



(12)发明专利申请

(10)申请公布号 CN 113918535 A

(43)申请公布日 2022.01.11

(21)申请号 202010650956.1

(22)申请日 2020.07.08

(71)申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 王少华

(74)专利代理机构 深圳市深佳知识产权代理事
务所(普通合伙) 44285

代理人 王兆林

(51) Int. Cl.

G06F 16/21(2019.01)

G06F 16/22(2019.01)

G06F 16/2453(2019.01)

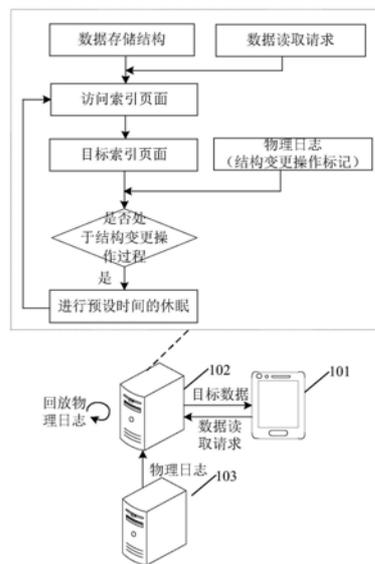
权利要求书2页 说明书12页 附图7页

(54)发明名称

一种数据读取方法、装置、设备及存储介质

(57)摘要

本申请实施例公开一种数据读取方法、装置、设备及存储介质,在基于物理复制的主从架构中,从机设备获取用于读取目标数据的数据读取请求,根据数据读取请求,按照数据存储结构依次访问索引页面。当访问到目标索引页面,从机设备根据物理日志中的结构变更操作标记确定目标索引页面是否处于结构变更操作过程中,若确定目标索引页面处于结构变更操作过程中,则进行预设时间的休眠。由于从机设备通过日志回放线程回放的速度非常快,在进行预设时间的休眠后,目标索引页面的结构变更操作回放已经完成,此时,重新执行根据数据读取请求,按照从机设备中的数据存储结构依次访问索引页面的步骤,以读取目标数据。从而实现数据的一致性读取,避免系统崩溃。



1. 一种数据读取方法,其特征在于,由基于物理复制的主从架构中的从机设备执行所述方法,所述方法包括:

在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;

根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面;

当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;

若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;

结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。

2. 根据权利要求1所述的方法,其特征在于,所述根据物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中,包括:

根据所述结构变更操作标记,从所述物理日志中确定结构变更操作日志,所述结构变更操作日志包括所述目标索引页面对应的目标日志和所述结构变更操作标记;

根据所述目标日志对应的第一日志序列号、所述结构变更操作标记对应的第二日志序列号和第三日志序列号之间的大小关系,确定所述目标索引页面是否处于结构变更操作过程中,所述第三日志序列号表征所述从机设备回放所述物理日志的进展位点。

3. 根据权利要求2所述的方法,其特征在于,根据所述目标日志对应的第一日志序列号、所述结构变更操作标记对应的第二日志序列号和第三日志序列号之间的大小关系,确定所述目标索引页面是否处于结构变更操作过程中,包括:

若所述第一日志序列号大于所述第二日志序列号,且所述第一日志序列号大于所述第三日志序列号,确定所述目标索引页面处于所述结构变更操作过程中;

否则,确定所述目标索引页面未处于所述结构变更操作过程中。

4. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

若根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面未处于结构变更操作过程中,读取所述目标数据,或,继续按照所述数据存储结构访问所述目标索引页面的下一个索引页面,直到读取到所述目标数据。

5. 根据权利要求1所述的方法,其特征在于,所述数据存储结构为B树,所述B树包括父节点、中间节点和叶子节点,所述父节点、所述中间节点和所述叶子节点之间具有连接关系,所述父节点、所述中间节点和所述叶子节点分别对应不同的索引页面。

6. 根据权利要求5所述的方法,其特征在于,若确定所述目标索引页面处于结构变更操作过程中,所述进行预设时间的休眠之前,所述方法还包括:

释放访问至所述目标索引页面的目标路径,所述目标路径为从所述父节点依次访问到所述目标索引页面所对应的所述中间节点之间的路径。

7. 根据权利要求5所述的方法,其特征在于,若确定所述目标索引页面处于结构变更操作过程中,所述进行预设时间的休眠之前,所述方法还包括:

若所述目标索引页面为所述叶子节点对应的索引页面,保存所述目标数据在所述目标索引页面中的位置信息。

8. 根据权利要求1-7任一项所述的方法,其特征在于,所述方法还包括:

向用户对应的终端设备返回所述目标数据。

9. 一种数据读取装置,其特征在於,所述装置部署在基于物理复制的主从架构中的从机设备中,所述装置包括获取单元、访问单元、确定单元、休眠单元和读取单元:

所述获取单元,用于在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;

所述访问单元,用于根据所述数据读取请求,按照所述从机设备中的数据存储空间依次访问索引页面;

所述确定单元,用于当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;

所述休眠单元,用于若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;

所述读取单元,用于结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储空间依次访问索引页面的步骤,以读取所述目标数据。

10. 根据权利要求9所述的装置,其特征在於,所述确定单元,用于:

根据所述结构变更操作标记,从所述物理日志中确定结构变更操作日志,所述结构变更操作日志包括所述目标索引页面对应的目标日志和所述结构变更操作标记;

根据所述目标日志对应的第一日志序列号、所述结构变更操作标记对应的第二日志序列号和第三日志序列号之间的大小关系,确定所述目标索引页面是否处于结构变更操作过程中,所述第三日志序列号表征所述从机设备回放所述物理日志的进展位点。

11. 根据权利要求10所述的装置,其特征在於,所述确定单元,用于:

若所述第一日志序列号大于所述第二日志序列号,且所述第一日志序列号大于所述第三日志序列号,确定所述目标索引页面处于所述结构变更操作过程中;

否则,确定所述目标索引页面未处于所述结构变更操作过程中。

12. 根据权利要求9所述的装置,其特征在於,所述读取单元还用于:

若根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面未处于结构变更操作过程中,读取所述目标数据,或,继续按照所述数据存储空间访问所述目标索引页面的下一个索引页面,直到读取到所述目标数据。

13. 根据权利要求9所述的装置,其特征在於,所述数据存储空间为B树,所述B树包括父节点、中间节点和叶子节点,所述父节点、所述中间节点和所述叶子节点之间具有连接关系,所述父节点、所述中间节点和所述叶子节点分别对应不同的索引页面。

14. 一种用于数据读取的电子设备,其特征在於,所述电子设备包括处理器以及存储器:

所述存储器用于存储程序代码,并将所述程序代码传输给所述处理器;

所述处理器用于根据所述程序代码中的指令执行权利要求1-8任一项所述的方法。

15. 一种计算机可读存储介质,其特征在於,所述计算机可读存储介质用于存储程序代码,所述程序代码用于执行权利要求1-8任一项所述的方法。

一种数据读取方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及计算机领域,特别是涉及一种数据读取方法、装置、设备及存储介质。

背景技术

[0002] 随着计算机技术的快速发展,数据库中存储的数据量越来越大。为此,数据库的主从架构应运而生,其中,主数据库一般用于写操作,从数据库一般用于读操作,实现读写分离,降低主数据库读写压力,提高性能。

[0003] 主从数据库基于物理复制实现主从一致性,然而,在基于物理复制的主从架构下,从数据库日志回放线程对结构变更操作(structure modification operation,SMO)进行回放时,会改变数据存储结构例如B树(B-tree)的结构,若在该过程读取数据,可能会导致并发的用户读线程结果不一致,即读取的数据不一致,或者系统崩溃。

