



(12) 发明专利

(10) 授权公告号 CN 111538719 B

(45) 授权公告日 2024.04.19

(21) 申请号 202010371096.8

(22) 申请日 2020.04.30

(65) 同一申请的已公布的文献号
申请公布号 CN 111538719 A

(43) 申请公布日 2020.08.14

(73) 专利权人 深圳前海微众银行股份有限公司
地址 518000 广东省深圳市前海深港合作
区前湾一路1号A栋201室(入驻深圳市
前海商务秘书有限公司)

(72) 发明人 邹普

(74) 专利代理机构 广州三环专利商标代理有限
公司 44202
专利代理师 赵蕊

(51) Int. Cl.
G06F 16/21 (2019.01)

(56) 对比文件

CN 105718570 A, 2016.06.29

CN 108628874 A, 2018.10.09

CN 110209653 A, 2019.09.06

CN 110851237 A, 2020.02.28

US 2016196324 A1, 2016.07.07

WO 2018149221 A1, 2018.08.23

审查员 宋娜娜

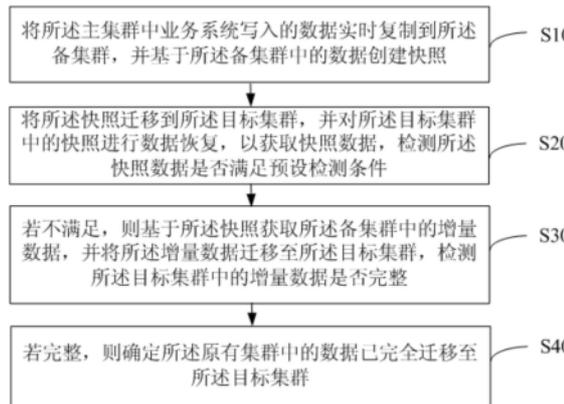
权利要求书2页 说明书12页 附图3页

(54) 发明名称

数据迁移方法、装置、设备及计算机存储介
质

(57) 摘要

本发明涉及金融科技(Fintech)技术领域,并公开了一种数据迁移方法,该方法包括:将所述主集群中业务系统写入的数据实时复制到所述备集群,并基于所述备集群中的数据创建快照;将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;若完整,则确定所述原有集群中的数据已完全迁移至所述目标集群。本发明还公开了一种数据迁移装置、设备和一种计算机存储介质。本发明实现了在业务不停服的同时,进行数据迁移。



1. 一种数据迁移方法,其特征在于,所述数据迁移方法应用于数据迁移系统,所述数据迁移系统包括原有集群和目标集群,所述原有集群包括主集群和备集群,所述数据迁移方法包括如下步骤:

在业务不停服的情况下,将所述主集群中业务系统写入的数据实时复制到所述备集群,具体为:开启所述主集群中的regplication创建dest集群表结构,并采用enable_table_replication的形式将source hbase集群的表结构同步至所述备集群,开启所述主集群和所述备集群的replication命令,并通过观察预设页面,观察是否有相应tps写入所述备集群,即检测是否有数据写入到所述备集群,若有,则读写此相应tps,并确定所述主集群中的regplication功能正常;若没有数据写入到所述备集群,则检测相应的regplication机制是否正常,不正常就进行修正;并在确定所述备集群中有写入数据后,清理hbase数据,并保留所述主集群中的原有数据分区;

基于所述备集群中的数据创建快照,以避免影响到所述业务系统和所述主集群之间的服务交流;

将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;

若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;

若完整,则确定所述原有集群中的数据已完全迁移至所述目标集群。

2. 如权利要求1所述的数据迁移方法,其特征在于,所述检测所述快照数据是否满足预设检测条件的步骤,包括:

获取所述备集群的存量数据,检测所述快照数据是否和所述存量数据匹配;

若所述快照数据和所述存量数据不匹配,则获取所述备集群中快照迁移前所在的历史文件,并获取所述目标集群中快照所在的线程文件,检测所述历史文件和所述线程文件是否匹配;

若所述历史文件和所述线程文件不匹配,则确定所述快照数据不满足预设检测条件。

3. 如权利要求1所述的数据迁移方法,其特征在于,所述将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复的步骤,包括:

控制所述备集群中具有所述数据的表处于禁止状态,并根据输入的迁移指令将所述快照迁移至所述目标集群,将所述快照中的数据恢复至所述目标集群的工作目录,将具有所述快照中的数据的工作目录作为快照数据。

4. 如权利要求1所述的数据迁移方法,其特征在于,所述检测所述目标集群中的增量数据是否完整的步骤,包括:

基于预设时间戳获取所述备集群的备集群数据,并获取所述目标集群的目标集群数据,校验所述备集群数据和所述目标集群数据之间的差异数据;

根据所述差异数据对所述备集群数据进行过滤,并检测过滤后的备集群数据是否和所述目标集群数据匹配;

若匹配,则确定所述目标集群中的增量数据完整。

5. 如权利要求1所述的数据迁移方法,其特征在于,所述基于所述备集群中的数据创建快照的步骤,包括:

对所述备集群中的数据进行加锁,以获取加锁数据,将所述加锁数据刷盘到所述备集群中的hfile文件,并对具有所述加锁数据的hfile文件创建引用指针,根据所述引用指针创建快照。

6.如权利要求1所述的数据迁移方法,其特征在于,所述基于所述快照获取所述备集群中的增量数据的步骤,包括:

确定所述快照在所述备集群中所有数据的数据节点,根据所述数据节点对所述备集群中的所有数据进行回放,以获取所述备集群中的增量数据。

7.如权利要求1-6任一项所述的数据迁移方法,其特征在于,所述将所述主集群中业务系统写入的数据实时复制到所述备集群的步骤,包括:

将所述主集群中的表结构复制到所述备集群,并检测所述备集群接收的表结构中是否存在业务系统写入的数据;

若存在,则执行所述基于所述备集群中的数据创建快照的步骤;

若不存在,则对所述主集群中的复制机制进行修正,并输出未复制成功的提示信息。

8.一种数据迁移装置,其特征在于,所述数据迁移装置包括:

