



(12) 发明专利申请

(10) 申请公布号 CN 113760862 A

(43) 申请公布日 2021. 12. 07

(21) 申请号 202110077442.6

G06F 9/46 (2006.01)

(22) 申请日 2021.01.20

G06F 9/54 (2006.01)

(71) 申请人 北京沃东天骏信息技术有限公司
地址 100176 北京市大兴区北京经济技术
开发区科创十一街18号院2号楼4层
A402室

申请人 北京京东世纪贸易有限公司

(72) 发明人 张师聪

(74) 专利代理机构 北京品源专利代理有限公司
11332

代理人 孟金喆

(51) Int. Cl.

G06F 16/21 (2019.01)

G06F 16/25 (2019.01)

G06F 16/27 (2019.01)

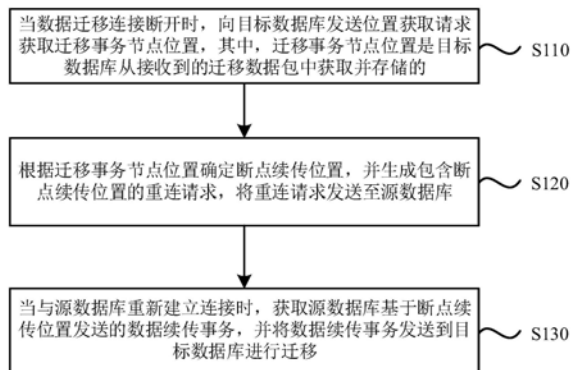
权利要求书2页 说明书9页 附图3页

(54) 发明名称

一种增量数据断点续传方法、装置、设备及存储介质

(57) 摘要

本发明实施例公开了一种增量数据断点续传方法、装置、设备及存储介质,所述方法包括:当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。本发明实施例提供的方法实现了数据库之间数据迁移时增量数据的断点续传,提高了数据迁移效率。



1. 一种增量数据断点续传方法,其特征在于,包括:

当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,所述迁移事务节点位置是所述目标数据库从接收到的迁移数据包中获取并存储的;

根据所述迁移事务节点位置确定断点续传位置,并生成包含所述断点续传位置的重连请求,将所述重连请求发送至源数据库;

当与所述源数据库重新建立连接时,获取所述源数据库基于所述断点续传位置发送的数据续传事务,并将所述数据续传事务发送到目标数据库进行迁移。

2. 根据权利要求1所述的方法,其特征在于,所述根据所述迁移事务节点位置确定断点续传位置,包括:

当所述源数据库故障时,将所述迁移事务节点位置作为所述断点续传位置。

3. 根据权利要求1所述的方法,其特征在于,所述根据所述迁移事务节点位置确定断点续传位置,包括:

当数据迁移工具故障时,获取所述源数据库中的当前迁移位置;

将所述当前迁移位置与所述迁移事务节点位置进行比对,当所述当前迁移位置在所述迁移事务节点位置之前时,将所述当前迁移位置作为所述断点续传位置。

4. 根据权利要求1所述的方法,其特征在于,还包括:

获取消息队列中的待迁移事务信息;

获取消息队列中末端待迁移事务在所述源数据库中的待迁移源位置;

基于所述待迁移源位置和所述待迁移事务信息生成迁移数据包;

将所述迁移数据包发送至目标数据库,以使所述目标数据库获取所述迁移数据包中的待迁移源位置,并将所述待迁移源位置作为迁移事务节点位置进行存储。

5. 根据权利要求4所述的方法,其特征在于,所述获取消息队列中末端待迁移事务在所述源数据库中的待迁移源位置,包括:

调用节点状态查看指令获取所述源数据库的二进制日志文件,从所述二进制日志文件中获取所述待迁移源位置。

6. 根据权利要求5所述的方法,其特征在于,所述待迁移源位置为所述末端待迁移事务在所述源数据库中的文件所在偏移量,或,所述末端待迁移事务在所述源数据库中的全局事务标识符;

所述全局事务标识符包括通用唯一识别码和线程控制符,所述通用位移识别码表示所述末端待迁移事务在所述源数据库中的编号,所述线程控制符表示所述末端待迁移事务提交至所述源数据库时所述源数据库为所述末端待迁移事务分配的序列号。

7. 根据权利要求4所述的方法,其特征在于,所述基于所述待迁移源位置和所述待迁移事务信息生成迁移数据包,包括:

将各所述待迁移事务信息拼装为结构化查询语言语句,将所述结构化查询语言语句和所述待迁移源位置添加至同一事务中,生成所述迁移数据包。

8. 一种增量数据断点续传装置,其特征在于,包括:

迁移事务节点位置模块,用于当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,所述迁移事务节点位置是所述目标数据库从接收到的迁移数据包中获取并存储的;

重连请求发送模块,用于根据所述迁移事务节点位置确定断点续传位置,并生成包含所述断点续传位置的重连请求,将所述重连请求发送至源数据库;

断点位置续传模块,用于当与所述源数据库重新建立连接时,获取所述源数据库基于所述断点续传位置发送的数据续传事务,并将所述数据续传事务发送到目标数据库进行迁移。

9. 一种计算机设备,其特征在于,所述设备包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序;

