



(12) 发明专利申请

(10) 申请公布号 CN 115348160 A

(43) 申请公布日 2022. 11. 15

(21) 申请号 202210829799.X

(22) 申请日 2022.07.15

(71) 申请人 深圳手回科技集团有限公司

地址 518000 广东省深圳市福田区梅林街道孖岭社区凯丰路10号翠林大厦20层01室

(72) 发明人 李鉴庭

(74) 专利代理机构 深圳中一联合知识产权代理有限公司 44414

专利代理师 任敏

(51) Int. Cl.

H04L 41/069 (2022.01)

H04L 47/62 (2022.01)

H04L 49/103 (2022.01)

H04L 49/25 (2022.01)

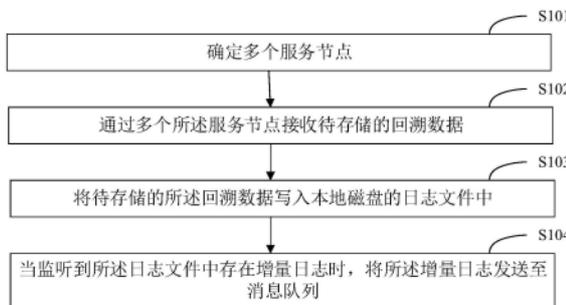
权利要求书2页 说明书10页 附图6页

(54) 发明名称

回溯数据存储方法、装置及计算机设备

(57) 摘要

本申请实施例适用于数据存储技术领域,提供了一种回溯数据存储方法、装置及计算机设备,所述方法可以应用于回溯数据服务端,包括:确定多个服务节点;通过多个所述服务节点接收待存储的回溯数据;将待存储的所述回溯数据写入本地磁盘的日志文件中;当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。采用上述方法,可以实现回溯数据的快速上传,并保证回溯数据的高可用性,满足对回溯数据实时检索的需求。



1. 一种回溯数据存储方法,其特征在于,应用于回溯数据服务端,所述方法包括:
 - 确定多个服务节点;
 - 通过多个所述服务节点接收待存储的回溯数据;
 - 将待存储的所述回溯数据写入本地磁盘的日志文件中;
 - 当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。
2. 根据权利要求1所述的方法,其特征在于,所述确定多个服务节点,包括:
 - 获取待存储的所述回溯数据的数据总量;
 - 根据所述数据总量确定多个已配置的所述服务节点的上报平均值;
 - 判断所述上报平均值是否大于设定阈值;
 - 若所述上报平均值大于所述设定阈值,则增加至少一个所述服务节点;
 - 若所述上报平均值小于所述设定阈值,则减少至少一个所述服务节点。
3. 根据权利要求2所述的方法,其特征在于,所述减少至少一个所述服务节点,包括:
 - 确定待减少的所述服务节点的数量;
 - 计算减少所述数量的所述服务节点后,剩余所述服务节点的上报平均值是否仍小于所述设定阈值,其中剩余的所述服务节点的数量大于或等于所述服务节点的最小配置数量;
 - 若剩余的所述服务节点的上报平均值仍小于所述设定阈值,则减少所述数量的所述服务节点;否则,调整待减少的所述服务节点的数量,重复执行计算减少所述数量的所述服务节点后,剩余的所述服务节点的上报平均值是否仍小于所述设定阈值的步骤。
4. 根据权利要求1所述的方法,其特征在于,所述当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,包括:
 - 监听所述日志文件中是否存在所述增量日志;
 - 若所述日志文件中存在所述增量日志,则通过轮询和事件并存的方式获取所述增量日志;
 - 将所述增量日志发送至所述消息队列。
5. 一种回溯数据存储方法,其特征在于,应用于检索数据库,所述方法包括:
 - 订阅消息队列,所述消息队列与远端的回溯数据服务端关联;
 - 若所述消息队列中存在增量日志,则获取并存储所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;
 - 定期清空所述检索数据库中的历史数据。
6. 一种回溯数据存储方法,其特征在于,应用于对象存储服务器,所述方法包括:
 - 订阅消息队列,所述消息队列与远端的回溯数据服务端关联;
 - 若所述消息队列中存在增量日志,则获取并压缩所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;
 - 对压缩后的所述增量日志进行存储;
 - 将已存储的所述增量日志写入物理盘中。

7. 根据权利要求6所述的方法,其特征在于,所述将已存储的所述增量日志写入物理盘中,包括:

通过异地复制技术,将已存储的所述增量日志同时写入多个不同的所述物理盘中,多个不同的所述物理盘位于不同的地域。

8. 一种回溯数据存储装置,其特征在于,应用于回溯数据服务端,所述装置包括:

服务节点确定模块,用于确定多个服务节点;

回溯数据接收模块,用于通过多个所述服务节点接收待存储的回溯数据;

回溯数据缓存模块,用于将待存储的所述回溯数据写入本地磁盘的日志文件中;

回溯数据存储模块,用于当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。

9. 一种计算机设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求1-7任一项所述的回溯数据存储方法。

10. 一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1-7任一项所述的回溯数据存储方法。

回溯数据存储方法、装置及计算机设备

技术领域

[0001] 本申请实施例属于数据存储技术领域，特别是涉及一种回溯数据存储方法、装置及计算机设备。

背景技术

[0002] 在网上保险交易过程中，为了保障用户利益和解决交易纠纷，网站前端会记录用户使用网站过程中的交互数据用于回溯交互过程，记录下来的数据便是回溯数据。通过在服务器回放回溯数据中存储的样式变更信息可以形成回溯视频，生成的回溯视频可以用于解决交易纠纷、在线同屏指导交易等场景。因此对回溯数据的快速上传和实时查询有一定的要求。

[0003] 在现有技术中，回溯数据多直接存储在数据库中，但数据库存在容量有限、维护成本高以及存储海量数据时检索速度慢等问题。利用数据库进行存储，无法满足对保险交易过程中产生的海量回溯数据进行快速存储和实时检索的需求。

发明内容

[0004] 有鉴于此，本申请实施例提供了一种回溯数据存储方法、装置及计算机设备，用以满足对保险交易过程中产生的海量回溯数据进行快速存储和实时检索的需求。

[0005] 本申请实施例的第一方面提供了一种回溯数据存储方法，应用于回溯数据服务端，所述方法包括：

[0006] 确定多个服务节点；

[0007] 通过多个所述服务节点接收待存储的回溯数据；

[0008] 将待存储的所述回溯数据写入本地磁盘的日志文件中；

[0009] 当监听到所述日志文件中存在增量日志时，将所述增量日志发送至消息队列，所述消息队列用于被远端的检索数据库和对象存储服务器订阅，以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。

[0010] 本申请实施例的第二方面提供了一种回溯数据存储方法，应用于检索数据库，所述方法包括：

[0011] 订阅消息队列，所述消息队列与远端的回溯数据服务端关联；

[0012] 若所述消息队列中存在增量日志，则获取并存储所述增量日志，所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时，将所述增量日志发送至所述消息队列的；

