



(12) 发明专利申请

(10) 申请公布号 CN 114356650 A

(43) 申请公布日 2022.04.15

(21) 申请号 202111397882.6

(22) 申请日 2021.11.19

(71) 申请人 科大讯飞股份有限公司

地址 230088 安徽省合肥市高新开发区望江西路666号

(72) 发明人 张致江 殷奎 张明 夏静霆
黄开元 王芝斌

(74) 专利代理机构 深圳市威世博知识产权代理
事务所(普通合伙) 44280

代理人 刘希

(51) Int. Cl.

G06F 11/14 (2006.01)

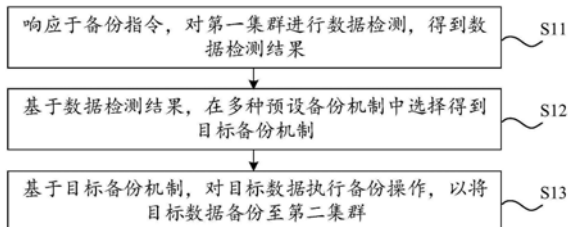
权利要求书2页 说明书12页 附图2页

(54) 发明名称

数据备份方法、装置、设备、系统及存储介质

(57) 摘要

本申请公开了一种数据备份方法、装置、设备、系统及存储介质,其中,数据设备方法包括:响应于备份指令,对第一集群进行数据检测,得到数据检测结果;其中,数据检测结果包括第一集群中是否存入目标数据;基于数据检测结果,在多种预设备份机制中选择得到目标备份机制;且多种预设备份机制至少包括快照机制、日志机制;基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。通过上述方式,本申请能够尽可能降低数据丢失对用户侧业务影响。



1. 一种数据备份方法,其特征在于,包括:

响应于备份指令,对第一集群进行数据检测,得到数据检测结果;其中,所述数据检测结果包括所述第一集群中是否存入目标数据;

基于所述数据检测结果,在多种预设备份机制中选择得到目标备份机制;其中,所述多种预设备份机制至少包括快照机制、日志机制;

基于所述目标备份机制,对所述目标数据执行备份操作,以将所述目标数据备份至第二集群。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述数据检测结果,在多种预设备份机制中选择得到目标备份机制,包括:

在所述检测结果包括所述第一集群中存入所述目标数据的情况下,选择所述快照机制作为所述目标备份机制;

和/或,在所述检测结果包括所述第一集群中未存入所述目标数据的情况下,选择所述日志机制作为所述目标备份机制。

3. 根据权利要求1所述的方法,其特征在于,所述目标备份机制为所述快照机制,所述基于所述目标备份机制,对所述目标数据执行备份操作,以将所述目标数据备份至第二集群,包括:

对所述第一集群中所述目标数据执行快照标记,得到所述目标数据的数据副本;

基于所述目标数据的数据副本恢复得到所述目标数据,并将所述目标数据发送至所述第二集群。

4. 根据权利要求3所述的方法,其特征在于,在所述基于所述目标数据的数据副本恢复得到所述目标数据,并将所述目标数据发送至所述第二集群之后,所述方法还包括:

删除所述第一集群中所述数据副本。

5. 根据权利要求1所述的方法,其特征在于,所述目标备份机制为所述日志机制,所述第一集群还存储有操作日志,所述操作日志用于存储数据操作记录,所述基于所述目标备份机制,对所述目标数据执行备份操作,以将所述目标数据备份至第二集群,包括:

从所述第一集群复制所述目标数据的数据操作记录,并基于所述目标数据的数据操作记录恢复出所述目标数据;

将所述目标数据发送至所述第二集群。

6. 根据权利要求5所述的方法,其特征在于,在将所述目标数据发送至所述第二集群之后,所述方法还包括:

删除所述第一集群中所述目标数据的数据操作记录。

7. 根据权利要求1所述的方法,其特征在于,所述备份指令是用户设备在存储目标数据至第一集群时由用户发送的;

或者,所述备份指令是用户设备在存储目标数据至第一集群时基于预设备份策略而自动发送的。

8. 根据权利要求1所述的方法,其特征在于,所述第一集群为主集群,且所述第二集群为备份集群,所述方法还包括:

响应于切换指令,切换所述第一集群为所述备份集群,并切换所述第二集群为所述主集群;

和/或,响应于回迁指令,基于所述快照机制从备份集群向主集群进行数据回迁。

9. 根据权利要求1所述的方法,其特征在于,在所述响应于备份指令,对第一集群进行数据检测,得到数据检测结果之前,所述方法还包括:

获取所述第一集群、所述第二集群的元数据;

其中,所述元数据包括所述第一集群的第一网络接口和所述第二集群的第二网络接口,所述第一网络接口用于实现与所述第一集群之间的数据传输,所述第二网络接口用于实现与所述第二集群之间的数据传输。