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1-7中任一所述的增量数据断点续传方法。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被处理器执行时实现如权利要求1-7中任一所述的增量数据断点续传方法。

一种增量数据断点续传方法、装置、设备及存储介质

技术领域

[0001] 本发明实施例涉及计算机技术领域,尤其涉及一种增量数据断点续传方法、装置、设备及存储介质。

背景技术

[0002] 随着互联网业务的不断发展,业务平台数据库的访问量不断增多、数据量持续累加,需要对现有的数据库进行扩容或者迁移。在扩容或者迁移过程中最关键的是数据库中数据的迁移。

[0003] 在实现本发明的过程中,发明人发现现有技术中至少存在以下技术问题:在数据同步工具进行数据同步的过程中,无法避免会遇到各种异常情况(源数据库故障、宕机、网络抖动、中断、迁移工具故障重启等),造成正在同步中的任务终止,无法继续进行同步。需要将目标库已同步的数据清空后,再重新做数据同步,如果碰到分库分表或者数据量十分庞大的情况将会大大增加人力和时间成本。

发明内容

[0004] 本发明实施例提供了一种增量数据断点续传方法、装置、设备及存储介质,以实现数据库中增量数据迁移时的断点续传。

[0005] 第一方面,本发明实施例提供了一种增量数据断点续传方法,包括:

[0006] 当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;

[0007] 根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;

[0008] 当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0009] 第二方面,本发明实施例还提供了一种增量数据断点续传装置,包括:

[0010] 迁移事务节点位置模块,用于当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;

[0011] 重连请求发送模块,用于根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;

[0012] 断点位置续传模块,用于当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0013] 第三方面,本发明实施例还提供了一种计算机设备,设备包括:一个或多个处理器;

[0014] 存储装置,用于存储一个或多个程序;

[0015] 当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现如本发

明任意实施例所提供的增量数据断点续传方法。

[0016] 第四方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如本发明任意实施例所提供的增量数据断点续传方法。

[0017] 本发明实施例通过当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移,通过目标数据库中已成功迁移的数据关联的迁移事务节点位置确定断点后的断点续传位置,以使源数据库基于断点续传位置确定续传数据并生成数据续传事务进行断点处的数据传输,实现了数据库之间数据迁移时增量数据的断点续传,提高了数据迁移效率。

附图说明

[0018] 图1是本发明实施例一所提供的一种增量数据断点续传方法的流程图;

[0019] 图2是本发明实施例二所提供的一种增量数据断点续传方法的流程示意图;

[0020] 图3是本发明实施例三所提供的一种增量数据断点续传装置的结构示意图;

[0021] 图4是本发明实施例四所提供的一种计算机设备的结构示意图。

具体实施方式

[0022] 下面结合附图和实施例对本发明作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释本发明,而非对本发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与本发明相关的部分而非全部结构。

[0023] 实施例一

[0024] 图1是本发明实施例一所提供的一种增量数据断点续传方法的流程图。本实施例可适用于数据库之间的数据迁移时的情形。该方法可以由增量数据断点续传装置执行,该增量数据断点续传装置可以采用软件和/或硬件的方式实现,例如,该增量数据断点续传装置可配置于计算机设备中。如图1所示,该方法包括:

[0025] S110、当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的。

[0026] 在本实施例中,通过增量数据断点续传装置将源数据库中的数据迁移至目标数据库中,即增量数据断点续传装置分别与源数据库和目标数据库建立了有效的通信连接。在数据传输时,源数据库将需要迁移的数据发送至增量数据断点续传装置,增量数据断点续传装置将接收到的数据发送至目标数据库,完成数据的迁移。

[0027] 需要说明的是,目标数据库在接收到增量数据断点续传位置发送的迁移数据包时,除将迁移数据包中的待迁移事务存储外,还会获取迁移数据包中的事务节点位置,并将迁移数据包中的事务节点位置作为已迁移成功的最新数据的位置进行存储,当源数据库与增量数据断点续传装置之间的数据迁移连接断开时,增量数据断点续传装置与目标数据库的连接未断开,增量数据断点续传装置可以向目标数据库发送位置获取请求,获取目标数据库中存储的位置信息作为迁移事务节点位置。

[0028] 一个实施例中,目标数据库中的位置存储方式可以为:获取消息队列中的待迁移事务信息;获取消息队列中末端待迁移事务在源数据库中的待迁移源位置;基于待迁移源位置和待迁移事务信息生成迁移数据包;将迁移数据包发送至目标数据库,以使目标数据库获取迁移数据包中的待迁移源位置,并将待迁移源位置作为迁移事务节点位置进行存储。可选的,增量数据断点续传装置获取源数据库发送的消息队列中的待迁移事务信息,调用源数据库接口获取消息队列中末端待迁移事务信息在源数据库中的位置信息作为待迁移源位置,将待提交事务信息和待迁移源位置添加至同一迁移数据包中,基于生成的迁移数据包生成待迁移事务发送至目标数据库,目标数据库接收到增量数据断点续传装置发送的待迁移事务后,将待迁移事务中迁移数据包中的事务进行存储,并将迁移数据包中的待迁移源位置作为迁移事务节点位置存储。可以理解的是,在将待迁移源位置进行存储时,需要将基于上一迁移数据包存储的迁移事务节点位置覆盖存储,以保证目标数据库中存储的迁移事务节点位置一直是最新存储的事务在源数据库中的位置。将待提交事务信息和待迁移源位置添加至同一事务进行提交,使得待提交事务信息和待迁移源位置的迁移状态统一,避免了待提交事务信息和待迁移源位置的迁移状态不统一导致的数据迁移丢失。