复制模块,用于在业务不停服的情况下,将主集群中业务系统写入的数据实时复制到备集群,具体为:开启所述主集群中的reglication创建dest集群表结构,并采用enable_table_replication的形式将source hbase集群的表结构同步至所述备集群,开启所述主集群和所述备集群的replication命令,并通过观察预设页面,观察是否有相应tps写入所述备集群,即检测是否有数据写入到所述备集群,若有,则读写此相应tps,并确定所述主集群中的reglication功能正常;若没有数据写入到所述备集群,则检测相应的reglication机制是否正常,不正常就进行修正;并在确定所述备集群中有写入数据后,清理hbase数据,并保留所述主集群中的原有数据分区;基于所述备集群中的数据创建快照,以避免影响到所述业务系统和所述主集群之间的服务交流;

恢复模块,用于将所述快照迁移到目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;

检测模块,用于若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;

确定模块,用于若完整,则确定原有集群中的数据已完全迁移至所述目标集群。

9.一种数据迁移设备,其特征在于,所述数据迁移设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的数据迁移程序,所述数据迁移程序被所述处理器执行时实现如权利要求1至7中任一项所述的数据迁移方法的步骤。

10.一种计算机存储介质,其特征在于,所述计算机存储介质上存储有数据迁移程序,所述数据迁移程序被处理器执行时实现如权利要求1至7中任一项所述的数据迁移方法的步骤。

数据迁移方法、装置、设备及计算机存储介质

技术领域

[0001] 本发明涉及金融科技 (Fintech) 的测试技术领域, 尤其涉及数据迁移方法、装置、设备及计算机存储介质。

背景技术

[0002] 随着计算机技术的发展, 越来越多的技术 (大数据、分布式、区块链Blockchain、人工智能等) 应用在金融领域, 传统金融业正在逐步向金融科技 (Fintech) 转变, 但由于金融行业的安全性、实时性要求, 也对技术提出了更高的要求。目前互联网的爆炸式数据增长会导致对海量数据的存储有迫切需求, 由于传统的mysql (关系型数据库管理系统)、oracle (关系型数据库) 在百万级的数据能提供很好的事务能力和读写性能, 但随着数据量的增加, 它的性能急剧下降。而Hbase (开源数据库) 是基于google bigtable (分布式存储系统) 的开源实现, 适合海量的数据存储, 和较好的随机读写能力, 故Hbase在海量数据的存储方面具有非常典型的应用场景。但由于机房的IDC (Internet Data Center, 互联网数据中心) 变化和业务需求引起集群迁移会给大数据的维护人员造成很大的挑战, 传统的数据迁移方案有的会造成业务停服待集群数据追平后再进行业务写入, 有的会涉及到客户端改造, 将迁移窗口期间的数据进行缓存, 待存量迁移完成后, 再由客户端回放数据追平。因此, 如何在业务不停服的同时, 进行数据的迁移成为了目前亟待解决的技术问题。

发明内容

[0003] 本发明的主要目的在于提出一种数据迁移方法、装置、设备及计算机存储介质, 旨在解决如何在业务不停服的同时, 进行数据的迁移的技术问题。

[0004] 为实现上述目的, 本发明提供一种数据迁移方法, 所述数据迁移方法应用于数据迁移系统, 所述数据迁移系统包括原有集群和目标集群, 所述原有集群包括主集群和备集群, 所述数据迁移方法包括如下步骤:

[0005] 将所述主集群中业务系统写入的数据实时复制到所述备集群, 并基于所述备集群中的数据创建快照;

[0006] 将所述快照迁移到所述目标集群, 并对所述目标集群中的快照进行数据恢复, 以获取快照数据, 检测所述快照数据是否满足预设检测条件;

[0007] 若不满足, 则基于所述快照获取所述备集群中的增量数据, 并将所述增量数据迁移至所述目标集群, 检测所述目标集群中的增量数据是否完整;

[0008] 若完整, 则确定所述原有集群中的数据已完全迁移至所述目标集群。

[0009] 可选地, 检测所述快照数据是否满足预设检测条件的步骤, 包括:

[0010] 获取所述备集群的存量数据, 检测所述快照数据是否和所述存量数据匹配;

[0011] 若所述快照数据和所述存量数据不匹配, 则获取所述备集群中快照迁移前所在的历史文件, 并获取所述目标集群中快照所在的线程文件, 检测所述历史文件和所述现存文件是否匹配;

[0012] 若所述历史文件和所述现存文件不匹配,则确定所述快照数据不满足预设检测条件。

[0013] 可选地,将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复的步骤,包括:

[0014] 控制所述备集群中具有所述数据的表处于禁止状态,并根据输入的迁移指令将所述快照迁移至所述目标集群,将所述快照中的数据恢复至所述目标集群的工作目录,将具有所述快照中的数据的工作目录作为快照数据。

[0015] 可选地,检测所述目标集群中的增量数据是否完整的步骤,包括:

[0016] 基于预设时间戳获取所述备集群的备集群数据,并获取所述目标集群的目标集群数据,校验所述备集群数据和所述目标集群数据之间的差异数据;

[0017] 根据所述差异数据对所述备集群数据进行过滤,并检测所述过滤后的备集群数据是否和所述目标集群数据匹配;

[0018] 若匹配,则确定所述目标集群中的增量数据完整。

[0019] 可选地,基于所述备集群中的数据创建快照的步骤,包括:

[0020] 对所述备集群中的数据进行加锁,以获取加锁数据,将所述加锁数据刷盘到所述备集群中的hfile文件,并对具有所述加锁数据的hfile文件创建引用指针,根据所述引用指针创建快照。

[0021] 可选地,基于所述快照获取所述备集群中的增量数据的步骤,包括:

[0022] 确定所述快照在所述备集群中所有数据的数据节点,根据所述数据节点对所述备集群中的所有数据进行回放,以获取所述备集群中的增量数据。

[0023] 可选地,将所述主集群中业务系统写入的数据实时复制到所述备集群的步骤,包括:

[0024] 将所述主集群中的表结构复制到所述备集群,并检测所述备集群接收的表结构中是否存在业务系统写入的数据;

[0025] 若存在,则执行所述基于所述备集群中的数据创建快照的步骤;