10. 根据权利要求1所述的方法,其特征在于,在执行所述备份操作的过程中,所述方法还包括:

基于已发送操作指令至所述第一集群和/或所述第二集群,发送询问指令至第一集群和/或第二集群;其中,所述询问指令用于询问所述第一集群和/或所述第二集群关于所述操作指令的完成情况;

接收所述第一集群和/或所述第二集群的反馈指令;其中,所述反馈指令包括以下任一者:基于已完成所述操作指令而反馈的完成指令、基于已收到所述操作指令但未完成所述操作指令的确认指令;

所述第一集群、第二集群基于判断模块判断所述备份操作过程中的每个接收指令是否已完成,若已完成回复完成指令;

若未完成回复确认指令。

11. 根据权利要求1所述的方法,其特征在于,所述第一集群包含以分布式布置的多个存储镜像,且数据均存入所述存储镜像中。

12. 一种数据备份装置,其特征在于,包括:

检测模块,用于响应备份指令,对第一集群进行数据检测,得到数据检测结果;其中,所述数据检测结果包括所述第一集群中是否存入目标数据;

选择模块,用于基于所述数据检测结果,在多种预设备份机制中选择得到目标备份机制;其中,所述多种预设备份机制至少包括快照机制、日志机制;

备份模块,用于基于所述目标备份机制,对所述目标数据执行备份操作,以将所述目标数据备份至第二集群。

13. 一种数据备份设备,其特征在于,包括存储器、通信电路和处理器,且所述存储器和所述通信电路耦接于所述处理器,所述存储器中存储有程序指令,所述处理器用于执行所述程序指令,以实现权利要求1至11任一项所述的数据备份方法。

14. 一种数据备份系统,其特征在于,包括数据备份设备和多个集群,所述数据备份设备分别与所述多个集群通信连接,所述数据备份设备用于执行权利要求1至11任一项所述的数据备份方法,以在所述多个集群之间实现数据备份。

15. 一种计算机可读存储介质,其特征在于,存储有能够被处理器运行的程序指令,所述程序指令用于实现权利要求1至11任一项所述的数据备份方法。

数据备份方法、装置、设备、系统及存储介质

技术领域

[0001] 本申请涉及数据备份技术领域,特别是涉及一种数据备份方法、装置、设备、系统及存储介质。

背景技术

[0002] 随着虚拟化技术的不断发展,用户对于虚拟化场景中云主机中关键数据的丢失问题尤其关注,与此同时,在混合多云环境中,不仅仅局限于数据备份,更需要考虑在备份的同时降低对于业务运行影响和主集群故障时候对于快速切换备机,同时尽可能不影响性能、简化管理、降低存储成本、更有效和高效的业务运营。

[0003] 目前,现有技术数据保护解决方案的主要目的是数据备份和恢复,但未着重考虑如何尽可能的降低数据丢失对于业务侧影响,从而在由于不可抗等因素导致数据意外丢失时,可能直接造成业务侧无法正常进行。有鉴于此,如何尽可能降低数据丢失对于用户侧业务影响成为亟待解决的问题。

发明内容

[0004] 本申请主要解决的技术问题是提供一种数据备份方法、装置、设备、系统及存储介质,能够尽可能降低数据丢失对于用户侧业务影响。

[0005] 为了解决上述技术问题,本申请第一方面提供了一种数据备份方法,包括:响应于备份指令,对第一集群进行数据检测,得到数据检测结果;且数据检测结果包括第一集群中是否存入目标数据;再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制;且多种预设备份机制至少包括快照机制、日志机制;再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。

[0006] 为了解决上述技术问题,本申请第二方面提供了一种数据备份装置,包括:检测模块,用于响应备份指令,对第一集群进行数据检测,得到数据检测结果;数据检测结果包括第一集群中是否存入目标数据;选择模块,用于基于数据检测结果,在多种预设备份机制中选择得到目标备份机制;多种预设备份机制至少包括快照机制、日志机制;备份模块,用于基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。

[0007] 为了解决上述技术问题,本申请第三方面提供了一种数据备份设备,包括存储器、通信电路和处理器,且存储器和通信电路耦接于处理器,存储器中存储有程序指令,处理器用于执行程序指令,以实现上述第一方面中的数据备份方法。

[0008] 为了解决上述技术问题,本申请第四方面提供了一种数据备份系统,包括数据备份设备和多个集群,数据备份设备分别与多个集群通信连接,数据备份设备用于执行上述第一方面中的数据备份方法,以在多个集群之间实现数据备份。