[0004] 为此,在基于物理复制的主从架构下,从数据库中一致性的读取数据是亟需解决的技术问题。

发明内容

[0005] 为了解决上述技术问题,本申请提供了一种数据读取方法、装置、设备及存储介质,避免由于数据存储结构的改变导致系统崩溃,实现数据的一致性读取。

[0006] 本申请实施例公开了如下技术方案:

[0007] 一方面,本申请实施例提供一种数据读取方法,由基于物理复制的主从架构中的从机设备执行所述方法,所述方法包括:

[0008] 在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;

[0009] 根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面;

[0010] 当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;

[0011] 若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;

[0012] 结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。

[0013] 一方面,本申请实施例提供一种数据读取装置,所述装置部署在基于物理复制的主从架构中的从机设备中,所述装置包括获取单元、访问单元、确定单元、休眠单元和读取单元:

[0014] 所述获取单元,用于在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;

[0015] 所述访问单元,用于根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面;

[0016] 所述确定单元,用于当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;

[0017] 所述休眠单元,用于若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;

[0018] 所述读取单元,用于结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。

[0019] 在一种可能的实现方式中,所述装置还包括释放单元:

[0020] 所述释放单元,用于若确定所述目标索引页面处于结构变更操作过程中,释放访问至所述目标索引页面的目标路径,所述目标路径为从所述父节点依次访问到所述目标索引页面所对应的所述中间节点之间的路径。

[0021] 在一种可能的实现方式中,所述装置还包括保存单元:

[0022] 所述保存单元,用于若确定所述目标索引页面处于结构变更操作过程中,若所述目标索引页面为所述叶子节点对应的索引页面,保存所述目标数据在所述目标索引页面中的位置信息。

[0023] 在一种可能的实现方式中,所述装置还包括返回单元:

[0024] 所述返回单元,用于向用户对应的终端设备返回所述目标数据。

[0025] 一方面,本申请实施例提供一种用于数据读取的电子设备,所述电子设备包括处理器以及存储器:

[0026] 所述存储器用于存储程序代码,并将所述程序代码传输给所述处理器;

[0027] 所述处理器用于根据所述程序代码中的指令执行任一所述的数据读取方法。

[0028] 一方面,本申请实施例提供一种计算机可读存储介质,所述计算机可读存储介质用于存储程序代码,所述程序代码用于执行任一所述的数据读取方法。

[0029] 由上述技术方案可以看出,在基于物理复制的主从架构中,从机设备可以通过回放物理日志以进行物理复制,实现主从设备的数据一致性,在从机设备回放物理日志以进行物理复制的过程中,若用户需要通过从机设备中读取目标数据,从机设备可以获取用于读取目标数据的数据读取请求,根据数据读取请求,按照从机设备中的数据存储结构依次访问索引页面。针对访问的每个索引页面例如目标索引页面,从机设备根据物理日志中添加的结构变更操作标记确定目标索引页面是否处于结构变更操作过程中,若确定目标索引页面处于结构变更操作过程中,则进行预设时间的休眠。由于从机设备通过日志回放线程回放的速度非常快,在进行预设时间的休眠后,目标索引页面的结构变更操作回放已经完成,此时,重新执行根据所述数据读取请求,按照从机设备中的数据存储结构依次访问索引页面的步骤,以读取目标数据。通过对访问到的每个索引页面例如目标索引页面实时进行是否处于结构变更操作过程中的判断,可以保证在目标索引页面的结构变更操作回放完成之后再访问,此时数据存储结构例如B树的结构不再变化,处于一致性状态,避免由于数据存储结构的改变导致系统崩溃,实现数据的一致性读取。

附图说明

[0030] 为了更清楚地说明本申请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简单地介绍,显而易见地,下面描述中的附图仅仅是本

申请的一些实施例,对于本领域普通技术成员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

- [0031] 图1为本申请实施例提供的一种数据读取方法的系统架构示意图;
- [0032] 图2为本申请实施例提供的一种数据读取方法的流程图;
- [0033] 图3为本申请实施例提供的一种B-tree的结构示意图;
- [0034] 图4为本申请实施例提供的一种完整的SMO日志的示意图;
- [0035] 图5为本申请实施例提供的SMO日志回放到SMO标记的示意图;
- [0036] 图6为本申请实施例提供的B-tree处于SMO回放过程中的结构示意图;
- [0037] 图7为本申请实施例提供的SMO日志回放完成后完整的SMO日志的示意图;
- [0038] 图8为本申请实施例提供的SMO日志回放完成后的B-tree的结构示意图;
- [0039] 图9为本申请实施例提供的一种数据读取方法的流程图;
- [0040] 图10为本申请实施例提供的一种数据读取装置的结构图;
- [0041] 图11为本申请实施例提供的一种终端设备的结构图;
- [0042] 图12为本申请实施例提供的一种服务器的结构图。

具体实施方式

[0043] 下面结合附图,对本申请的实施例进行描述。

[0044] 物理复制:是指基于数据库的物理日志(redo log)进行主从复制的方法。主机设备传送物理日志到从机设备,从机设备应用物理日志,从而达到主从数据一致。

[0045] 在基于物理复制的主从架构下,以从数据库中数据存储结构是B树(B-tree)为例,若从数据库安装在从机设备上,当从机设备通过日志回放线程对结构变更操作(structure modification operation,SMO)进行回放时,会改变B-tree的结构,若在SMO过程中读取数据,可能会由于B-tree结构的改变,导致并发的用户读线程读取的数据不一致,或者系统崩溃。其中,B-tree是常见的数据物理组织方式,也是一种数据的索引方式;SMO是指结构变更操作,包括分裂或合并操作,一般涉及到三个及以上索引页面的修改,例如索引页面A分裂为索引页面A和索引页面B,即为分裂操作,该分裂操作改变了数据存储结构。

[0046] 为了解决上述技术问题,本申请实施例提供一种数据读取方法,该方法对访问到的每个索引页面例如目标索引页面实时进行是否处于SMO过程中的判断,可以保证在目标索引页面的SMO回放完成之后再访问,此时数据存储结构例如B-tree的结构不再变化,处于一致性状态,避免由于数据存储结构的改变导致系统崩溃,实现数据的一致性读取。

[0047] 因为在读取数据时,对数据存储结构可能存在并发的增删改查操作发生,一致性读取是指读取时,基于数据存储结构进行访问时例如进行B-tree扫描时,保证数据存储结构是完整的,不发生结构改变,从而读取的数据是一致的。不存在多读记录,和少读记录的情况。

[0048] 其中,从机设备与主机设备构成基于物理复制的主从架构,从机设备上可以安装有从数据库,主机设备上可以安装有主数据库,从机设备和主机设备通常是服务器。本申请实施例提供的的数据读取方法可以由基于物理复制的主从架构中的从机设备执行。

[0049] 本申请实施例所提供的方法涉及到云技术领域,例如涉及云技术领域中的数据库(Database)。数据库是存储电子文件的处所,用户可以对文件中的数据进行新增、查询、更

新、删除等操作。所谓“数据库”是以一定方式储存在一起、能与多个用户共享、具有尽可能小的冗余度、与应用程序彼此独立的数据集合。

[0050] 本申请实施例所提供的方法例如涉及数据库中的数据管理,数据管理可以通过数据库管理系统(Database Management System,DBMS)实现,DBMS是为管理数据库而设计的电脑软件系统,一般具有存储、截取、安全保障、备份等基础功能。数据库管理系统可以依据它所支持的数据库模型来作分类,例如关系式、可扩展标记语言(Extensible Markup Language,XML);或依据所支持的计算机类型来作分类,例如服务器群集、移动电话;或依据所用查询语言来作分类,例如结构化查询语言(Structured Query Language,SQL)、XQuery;或依据性能冲量重点来作分类,例如最大规模、最高运行速度;亦或其他分类方式。不论使用哪种分类方式,一些DBMS能够跨类别,例如,同时支持多种查询语言。