[0026] 若不存在,则对所述主集群中的复制机制进行修正,并输出未复制成功的提示信息。

[0027] 此外,为实现上述目的,本发明还提供一种数据迁移装置,所述数据迁移装置包括:

[0028] 复制模块,用于将主集群中业务系统写入的数据实时复制到备集群,并基于所述备集群中的数据创建快照;

[0029] 恢复模块,用于将所述快照迁移到目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;

[0030] 检测模块,用于若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;

[0031] 确定模块,用于若完整,则确定原有集群中的数据已完全迁移至所述目标集群。

[0032] 此外,为实现上述目的,本发明还提供一种数据迁移设备,所述数据迁移设备包括:存储器、处理器及存储在所述存储器上并可在所述处理器上运行的数据迁移程序,所述数据迁移程序被所述处理器执行时实现如上所述的数据迁移方法的步骤。

[0033] 此外,为实现上述目的,本发明还提供一种计算机存储介质,所述计算机存储介质上存储有数据迁移程序,所述数据迁移程序被处理器执行时实现如上所述的数据迁移方法的步骤。

[0034] 本发明通过将所述主集群中业务系统写入的数据实时复制到所述备集群,并基于所述备集群中的数据创建快照;将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;若完整,则确定所述原有集群中的数据已完全迁移至所述目标集群。通过原有集群划分为主集群和备集群,再将主集群中业务系统写入的数据实时复制到备集群,并对备集群创建快照,再将快照迁移到目标集群,从而可以避免影响到业务系统和主集群之间的服务交流,保障了在业务系统写入原有集群的数据不间断的情况下,又能将原有集群中的数据迁移到目标集群,并且在将快照复制到目标集群后,还会获取原有集群中的增量数据,并将增量数据迁移到目标集群,从而保障了将原有集群中的所有数据都迁移到目标集群,使得业务系统在原有集群将所有数据迁移到目标集群后,能无缝直接切换到目标集群,避免了现有技术中在进行数据迁移时,会造成业务系统停止服务的现象发生。

附图说明

[0035] 图1是本发明实施例方案涉及的硬件运行环境的数据迁移设备结构示意图;

[0036] 图2为本发明数据迁移方法第一实施例的流程示意图;

[0037] 图3为本发明数据迁移装置的装置模块示意图;

[0038] 图4为本发明数据迁移方法中快照创建的流程示意图;

[0039] 图5为本发明数据迁移方法的流程示意图。

[0040] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0041] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0042] 如图1所示,图1是本发明实施例方案涉及的硬件运行环境的数据迁移设备结构示意图。

[0043] 本发明实施例数据迁移设备可以是PC机或服务器设备,其上运行有Java虚拟机。

[0044] 如图1所示,该数据迁移设备可以包括:处理器1001,例如CPU,网络接口1004,用户接口1003,存储器1005,通信总线1002。其中,通信总线1002用于实现这些组件之间的连接通信。用户接口1003可以包括显示屏(Display)、输入单元比如键盘(Keyboard),可选用户接口1003还可以包括标准的有线接口、无线接口。网络接口1004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器1005可以是高速RAM存储器,也可以是稳定的存储器(non-volatile memory),例如磁盘存储器。存储器1005可选的还可以是独立于前述处理器1001的存储装置。

[0045] 本领域技术人员可以理解,图1中示出的数据迁移设备结构并不构成对设备的限定,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。

[0046] 如图1所示,作为一种计算机存储介质的存储器1005中可以包括操作系统、网络通信模块、用户接口模块以及数据迁移程序。

[0047] 在图1所示的数据迁移设备中,网络接口1004主要用于连接后台服务器,与后台服务器进行数据通信;用户接口1003主要用于连接客户端(用户端),与客户端进行数据通信;而处理器1001可以用于调用存储器1005中存储的数据迁移程序,并执行下述数据迁移方法中的操作。

[0048] 基于上述硬件结构,提出本发明数据迁移方法实施例。

[0049] 参照图2,图2为本发明数据迁移方法第一实施例的流程示意图,所述方法包括:

[0050] 步骤S10,将所述主集群中业务系统写入的数据实时复制到所述备集群,并基于所述备集群中的数据创建快照;

[0051] 在本实施例中,数据迁移方法应用于数据迁移系统,数据迁移系统包括原有集群和目标集群。原有集群包括主集群和备集群。并且本实施例中的相关术语解释如下:Hadoop是一个能够对大量数据进行分布式处理的软件框架。并且Hadoop包括四个模块,即Common:支持其他模块的公共工具;HDFS:提供高吞吐访问性能的分布式文件系统;YARN:提供作业调度和集群资源管理的框架;MapReduce:大数据并行计算框架,简称MR。Hbase是一个开源的非关系型分布式数据库(NoSQL)。并且Hbase对数据存储管理包括两方面:一是元数据的管理、二是数据的管理。其中,元数据是Hbase将相应的region信息存储在meta表中,客户端的读写每次都会先读meta表找到响应region所在的server,进而通过RPC直接进行读写操作。并且Hbase的数据存储在HDFS中,以严格的字典序、以hfile为最小文件单元存储。

[0052] 在本实施例中,再将主集群中业务系统写入的数据实时复制到备集群时,需要先开启主集群中的regplication(复制)创建dest(目的)集群表结构,并采用enable_table_replication的形式将source hbase集群的表结构同步(即复制)至备集群。再开启主集群和备集群的replication(add peer)命令,并通过观察预设页面(如hbase web 60010页面),观察是否有相应tps(系统吞吐量)写入备集群,即检测是否有数据写入到备集群,若有,则读写此相应tps,并确定主集群中的regplication功能正常。若没有数据写入到备集群,则检测相应的regplication机制是否正常,不正常就进行修正。并在确定备集群中有写入数据后,需要清理hbase数据,并保留主集群中的原有数据分区。

[0053] 再在备集群中通过snapshot快照创建source集群表快照,并通过对集群表进行如图4所示的加锁,刷判断,创建指针等操作,来创建海量(可达到PB级别)数据的快照,并保留截止当前时刻的存量数据,即确定备集群中快照对应的数据节点。

[0054] 步骤S20,将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;