[0009] 为了解决上述技术问题,本申请第五方面提供了一种计算机可读存储介质,存储有能够被处理器运行的程序指令,程序指令用于实现上述第一方面中的数据备份方法。

[0010] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据

检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽可能降低数据丢失对于用户侧业务影响。

附图说明

- [0011] 图1是本申请数据备份方法一实施例的流程示意图;
- [0012] 图2是本申请数据备份方法一实施例的架构图;
- [0013] 图3是图1中步骤S12一实施例的架构图;
- [0014] 图4是图1中步骤S12另一实施例的架构图;
- [0015] 图5是本申请数据备份装置一实施例的框架示意图;
- [0016] 图6是本申请数据备份设备一实施例的框架示意图;
- [0017] 图7是本申请计算机可读存储介质一实施例的框架示意图。

具体实施方式

[0018] 下面结合说明书附图,对本申请实施例的方案进行详细说明。

[0019] 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、接口、技术之类的具体细节,以便透彻理解本申请。

[0020] 本文中术语“系统”和“网络”在本文中常被可互换使用。本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。此外,本文中的“多”表示两个或者多于两个。

[0021] 请参阅图1,图1是本申请数据备份方法一实施例的流程示意图。

[0022] 具体而言,可以包括如下步骤:

[0023] 步骤S11:响应于备份指令,对第一集群进行数据检测,得到数据检测结果。

[0024] 在一个实施场景中,备份指令可以是用户设备在存储目标数据至第一集群时由用户发送的;或者,备份指令也可以是由用户设备在存储目标数据至第一集群时基于预设备份策略而自动发送的。当备份指令由用户设备在存储目标数据至第一集群时发送,也即是用户在对数据执行写入操作;当备份指令是由用户设备在存储目标数据至第一集群200时基于预设备份策略而自动发送时,也即是在数据备份设备208中通过点击备份按钮,自动生成备份策略进行自动发送。备份策略可以确定备份时间和方式,例如,可以将备份策略设置为实时备份,也就是说数据若执行写入操作即对数据执行备份。上述方式,两种备份指令的发送可以避免用户在数据存入的过程中,遗忘对数据执行备份操作而导致的数据丢失问题,进一步降低了数据的丢失情况。

[0025] 在一个具体实施场景中,用户对数据执行的写入操作可以分为添加操作、更改操作和删除操作。其中,添加操作即为用户对新数据的记录过程,在用户对数据添加完成后,

只会对当前存入的数据执行备份操作。更改操作可能发生在数据备份之后,因此,在更改后的数据执行备份的过程后,若需要对更改之前的数据留底,可以选择重新保存,重新保存的数据不会对更改之前的数据进行覆盖,当需要对更改前的数据进行查找或者恢复时,不会影响用户的正常使用和查询。删除操作也可能发生在数据备份之后,因此可以选择直接保存进而覆盖之前的数据,或者重新保存进而对删除之前的数据进行留底。具体情况在此不做限定,可以根据实际应用中的需求进行选择。

[0026] 在一个具体实施场景中,数据备份设备可以包括云管平台、备份代理、备份组件,云管平台中的显示界面可以通过网络连接在不同服务器处分别予以显示,且云管平台可以安装在集群中的任意一台服务器中,或者任何与集群连接的服务器设备中,具体情况在此不做限定,可以根据实际应用中的需求进行选择。云管平台发出的指令经过备份代理发送出去,并且备份代理内部会对收到的消息执行检测,若备份代理收到云管平台发送的消息为重复时,备份代理经过检测可以判定收到消息是重新发送消息,或者是重新发送新消息。备份组件具有检测功能,对第一集群、第二集群进行检测。云管平台、备份代理、备份组件可以安装在同一台服务器中,也可以安装在不同的服务器中,具体情况在此不做限定,可以根据实际情况进行设置。

[0027] 在一个实施场景中,在响应于备份指令,对第一集群进行数据检测,得到数据检测结果之前,还可以获取第一集群、第二集群的元数据,且元数据可以包括第一集群的第一网络接口和第二集群的第二网络接口,第一网络接口用于实现与第一集群之间的数据传输,第二网络接口用于实现与第二集群之间的数据传输。上述方式,一方面可以实现数据备份设备与集群之间的数据传输,另一方面可以实现在一个集群重新连接时,不会对集群中的任一存储系统发生遗漏。

[0028] 在一个具体实施场景中,元数据可以包括但不限于名称、规格、网络、安全组、卷挂载点等信息,元数据主要记录第一集群、第二集群模型的定义、各层级间的映射关系、监控数据仓库的数据状态及任务运行状态,获取元数据使得第一集群、第二集群部署、操作和管理能达成协同和一致。其中,元数据中卷挂载点为磁盘文件系统的入口目录,入口目录可以是第一网络接口或第二网络接口,即第一网络接口、第二网络接口可以对第一集群、第二集群中的数据进行访问和传输,进而实现数据备份设备与第一集群、第二集群之间的数据传输。