[0013] 定期清空所述检索数据库中的历史数据。

[0014] 本申请实施例的第三方面提供了一种回溯数据存储方法，应用于对象存储服务器，所述方法包括：

[0015] 订阅消息队列，所述消息队列与远端的回溯数据服务端关联；

[0016] 若所述消息队列中存在增量日志，则获取并压缩所述增量日志，所述增量日志是

所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;

[0017] 对压缩后的所述增量日志进行存储;

[0018] 将已存储的所述增量日志写入物理盘中。

[0019] 本申请实施例的第四方面提供了一种回溯数据存储装置,应用于回溯数据服务端,所述装置包括:

[0020] 服务节点确定模块,用于确定多个服务节点;

[0021] 回溯数据接收模块,用于通过多个所述服务节点接收待存储的回溯数据;

[0022] 回溯数据缓存模块,用于将待存储的所述回溯数据写入本地磁盘的日志文件中;

[0023] 回溯数据存储模块,用于当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。

[0024] 本申请实施例的第五方面提供了一种回溯数据存储装置,应用于检索数据库,所述装置包括:

[0025] 订阅模块,用于订阅消息队列,所述消息队列与远端的回溯数据服务端关联;

[0026] 增量日志存储模块,用于若所述消息队列中存在增量日志,则获取并存储所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;

[0027] 数据清空模块,用于定期清空所述检索数据库中的历史数据。

[0028] 本申请实施例的第六方面提供了一种回溯数据存储装置,应用于对象存储服务器,所述装置包括:

[0029] 订阅模块,用于订阅消息队列,所述消息队列与远端的回溯数据服务端关联;

[0030] 增量日志获取模块,用于若所述消息队列中存在增量日志,则获取并压缩所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;

[0031] 增量日志存储模块,用于对压缩后的所述增量日志进行存储;

[0032] 增量日志写入模块,用于将已存储的所述增量日志写入物理盘中。

[0033] 本申请实施例的第七方面提供了一种计算机设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现如上述第一方面、第二方面和/或第三方面所述的回溯数据存储方法。

[0034] 本申请实施例的第八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存储有计算机程序,所述计算机程序被处理器执行时实现如上述第一方面、第二方面和/或第三方面所述的回溯数据存储方法。

[0035] 本申请实施例的第九方面提供了一种计算机程序产品,当所述计算机程序产品在计算机设备上运行时,使得所述计算机设备执行上述第一方面、第二方面和/或第三方面所述的回溯数据存储方法。

[0036] 与现有技术相比,本申请实施例具有以下优点:

[0037] 在本申请实施例中,在通过服务节点将待存储的回溯数据写入回溯数据服务端的本地磁盘的日志文件后,新写入的回溯数据可以作为日志文件中的增量日志被缓存至消息

队列,检索数据库和对象存储服务器通过订阅消息队列可以获取上述增量日志,从而实现回溯数据的存储。存储至检索数据库中的回溯数据可以满足实时检索的需求,存储至对象存储服务器的回溯数据可以被长期存储。由于在回溯数据的上报过程中可以通过容器集群管理系统管理服务节点数量,且回溯数据上报后直接写入本地磁盘,因此不需要依赖外部条件便可实现大规模回溯数据的快速上报,实现水平扩容。此外本申请实施例通过实时采集新增日志,并将采集到的新增日志通过分布式消息队列写入检索数据库中,实现了短期热数据的实时检索。最终,本申请实施例将数据压缩后写入机械硬盘,确保数据可以长期存储,降低了数据的存储成本。

附图说明

[0038] 为了更清楚地说明本申请实施例中的技术方案,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的介绍。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。

[0039] 图1是本申请实施例提供的一种回溯数据存储方法的示意图;

[0040] 图2是本申请实施例提供的一种回溯数据存储方法中S101的一种实现方式的示意图;

[0041] 图3是本申请实施例提供的另一种回溯数据存储方法的示意图;

[0042] 图4是本申请实施例提供的又一种回溯数据存储方法的示意图;

[0043] 图5是本申请实施例提供的一种将回溯数据存储至回溯数据服务端的过程示意图;

[0044] 图6是本申请实施例提供的一种将回溯数据存储至对象存储服务器和检索数据库的过程示意图;

[0045] 图7是本申请实施例提供的一种回溯数据存储装置的示意图;

[0046] 图8是本申请实施例提供的另一种回溯数据存储装置的示意图;

[0047] 图9是本申请实施例提供的又一种回溯数据存储装置的示意图;

[0048] 图10是本申请实施例提供的一种计算机设备的示意图。

具体实施方式

[0049] 以下描述中,为了说明而不是为了限定,提出了诸如特定系统结构、技术之类的具体细节,以便透彻理解本申请实施例。然而,本领域技术人员应当清楚,在没有这些具体细节的其他实施例中也可以实现本申请。在其他情况中,省略对众所周知的系统、装置、电路以及方法的详细说明,以免不必要的细节妨碍本申请的描述。

[0050] 下面通过具体实施例来说明本申请的技术方案。

[0051] 参照图1,示出了本申请实施例提供的一种回溯数据存储方法的示意图,该方法应用于回溯数据服务端,具体可以包括如下步骤:

[0052] S101、确定多个服务节点。

[0053] 需要说明的是,本申请实施例中的服务节点可以用于为待存储的回溯数据提供回溯数据服务端的访问入口。

[0054] 本申请实施例中的回溯数据可以是一种结构化的json数据,这些结构化的json数据中可以存储样式变更等信息,通过服务器可以回放回溯数据中存储的样式变更信息从而形成回溯视频。由于得到的回溯视频展示了用户使用网站进行交易的完整交互过程和展示快照,因此,由回溯数据得到的回溯视频可以用于解决交易纠纷和用于在线同屏指导用户进行操作。

[0055] 在一些示例中,回溯数据可以是在用户进行网上保险交易过程所生成的,也可以是在用户网上购物过程所生成的,即本申请实施例中的回溯数据可以是保险交易回溯数据,也可以是网购回溯数据或其他类型的回溯数据,本申请实施例对回溯数据的具体类型不作限定。

[0056] 在本申请实施例中,回溯数据通过服务节点上报到服务端,因此可以根据回溯数据的上报量来确定需要的服务节点。在确定回溯数据的上报量后,可以通过容器集群管理系统,如Kubernetes技术,来实现服务节点的动态伸缩。

[0057] 在本申请实施例的一种可能的实现方式中,如图2所示,S101中确定多个服务节点具体可以包括如下步骤S1011-S1015:

[0058] S1011、获取待存储的所述回溯数据的数据总量。

[0059] 在本申请实施例中,可以通过定时统计需要上报的回溯数据的方式来获取单位时间内待存储的回溯数据的数据总量

[0060] S1012、根据所述数据总量确定多个已配置的所述服务节点的上报平均值。

[0061] 在本申请实施例中,已配置的服务节点可以是当前用于回溯数据上报服务的服务节点。可以通过单位时间内待存储的回溯数据的数据总量除以现有已配置的服务节点的数量的方式来得到已配置的服务节点的上报平均值。