[0055] 当快照创建成功后,可以通过hadoop MR exportSnaphost(迁移快照)将快照导出到dest集群(即目标集群)。并且,exportSnaphost是底层采用distcp的形式传输hdfs文件,因此可以通过控制mappers(映射器)和bandwidth(带宽)来控制相应的hdfs数据传输速度,并且由于是采用分布式传输的方式,因此可以去掉-no-checksum-verify和-no-target-verify,并检查传输前后的checksum(校验和),以确保传输过程中不存在数据丢失或者数据错乱的现象发生。

[0056] 当目标集群收到备集群发送的快照后,可以采用恢复快照技术来恢复快照的快照

内容,即对快照进行数据恢复,以获取快照数据。并且恢复快照技术为创建快照的逆操作,并且在恢复快照时,需要将提前在目标集群中创建的表设置为diabile状态,并通过restore_snapshot将快照中的元数据信息恢复至当前hbase集群,待hbase进行合并时,才会将相应快照中的hfile逐步的移动至当前hbase中的工作目录。至此,存量数据恢复完成,由于存量数据的迁移是保证当前原集群正常业务运行做操的,故不会影响业务。

[0057] 当获取到大量的数据迁移,且不需要业务进行暂停服务后,还需要进一步验证传递过来的快照数据的完整性,即检测快照数据是否满足预设检测条件(可以为用户设置的任意条件,在本实施例中优选的,是检测快照数据的完整性)。并且由于hbase(集群)中的数据是来自于hive(数据仓库),因此可以通过对比hive数据仓库中的数据和目标集群中的数据进行校验,并由于快照在进行迁移完成后也会校验迁移前后的hdfs的文件(即快照所在文件)完整性,并在这两个校验都通过时,则可以确定快照数据满足预设检测条件。其中,在hive中通过统计总行数和hbase通过以下MR任务统计行数进行比对:

[0058] hbase org.apache.hadoop.hbase.mapreduce.RowCounter 'table' :

[0059] Usage:RowCounter[options]<tablename>[--starttime=[start]--endtime=[end] [--range=[startKey],[endKey]] [<column1><column2>...]

[0060] For performance consider the following options:

[0061] -Dhbase.client.scanner.caching=100

[0062] -Dmapreduce.map.speculative=false

[0063] 步骤S30,若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;

[0064] 当经过判断发现快照数据不满足预设检测条件时,也就是在原有集群中仍然有业务系统写入数据时,就可以在flush(清空)到磁盘hdfs wallog时,在zk中记录相应的目录和hdfs wallog文件ID,由于已将原有集群中的大部分数据通过快照的方式迁移到目标集群中,因此在快速时间节点之后产生的数据并不会很多,此时就可以按照相应的wallog格式标准将数据序列化顺序存放在hdfs上(即确定增量数据),并通过回放的方式在目标hbase集群中将数据重写,即将增量数据迁移到目标集群。并且在source原有集群查找相应的主备replication peer,并enable(允许),回放相应堆积在source集群的wallog数据文件,相应wallog文件存储在hdfs上,具体的文件目录管理是存放在zookeeper中相应的peer目录下。再通过观察hbase web 60010观察是否有相应的数据在正常回放至目标hbase集群,待目标集群tps和原有集群tps相差无误时,则意味着增量数据已和原有集群相同。并在目标集群获取到增量数据后,还需要检测目标集群中的增量数据是否完整,并根据不同的检测结果执行不同的操作。但是当快照数据满足预设检测条件时,则可以确定已将原有集群中的所有数据都迁移到目标集群了。

[0065] 其中,可以通过改进的WbVerifyReplication MR工具验证增量数据完整性。即WbVerifyReplication可指定-slaveClusterKey target slaveClusterKey可直接忽略掉-peer信息,直接设置另一被比较集群的zookeeper地址,解决不能比对非replication的集群。而且对于设置了TTL的数据比对源码缺陷,可通过二次重校验的形式避免,由于在scan的时候根据时间进行扫描,会过滤掉最终数据在输出时的TTL校验,所以在源码中新增了如果在校验失败的情况下,再次进行二次校验,并采用get的形式。因为get会在最终返回结果

的时候,在服务端内存中根据表的TTL会进行过滤。并且改进后的WbVerifyReplication可以通过时间很好的校验当前迁移前和迁移后的增量数据完整性。也就是在本实施例中是直接目标集群中的增量数据和备集群中的增量数据进行对比,以确定备集群中的数据是否全部迁移到目标集群。

[0066] 步骤S40,若完整,则确定所述原有集群中的数据已完全迁移至所述目标集群。

[0067] 当经过判断发现目标集群中的增量数据完整时,就可以直接确定原有集群中的数据已完全迁移至目标集群,可以通知业务系统直接无缝切换到新集群,正常提供服务。但是若目标集群中的增量数据不完整,则需要继续进行数据迁移操作。

[0068] 另外,为辅助理解本实施例中的数据迁移的原理的理解,下面进行举例说明。

[0069] 例如,如图5所示,数据迁移系统包括原有集群HBASE集群source和目标集群HBASE集群destination,及多zookeeper(分布式应用程序)协调集群,并且在采用enable_table_replication的形式同步表结构是,确定peer(身份)01,wallog(用户)或peer02等,并观察是否有相应tps(系统吞吐量)写入备集群,若无,则暂停同步,并对主集群中的replication机制进行修正,再清理目标集群中htable(线程)的hbase数据。再创建快照(即加锁(禁止更新、添加等操作)、Hdfs创建目录,并生成相应描述信息,再进行刷盘操作,即刷新memstore中的数据到hfile,并为hfile文件创建引用指针,释放表级锁)。再通过export导出快照到目标集群,并在目标集群中恢复快照,即先disable(禁用)表,再进行restore_snapshot(快照恢复),并自动启动enable表。并在快照数据导出后,验证存量快照数据,若快照数据完整,且备集群中存在增量数据,则回放增量wallog到目标集群,并验证增量数据。

