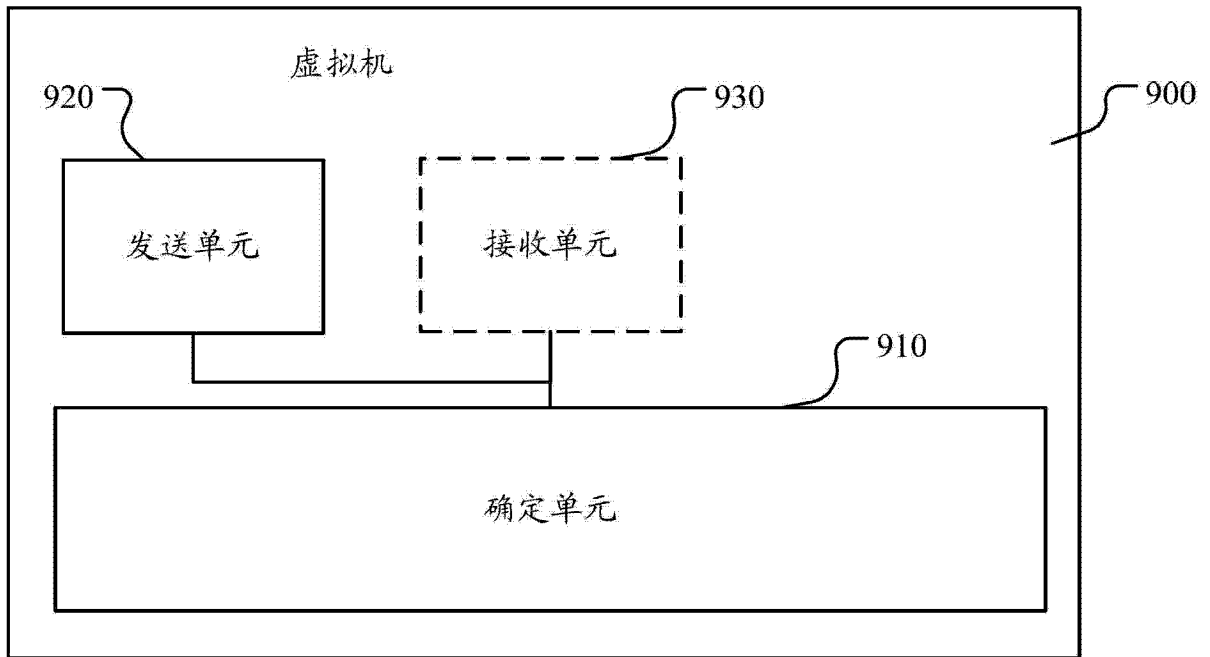


图 9

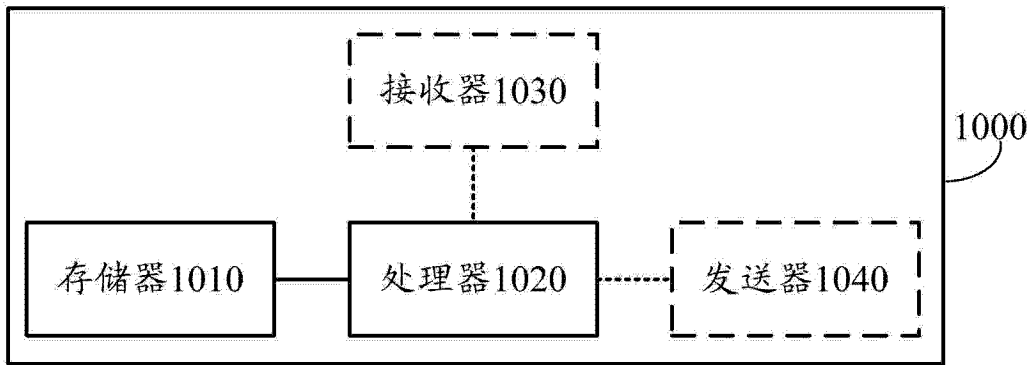


图 10

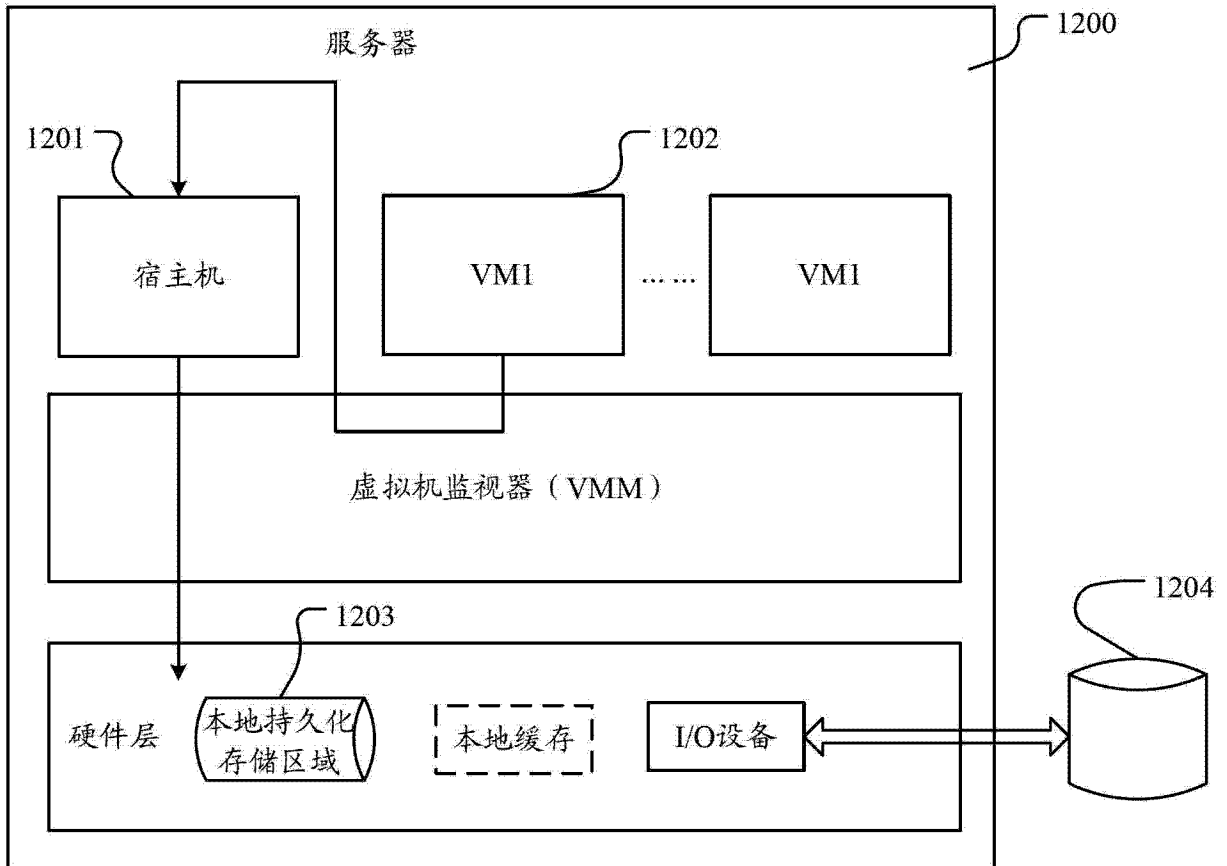


图 11