[0062] S1013、判断所述上报平均值是否大于设定阈值。

[0063] 在本申请实施例中,设定阈值可以是根据实际需要任意设定的一个数值。在一个示例中,设定阈值可以为服务节点的最佳上报量,即当服务节点的上报量超过该设定阈值时,将回溯数据的上报到服务器端的速度会下降。用户在浏览网站的过程中会产生大量的回溯数据,因此为了实现对回溯数据的实时查询功能,需要将网站前端采集到的回溯数据快速上传到服务端。通过将服务节点的上报平均值维持在设定阈值之下,可以实现回溯数据的快速上报。

[0064] 若上报平均值大于设定阈值,可以执行S1014,增加至少一个服务节点;若上报平均值小于设定阈值,可以执行S1015,减少至少一个服务节点。

[0065] S1014、增加至少一个所述服务节点。

[0066] 在本申请实施例中,当已配置服务节点的上报平均值大于设定的服务节点上报阈值时,可以通过容器集群管理系统增加服务节点来减少每个服务节点的上报平均值。服务节点增加的数量可以根据具体的回溯数据上报数量来确定,若增加服务节点后,每个服务节点的上报平均值仍大于设定的服务节点上报阈值,则可以继续增加服务节点,直至每个服务节点的上报平均值小于或等于设定的服务节点上报阈值。

[0067] 在本申请实施例中,若判断发现当前已配置服务节点的上报平均值等于设定阈值,即代表当前已配置的服务节点处于最佳上报量,则无需调整服务节点数量。

[0068] S1015、减少至少一个所述服务节点。

[0069] 在本申请实施例中,若判断发现当前已配置服务节点的上报平均值小于设定阈值,则可能存在服务节点冗余的情况,可能需要减少服务节点数量。因此,当发现当前已配置服务节点上报平均值小于设定阈值时,需要首先判断服务节点是否冗余。

[0070] 在本申请实施例中,减少服务节点后,剩余的服务节点的数量需要大于或等于服务节点的最小配置数量。若服务节点的数量已经是服务节点的最小配置数量,即服务节点不存在冗余的情况,则即使判断发现当前已配置的服务节点的上报平均值小于设定的服务节点上报阈值也无需减少服务节点的数量。

[0071] 在本申请实施例中,可以通过计算减少特定数量的服务节点后剩余服务节点的上报平均值是否小于设定阈值来确定需要减少的服务节点的数量。在本申请实施例中,通过将单位时间内待存储的回溯数据的数据总量除以剩余的服务节点数量,可以得到剩余服务节点的上报平均值。若得到的剩余服务节点的上报平均值小于或等于设定阈值,且减少后剩余服务节点的数量大于或等于服务节点的最小配置数量,则服务节点存在冗余,可以减少该特定数量的服务节点。具体的待减少的服务节点的特定数量可由开发人员依据实际情况设定。若得到的剩余服务节点的上报平均值少于设定阈值,但减少后剩余服务节点的数量小于服务节点的最小配置数量,则将服务节点数量减少至服务节点的最小配置数量。若得到的剩余服务节点的上报平均值大于设定阈值,则可以将剩余的服务节点数量加一后,再次计算剩余服务节点的上报平均值,直至得到的剩余服务节点上报平均值满足小于或等于设定阈值的条件且减少后剩余服务节点的数量大于或等于服务节点的最小配置数量,那么此时减去的服务节点的数量则为已配置服务节点的实际冗余量,则减去已配置服务节点中冗余的服务节点。

[0072] 在本申请实施例中,在回溯数据服务端通过容器集群管理系统对服务节点数量进行动态伸缩,可以保证回溯数据的上报过程不需要依赖外部条件便可进行,即仅通过回溯数据服务端自身便可实现回溯数据的上报。此外,由于容器集群管理系统可以通过调整服务节点的数量使得已配置服务节点上报平均值始终接近设定的服务节点上报阈值,因此可以将大量回溯数据快速上传到回溯数据服务端。

[0073] S102、通过多个所述服务节点接收待存储的回溯数据。

[0074] 在本申请实施例中,调整好服务节点数量,确定各个服务节点负载均衡后,回溯数据服务端可以通过服务节点接收待存储的回溯数据。

[0075] S103、将待存储的所述回溯数据写入本地磁盘的日志文件中。

[0076] 回溯数据服务端在接收到服务节点上报的待存储的回溯数据后,可以将待存储的回溯数据写入本地磁盘的日志文件中,以实现对接回溯数据的缓存。

[0077] S104、当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。

[0078] 在本申请实施例中,可以使用日志采集代理,如Logtail采集,来对新增的日志文件进行异步实时采集。通过日志采集代理可以动态监控回溯数据服务端本地磁盘中的日志文件,并对新增日志进行实时读取和解析。当Logtail监听到本地磁盘的日志文件中存在新增日志时,可以通过轮询和事件并存的方式对新增日志进行异步实时的采集并将采集到的新增日志存储到消息队列,如Kafka消息队列。在日志采集系统中,轮询指通过定期检测日

志文件有无更新来进行日志采集,事件指依赖于操作系统的事件通知来进行日志采集。日志采集代理可以在系统内部以事件的方式触发日志读取行为,再通过独立的轮询模块定期遍历文件夹并生成增量日志。将两种日志采集技术结合在一起对本地磁盘的新增日志进行采集,可以保证日志采集过程的高效性、可靠性和日志顺序的稳定性。在本申请实施例中,消息队列的功能相当于中转站,可以将缓存在回溯数据服务端本地磁盘上的增量日志写入到消息队列上,后续检索数据库和对象存储服务器可以通过订阅消息队列的方式来获得存储在消息队列上的增量日志。

[0079] 在本申请实施例中,回溯数据的上报过程可以仅通过回溯数据服务端的本地磁盘即可完成,同时在上报过程中可以通过容器集群管理系统对上报回溯数据的服务节点的数量进行动态伸缩,因此可以应对同时大规模回溯数据的情况。待存储的回溯数据可以同时存储在回溯数据服务端的本地磁盘、消息队列、检索数据库和对象存储服务器中,由于数据同时存储在多个不同服务器中,因此即使业务在短时间内停止运转,业务恢复后仍可恢复以往存储的回溯数据,提高了回溯数据存储系统的健壮性。此外,通过日志采集系统和消息队列,可以将待存储的回溯数据高效写入检索数据库中,因此实现了回溯数据的实时检索功能,使得该回溯数据存储方法可以应用于需要对回溯数据进行实时检索的情况,扩大了回溯数据存储系统的适用性。在将数据写入检索数据库和对象存储服务器时,本申请实施例可以通过冷热分层的存储技术,将短期热数据写入检索数据库中,保证短期数据的实时检索,同时可以将需要长期存储的数据压缩后写入对象存储服务器的机械硬盘上进行长期储存,从而大大降低了回溯数据的存储成本。