[0070] 在本实施例中,通过将所述主集群中业务系统写入的数据实时复制到所述备集群,并基于所述备集群中的数据创建快照;将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;若完整,则确定所述原有集群中的数据已完全迁移至所述目标集群。通过原有集群划分为主集群和备集群,再将主集群中业务系统写入的数据实时复制到备集群,并对备集群创建快照,再将快照迁移到目标集群,从而可以避免影响到业务系统和主集群之间的服务交流,保障了在业务系统写入原有集群的数据不间断的情况下,又能将原有集群中的数据迁移到目标集群,并且在将快照复制到目标集群后,还会获取原有集群中的增量数据,并将增量数据迁移到目标集群,从而保障了将原有集群中的所有数据都迁移到目标集群,使得业务系统在原有集群将所有数据迁移到目标集群后,能无缝直接切换到目标集群,避免了现有技术中在进行数据迁移时,会造成业务系统停止服务的现象发生。

[0071] 进一步地,基于本发明数据迁移方法第一实施例,提出本发明数据迁移方法第二实施例。本实施例是本发明第一实施例的步骤S20,检测所述快照数据是否满足预设检测条件的步骤之后,包括:

[0072] 步骤a,获取所述备集群的存量数据,检测所述快照数据是否和所述存量数据匹配;

[0073] 在本实施例中,当将快照已迁移到目标集群,并进行快照恢复后,需要对恢复的快照数据进行检测,即检测备集群中的数据是否已全部迁移到目标集群。因此可以获取备集

群对应的数据仓库中的存量数据(即当前时刻备集群已存储的所有数据),并检测快照数据是否和存量数据匹配,再基于不同的检测结果执行不同的操作。

[0074] 步骤b,若所述快照数据和所述存量数据不匹配,则获取所述备集群中快照迁移前所在的历史文件,并获取所述目标集群中快照所在的线程文件,检测所述历史文件和所述现存文件是否匹配;

[0075] 当经过判断发现快照数据和存量数据不匹配,则获取备集群中快照在进行迁移前所在的hdfs文件(即历史文件)。并获取目标集群中快照所在的hdfs文件(即现存文件),再检测历史文件和现存文件是否匹配,并根据不同的检测结果执行不同的操作。但是若快照数据和存量数据匹配,则确定备集群中的所有数据已全部迁移到目标集群,此时还需要检测历史文件和现存文件是否匹配。

[0076] 步骤c,若所述历史文件和所述现存文件不匹配,则确定所述快照数据满足预设检测条件。

[0077] 当经过判断发现历史文件和现存文件不匹配时,则可以确定快照数据不满足预设检测条件。但是若历史文件和现存文件匹配,也确定快照数据不满足预设检测条件。也就是在本实施例中,只有当快照数据和存量数据匹配,且历史文件和现存文件也匹配时,才能够确定快照数据满足预设检测条件。若快照数据和存量数据不匹配,或者历史文件和现存文件也不匹配时,则确定快照数据不满足预设检测条件。

[0078] 在本实施例中,通过在备集群的存量数据和快照数据不匹配时,检测快照迁移前所在的历史文件和快照迁移后所在的现存文件是否匹配,并在匹配时,确定快照数据不满足预设检测条件,从而保障了迁移到目标集群中的快照数据的完整性。

[0079] 进一步地,将所述快照迁移到所述目标集群,并对所述目标集群中的快照进行数据恢复的步骤,包括:

[0080] 步骤d,控制所述备集群中具有所述数据的表处于禁止状态,并根据输入的迁移指令将所述快照迁移至所述目标集群,将所述快照中的数据恢复至所述目标集群的工作目录,将具有所述快照中的数据的工作目录作为快照数据。

[0081] 在本实施例中,当准备将备集群中的快照迁移至目标集群时,可以先控制备集群中具有数据的表处于禁止状态(即禁止备集群中的读操作或写操作)。并根据用户手动输入或根据其它设备输入的迁移指令将备集群中的快照迁移至目标集群,并且在迁移的过程中,exportSnaphost是底层采用distcp的形式传输hdfs文件,因此可以通过控制mappers和bandwidth来控制相应的hdfs数据传输速度。在将快照迁移至目标集群后,可以通过快照恢复技术将快照中的数据恢复至目标集群的工作目标,并将具有快照中的数据的工作目录作为快照数据。

[0082] 在本实施例中,通过进行快照迁移前,先控制备集群处于禁止状态,再根据迁移指令将快照迁移至目标集群,并进行数据恢复,得到快照数据,从而保障了获取到的快照数据的准确性。

[0083] 进一步地,检测所述目标集群中的增量数据是否完整的步骤,包括:

[0084] 步骤p,基于预设时间戳获取所述备集群的备集群数据,并获取所述目标集群的目标集群数据,校验所述备集群数据和所述目标集群数据之间的差异数据;

[0085] 在本实施例中,在对增量数据进行检测时,可以通过改进的WbVerifyReplication

MR工具验证增量数据完整性,通过提前制定时间范围进行验证备集群的数据(即备集群数据)和目标集群的数据(即目标集群数据)是否匹配。也就是通过预设时间戳来获取备集群中的备集群数据和目标集群的目标集群数据,然后再检测备集群数据和目标集群数据是否匹配。若不匹配,则可以确定备集群数据中存在各种无用数据、过期数据等,再获取备集群数据和目标集群数据之间的差异数据。

[0086] 步骤x,根据所述差异数据对所述备集群数据进行过滤,并检测所述过滤后的备集群数据是否和所述目标集群数据匹配;

[0087] 当获取到差异数据后,就可以对增量数据进行二次校验,即根据此差异数据对备集群数据进行过滤处理,以过滤掉备集群数据中的过期数据、无用数据等,然后再将过滤后的备集群数据和目标集群数据进行匹配,并基于不同的匹配结果执行不同的操作。需要说明的是,在进行二次校验时,会采用get的形式,并通过备集群与改进的WbVerifyReplication MR工具的接口和差异数据再次获取过滤后的备集群数据。