[0029] 在一个实施场景中,第一集群可以用作主集群,第二集群可以用作备份集群;当然,也可以根据业务需要,第一集群用作备份集群,第二集群用作主集群,对第一集群、第二集群的应用在此不做具体限定,可以根据实际应用中的需要进行设置。在这里,主集群即对数据执行保存的集群,备份集群即对保存后的数据执行备份保存的集群,第一集群、第二集群中底层存储均采用分布式存储,分布式存储相较于现有技术中底层存储采用集中式存储,运维部署更简单,并且可以更好地统一利用低成本的磁盘,最终提供更高容量的存储池,分布式存储的可移植性较好,进而可以降低节点死机对于存储系统的影响,提高用户的使用体验。

[0030] 在一个实施场景中,对第一集群进行数据检测是用过数据备份设备中的备份组件对第一集群中存储数据进行检测,备份组件对第一集群中数据进行检测是为了保证在进行数据备份的过程中不会对未备份数据发生遗漏,进而减少数据的丢失情况。

[0031] 在本公开实施例中,数据检测结果包括第一集群中是否存入目标数据。具体地,目标数据为需要进行备份的数据,也即是说,若数据检测结果存入目标数据,则表明之前存储有未进行备份的数据,则目标数据包括未进行备份的所有数据;若数据检测结果未存入目标数据,则表明之前未存储有未进行备份的数据,则目标数据为写入数据。

[0032] 步骤S12:基于数据检测结果,在多种预设备份机制中选择得到目标备份机制。

[0033] 在一个实施场景中,预设备份机制可以根据实际情况进行设置,例如:若数据备份方法用于服务器时,可以设置为完全备份方式,完全备份方式可以将所有的数据进行系统的备份;若数据备份方法用于大型交易市场,可以设置为增量备份,增量备份即备份自从上次备份操作以来新改变的数据,这些新改变的数据或者是新产生的数据,或者是更新的数据,增量备份所要求的备份事件最短,采用增量备份的方式,将有效节省存储空间,同时在数据发生丢失的时候,很快可以从备份数据中还原;还可以选择增量备份,增量备份需要备份的数据较多,是前后两次比较的结果,将不同的部分数据进行备份;或者设置为有选择的备份或即时备份,它可以按照所定义的时间安排,自动地执行上面列出的备份操作,但是,数据中心运维的人员有时也需要根据情形即时地备份数据。要经常对备份数据进行检查,当发现数据缺少或者不对时,要主动进行即时备份,刷新备份的数据,确保和正在运行的数据中心里的数据保持一致。

[0034] 在一个具体实施场景中,预设的备份机制可以为快照机制,快照是数据存储的某一时刻的状态记录,即将某一时刻的存储状态记录,并且进行恢复得到存储数据,利用快照的镜像可以在数秒钟内把数据恢复到做快照的时间点,还允许系统管理员选择性地迅速恢复受损或被删文件。数据快照的功能还有很多用处,比如现在需要一份最新的生产数据来做新系统的测试或者提供决策支持和数据分析所用,而系统又不能停机,使用磁带备份恢复一份数据时间又很长,这样的情况可以利用数据快照的备份功能在任一时间点建立快照拷贝,利用拷贝的数据进行测试和分析,不会影响系统的正常使用。

[0035] 在一个具体实施场景中,预设的备份机制可以为日志机制,日志备份机制即将对存储数据的操作记录,依据操作记录可以将存储的数据内容进行恢复,并且由于对数据的所有操作记录都被存储,尽可能的降低了数据在备份的过程中出现的数据丢失情况,提高数据备份的效率。

[0036] 在一个实施场景中,可以基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,例如在检测结果包括第一集群中存入目标数据的情况下,可以选择快照机制作为目标备份机制;在检测结果包括第一集群中未存入目标数据的情况下,可以选择日志机制作为目标备份机制。上述方式,通过对不同的检测结果选择不同的备份机制,将集群中存储的数据快速进行备份,进一步减少了数据混乱造成的数据未备份,进而造成数据丢失的情况。

[0037] 步骤S13:基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。

[0038] 请参阅图2,图2是本申请数据备份方法一实施例的架构图。如图2所示,此时,第一集群200为主集群,第二集群212为备份集群,第一集群200与第二集群212通过数据备份设备208进行数据交互并实现数据的备份。第一集群200包括云主机(1) 201、Qemu-kvm(1) 202、第一集群存储系统203,云主机(1) 201为虚拟机,例如:一个虚拟机为云主机,用户登陆之后