[0080] 参照图3,示出了本申请实施例提供的一种回溯数据存储方法的示意图,该方法应用于检索数据库,具体可以包括如下步骤:

[0081] S301、订阅消息队列,所述消息队列与远端的回溯数据服务端关联。

[0082] 在本申请实施例中,检索数据库可以是分析型列式数据库ClickHouse或者其他数据库,检索数据库可以通过订阅消息队列获取增量日志。

[0083] S302、若所述消息队列中存在增量日志,则获取并存储所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的。

[0084] 在本申请实施例中,检索数据库在订阅消息队列后,若消息队列中存在增量日志,那么检索数据库可以通过实时消费消息队列中的新增日志的方式来获取待存储的回溯数据。

[0085] 在本申请实施例中,可以在检索数据库建表时通过任意合法的日期表达式指定检索数据库将回溯数据按存储时间进行分区存储,以便后续清除检索数据库中的历史数据。如在分析型列式数据库ClickHouse中通过分区函数,toYYYYMM()将数据按月进行分区。

[0086] S303、定期清空所述检索数据库中的历史数据。

[0087] 在本申请实施例中,将写入检索数据库的回溯数据按存储时间进行分区储存后,可以定期清空检索数据库中历史分区的数据。如在分析型列式数据库ClickHouse中,可以通过对存储分区进行TTL管理以淘汰过期的分区数据。

[0088] 在本申请实施例中,检索数据库通过订阅消息队列实时获得缓存在回溯数据服务端的本地磁盘上的新增日志,可以满足生成回溯视频时,回溯数据进行实时检索的需求。此

外,由于检索数据库中存储的回溯数据进行了分区存储并定期清空过期的分区数据,所以能有效避免检索数据时因为检索数据库存储的数据过多导致检索速度下降的情况,同时,由于定期情况过期的历史数据,因此能使得检索数据库中存储的数据量保持在合理范围,无需对检索数据库进行扩容等操作。

[0089] 参照图4,示出了本申请实施例提供的一种回溯数据存储方法的示意图,该方法应用于对象存储服务器,具体可以包括如下步骤:

[0090] S401、订阅消息队列,所述消息队列与远端的回溯数据服务端关联。

[0091] 在本申请实施例中,对象存储服务器,如OSS,可以通过订阅消息队列获取消息队列中的增量日志。

[0092] S402、若所述消息队列中存在增量日志,则获取并压缩所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的。

[0093] 在本申请实施例中,对象存储服务器在订阅消息队列后,若消息队列中存在增量日志,那么对象存储服务器可以通过实时消费消息队列中的新增日志的方式来获取待存储的回溯数据。

[0094] S403、对压缩后的所述增量日志进行存储。

[0095] 对象存储服务器在获得压缩后的增量日志后,可以将压缩后的增量日志写入对象存储服务器进行缓存。

[0096] S404、将已存储的所述增量日志写入物理盘中。

[0097] 在本申请实施例中,可以将写入对象存储服务器的新增日志通过异地复制技术将其写入对象存储服务器的机械硬盘中。通过异地复制技术,可以将新增日志写入位于不同地区的物理盘中进行长期存储。如可以设置位于A城市的机械硬盘作为第一节点来存储回溯数据,同时设置位于B城市的机械硬盘作为第二个存储节点来对同一批回溯数据进行长期存储。

[0098] 在本申请实施例中,通过异地复制技术选择不同地域的物理盘来对回溯数据进行长期存储,可以减少由于火灾、停电等极端情况造成的数据丢失的概率,因此可以在降低数据存储成本的同时,提高回溯数据的可用性。

[0099] 为了便于理解,下面结合一个具体的示例,对本申请实施例提供的回溯数据存储方法作一介绍。

[0100] 如图5所示,当待存储的回溯数据开始上报时,容器集群管理系统会首先根据待存储的回溯数据的数据量和已配置的服务节点数量判断是否满足对服务节点进行伸缩容的条件,即是否需要调整服务节点的数量。若满足伸缩容条件,即服务节点数量存在冗余或服务节点数量不足导致服务节点负载过重,则通过容器集群管理系统对服务节点进行动态伸缩容,以使服务节点负载均衡。在服务节点负载均衡后,通过服务节点对待存储的回溯数据进行上报,将待存储的回溯数据上报至回溯数据服务端,并将待存储的回溯数据以日志文件的形式写入回溯数据服务端的本地磁盘中。若不满足伸缩容条件,即服务节点已达到负载均衡,则直接开始回溯数据上报服务,待存储的回溯数据以日志文件的形式写入回溯数据服务端的本地磁盘中。

[0101] 如图6所示,采用logtail技术,通过代理的方式异步实时地采集回溯数据服务端

的本地磁盘上的增量日志,并将采集到的增量日志上报到消息队列进行缓存。缓存在消息队列的增量日志同时被对象存储服务器和分析性列式数据库通过订阅消息队列的方式进行消费,以获得增量日志。被消费至对象存储服务器的增量日志,在对象存储服务器中被压缩并归档至本地硬盘中进行长期存储。被消费至分析性列式数据库的增量日志则在数据库中以时间为依据进行分区保存。数据库会定时删除过期分区中的回溯数据,以维持数据库的检索速度和存储的数据量。通过分析性列式数据库,可以对数据库中的回溯数据进行实时检索以生成回溯视频。上述分析性列式数据库可以是前述各个实施例中的检索数据库。

[0102] 参照图7,示出了本申请实施例提供的一种回溯数据存储装置的示意图,具体可以包括服务节点确定模块701、回溯数据接收模块702、回溯数据缓存模块703、回溯数据存储模块704,其中:

[0103] 服务节点确定模块701,用于确定多个服务节点;

[0104] 回溯数据接收模块702,用于通过多个所述服务节点接收待存储的回溯数据;

[0105] 回溯数据缓存模块703,用于将待存储的所述回溯数据写入本地磁盘的日志文件中;

[0106] 回溯数据存储模块704,用于当监听到所述日志文件中存在增量日志时,将所述增量日志发送至消息队列,所述消息队列用于被远端的检索数据库和对象存储服务器订阅,以将所述增量日志存储至所述检索数据库和所述对象存储服务器中。

[0107] 在本申请实施例的一种可能实现方式中,服务节点确定模块501具体可以用于:获取待存储的回溯数据的数据总量;根据数据总量确定多个已配置的服务节点的上报平均值;判断上报平均值是否大于设定阈值;若上报平均值大于设定阈值,则增加至少一个服务节点;若上报平均值小于设定阈值,则减少至少一个服务节点。

[0108] 在本申请实施例的另一种可能实现方式中,所述服务节点确定模块501还可以用于:确定待减少的服务节点的数量;计算减少数量的所述服务节点后,剩余服务节点的上报平均值是否仍小于设定阈值,其中剩余的服务节点的数量大于或等于服务节点的最小配置数量;若剩余的服务节点的上报平均值仍小于设定阈值,则减少所述数量的服务节点;否则,调整待减少的服务节点的数量,重复执行计算减少数量的所述服务节点后,剩余的服务节点的上报平均值是否仍小于设定阈值的步骤。