[0088] 步骤y,若匹配,则确定所述目标集群中的增量数据完整。

[0089] 当经过判断发现过滤后的备集群数据和目标集群数据匹配时,则确定目标集群中的增量数据完整。若不匹配,则确定目标集群中的增量数据不完整。另外,对增量数据进行两次校验的代码程序可以采用如下方式进行:

```

private boolean checkRowkeyInSourceAndPeer(Result row) {
    Get get = new Get(row.getRow());
    Result masterResult = null;
    Result slaveResult = null;
    try {
        masterResult = masterHtable.get(get);
        LOG.info("get master rowkey: " + Bytes.toString(row.getRow()));
    } catch (IOException e) {
        LOG.error("fail to get master rowkey =" +
[0090] Bytes.toString(row.getRow()), e);
    }
    try {
        slaveResult = slaveHtable.get(get);
        LOG.info("get slave rowkey: " + Bytes.toString(row.getRow()));
    } catch (IOException e) {
        LOG.error("fail to get peer slave rowkey=" +
[0091] Bytes.toString(row.getRow()), e);
    }
    if (masterResult != null && slaveResult != null) {
        try {
            Result.compareResults(masterResult, slaveResult);
            LOG.info("rowkey update after endtimestamp, check point: " +
                new SimpleDateFormat("yyyy-MM-dd
[0092] HH:mm:ss").format(new Date()) + ", rowkey:" + Bytes.toString(row.getRow()));
            return true;
        } catch (Exception e) {
            LOG.error("Exception while comparing row : " + e);
            return false;
        }
    }
    return false;
}

```

[0092] 在本实施例中,通过根据预设时间戳校验备集群和目标集群数据之间的差异数据,并根据差异数据对备集群数据进行过滤,并在过滤后的备集群数据和目标集群数据匹

配时,确定目标集群中的增量数据完整。

[0093] 进一步地,基于所述备集群中的数据创建快照的步骤,包括:

[0094] 步骤e,对所述备集群中的数据进行加锁,以获取加锁数据,将所述加锁数据刷盘到所述备集群中的hfile文件,并对具有所述加锁数据的hfile文件创建引用指针,根据所述引用指针创建快照。

[0095] 在本实施例中,可以通过snapshot快照创建source集群表快照,并通过对备集群中的数据进行加锁,得到加锁后的数据(即加锁数据),将加锁数据同步到备集群中的hfile文件(即将加锁数据刷盘到备集群中的hfile文件),再对具有加锁数据的hfile文件创建引用指针,并释放表级锁,以生成快照。即将SnapshotDescription对象序列化写入到hdfs://parent/hbase/.hbase-snapshot/.tmp/PhotoSnapshot目录的.snapshotinfo文件中,进而Photo表的最新的.tableinfo拷到hdfs://parent/hbase/.hbase-snapshot/.tmp/PhotoSnapshot/.tabledesc/目录下,名字命名为.tableinfo.0000000001,然后会通过从meta region中获取photo表的所有region信息和所在regionserver信息并直至结束,最后会调用completeSnapshot将真正工作目录working_dir改正hdfs://parent/hbase/.hbase-snapshot/PhotoSnapshot。至此快照建立完成。

[0096] 在本实施例中,通过对备集群中的数据进行加锁,刷盘,创建引用指针,并根据引用指针创建快照,从而保障了快照的准确性。

[0097] 进一步地,基于所述快照获取所述备集群中的增量数据的步骤,包括:

[0098] 步骤f,确定所述快照在所述备集群中所有数据的数据节点,根据所述数据节点对所述备集群中的所有数据进行回放,以获取所述备集群中的增量数据。

[0099] 在本实施例中,当将快照迁移到目标集群,且对快照的完整性检测成功后,还需要在备集群中确定快照在备集群中所有数据的数据节点(即可以根据获取快照的时间来确定进行快照操作的数据节点),并在确定数据节点后,确定自此数据节点之后是否有产生新的数据,并可以通过对备集群中的所有数据进行回放的方式,将数据节点之后产生的新的数据(即增量数据)迁移到目标集群。即可以按照相应的wallog格式标准将数据序列化顺序存放在hdfs上,并通过回放的方式在目标hbase集群将数据重写如集群,所以在source原有集群查找相应的主备replication peer,并enable(允许),回放相应堆积在source集群的wallog数据文件,相应wallog文件存储在hdfs上,具体的文件目录管理是存放在zookeeper中相应的peer目录下。

[0100] 在本实施例中,通过确定快照的数据节点,并根据数据节点进行回放,以获取增量数据,从而保障了获取到的增量数据的准确性。

[0101] 进一步地,将所述主集群中业务系统写入的数据实时复制到所述备集群的步骤,包括:

[0102] 步骤h,将所述主集群中的表结构复制到所述备集群,并检测所述备集群接收的表结构中是否存在业务系统写入的数据;

[0103] 在本实施例中,由于原有集群包括主集群和备集群,并且是将主集群的数据复制到备集群,并将备集群中的数据迁移到目标集群的,从而保障,业务系统和原有集群进行正常的服务的同时,又能完成原有集群和目标集群的数据迁移操作。因此,需要先将主集群中的表结构复制到备集群,并在业务系统将数据写入主集群时,会同步复制此数据到备集群

对应的表结构中。因此需要在主集群中的表结构迁移到备集群,且业务系统已开始往主集群中写入数据时,检测备集群中的表结构是否存在业务系统写入的数据,并根据不同的检测结果执行不同的操作。

[0104] 步骤m,若存在,则执行所述基于所述备集群中的数据创建快照的步骤;

[0105] 当经过判断发现备集群接收的表结构中存在业务系统写入的数据时,则确定主集群已成功将表结构复制到备集群,且成功复制到主集群中的数据,此时就可以执行后续的基于备集群中的数据创建快照的操作。

[0106] 步骤n,若不存在,则对所述主集群中的复制机制进行修正,并输出未复制成功的提示信息。

[0107] 若备集群中接收到的表结构中不存在业务系统写入的数据时,则需要检测主集群中的replication(复制)机制是否正常,若不正常,则对其进行修正,并输出未复制成功的提示信息。

[0108] 在本实施例中,通过在备集群中的表结构中不存在业务系统写入的数据时,对复制机制进行修正,并输出相应的提示信息,从而保障了备集群中能获取到准确的数据,方便后续的数据迁移操作。

[0109] 本发明还提供一种数据迁移装置,参照图3,所述数据迁移装置包括:

[0110] 复制模块A10,用于将主集群中业务系统写入的数据实时复制到备集群,并基于所述备集群中的数据创建快照;

[0111] 恢复模块A20,用于将所述快照迁移到目标集群,并对所述目标集群中的快照进行数据恢复,以获取快照数据,检测所述快照数据是否满足预设检测条件;