[0029] 具体的,基于待迁移源位置和待迁移事务信息生成迁移数据包,包括:将各待迁移事务信息拼装为结构化查询语言语句,将结构化查询语言语句和待迁移源位置添加至同一事务中,生成迁移数据包。示例性的,每读取一个待迁移事务,都需要将待迁移事务信息的Event事件中的信息拼装成具体的SQL语句,然后将各待迁移事务对应的SQL语句和待迁移源位置添加至同一事务中,得到迁移数据包。拼装的过程需要注意幂等性,需要将Insert类型的语句转换为Replace into类型的语句,Update类型和Delete类型的语句是幂等的,无需转换。

[0030] 可选的,获取消息队列中末端待迁移事务在源数据库中的待迁移源位置,包括:调用节点状态查看指令获取源数据库的二进制日志文件,从二进制日志文件中获取待迁移源位置。示例性的,可以通过show master status语句获取待迁移源位置。其中,待迁移源位置为末端待迁移事务在源数据库中的文件所在偏移量(File&Position),或,末端待迁移事务在源数据库中的全局事务标识符(Global Transaction ID,GTID)。其中,全局事务标识符包括通用唯一识别码和线程控制符,通用位移识别码表示末端待迁移事务在源数据库中的编号,线程控制符表示末端待迁移事务提交至源数据库时源数据库为末端待迁移事务分配的序列号,具体形式可以根据源数据库的版本确定。可选的,当源数据库为低版本时,待迁移源位置为File&Position模式,当源数据为高版本时,待提交位置为GTID模式。

[0031] S120、根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库。

[0032] 在本实施例中,获取增量数据断点续传装置中存储的迁移事务节点位置后,基于迁移事务节点位置确定断点续传位置确定断点续传位置,携带断点续传位置向源数据库发送重连请求。

[0033] 可选的,可以直接将迁移事务节点位置作为断点续传位置,也可以根据连接断开的类型判断是否需要获取其他位置信息作为参考位置信息以确定断点续传位置。一般的,数据迁移断开的原因可能为源数据库故障、增量数据断点续传装置故障以及目标数据库故障,不同故障端需要采用不同的方式确定断点续传位置。

[0034] 可选的,根据迁移事务节点位置确定断点续传位置,包括:当源数据库故障时,将迁移事务节点位置作为断点续传位置。当源数据故障时,断开连接后源数据无法向增量数据断点续传装置继续传输数据,因此可以直接将迁移事务节点位置作为断点续传位置。另外,当目标数据库故障时,不会影响源数据库与增量数据断点续传装置之间的数据传输,因此也可以直接将迁移事务节点位置作为断点续传位置。

[0035] 可选的,根据迁移事务节点位置确定断点续传位置,包括:当数据迁移工具故障时,获取源数据库中的当前迁移位置;将当前迁移位置与迁移事务节点位置进行比对,当当前迁移位置在迁移事务节点位置之前时,将当前迁移位置作为断点续传位置。当增量数据断点续传装置故障时,源数据库与增量数据断点续传装置之间的数据传输可能存在数据丢失的现象,因此不能直接将迁移事务节点位置作为断点续传位置,需要获取源数据库中的当前迁移位置,将当前迁移位置与迁移事务节点位置进行比对,判断是否有数据丢失。具体的,当当前迁移位置在迁移事务节点位置之前时,说明在故障发生前,源数据有一部分事务提交了但是还没有同步到增量数据断点续传装置中,增量数据断点续传装置也就没有解析到对应的坐标信息,这一部分数据可以理解为未知事务提交数据;为了不丢失这一部分数据,需要使用增量数据断点续传装置最后记录的位点坐标信息(即迁移事务节点位置)去重建连接。其中,源数据库中的当前迁移位置可以通过show master status语句获取。

[0036] S130、当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0037] 在本实施例中,当增量数据断点续传装置与源数据库重新建立连接后,源数据库获取重连请求中的断点续传位置,将与断点续传位置相邻,且位于断点续传位置之后的位置处的数据作为续传数据,基于续传数据生成数据续传事务,将数据续传事务发送至增量数据断点续传装置,增量数据断点续传装置接收到源数据库发送的数据续传事务后,将数据续传事务发送到目标数据库继续进行增量数据的断点续传。

[0038] 本发明实施例通过当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移,通过目标数据库中已成功迁移的数据关联的迁移事务节点位置确定断点后的断点续传位置,以使源数据库基于断点续传位置确定续传数据并生成数据续传事务进行断点处的数据传输,实现了数据库之间数据迁移时增量数据的断点续传,提高了数据迁移效率。