创建了文件,创建文件之后写入数据,数据写入完成之后需要对数据进行存储,数据可以有选择的进行存储,即将数据存储在不同的磁盘。Qemu-kvm(1) 202基于硬件虚拟化技术,并结合了Qemu提供设备虚拟化功能,实现整个系统虚拟化;第一集群存储系统203在进行数据备份时,接收到备份指令之后,让底层系统执行数据备份。第一集群存储系统203包括控制器(1) 204、操作日志205、存储盘206、存储池I(1) 207和存储池II(1) 207,控制器(1) 204可以作为一个协议,若有新的数据写入,可以将新的数据写入指定的存储盘进行存储;操作日志205存储有操作日志,备份机制选择用日志机制时,操作日志用于存储数据的操作记录,即在新的数据写入、更改之前存储的数据或者是对已存储的数据进行删除等记录,通过操作记录可以将数据重新恢复;存储盘206可以作为存储系统,即有数据写入时,在没有备份指令的情况下,将数据存储存储在存储盘206,当存储盘206中存储空间不充足时,可以通过存储池I(1) 207或存储池II(1) 207对存储空间进行扩展。数据备份设备208包括云管平台209、备份代理210和备份组件211,云管平台209中包括显示界面,显示界面支持用户登陆,登陆成功之后,显示界面呈现不同的功能选择建,例如:若选择点击备份按钮后,可以跳出不同的备份机制可供选择;若在第一集群200和第二集群212进行集群切换后,显示界面有选择回迁按钮,选择回迁即可将数据进行备份,尽可能减低数据丢失的情况。云管平台209发送的所有指令均通过备份代理210执行发送,并且备份代理210可以接收第一集群200和第二集群212的交互信息;备份组件211在接收到备份代理210的备份指令后,对数据进行检测并且执行不同的备份机制。第二集群212和第一集群200的构造类似,不同之处在于当第二集群212为备份集群时,对数据的存储与第一集群200略有不同。第二集群212包括云主机(2) 213、Qemu-kvm(2) 214、第二集群存储系统215,第二集群存储系统215包括控制器(2) 216、存储盘I217、存储盘II218、基础存储盘219、存储池I(2) 220和存储池II(2) 220,在这里,控制器(2) 216可以作为一个协议,存储盘I217、存储盘II218和基础存储盘219均为存储系统,例如:存储盘I217、存储盘II218代表的是后续重新增量写入的数据,存储盘I217写入数据为“ABC”,存储盘II218写入数据为“def”,不同的写入数据可以存储在不同的存储系统中。数据备份设备208对于不同的备份机制,选择不同的数据存储方式,将数据发送至第二集群212,并进行存储。

[0039] 在一个具体实施场景中,目标备份机制为快照机制,基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群212,包括对第一集群200中目标数据执行快照标记,得到目标数据的数据副本;再基于目标数据的数据副本恢复得到目标数据,并将目标数据发送至第二集群212。上述方式,可以在短时间内将目标数据快速备份,快照备份机制是源于某一个时间点对存储空间的数据内容进行快照备份,可以尽可能减少在备份的同时进行数据写入造成的数据遗漏问题,并且快照机制不需要完整的数据空间,进而达到不仅可以快速备份并且节约了存储空间的目的。

[0040] 请参阅图2和图3,图3是图1中步骤S12一实施例的架构图。云主机(1)为第一集群200中虚拟机,用户可通过云主机(1)写入数据,当用户设备在存储目标数据至第一集群200时发送备份指令或者基于预设备份策略自动发送备份指令时,数据备份设备208通过备份代理210发送备份指令至备份组件211,备份组件211接收到备份指令后对第一集群200进行数据检测,备份组件211对第一集群200中存储盘206中存储数据进行检测,若检测结果存储盘206中并存入目标数据,则选择快照机制进行备份。当对数据备份机制选择为快照机制

时,控制器(1)204根据日志快照将用户写入数据保存至存储盘206中,保存成功后,存储盘206返回确认指令至控制器(1)204,控制器(1)204接收到确认指令后返回确认指令至用户设备,呈现给用户设备的确认指令可以为“已保存至**”,具体内容可以根据实际情况进行设置。控制器(1)204返回确认指令至用户设备后,备份组件211基于快照备份机制对第一集群200中存储盘206进行检测,检测到存储盘206中有未备份的数据卷,对数据卷内容进行快照备份,即对数据存储的这一时刻的状态进行快照标记,进而得到目标数据的数据副本,备份组件211基于数据副本将数据恢复并且发送至第二集群212,备份组件211发送数据至第二集群212之后,将第一集群200中数据副本删除;第二集群212在接收到备份组件211发送的数据内容后,将数据存储于存储盘II等任一个存储磁盘中。数据备份设备208执行快照备份机制将数据进行备份后,若备份指令未被中断,用户再次写入数据,可以继续执行日志备份机制进行备份,尽可能减少数据丢失的情况。