[0112] 检测模块A30,用于若不满足,则基于所述快照获取所述备集群中的增量数据,并将所述增量数据迁移至所述目标集群,检测所述目标集群中的增量数据是否完整;

[0113] 确定模块A40,用于若完整,则确定原有集群中的数据已完全迁移至所述目标集群。

[0114] 可选地,所述恢复模块A20,还用于:

[0115] 获取所述备集群的存量数据,检测所述快照数据是否和所述存量数据匹配;

[0116] 若所述快照数据和所述存量数据不匹配,则获取所述备集群中快照迁移前所在的历史文件,并获取所述目标集群中快照所在的线程文件,检测所述历史文件和所述现存文件是否匹配;

[0117] 若所述历史文件和所述现存文件不匹配,则确定所述快照数据不满足预设检测条件。

[0118] 可选地,所述恢复模块A20,还用于:

[0119] 控制所述备集群中具有所述数据的表处于禁止状态,并根据输入的迁移指令将所述快照迁移至所述目标集群,将所述快照中的数据恢复至所述目标集群的工作目录,将具有所述快照中的数据的工作目录作为快照数据。

[0120] 可选地,所述检测模块A30,还用于:

[0121] 基于预设时间戳获取所述备集群的备集群数据,并获取所述目标集群的目标集群数据,校验所述备集群数据和所述目标集群数据之间的差异数据;

[0122] 根据所述差异数据对所述备集群数据进行过滤,并检测所述过滤后的备集群数据

是否和所述目标集群数据匹配；

[0123] 若匹配,则确定所述目标集群中的增量数据完整。

[0124] 可选地,所述复制模块A10,还用于:

[0125] 对所述备集群中的数据进行加锁,以获取加锁数据,将所述加锁数据刷盘到所述备集群中的hfile文件,并对具有所述加锁数据的hfile文件创建引用指针,根据所述引用指针创建快照。

[0126] 可选地,所述检测模块A30,还用于:

[0127] 确定所述快照在所述备集群中所有数据的数据节点,根据所述数据节点对所述备集群中的所有数据进行回放,以获取所述备集群中的增量数据。

[0128] 可选地,所述复制模块A10,还用于:

[0129] 将所述主集群中的表结构复制到所述备集群,并检测所述备集群接收的表结构中是否存在业务系统写入的数据;

[0130] 若存在,则执行所述基于所述备集群中的数据创建快照的步骤;

[0131] 若不存在,则对所述主集群中的复制机制进行修正,并输出未复制成功的提示信息。

[0132] 上述各程序模块所执行的方法可参照本发明数据迁移方法各个实施例,此处不再赘述。

[0133] 本发明还提供一种计算机存储介质。

[0134] 本发明计算机存储介质上存储有数据迁移程序,所述数据迁移程序被处理器执行时实现如上所述的数据迁移方法的步骤。

[0135] 其中,在所述处理器上运行的数据迁移程序被执行时所实现的方法可参照本发明数据迁移方法各个实施例,此处不再赘述。

[0136] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者系统不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者系统所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者系统中还存在另外的相同要素。

[0137] 上述本发明实施例序号仅仅为了描述,不代表实施例的优劣。

[0138] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳实施方式。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如上所述的一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本发明各个实施例所述的方法。