[0051] 参见图1,图1为本申请实施例提供的数据读取方法的系统架构示意图。该系统架构中包括终端设备101、从机服务器102和主机服务器103,从机服务器102和主机服务器103构成基于物理复制的主从架构。从机服务器102可以是本申请实施例所涉及的从机设备,从机服务器102上部署有从数据库;主机服务器103可以是本申请实施例所涉及的主机设备,主机服务器103上部署有主数据库。从机服务器102和主机服务器103通过物理复制实现主从数据一致。

[0052] 终端设备101通过从机服务器102访问其上的从数据库,读取到用户所需的数据,与从机服务器102连接的终端设备101可以包括一个或多个,当包括多个终端设备101时,多个终端设备101可能并行读取从数据库中的数据。

[0053] 从机服务器102和主机服务器103都是服务器,服务器可以是独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云计算服务的云服务器。终端设备101可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表等,但并不局限于此。终端设备101以及从机服务器102可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。

[0054] 当用户需要读取某一数据例如目标数据时,终端设备101可以向从机服务器102发送数据读取请求,该数据读取请求用于请求读取目标数据,例如可以包括目标数据的标识。

[0055] 若从机服务器102是在通过日志回放线程对物理日志进行回放过程中,获取数据读取请求,从机服务器102可以根据数据读取请求,按照从机设备(从机服务器102)中的数据存储结构依次访问索引页面。由于物理日志中可能包括SMO日志,即从机服务器102可能会在回放SMO日志,改变B-tree结构的过程中,访问某个索引页面,因此,为了避免回放SMO日志导致读取的数据不一致,在访问到每个索引页面例如目标索引页面时,从机服务器102可以根据物理日志中添加的结构变更操作标记确定目标索引页面是否处于结构变更操作过程中。若确定目标索引页面处于结构变更操作过程中,进行预设时间的休眠。

[0056] 由于从机设备(从机服务器102)通过日志回放线程回放的速度非常快,在进行预设时间的休眠后,目标索引页面的结构变更操作回放已经完成,此时,重新执行根据数据读取请求,按照从机设备(从机服务器102)中的数据存储结构依次访问索引页面的步骤,以读取目标数据。

[0057] 由于每个索引页面都是在结构变更操作回放完成之后再访问的,数据存储结构不再变化,处于一致性状态,避免由于数据存储结构的改变导致系统崩溃,实现数据的一

致性读取。

[0058] 接下来,将结合附图对本申请实施例提供的数据读取方法进行详细介绍。

[0059] 参见图2,图2示出了一种数据读取方法的流程图,该方法可以应用于基于物理复制的主从架构中的从机设备,所述方法包括:

[0060] S201、在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求。

[0061] 在基于物理复制的主从架构中,主机设备可以将物理日志发送至从机设备,从机设备通常需要通过回放物理日志来实现主从数据一致。这样,当用户需要使用某数据例如目标数据时,用户可以通过对应的终端设备向从机设备发送数据读取请求,以从从机设备获取目标数据。

[0062] 物理日志记录了每个索引页面上存储的数据的修改情况,表征每个索引页面对应的物理日志在回放过程中的回放顺序可以通过日志序列号(log sequence number,lsn)表示。

[0063] S202、根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面。

[0064] 从机设备可能在回放物理日志的过程中,获取到数据读取请求,此时,从机设备可以按照数据存储结构依次访问索引页面,以定位并读取目标数据。

[0065] 数据存储结构可以包括多种结构,例如B-tree、R树(R-tree)等。本申请实施例对数据存储结构不做限定,后续将主要以数据存储结构为B-tree进行介绍。

[0066] B-tree的结构可以参见图3所示,B-tree包括父节点、中间节点和叶子节点,父节点、中间节点和叶子节点之间具有连接关系,连接关系可以通过图3中带有箭头的连接线表示。父节点、中间节点和叶子节点分别对应不同的索引页面,其中,图3中标有1的节点表示索引页面1、标有2的节点表示索引页面2、标有3的节点表示索引页面3、标有4的节点表示索引页面4、标有5的节点表示索引页面5、标有6的节点表示索引页面6,第一层节点(例如标有1的节点)为父节点,第二层节点(例如标有2、3的节点)表示中间节点,第三层节点(例如标有4、5、6的节点)为叶子节点。

[0067] 基于图3所示的B-tree,S202中按照数据存储结构依次访问索引页面的方式可以是B-tree的游标打开,按查找条件从父节点向下层节点查找,到叶子节点结束。游标的移动在叶子节点之间向前向后进行,直到碰到不符合查找条件的记录为止。游标是基于数据存储结构查询数据的一种机制,游标打开表示开始从父节点扫描B-tree,游标扫描所经过的路径可以被称为游标路径,该游标路径可以用于指示查找目标数据的路径。

[0068] S203、当访问到目标索引页面时,根据物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中。

[0069] 在物理日志回放过程中,物理日志中可能包括SMO日志,从机服务器可能会在回放SMO日志,由于SMO是指对数据存储结构的结构变更操作,回放SMO日志会改变数据存储结构,因此,当访问到每个索引页面例如目标索引页面时,为了避免在SMO过程访问索引页面以读取数据,导致并发的用户读线程读取的数据不一致,或者系统崩溃,在本实施例中,可以判断目标索引页面是否处于SMO过程中。

[0070] 在本实施例中,SMO日志是带有SMO标记的物理日志,若物理日志记录的是结构变

更操作,则主机设备在针对SMO的物理日志的开始增加一种新的逻辑日志类型 (LOG_TYPE_SMO) 来标记SMO,即SMO标记,日志内容为 (table id,index id),表示在哪个索引表(table)的哪个索引(index)上发生了SMO。以数据存储结构是B-tree为例,B-tree也是一种数据的索引方式,因此,一个B-tree可以认为是一个index,数据库中包括多个table,一个table中包括多个B-tree,通过该SMO标记可以确定哪个B-tree发生了SMO。

[0071] 若从机设备在回放物理日志过程中,回放到带有SMO标记的物理日志,且该带有SMO标记的物理日志包括对目标索引页面进行修改操作的日志,那么,在访问该目标索引页面时,目标索引页面可能处于SMO过程中,因此,从机设备可以根据SMO标记确定目标索引页面是否处于SMO过程中。

[0072] 在本申请实施例中,提供了确定目标索引页面是否处于SMO过程中的方式,一种方式是根据结构变更操作标记,从物理日志中确定结构变更操作日志,即SMO日志,SMO日志为带有SMO标记的物理日志。若该SMO日志中不包括对目标索引页面进行修改操作的日志,则说明目标索引页面并未处于SMO过程中,可以访问该目标索引页面;若该SMO日志中包括对目标索引页面进行修改操作的日志,则认为目标索引页面处于SMO过程中。

[0073] 基于图3所示的B-tree,一条完整的SMO日志的示意图可以参见图4所示,该SMO日志包括物理日志头、SMO标记、修改索引页面6、修改索引页面7、修改索引页面3和物理日志尾。SMO标记表示新增的逻辑日志 (LOG_TYPE_SMO,table id,index id);该SMO日志表示一个分裂操作,索引页面6分裂为索引页面6和索引页面7。

[0074] 若目标索引页面为索引页面6,通过SMO标记可以确定出SMO日志,而该SMO日志中包括修改索引页面6的物理日志,因此,可以认为目标索引页面处于SMO过程中。

[0075] 在一些可能的实施例中,SMO日志中虽然包括目标索引页面对应的物理日志,但是由于SMO日志中可能还包括其他索引页面对应的物理日志,从机设备在回放SMO日志的过程中,会按照图4所示的从物理日志头到物理日志尾的顺序进行回放,也就是说,目标索引页面可能已经结束SMO过程,或者还未进行SMO过程。因此,为了提高目标索引页面是否处于SMO过程的准确性,另一种方式可以是根据结构变更操作标记,从物理日志中确定结构变更操作日志,结构变更操作日志包括目标索引页面对应的目标日志和结构变更操作标记。从机设备可以根据目标日志对应的第一日志序列号、结构变更操作标记对应的第二日志序列号和第三日志序列号之间的大小关系,确定目标索引页面是否处于结构变更操作过程中,第三日志序列号表征从机设备回放物理日志的进展位点。