[0039] 实施例二

[0040] 图2是本发明实施例二所提供的一种增量数据断点续传方法的流程图。本实施例在上述方案的基础上,提供了一种优选实施例。

[0041] 本实施例提供的增量数据断点续传方法可以通过迁移工具实现。在本实施例中,迁移工具分别建立与源数据库和目标数据库的连接,在历史数据迁移完成后进行增量数据的迁移。整体来说,迁移工具通过读取源数据库中的二进制日志(Binlog)文件获取增量数据的坐标(File&Position或GTID),然后将该坐标数据持久化到目标数据库中;当迁移工具与源数据库连接断开重新建立连接后,重新从Binlog的该坐标位置读取增量数据。

[0042] 在本实施例中,断点续传的最小粒度为事务,其体现在Binlog中是多个Event事件的集合,在默认开启GTID的情况下会以GTIDEvent开始,XIDEvent结束,中间可能会有多个QueryEvent等组成;这里需要对每一个Event事件逐一解析并记录。

[0043] 在数据迁移任务开始同步增量数据时,执行show master status语句获取到File&Position的信息,同时通过读取GTIDEvent事件的值获取GTID信息,将上述信息整理后进行持久化,实现首次记录坐标,每读取到一个事务后,都需要将Event事件中的信息拼装成具体的SQL语句,然后到目标数据库去执行;拼装的过程需要保证幂等性,需要将Insert类型的语句转换为Replace into类型的语句。到目标数据库执行也需要启动一个新的事物,为了保证数据一致性,需要将转换后的SQL和已经获取到的数据在源数据库的位点坐标信息放在同一个事物中进行提交,保证提交状态的一致性。

[0044] 示例性的,假设执行到“最后一次已成功提交事务坐标”时,迁移工具和源数据库连接出现异常(具体异常为源数据库故障或者宕机),连接中断无法再继续同步数据;利用迁移工具自身的重试机制,尝试重新与源数据库建立连接,并携带已经持久化的Binlog位置的坐标(File&Position或GTID)”。

[0045] 示例性的,假设执行到“最后一次已成功提交事务坐标”时,迁移工具自身异常导致宕机重启,重新启动后的迁移工具优先检查该迁移任务曾经是否记录过同步位点坐标,若获取到“最后一次已成功提交事务坐标”(即迁移事务节点位置),则去源数据库上执行show master status语句获取源数据库当前位点坐标(即当前源位置),并对两个坐标进行比较。可以理解的是,假设迁移工具宕机重启的过程中,源数据库依然会有事务的提交,此时上述两个坐标会不一致,相差的部分认为是“未知事务提交”;此时,需要迁移工具在“最后一次已成功提交事务坐标”位置重建与源数据库的连接,并且由于从转换SQL到去目标库执行是幂等操作,就不会丢失这部分“未知事务提交”。可选的,迁移工具还可以增加线程定期获取源数据库“当前位点坐标”并与当前记录的坐标进行比较,结果相同则代表当前时间点增量数据同步完成。

[0046] 具体的,可以通过JTransfer迁移工具实现上述方法。图2是本发明实施例二所提供的一种增量数据断点续传方法的流程示意图。如图2所示,JTransfer首先通过第三方包gomysql向源数据库发送符合MySQL Binlog Dump协议的请求,建立源数据库与迁移工具之间的连接,连接建立成功后将自己伪装成源数据库的一个从库。源数据库就会不断向JTransfer发送标准的MySQL Binlog Event事件,JTransfer则顺序解析一个一个的Event事件,同时以事务的维度解析并记录同步位点坐标,一个事务是由多个Event事件组成的,例如:GTID Event->QUERY Event with“BEGIN”->...->XID Event|QUERY Event with“COMMIT”or“ROLLBACK”。当JTransfer解析到GTID Event和XID Event事件时会持久化当前事件里面的位点坐标信息到目标数据库中,为了提高效率可采用了异步方式持久化。

[0047] 同时为了保证事务的一致性,会将记录位点坐标信息的SQL和同一个事务中解析出需要到目标库应用的SQL放到同一个新的事物中去目标库提交;确保位点坐标信息和业务数据要么都提交成功,要么都失败。避免了成功记录位点坐标却没有业务数据造成丢失业务数据的情况,也避免了成功同步业务数据却没同步其位点坐标信息造成重复同步数据的情况。当源数据库故障或者JTransfer本身故障需要重新建立MySQL Binlog Dump连接时,会首先从JTransfer的元数据库中读取最后一次记录的同步坐标信息,同时到源数据库

获取当前源数据库的Binlog位点坐标信息,将两者进行比较;如果JTransfer记录的位点坐标信息早于源数据库的binlog位点坐标信息,说明在故障发生前,源数据库有一部分事务提交了但是还没有同步到JTransfer中去,JTransfer也就没有解析到对应的Event事件也就没有记录最新的位点坐标信息,这一部分数据即上文中的“未知事务提交”;为了不丢失这一部分数据,需要使用JTransfer最后记录的位点坐标信息去重建连接(依然利用gomysql三方包);如果两者位点坐标相同,则选取两者中的任意一个建立连接即可。