[0041] 在一个具体实施场景中,在基于目标数据的数据副本恢复得到目标数据,并将目标数据发送至第二集群212之后,可以删除第一集群200中数据副本,上述方式,通过删除数据副本,避免了数据再次被恢复并且重新发送至第二集群212的情况,提高了数据备份的高效性。

[0042] 在一个具体实施场景中,目标备份机制为日志机制,第一集群200还存储有操作日志,操作日志用于存储数据操作记录,基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群212,包括从第一集群200复制目标数据的数据操作记录,并基于目标数据的数据操作记录恢复出目标数据,将目标数据发送至第二集群212。上述方式,通过日志备份机制将数据进行备份,能够快速地将数据恢复出来,并且尽可能减少了还原数据所需要的时间,进一步减少数据丢失的情况。

[0043] 请参阅图2和图4,图4是图1中步骤S12另一实施例的架构图。云主机(1)为第一集群200中虚拟机,用户可通过云主机(1)写入数据,当用户设备在存储目标数据至第一集群200时发送备份指令或者基于预设备份策略自动发送备份指令时,数据备份设备208通过备份代理210发送备份指令至备份组件211,备份组件211接收到备份指令后对第一集群200进行数据检测,备份组件211对第一集群200中存储盘206中存储数据进行检测,若检测结果存储盘206中并未存入目标数据,则选择日志机制进行备份。当对数据备份机制选择为日志机制时,控制器(1)204根据日志机制首先将用户存储数据操作记录在操作日志中,操作日志中内容记录完成后会返回确认指令至控制器(1)204,控制器(1)204接收到确认指令后返回确认指令至用户设备,呈现给用户设备的确认指令可以为“已保存至**”,具体内容可以根据实际情况进行设置。控制器(1)204返回确认指令至用户设备后,再将用户写入的原始数据存储至存储盘206中;备份组件211基于日志备份机制对操作日志进行检测,即检测操作日志205中是否有操作记录,若有操作记录,备份组件211获取操作记录内容,并依据操作记录将数据内容恢复出来,再将恢复出的数据发送至第二集群212,备份组件211发送恢复出的数据至第二集群212之后,将第一集群200中操作日志205的操作记录删除,确保不会对同一操作记录进行多次恢复;第二集群212在接收到备份组件211发送的数据内容后,将数据存储于存储盘I等任一个存储磁盘中。在用户设备将数据不断地写入时,数据备份设备208可以不断地将数据恢复至第二集群212中,尽可能的减少数据的丢失。

[0044] 在一个具体实施场景中,在将目标数据发送至第二集群212之后,方法还包括删除

第一集群200中目标数据的数据操作记录,上述方式,确保将目标数据备份至第二集群212后,数据备份设备208不会对目标数据再次进行恢复,避免了数据的重复恢复,并且减少了数据多次存储占用的空间,尽可能提高了数据备份的效率。

[0045] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽可能降低数据丢失对于用户侧业务影响。

[0046] 在一些公开实施例中,第一集群200为主集群,且第二集群212为备份集群,方法还包括响应于切换指令,切换第一集群200为备份集群,并切换第二集群212为主集群,或者,响应于回迁指令,基于快照机制从备份集群向主集群进行数据回迁,上述方式,通过将第一集群200与第二集群212进行切换,尽可能减少对于用户侧的影响,提升用户侧的使用体验。

[0047] 在一个实施场景中,第一集群200为主集群,第二集群212为备份集群,当第一集群200出现故障时,可以进行主备切换,即将第一集群200切换为备份集群,将第二集群212切换为主集群,对于集群之间的切换,不仅用户设备可以通过云管平台209对主备集群进行切换,并且基于数据的存储过程,当数据备份设备208对集群的访问或者对数据的读取出现故障时,数据备份设备208可以基于云管平台209发出切换指令,备份代理210可以对第一集群200与第二集群212进行切换。当第一集群200与第二集群212进行切换时,云管平台209将调用备份代理210恢复虚拟机磁盘,该过程中会对虚拟机系统盘、数据盘以及存储元数据(卷ID、名称、大小、类型等)进行恢复操作,进而减少对于用户侧的正常使用。