[0076] 由于每个索引页面对应的物理日志在SMO日志中具有对应的lsn;结构变更操作标记也具有对应的日志序列号,表示为sync lsn,通过sync lsn可以知晓何时开始改变数据存储结构的结构;而回放过程中,从机设备回放物理日志的进展位点也可以通过应用日志序列号 (Applied lsn) 表示,Applied lsn在完成整条SMO日志回放后才会更新到SMO日志的物理日志尾,当回放该SMO日志时,SMO日志回放到SMO标记的示意图参见图5所示,此时Applied lsn为100,结构变更操作标记的日志序列号sync lsn为110。基于此,从机设备可以根据目标日志对应的第一日志序列号 (即lsn)、结构变更操作标记对应的第二日志序列号 (即sync lsn) 和第三日志序列号 (即Applied lsn) 之间的大小关系,确定目标索引页面是否处于结构变更操作过程中。

[0077] 由于SMO日志中每个索引页面对应的物理日志被回放才会更新其lsn到SMO日志的

物理日志尾,回放该索引页面对应的物理日志,说明该索引页面处于SMO过程中,若第一日志序列号大于第二日志序列号,说明回放目标索引页面对应的物理日志,且第一日志序列号大于第三日志序列号,说明Applied lsn还未更新到物理日志尾,并未完成整条SMO日志的回放,从而确定目标索引页面处于SMO过程中。否则,确定目标索引页面未处于SMO过程中。

[0078] 参见图6,图6示出了B-tree处于SMO回放过程中的结构示意图,基于图4所示的SMO日志,虚线方框表示本次SMO过程影响的索引页面,包括索引页面3、索引页面6和索引页面7。灰色方框所表示的索引页面6和索引页面7表示对应的SMO日志已经应用,索引页面的lsn也更新到了本条SMO日志的物理日志尾,即索引页面6和索引页面7的lsn为160。若目标索引页面为索引页面6,第二日志序列号sync lsn为110,第三日志序列号Applied lsn为100,此时,目标索引页面对应的第一日志序列号160大于sync lsn,且第一日志序列号160大于Applied lsn,说明目标索引页面处于SMO过程中。

[0079] S204、若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠。

[0080] 若从机设备确定目标索引页面处于结构变更操作过程中,则暂停读取数据,控制用户读线程进行预设时间的休眠。由于从机设备通过日志回放线程回放的速度非常快,该预设时间内目标索引页面的结构变更操作回放已经完成。

[0081] 需要说明的是,针对不同的数据存储结构,确定目标索引页面处于结构变更操作过程中后,在进行休眠之前从机设备所执行的处理操作可能有所不同,本实施例主要以数据存储结构是B-tree为例进行介绍。

[0082] 对于B-tree来说,在访问目标索引页面时,用于扫描B-tree的游标可能处于不同的阶段,包括打开游标的过程中和游标移动的过程中。其中,打开游标表示根据游标处于不同的阶段,在进行预设时间的休眠之前,可以进行不同的处理。

[0083] 若访问目标索引页面时,游标处于打开游标的过程中,此时,目标索引页面通常是父节点或中间节点所对应的索引页面,则在进行预设时间的休眠之前,可以释放访问至目标索引页面的目标路径(目标路径可以是游标路径),目标路径为从父节点依次访问到所述目标索引页面所对应的中间节点之间的路径。

[0084] 若访问目标索引页面时,游标处于游标移动的过程中,此时,目标索引页面通常是叶子节点所对应的索引页面,则在进行预设时间的休眠之前,可以保存目标数据在目标索引页面中的位置信息(该位置信息可以是游标位置),以便后续可以根据该位置信息重新定位目标数据。

[0085] S205、结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。

[0086] 在预设时间的休眠后,目标索引页面的结构变更操作回放已经完成,Applied lsn更新到物理日志尾,即Applied lsn为160,参见图7所示。B-tree的SMO日志回放完成后的B-tree的结构示意图可以参见图8所示,此时,各个索引页面的lsn也更新为160,此时,若目标索引页面是索引页面6,此时,目标索引页面lsn并不满足大于sync lsn,且大于Applied lsn的条件,故可以重新执行S202中根据数据读取请求,按照从机设备中的数据存储结构依次访问索引页面的步骤,以读取到目标数据。

[0087] 在一些可能的实施例中,若通过S203确定目标索引页面未处于结构变更操作过程

中,则可以访问该目标索引页面。由于目标索引页面可以是B-tree中不同的索引页面,例如目标索引页面可以是保存目标数据的索引页面,此时,若目标索引页面未处于结构变更操作过程中,可以直接读取目标数据。当然,若目标索引页面为父节点或中间节点或叶子节点中未保存目标数据的索引页面,此时,若目标索引页面未处于结构变更操作过程中,可以继续按照数据存储结构(例如B-tree)访问目标索引页面的下一个索引页面,直到读取到目标数据。

[0088] 在从机设备通过上述方法读取到目标数据后,可以向用户对应的终端设备返回目标数据,以便终端设备使用。

[0089] 由上述技术方案可以看出,在基于物理复制的主从架构中,从机设备可以通过回放物理日志以进行物理复制,实现主从设备的数据一致性,在从机设备回放物理日志以进行物理复制的过程中,若用户需要通过从机设备中读取目标数据,从机设备可以获取用于读取目标数据的数据读取请求,根据数据读取请求,按照从机设备中的数据存储结构依次访问索引页面。针对访问的每个索引页面例如目标索引页面,从机设备根据物理日志中添加的结构变更操作标记确定目标索引页面是否处于结构变更操作过程中,若确定目标索引页面处于结构变更操作过程中,则进行预设时间的休眠。由于从机设备通过日志回放线程回放的速度非常快,在进行预设时间的休眠后,目标索引页面的结构变更操作回放已经完成,此时,重新执行根据所述数据读取请求,按照从机设备中的数据存储结构依次访问索引页面的步骤,以读取目标数据。通过对访问到的每个索引页面例如目标索引页面实时进行是否处于结构变更操作过程中的判断,可以保证在目标索引页面的结构变更操作回放完成之后再访问,此时数据存储结构例如B树的结构不再变化,处于一致性状态,避免由于数据存储结构的改变导致系统崩溃,实现数据的一致性读取,极大提高了从机读性能。

[0090] 与相关技术中,缓存多版本的索引页面,以保证一致性读取的方法相比,该方法无需通过缓冲区缓存多版本的索引页面,降低了从机设备的存储压力,不影响存储层索引页面淘汰,系统更稳健。

[0091] 另外,该方法以索引页面为粒度进行是否处于SMO过程中的判断,粒度更细,适用范围更广。粒度上只影响SMO过程相关的索引页面,其他索引页面不受影响;适应范围,游标打开过程中和游标移动过程中,均适用。且没有任何针对索引页面上锁操作,不会阻塞数据读取请求进入,延迟小,效率高。

[0092] 本申请实施例提供的数据读取方法实现一致性读,是一写多读的主从架构中的关键技术,直接影响从机数据读取的一致性和性能。

[0093] 接下来,将结合实际应用场景对本申请实施例提供的数据读取方法进行介绍。在该应用场景中,以B-tree为读取数据的索引方式,当从机设备接收到用户通过终端设备发送的数据读取请求后,从机设备可以打开游标,开始对B-tree从父节点开始进行扫描,从上至下访问索引页面,直到在叶子节点所对应的索引页面中获取到目标数据。在扫描B-tree的过程中,即按照B-tree访问索引页面的过程中,对访问的索引页面进行是否处于SMO过程进行判断。通常情况下,通过游标扫描B-tree可以包括两个过程,即打开游标,从父节点到中间节点这一过程,以及游标移动过程。在这两个过程中,所访问的索引页面都需要进行其是否处于SMO过程进行判断。基于此,本申请实施例提供一种数据读取方法,参见图9,所述方法包括:

- [0094] S901、进入打开游标过程。
- [0095] S902、判断是否访问到处于SM0过程中的索引页面,若是,执行S906;若否,执行S903。
- [0096] S903、进入游标左右移动过程。
- [0097] S904、判断是否访问到处于SM0过程中的索引页面,若是,执行S907,若否,执行S905。
- [0098] S905、读取目标数据。
- [0099] S906、释放游标路径。
- [0100] S907、保存游标位置。
- [0101] S908、进入预设时间的休眠,以在休眠结束后,重新执行S901。
- [0102] 基于图2对应实施例提供的数据读取方法,本申请实施例还提供一种数据读取装置,参见图10,所述装置部署在基于物理复制的主从架构中的从机设备中,所述装置包括获取单元1001、访问单元1002、确定单元1003、休眠单元1004和读取单元1005:
- [0103] 所述获取单元1001,用于在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;
- [0104] 所述访问单元1002,用于根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面;
- [0105] 所述确定单元1003,用于当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;
- [0106] 所述休眠单元1004,用于若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;
- [0107] 所述读取单元1005,用于结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。
- [0108] 在一种可能的实现方式中,所述确定单元1003,用于:
- [0109] 根据所述结构变更操作标记,从所述物理日志中确定结构变更操作日志,所述结构变更操作日志包括所述目标索引页面对应的目标日志和所述结构变更操作标记;
- [0110] 根据所述目标日志对应的第一日志序列号、所述结构变更操作标记对应的第二日志序列号和第三日志序列号之间的大小关系,确定所述目标索引页面是否处于结构变更操作过程中,所述第三日志序列号表征所述从机设备回放所述物理日志的进展位点。
- [0111] 在一种可能的实现方式中,所述确定单元1003,用于:
- [0112] 若所述第一日志序列号大于所述第二日志序列号,且所述第一日志序列号大于所述第三日志序列号,确定所述目标索引页面处于所述结构变更操作过程中;
- [0113] 否则,确定所述目标索引页面未处于所述结构变更操作过程中。
- [0114] 在一种可能的实现方式中,所述读取单元1005还用于:
- [0115] 若根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面未处于结构变更操作过程中,读取所述目标数据,或,继续按照所述数据存储结构访问所述目标索引页面的下一个索引页面,直到读取到所述目标数据。
- [0116] 在一种可能的实现方式中,所述数据存储结构为B树,所述B树包括父节点、中间节点和叶子节点,所述父节点、所述中间节点和所述叶子节点之间具有连接关系,所述父节

点、所述中间节点和所述叶子节点分别对应不同的索引页面。

[0117] 在一种可能的实现方式中,所述装置还包括释放单元:

[0118] 所述释放单元,用于若确定所述目标索引页面处于结构变更操作过程中,释放访问至所述目标索引页面的目标路径,所述目标路径为从所述父节点依次访问到所述目标索引页面所对应的所述中间节点之间的路径。

[0119] 在一种可能的实现方式中,所述装置还包括保存单元:

[0120] 所述保存单元,用于若确定所述目标索引页面处于结构变更操作过程中,若所述目标索引页面为所述叶子节点对应的索引页面,保存所述目标数据在所述目标索引页面中的位置信息。

[0121] 在一种可能的实现方式中,所述装置还包括返回单元:

[0122] 所述返回单元,用于向用户对应的终端设备返回所述目标数据。

[0123] 本申请实施例还提供了一种终端设备,下面结合附图对该终端设备进行介绍。请参见图11所示,以终端设备为智能手机为例进行介绍:

[0124] 图11示出的是与本申请实施例提供的终端设备相关的智能手机的部分结构的框图。参考图11,智能手机包括:射频(英文全称:Radio Frequency,英文缩写:RF)电路1110、存储器1120、输入单元1130、显示单元1140、传感器1150、音频电路1160、无线保真(英文全称:wireless fidelity,英文缩写:WiFi)模块1170、处理器1180、以及电源1190等部件。本领域技术人员可以理解,图11中示出的智能手机结构并不构成对智能手机的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。

[0125] 存储器1120可用于存储软件程序以及模块,处理器1180通过运行存储在存储器1120的软件程序以及模块,从而执行智能手机的各种功能应用以及数据处理。存储器1120可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据智能手机的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器1120可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0126] 处理器1180是智能手机的控制中心,利用各种接口和线路连接整个智能手机的各个部分,通过运行或执行存储在存储器1120内的软件程序和/或模块,以及调用存储在存储器1120内的数据,执行智能手机的各种功能和处理数据,从而对智能手机进行整体监控。可选的,处理器1180可包括一个或多个处理单元;优选的,处理器1180可集成应用处理器和调制解调处理器,其中,应用处理器主要处理操作系统、用户界面和应用程序等,调制解调处理器主要处理无线通信。可以理解的是,上述调制解调处理器也可以不集成到处理器1180中。

[0127] 在本实施例中,由终端设备所实现的步骤可以基于图11所示的终端设备的结构实现。

[0128] 本申请实施例提供一种用于数据读取的电子设备,该电子设备即前述实施例提到的从机设备,用于执行本申请实施例提供的数据读取方法。该电子设备可以是服务器,请参见图12所示,图12为本申请实施例提供的服务器1200的结构图,服务器1200可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上中央处理器(Central Processing

Units,简称CPU) 1222 (例如,一个或一个以上处理器) 和存储器1232,一个或一个以上存储应用程序1242或数据1244的存储介质1230 (例如一个或一个以上海量存储设备)。其中,存储器1232和存储介质1230可以是短暂存储或持久存储。存储在存储介质1230的程序可以包括一个或一个以上模块 (图示没标出),每个模块可以包括对服务器中的一系列指令操作。更进一步地,中央处理器1222可以设置为与存储介质1230通信,在服务器1200上执行存储介质1230中的一系列指令操作。

[0129] 服务器1200还可以包括一个或一个以上电源1226,一个或一个以上有线或无线网络接口1250,一个或一个以上输入输出接口1258,和/或,一个或一个以上操作系统1241,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™等等。

[0130] 在本实施例中,所述服务器1200中的中央处理器1222可以执行以下步骤:

[0131] 在所述从机设备回放物理日志的过程中,获取用于读取目标数据的数据读取请求;

[0132] 根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面;

[0133] 当访问到目标索引页面时,根据所述物理日志中添加的结构变更操作标记确定所述目标索引页面是否处于结构变更操作过程中;

[0134] 若确定所述目标索引页面处于结构变更操作过程中,进行预设时间的休眠;

[0135] 结束所述休眠后,重新执行所述根据所述数据读取请求,按照所述从机设备中的数据存储结构依次访问索引页面的步骤,以读取所述目标数据。

[0136] 根据本申请的一个方面,提供了一种计算机可读存储介质,所述计算机可读存储介质用于存储程序代码,所述程序代码用于执行前述各个实施例所述的数据读取方法。

[0137] 根据本申请的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述实施例各种可选实现方式中提供的方法。

[0138] 本申请的说明书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等 (如果存在) 是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例例如能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0139] 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。

[0140] 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。

[0141] 另外,在本申请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。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。

[0142] 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,本申请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)执行本申请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(Read-Only Memory,简称ROM)、随机存取存储器(Random Access Memory,简称RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0143] 以上所述,以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术成员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