[0048] 本发明实施例借助源数据库本身的数据位点坐标,在同步增量数据时,逐一按单一事务维度解析Event事件获取当前数据位点坐标并按照特定规则进行持久化至目标数据库中;在数据迁移过程中遇到源数据库宕机连接中断或者迁移工具故障重启时,利用最后同步的数据所记录的位点坐标重建与源数据库的连接来达到断点续传的目的;避免了大数据量迁移遇到故障需要重新迁移的繁琐操作,节省了人力和时间成本。

[0049] 实施例三

[0050] 图3是本发明实施例四所提供的一种增量数据断点续传装置的结构示意图。该增量数据断点续传装置可以采用软件和/或硬件的方式实现,例如该增量数据断点续传装置可以配置于计算机设备中。如图3所示,该装置包括迁移事务节点位置模块310、重连请求发送模块320和断点位置续传模块330,其中:

[0051] 迁移事务节点位置模块310,用于当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;

[0052] 重连请求发送模块320,用于根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;

[0053] 断点位置续传模块330,用于当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0054] 本发明实施例通过迁移事务节点位置模块当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;重连请求发送模块根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;重连请求发送模块当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移,通过目标数据库中已成功迁移的数据关联的迁移事务节点位置确定断点后的断点续传位置,以使源数据库基于断点续传位置确定续传数据并生成数据续传事务进行断点处的数据传输,实现了数据库之间数据迁移时增量数据的断点续传,提高了数据迁移效率。

[0055] 可选的,在上述方案的基础上,重连请求发送模块320具体用于:

[0056] 当源数据库故障时,将迁移事务节点位置作为断点续传位置。

[0057] 可选的,在上述方案的基础上,重连请求发送模块320具体用于:

[0058] 当数据迁移工具故障时,获取源数据库中的当前迁移位置;

[0059] 将当前迁移位置与迁移事务节点位置进行比对,当当前迁移位置在迁移事务节点位置之前时,将当前迁移位置作为断点续传位置。

[0060] 可选的,在上述方案的基础上,装置还包括迁移数据包发送模块,用于:

- [0061] 获取消息队列中的待迁移事务信息；
- [0062] 获取消息队列中末端待迁移事务在源数据库中的待迁移源位置；
- [0063] 基于待迁移源位置和待迁移事务信息生成迁移数据包；
- [0064] 将迁移数据包发送至目标数据库，以使目标数据库获取迁移数据包中的待迁移源位置，并将待迁移源位置作为迁移事务节点位置进行存储。
- [0065] 可选的，在上述方案的基础上，迁移数据包发送模块具体用于：
- [0066] 调用节点状态查看指令获取源数据库的二进制日志文件，从二进制日志文件中获取待迁移源位置。
- [0067] 可选的，在上述方案的基础上，待迁移源位置为位点坐标或全局事务标识符。
- [0068] 可选的，在上述方案的基础上，迁移数据包发送模块具体用于：
- [0069] 将各待迁移事务信息拼装为结构化查询语言语句，将结构化查询语言语句和待迁移源位置添加至同一事务中，生成迁移数据包。
- [0070] 本发明实施例所提供的增量数据断点续传装置可执行本发明任意实施例所提供的增量数据断点续传方法，具备执行方法相应的功能模块和有益效果。
- [0071] 实施例四
- [0072] 图4是本发明实施例四所提供的一种计算机设备的结构示意图。图4示出了适于用来实现本发明实施方式的示例性计算机设备412的框图。图4显示的计算机设备412仅仅是一个示例，不应对本发明实施例的功能和使用范围带来任何限制。
- [0073] 如图4所示，计算机设备412以通用计算设备的形式表现。计算机设备412的组件可以包括但不限于：一个或者多个处理器416，系统存储器428，连接不同系统组件（包括系统存储器428和处理器416）的总线418。
- [0074] 总线418表示几类总线结构中的一种或多种，包括存储器总线或者存储器控制器，外围总线，图形加速端口，处理器416或者使用多种总线结构中的任意总线结构的局域总线。举例来说，这些体系结构包括但不限于工业标准体系结构（ISA）总线，微通道体系结构（MAC）总线，增强型ISA总线、视频电子标准协会（VESA）局域总线以及外围组件互连（PCI）总线。
- [0075] 计算机设备412典型地包括多种计算机系统可读介质。这些介质可以是任何能够被计算机设备412访问的可用介质，包括易失性和非易失性介质，可移动的和不可移动的介质。
- [0076] 系统存储器428可以包括易失性存储器形式的计算机系统可读介质，例如随机存取存储器（RAM）430和/或高速缓存存储器432。计算机设备412可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例，存储装置434可以用于读写不可移动的、非易失性磁介质（图4未显示，通常称为“硬盘驱动器”）。尽管图4中未示出，可以提供用于对可移动非易失性磁盘（例如“软盘”）读写的磁盘驱动器，以及对可移动非易失性光盘（例如CD-ROM，DVD-ROM或者其它光介质）读写的光盘驱动器。在这些情况下，每个驱动器可以通过一个或者多个数据介质接口与总线418相连。存储器428可以包括至少一个程序产品，该程序产品具有一组（例如至少一个）程序模块，这些程序模块被配置以执行本发明各实施例的功能。
- [0077] 具有一组（至少一个）程序模块442的程序/实用工具440，可以存储在例如存储器