[0109] 在本申请实施例的一种可能实现方式中,回溯数据存储模块504具体可以用于:监听所述日志文件中是否存在所述增量日志;若所述日志文件中存在所述增量日志,则通过轮询和/或事件并存的方式获取所述增量日志;将所述增量日志发送至所述消息队列。

[0110] 参照图8,示出了本申请实施例提供的另一种回溯数据存储装置的示意图,具体可以包括订阅模块801、增量日志存储模块802和数据清空模块803,其中:

[0111] 订阅模块801,用于订阅消息队列,所述消息队列与远端的回溯数据服务端关联;

[0112] 增量日志存储模块802,用于若所述消息队列中存在增量日志,则获取并存储所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;

[0113] 数据清空模块803,用于定期清空所述检索数据库中的历史数据。

[0114] 参照图9,示出了本申请实施例提供的又一种回溯数据存储装置的示意图,具体可以包括订阅模块901、增量日志获取模块902、增量日志存储模块903和增量日志写入模块

904,其中:

[0115] 订阅模块901,用于订阅消息队列,所述消息队列与远端的回溯数据服务端关联;

[0116] 增量日志获取模块902,用于若所述消息队列中存在增量日志,则获取并压缩所述增量日志,所述增量日志是所述回溯数据服务端在监听到本地磁盘的日志文件中存在所述增量日志时,将所述增量日志发送至所述消息队列的;

[0117] 增量日志存储模块903,用于对压缩后的所述增量日志进行存储;

[0118] 增量日志写入模块904,用于将已存储的所述增量日志写入物理盘中。

[0119] 在本申请实施例的一种可能实现方式中,增量日志写入模块904还可以用于:通过异地复制技术,将已存储的所述增量日志同时写入多个不同的所述物理盘中,多个不同的所述物理盘位于不同的地域。

[0120] 对于装置实施例而言,由于其与方法实施例基本相似,所以描述得比较简单,相关之处参见方法实施例部分的说明即可。

[0121] 参照图10,示出了本申请实施例提供的一种计算机设备的示意图。如图10所示,本申请实施例中的终端设备1000包括:处理器1010、存储器1020以及存储在所述存储器1020中并可在所述处理器1010上运行的计算机程序1021。所述处理器1010执行所述计算机程序1021时实现上述回溯数据存储方法各个实施例中的步骤,例如图1所示的步骤S101至S104。或者,所述处理器1010执行所述计算机程序1021时实现上述各装置实施例中各模块/单元的功能,例如图7所示模块701至704的功能。

[0122] 示例性的,所述计算机程序1021可以被分割成一个或多个模块/单元,所述一个或者多个模块/单元被存储在所述存储器1020中,并由所述处理器1010执行,以完成本申请。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段,该指令段可以用于描述所述计算机程序1021在所述终端设备1000中的执行过程。例如,所述计算机程序1021可以被分割成服务节点确定模块、回溯数据接收模块、回溯数据缓存模块和回溯数据存储模块,各模块具体功能如下:

[0123] 服务节点确定模块,用于确定多个服务节点;

[0124] 回溯数据接收模块,用于通过多个服务节点接收待存储的回溯数据;

[0125] 回溯数据缓存模块,用于将待存储的回溯数据写入本地磁盘的日志文件中;

[0126] 回溯数据存储模块,用于当监听到日志文件中存在增量日志时,将增量日志发送至消息队列,消息队列用于被远端的检索数据库和对象存储服务器订阅,以将增量日志存储至检索数据库和对象存储服务器中。

[0127] 所述计算机设备1000可以是用于实现前述各个方法实施例的计算机设备,该计算机设备1000可以是桌上型计算机、云端服务器等计算设备。所述计算机设备1000可包括,但不限于,处理器1010、存储器1020。本领域技术人员可以理解,图10仅仅是计算机设备1000的一种示例,并不构成对计算机设备1000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,例如所述计算机设备1000还可以包括输入输出设备、网络接入设备、总线等。

[0128] 所述处理器1010可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-

Programmable Gate Array, FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0129] 所述存储器1020可以是所述计算机设备1000的内部存储单元,例如计算机设备1000的硬盘或内存。所述存储器1020也可以是所述计算机设备1000的外部存储设备,例如所述计算机设备1000上配备的插接式硬盘,智能存储卡(Smart Media Card, SMC),安全数字(Secure Digital, SD)卡,闪存卡(Flash Card)等等。进一步地,所述存储器1020还可以既包括所述计算机设备1000的内部存储单元也包括外部存储设备。所述存储器1020用于存储所述计算机程序1021以及所述计算机设备1000所需的其他程序和数据。所述存储器1020还可以用于暂时地存储已经输出或者将要输出的数据。

[0130] 本申请实施例还公开了一种计算机设备,包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现如前述各个实施例所述的回溯数据存储方法。

[0131] 本申请实施例还公开了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,所述计算机程序被处理器执行时实现如前述各个实施例所述的回溯数据存储方法。

[0132] 本申请实施例还公开了一种计算机程序产品,当所述计算机程序产品在计算机设备上运行时,使得所述计算机设备执行前述各个实施例所述的回溯数据存储方法。

[0133] 以上所述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,本领域的普通技术人员应当理解: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,均应包含在本申请的保护范围之内。