[0139] 以上仅为本发明的优选实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

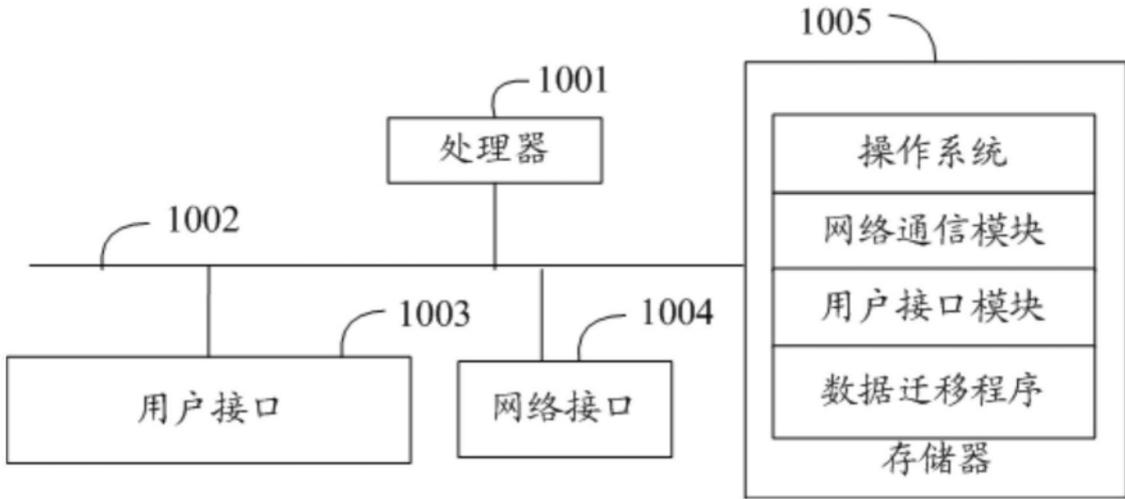


图1

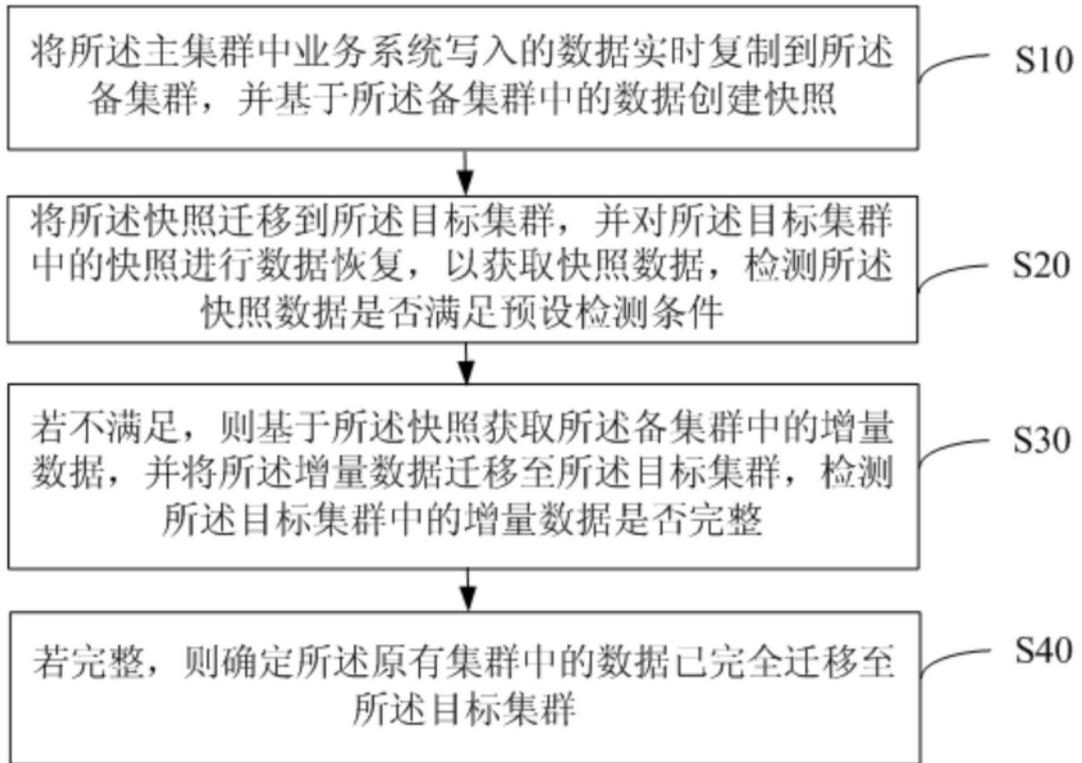


图2

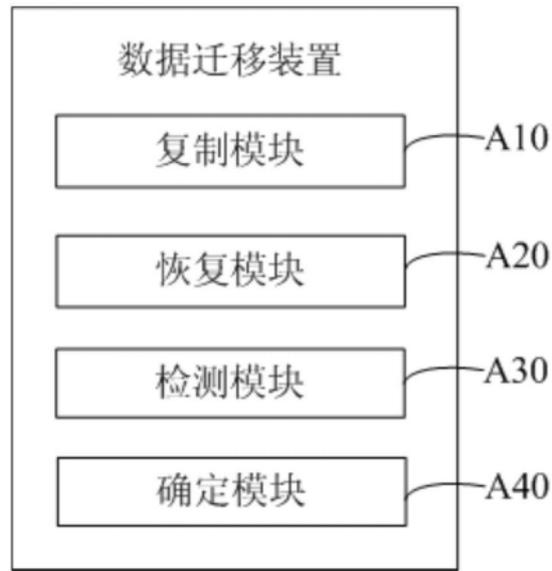


图3



图4

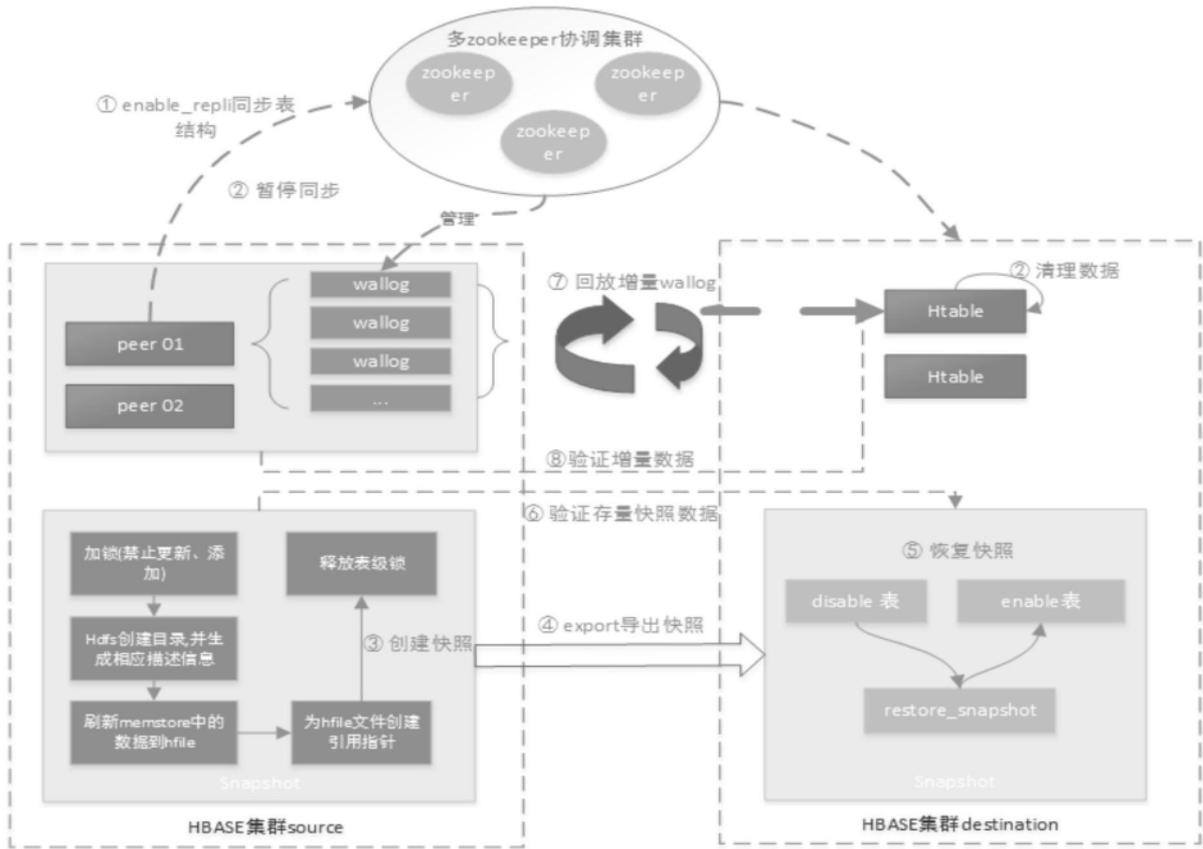


图5