428中,这样的程序模块442包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块442通常执行本发明所描述的实施例中的功能和/或方法。

[0078] 计算机设备412也可以与一个或多个外部设备414(例如键盘、指向设备、显示器424等)通信,还可与一个或者多个使得用户能与该计算机设备412交互的设备通信,和/或与使得该计算机设备412能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口422进行。并且,计算机设备412还可以通过网络适配器420与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共网络,如因特网)通信。如图所示,网络适配器420通过总线418与计算机设备412的其它模块通信。应当明白,尽管图中未示出,可以结合计算机设备412使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0079] 处理器416通过运行存储在系统存储器428中的程序,从而执行各种功能应用以及数据处理,例如实现本发明实施例所提供的增量数据断点续传方法,该方法包括:

[0080] 当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;

[0081] 根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;

[0082] 当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0083] 当然,本领域技术人员可以理解,处理器还可以实现本发明任意实施例所提供的增量数据断点续传方法的技术方案。

[0084] 实施例五

[0085] 本发明实施例五还提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现本发明实施例所提供的增量数据断点续传方法,该方法包括:

[0086] 当数据迁移连接断开时,向目标数据库发送位置获取请求获取迁移事务节点位置,其中,迁移事务节点位置是目标数据库从接收到的迁移数据包中获取并存储的;

[0087] 根据迁移事务节点位置确定断点续传位置,并生成包含断点续传位置的重连请求,将重连请求发送至源数据库;

[0088] 当与源数据库重新建立连接时,获取源数据库基于断点续传位置发送的数据续传事务,并将数据续传事务发送到目标数据库进行迁移。

[0089] 当然,本发明实施例所提供的一种计算机可读存储介质,其上存储的计算机程序不限于如上的方法操作,还可以执行本发明任意实施例所提供的增量数据断点续传方法的相关操作。

[0090] 本发明实施例的计算机存储介质,可以采用一个或多个计算机可读的介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。计算机可读存储介质例如可以是一—但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机存取存储器(RAM)、只读存储器

(ROM)、可擦式可编程只读存储器 (EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本文件中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

[0091] 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0092] 计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括——但不限于无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0093] 可以以一种或多种程序设计语言或其组合来编写用于执行本发明操作的计算机程序代码,程序设计语言包括面向对象的程序设计语言——诸如Java、Smalltalk、C++,还包括常规的过程式程序设计语言——诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络——包括局域网 (LAN) 或广域网 (WAN) ——连接到用户计算机,或者,可以连接到外部计算机 (例如利用因特网服务提供商来通过因特网连接)。

[0094] 注意,上述仅为本发明的较佳实施例及所运用技术原理。本领域技术人员会理解,本发明不限于这里的特定实施例,对本领域技术人员来说能够进行各种明显的变化、重新调整和替代而不会脱离本发明的保护范围。因此,虽然通过以上实施例对本发明进行了较为详细的说明,但是本发明不仅仅限于以上实施例,在不脱离本发明构思的情况下,还可以包括更多其他等效实施例,而本发明的范围由所附的权利要求范围决定。