[0048] 在一个实施场景中,第一集群200与第二集群212中对虚拟机磁盘恢复完成后,云管平台209将根据存储的实例云数据信息创建容灾虚拟机,在此创建过程中涉及一系列前置资源的检查、创建与配置(如首先检查实例网络是否存在,若存在检查参数是否一致,若不存在将进行网络创建操作),最终云管平台209会调用API创建容灾虚拟机,完成网卡和磁盘挂载,恢复出参数一致的容灾虚拟机,主集群恢复之后,由用户设备发起回迁操作,按照类似切换第一集群200与第二集群212功能进行虚拟机回迁操作,云管平台209中有点击按钮,点击按钮之后,将会显示是否选择回迁选择页面,用户对数据内容进行评估并且选择,若选择不回迁,则第一集群200与第二集群212再次进行切换,切换后的第一集群200为主集群,第二集群212为备份集群;若选择回迁,即对集群中数据进行保存,再次对第一集群200与第二集群212执行备份操作,将第二集群212中的数据备份至第一集群200,第一集群200与第二集群212再进行切换,切换后的第一集群200为主集群,第二集群212为备份集群。

[0049] 在一些公开实施例中,在执行备份操作的过程中,还可以基于已发送操作指令至第一集群200或第二集群212,发送询问指令至第一集群200或第二集群212,且询问指令用于询问第一集群200或第二集群212关于操作指令的完成情况,接收第一集群200或第二集群212的反馈指令,且反馈指令包括以下任一者:基于已完成操作指令而反馈的完成指令、基于已收到操作指令但未完成操作指令的确认指令,第一集群200、第二集群212基于判断

模块判断备份操作过程中的每个接收指令是否已完成,若已完成回复完成指令,若未完成回复确认指令,上述方式,第一集群200与第二集群212通过对问询指令的处理,得到是否已经执行过问询指令内容,进一步避免重复相同的操作指令而导致备份效率降低,提高了用户侧的使用体验。

[0050] 在一个实施场景中,发送操作指令至第一集群200或第二集群212之后,第一集群200或第二集群212基于收到的操作指令进行确认,对于已经完成的操作指令,第一集群200或第二集群212不会再次进行操作,并发送已完成操作指令至数据备份设备208中的备份代理210,使得数据备份设备208中对操作指令不会再次进行发送;若第一集群200或第二集群212基于收到的操作指令进行确认,第一集群200或第二集群212未完成操作指令,第一集群200或第二集群212发送确认指令至数据备份设备208,使备份指令中备份代理210确认指令发送成功,第一集群200或第二集群212发送确认指令后,执行操作指令。

[0051] 在一个实施场景中,备份代理210基于收到第一集群200和第二集群212的反馈指令进行确认,若第一集群200或第二集群212发送的指令为完成指令,备份代理210接收到的反馈指令内容经处理后发送至云管平台209,若第一集群200或第二集群212发送的指令为确认指令,云管平台209再次确认指令并且处理反馈至云管平台209。云管平台209再次发送指令至备份代理210时,备份代理210需对接收指令进行确认,若接收指令为当前重复指令,返回当前反馈指令至云管平台209,若接收指令为新的命令指令,备份代理210将命令指令进行处理并且发送至第一集群200或第二集群212。

[0052] 在一些公开实施例中,第一集群200包含以分布式布置的多个存储镜像,且数据均存入存储镜像中,上述方式,通过在底层存储采用分布式存储,运维部署更加简单,并且分布式存储可以更好地统一利用低成本的磁盘,最终对外提供更高容量的存储池,并且降低节点死机对于存储系统影响,在不需要仲裁机制情况下,满足云主机主备集群数据备份,并尽可能减少数据丢失。

[0053] 具体地,存储镜像用于存储数据,如图2所述,第一集群200中存储镜像可以包括但不限于存储盘206,第二集群212中存储镜像可以包括但不限于存储盘I217、存储盘II218和基础存储盘219。

[0054] 请参阅图5,图5是本申请数据备份装置30一实施例的框架示意图。数据备份装置30,包括检测模块31、选择模块32和备份模块33,其中,检测模块31用于响应备份指令,对第一集群200进行数据检测,得到数据检测结果,且数据检测结果包括第一集群200中是否存入目标数据;选择模块32用于基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制;备份模块33用于基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群212。

[0055] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽

可能降低数据丢失对于用户侧业务影响。

[0056] 在一些公开实施例中,基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,包括在检测结果包括第一集群200中存入目标数据的情况下,选择快照机制作为目标备份机制;在检测结果包括第一集群200中未存入目标数据的情况下,选择日志机制作为目标备份机制。

[0057] 因此,通过对不同的检测结果选择不同的备份机制,将集群中存储的数据快速进行备份,进一步减少了数据混乱造成的数据未备份,进而造成数据丢失的情况。

[0058] 在一些公开实施例中,目标备份机制为快照机制,基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群212,包括对第一集群200中目标数据执行快照标记,得到目标数据的数据副本;再基于目标数据的数据副本恢复得到目标数据,并将目标数据发送至第二集群212。

[0059] 因此,可以在短时间内将目标数据快速备份,快照备份机制是源于某一个时间点对存储空间的数据内容进行快照备份,可以尽可能减少在备份的同时进行数据写入造成的数据遗漏问题,并且快照机制不需要完整的数据空间,进而达到不仅可以快速备份并且节约了存储空间的目的。