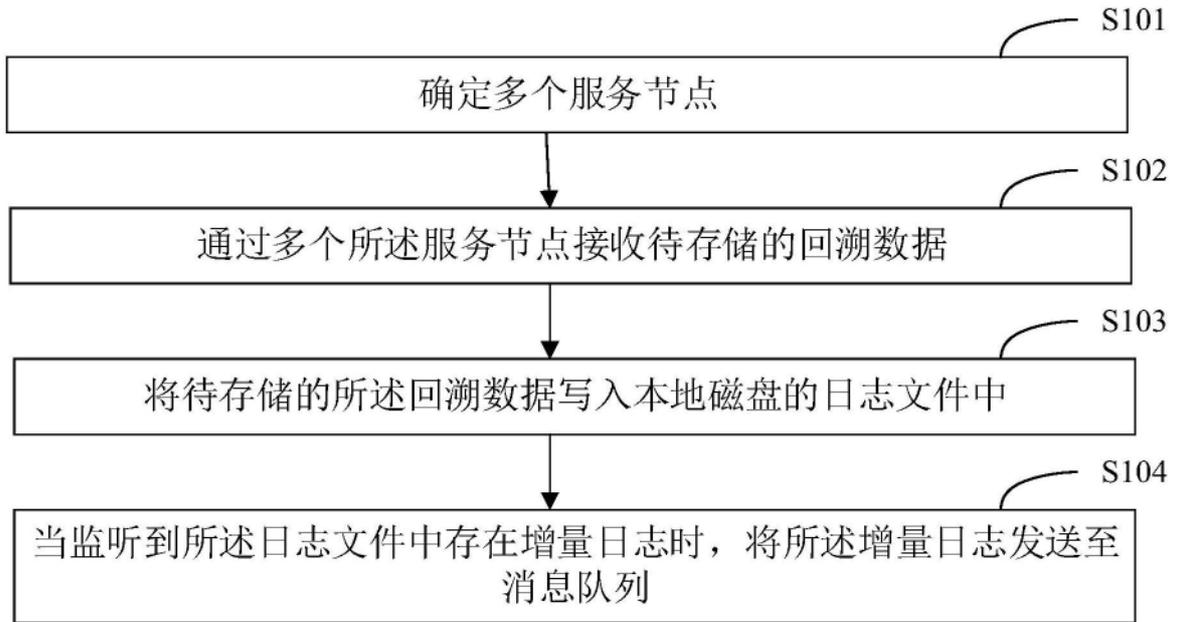


图1

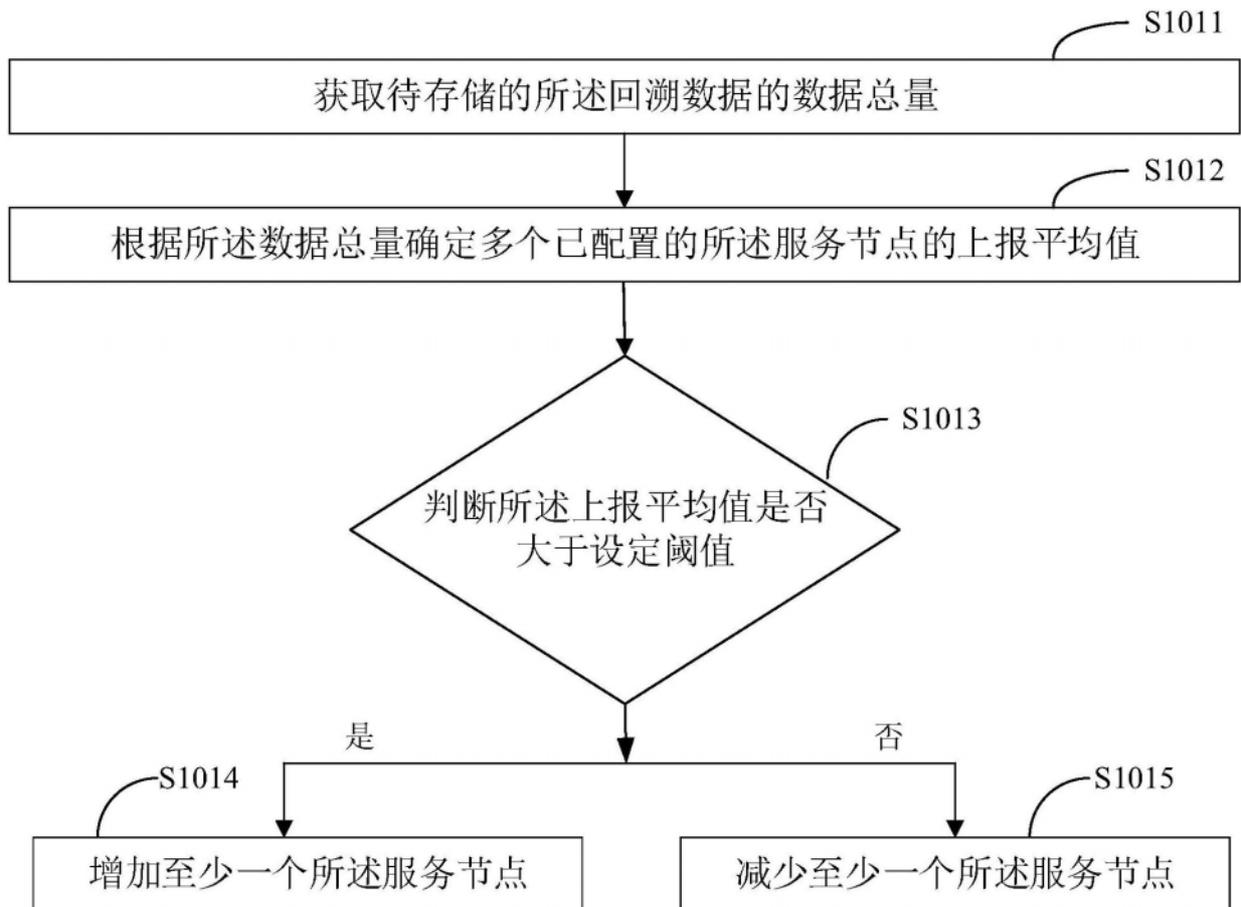


图2

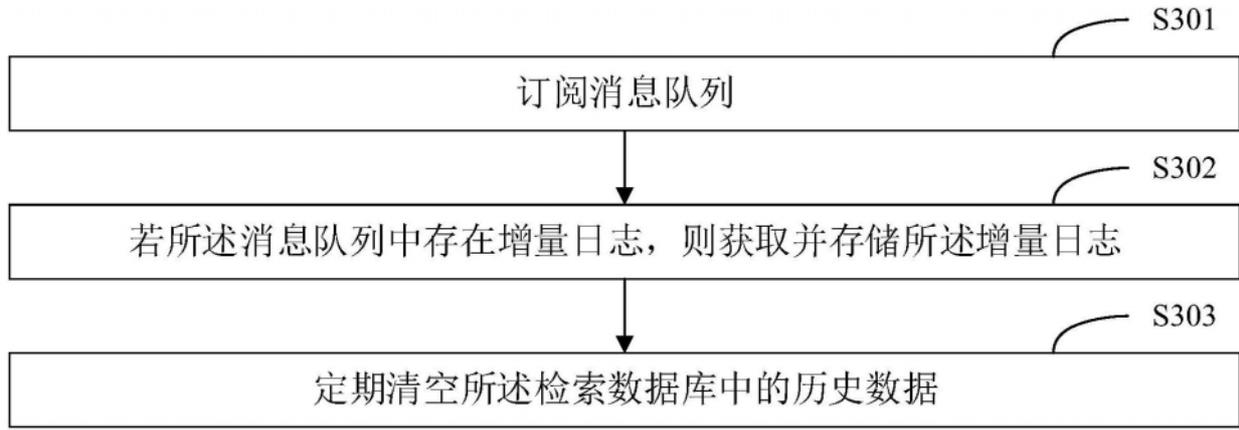


图3