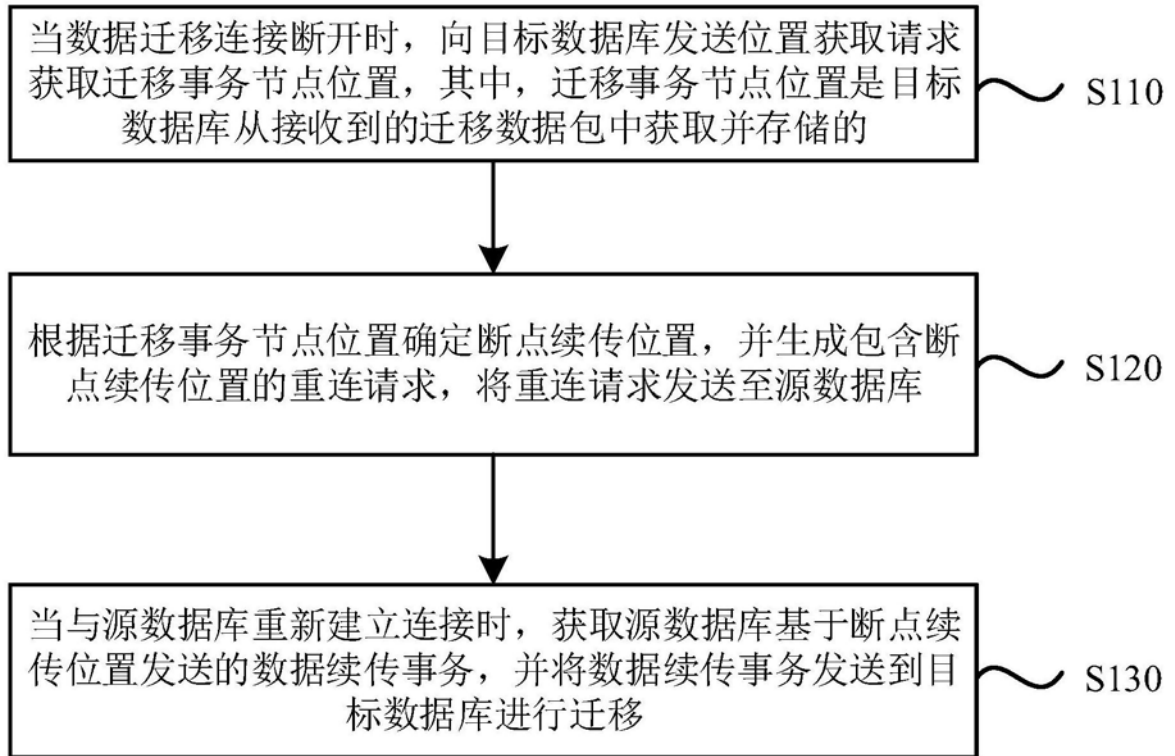


图1

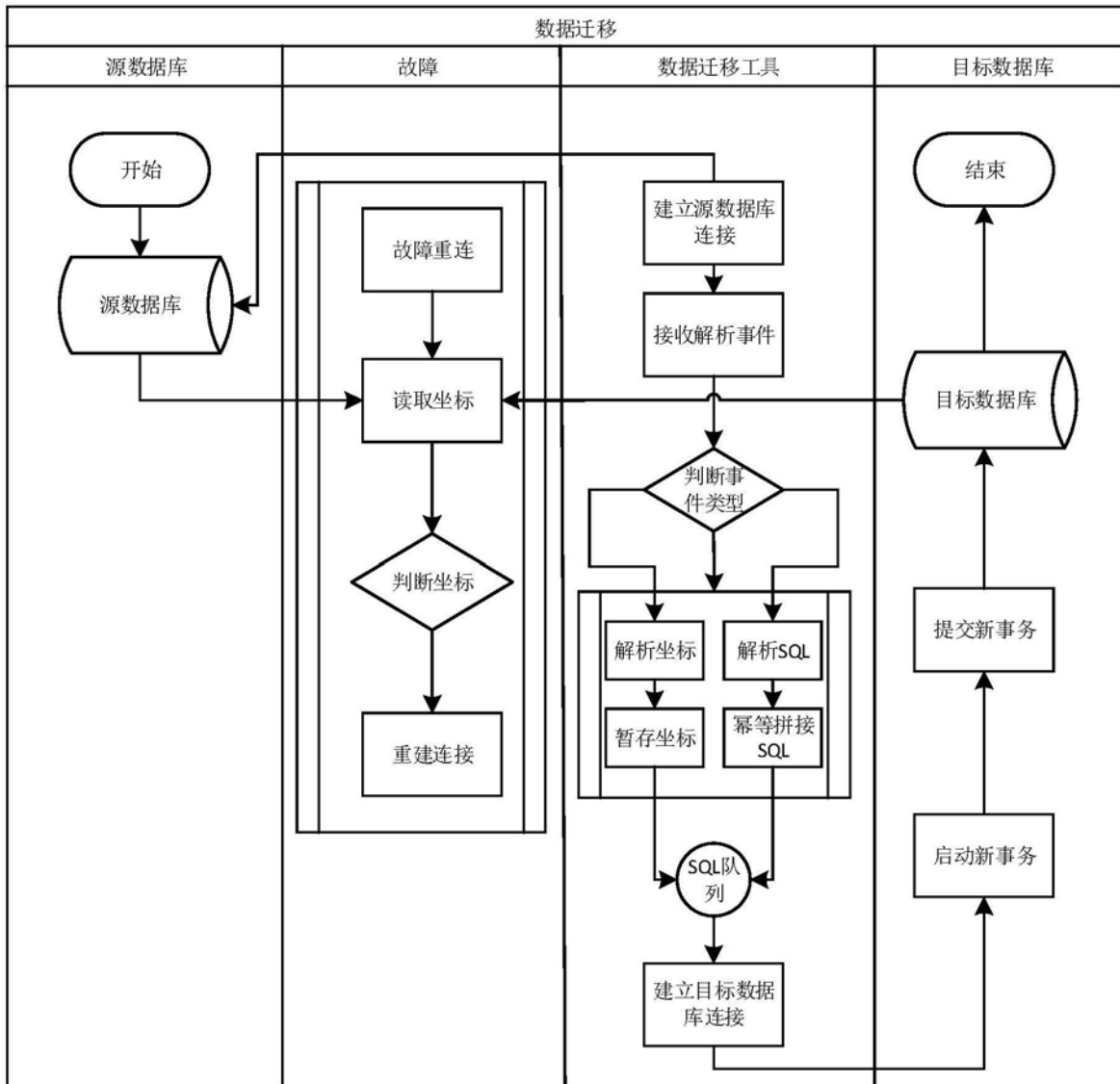


图2