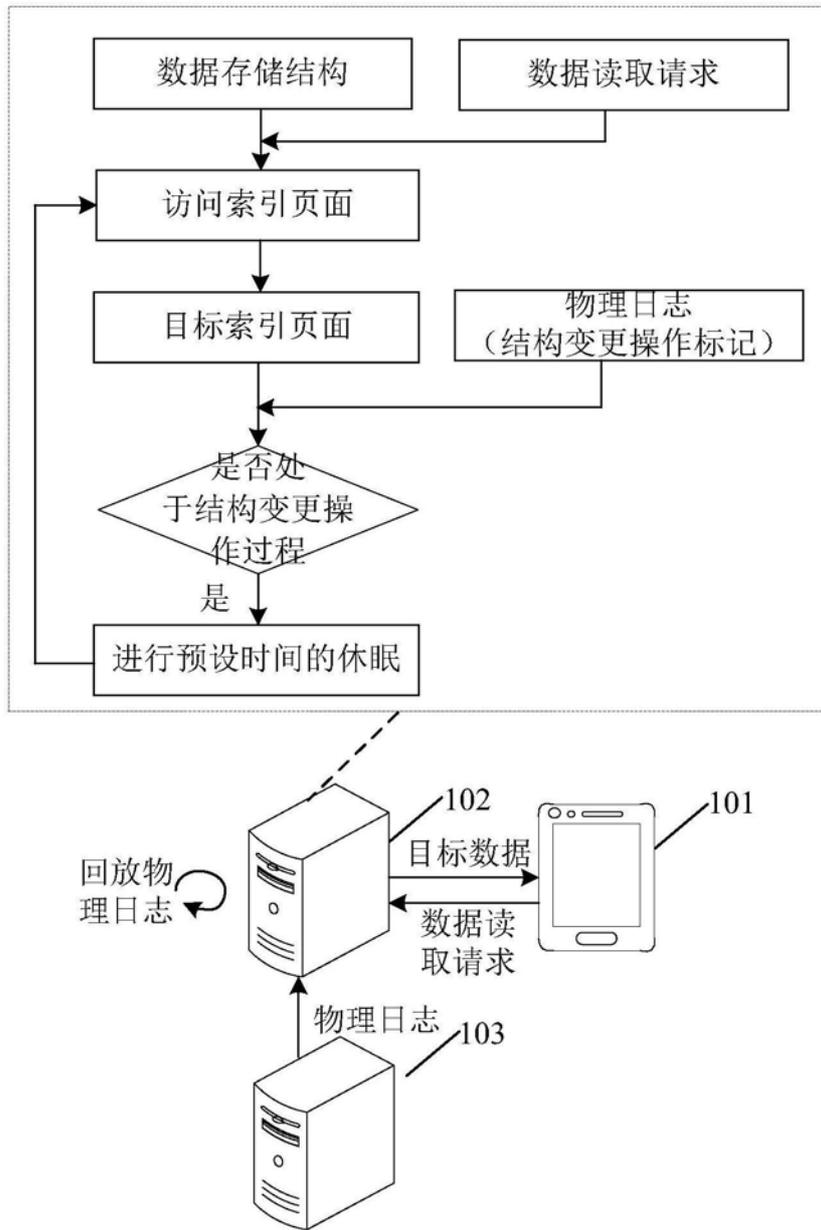


图1

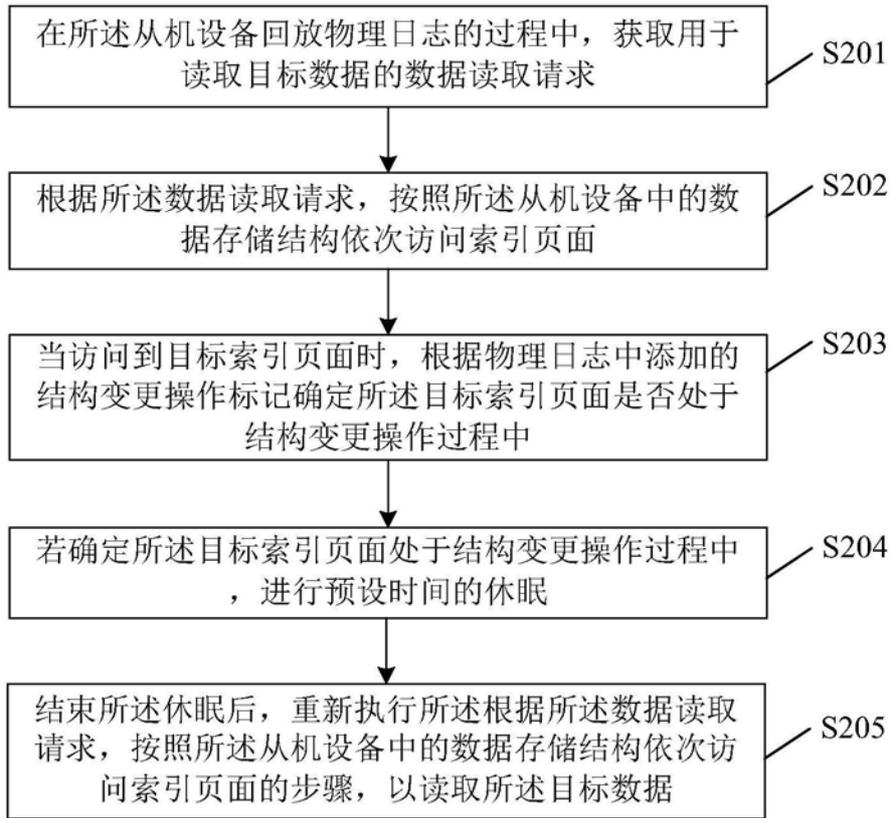


图2

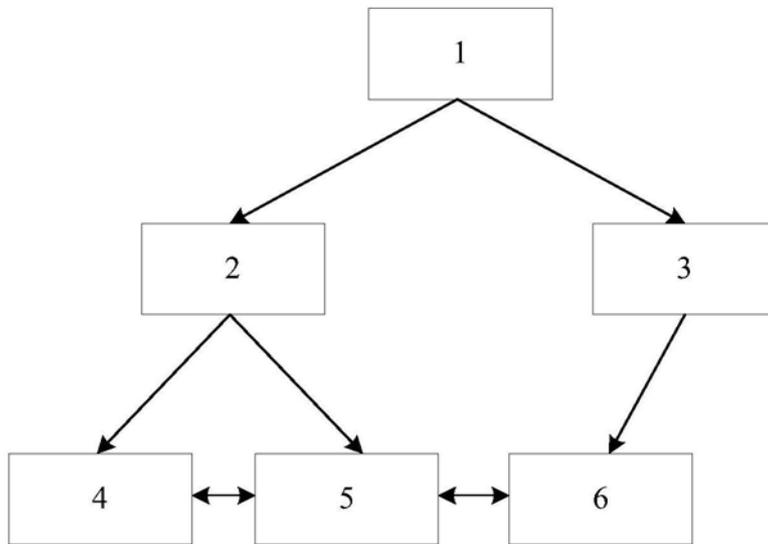


图3

物理日志头	SMO标记	修改索引页面6	修改索引页面7	修改索引页面3	物理日志尾
-------	-------	---------	---------	---------	-------