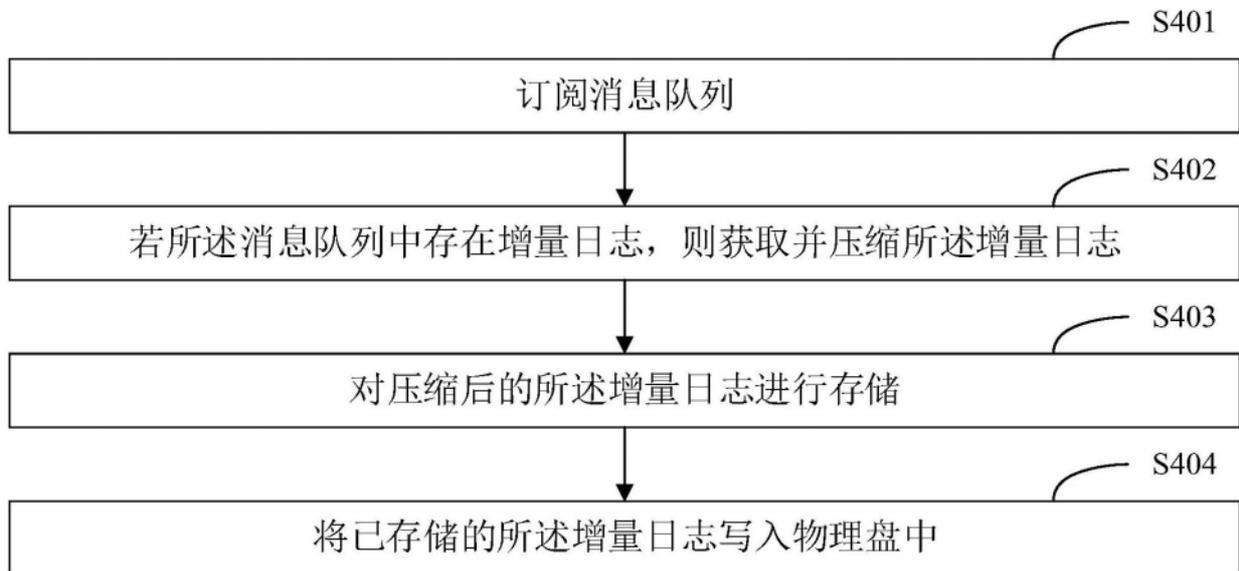


图4

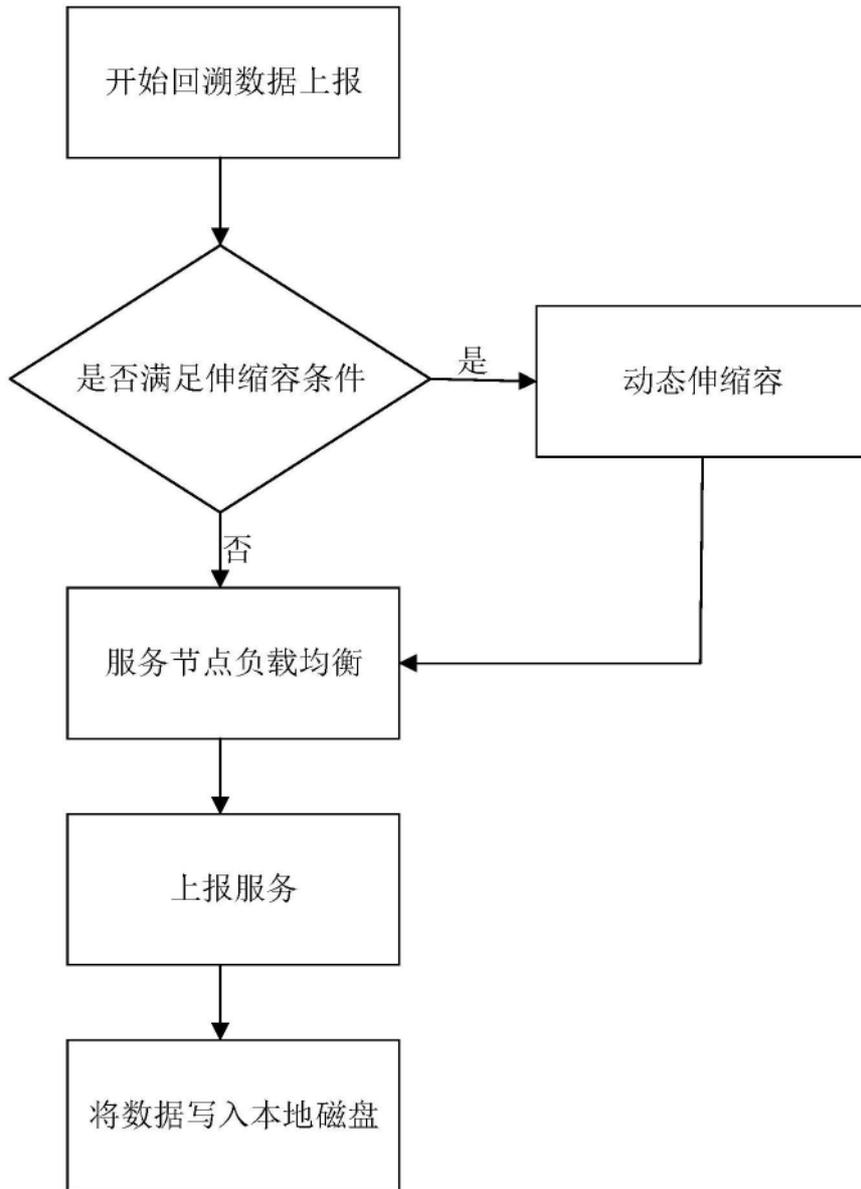


图5

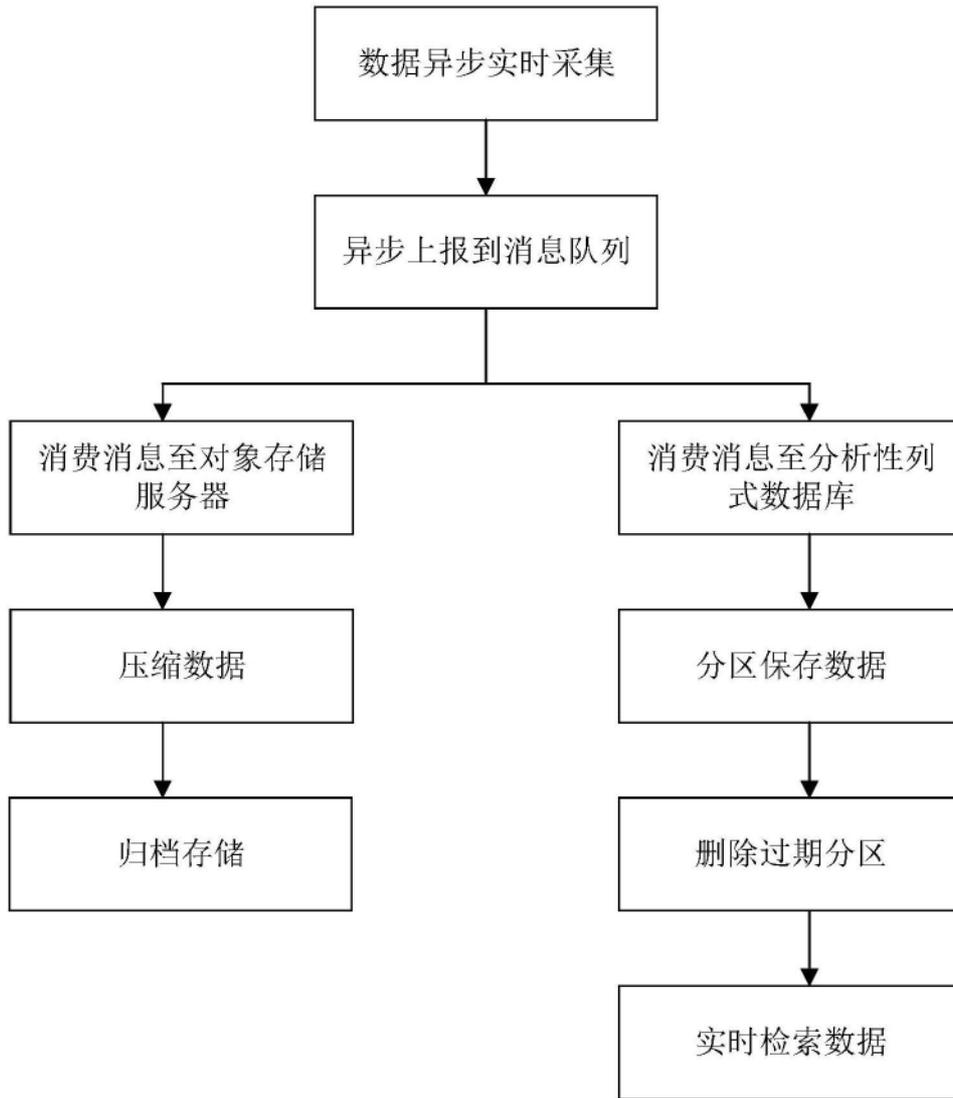


图6

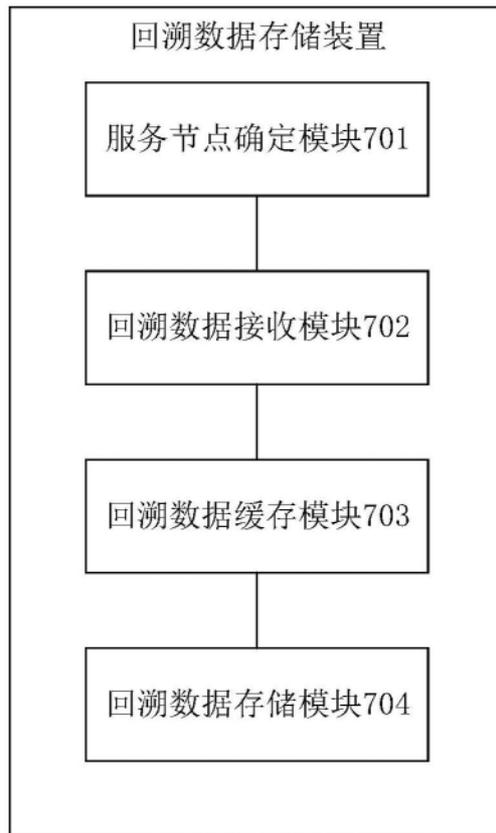