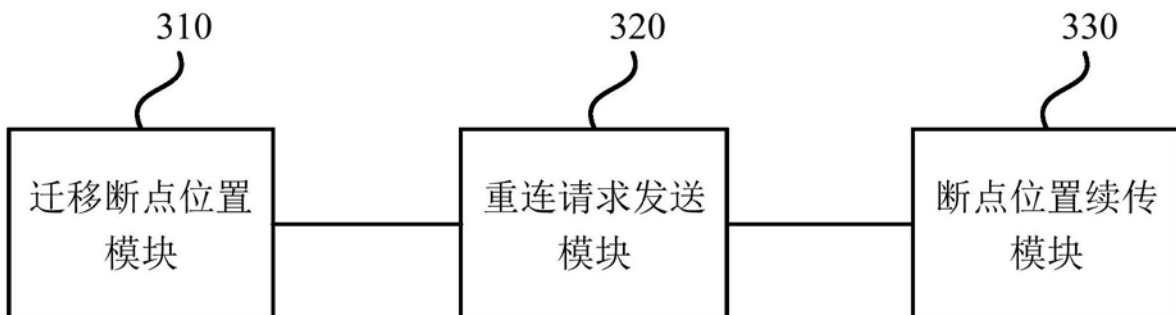


图3

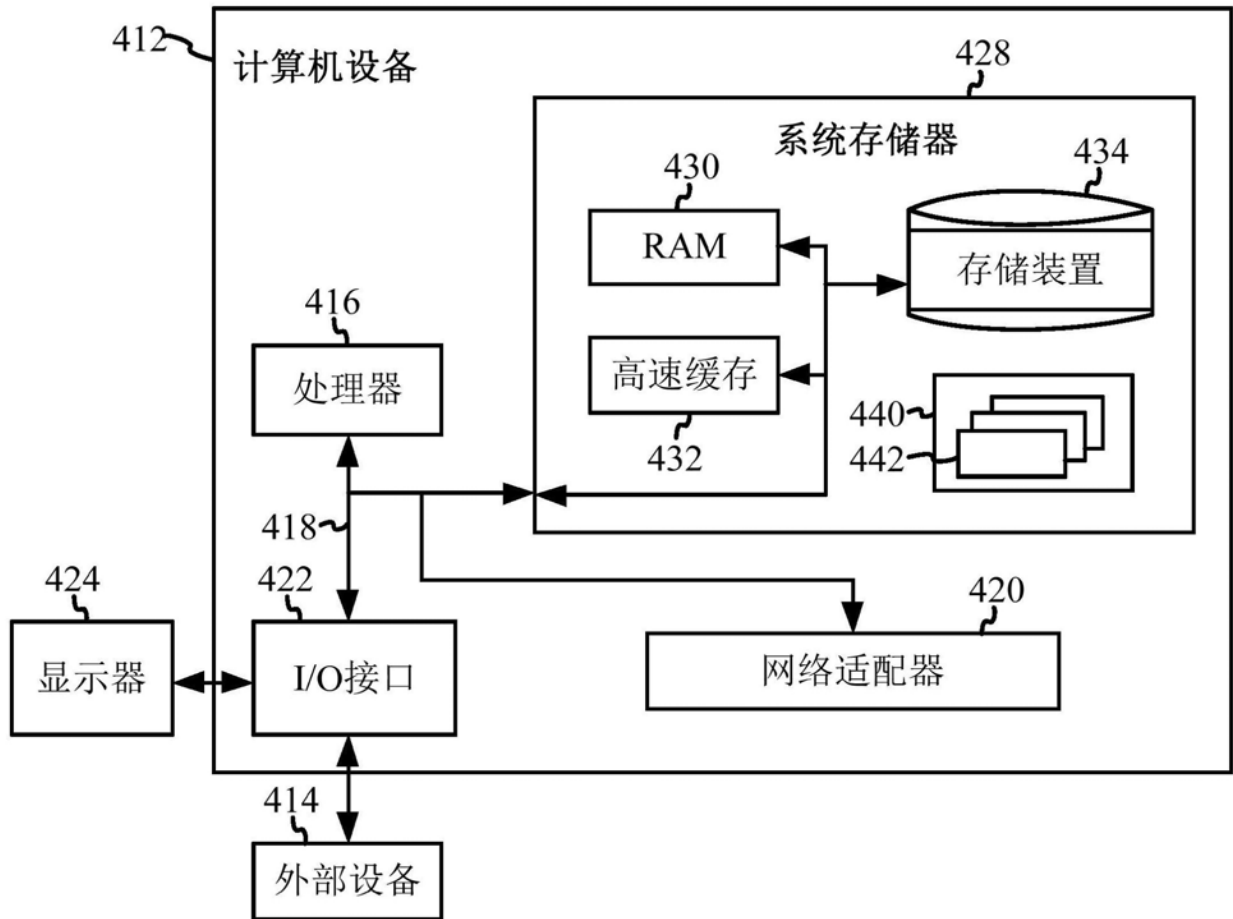


图4