图4



图5

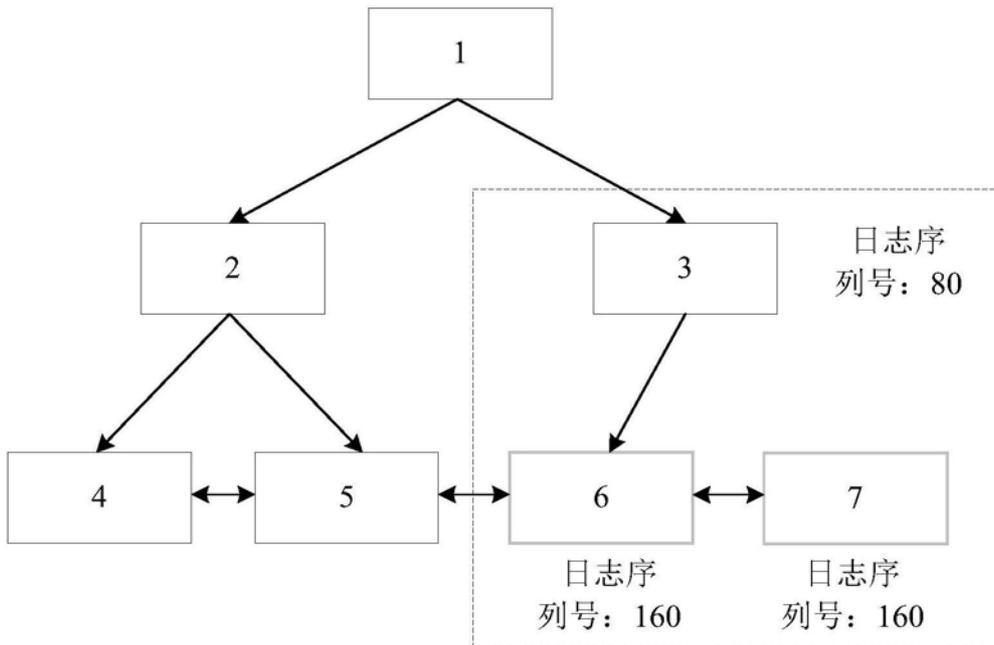


图6



图7

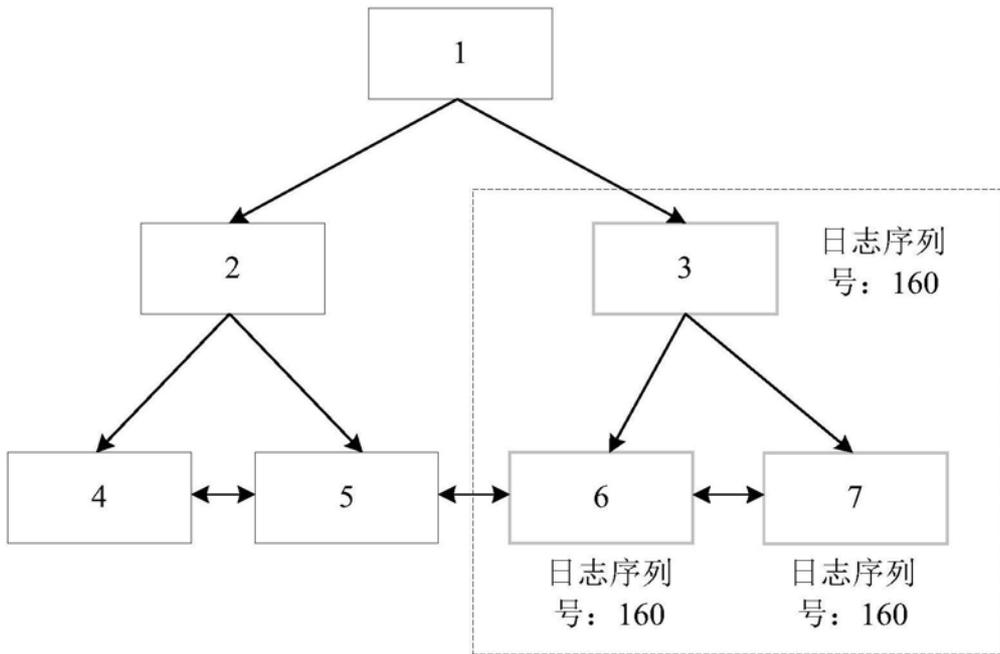


图8

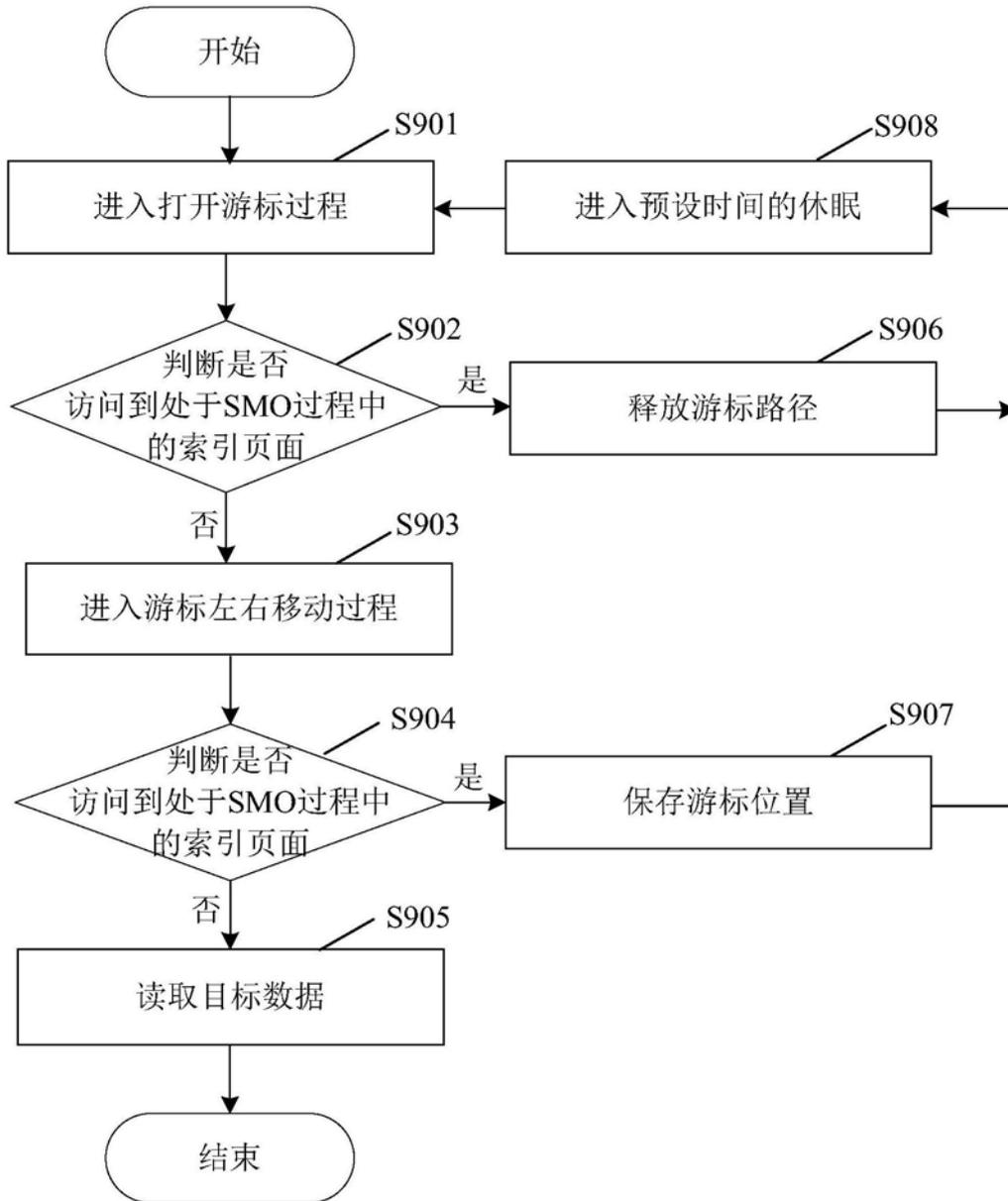