图7

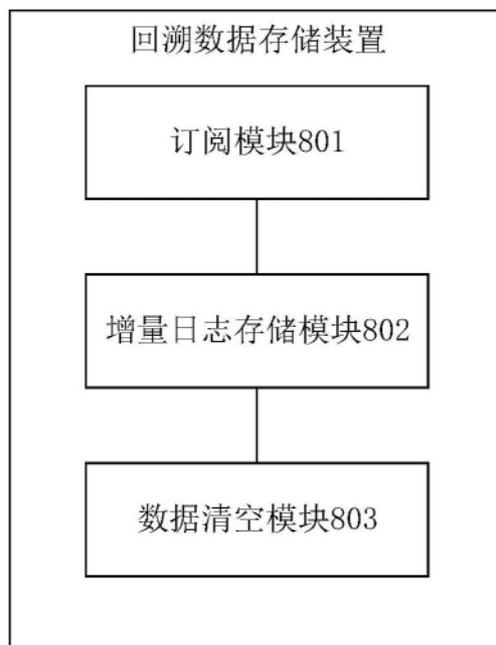


图8

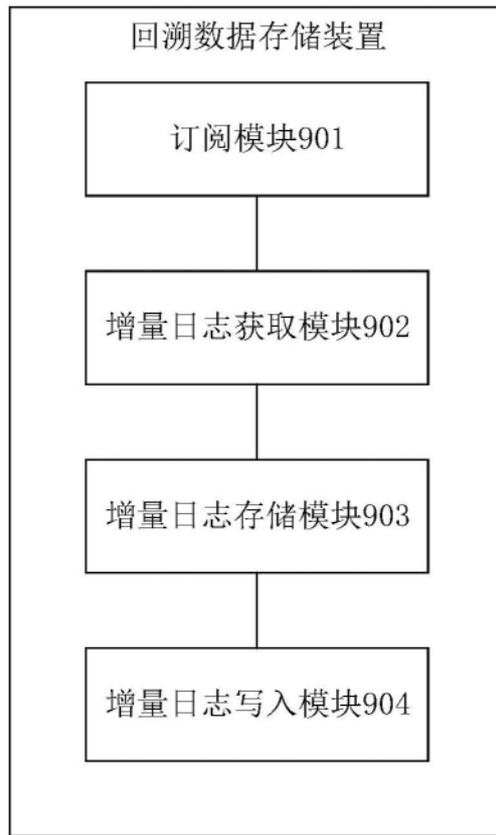


图9

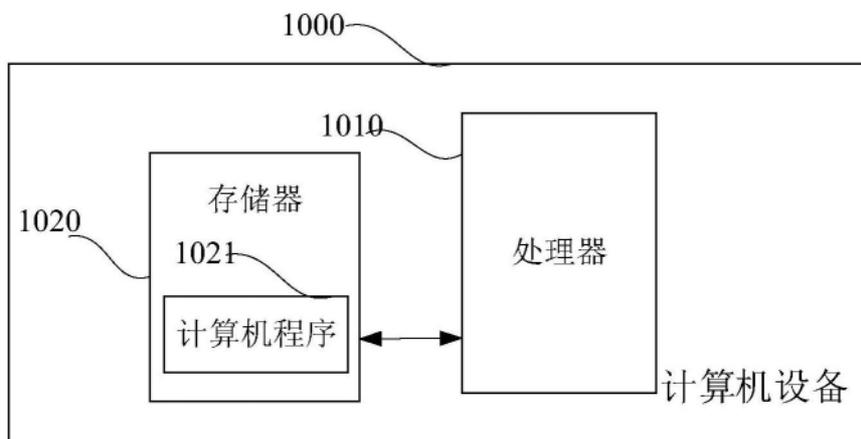


图10