图9

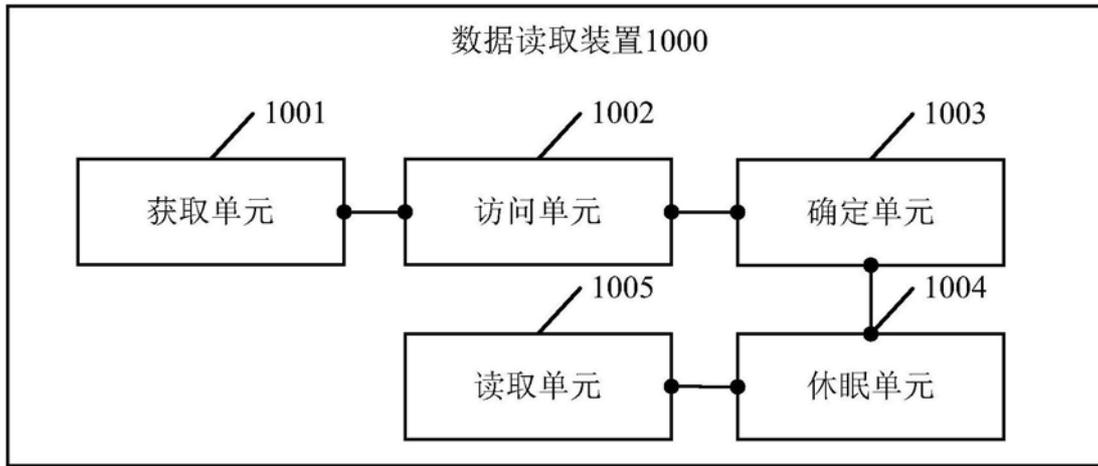


图10

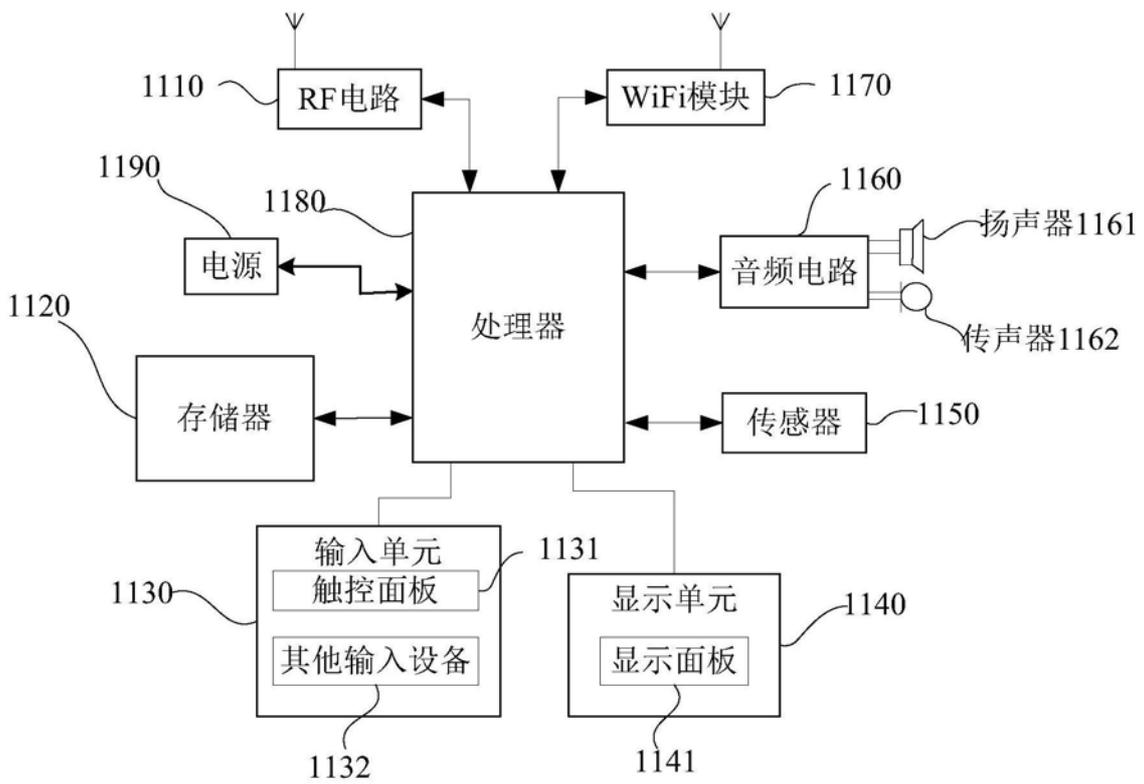


图11

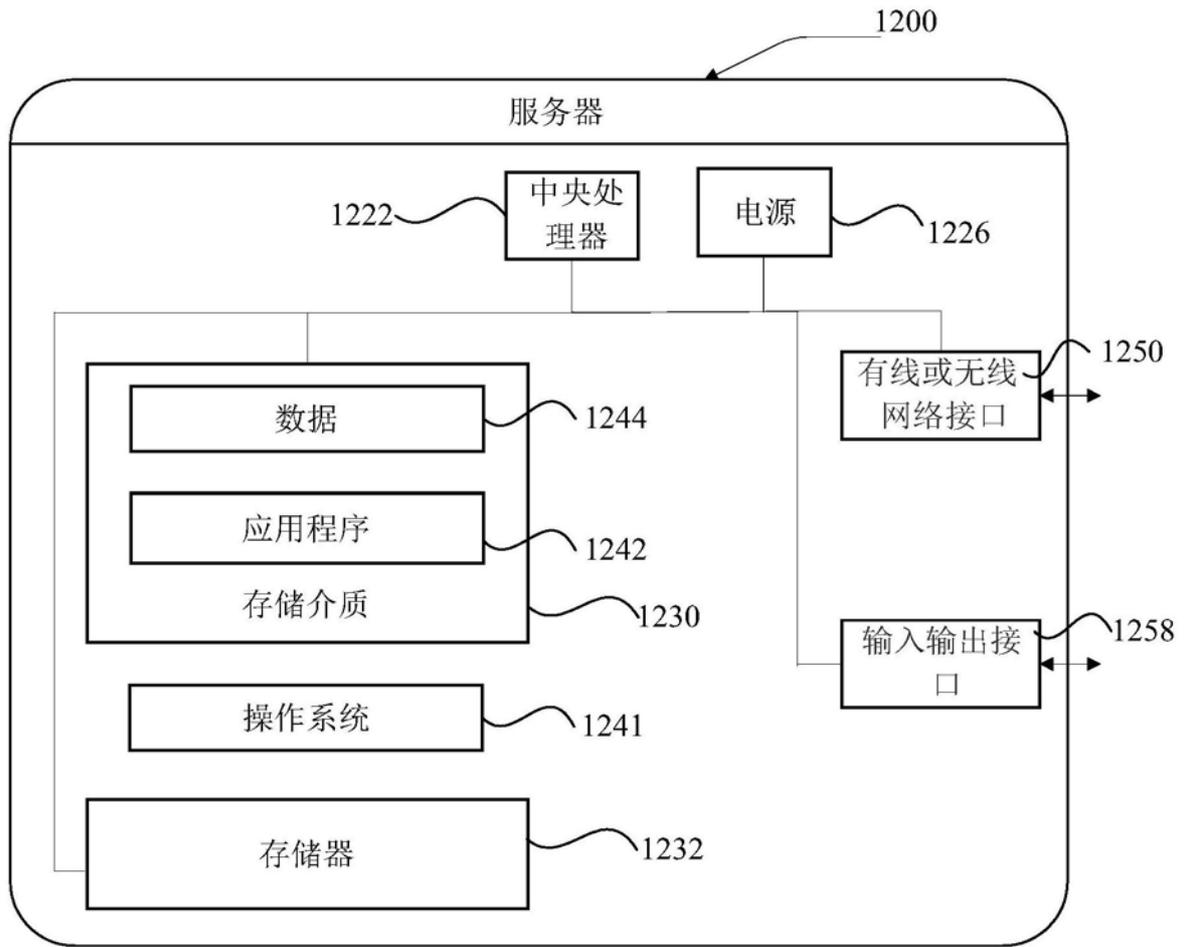


图12