[0060] 在一个公开实施例中,在基于目标数据的数据副本恢复得到目标数据,并将目标数据发送至第二集群212之后,方法还包括删除第一集群200中数据副本。

[0061] 因此,通过删除数据副本,避免了数据再次被恢复并且重新发送至第二集群212的情况,提高了数据备份的高效性。

[0062] 在一个公开实施例中,目标备份机制为日志机制,第一集群200还存储有操作日志,操作日志用于存储数据操作记录,基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群212,包括从第一集群200复制目标数据的数据操作记录,并基于目标数据的数据操作记录恢复出目标数据,将目标数据发送至第二集群212。

[0063] 因此,通过日志备份机制将数据进行备份,能够快速地将数据恢复出来,并且尽可能减少了还原数据所需要的时间,进一步减少数据丢失的情况。

[0064] 在一个公开实施例中,在将目标数据发送至第二集群212之后,方法还包括删除第一集群200中目标数据的数据操作记录。

[0065] 因此,确保将目标数据备份至第二集群212后,数据备份设备208不会对目标数据再次进行恢复,避免了数据的重复恢复,并且减少了数据多次存储占用的空间,尽可能提高了数据备份的效率。

[0066] 在一个公开实施例中,备份指令是用户设备在存储目标数据至第一集群200时由用户发送的;或备份指令是用户设备在存储目标数据至第一集群200时基于预设备份策略而自动发送的。

[0067] 因此,两种备份指令的发送可以避免用户在数据存入的过程中,遗忘对数据执行备份操作而导致的数据丢失问题,进一步降低了数据的丢失情况。

[0068] 在一个公开实施例中,第一集群200为主集群,且第二集群212为备份集群,方法还包括响应于切换指令,切换第一集群200为备份集群,并切换第二集群212为主集群;响应于回迁指令,基于快照机制从备份集群向主集群进行数据回迁。

[0069] 因此,通过将第一集群200与第二集群212进行切换,尽可能减少对于用户侧的影

响,提升用户侧的使用体验。

[0070] 在一个公开实施例中,在响应于备份指令,对第一集群200进行数据检测,得到数据检测结果之前,方法还包括获取第一集群200、第二集群212的元数据;且元数据包括第一集群200的第一网络接口和第二集群212的第二网络接口,第一网络接口用于实现与第一集群200之间的数据传输,第二网络接口用于实现与第二集群212之间的数据传输。

[0071] 因此,可以实现数据备份设备208与集群之间的数据传输,还可以实现在一个集群重新连接时,不会对集群中的任一存储系统发生遗漏。

[0072] 在一个公开实施例中,在执行备份操作的过程中,方法还包括基于已发送操作指令至第一集群200和/或第二集群212,发送问询指令至第一集群200和/或第二集群212;且问询指令用于问询第一集群200和/或第二集群212关于操作指令的完成情况;接收第一集群200和/或第二集群212的反馈指令;且反馈指令包括以下任一者:基于已完成操作指令而反馈的完成指令、基于已收到操作指令但未完成操作指令的确认指令,第一集群200、第二集群212基于判断模块判断备份操作过程中的每个接收指令是否已完成,若已完成回复完成指令;若未完成回复确认指令。

[0073] 因此,第一集群200与第二集群212通过对问询指令的处理,得到是否已经执行过问询指令内容,进一步避免重复相同的操作指令而导致备份效率降低,提高了用户侧的使用体验。

[0074] 在一个公开实施例中,第一集群200包含以分布式布置的多个存储镜像,且数据均存入存储镜像中。

[0075] 因此,通过在底层存储采用分布式存储,运维部署更加简单,并且分布式存储可以更好地统一利用低成本的磁盘,最终对外提供更高容量的存储池,并且降低节点死机对于存储系统影响,在不需要仲裁机制情况下,满足云主机主备集群数据备份,并尽可能减少数据丢失。

[0076] 请参阅图6,图6是本申请数据备份设备一实施例的框架示意图。数据备份设备208包括存储器41、通信电路43和处理器42,且存储器41和通信电路43耦接于处理器42,存储器41中存储有程序指令,处理器42用于执行程序指令,以实现上述任一数据备份方法实施例中的步骤。关于数据备份设备208的具体含义,可以结合参阅前述公开实施例以及图2,在此不再赘述。

[0077] 具体而言,处理器42用于控制其自身以及存储器41以实现上述任一数据备份方法实施例中的步骤。处理器42还可以称为CPU (Central Processing Unit,中央处理单元)。处理器42可能是一种集成电路芯片,具有信号的处理能力。处理器42还可以是通用处理器、数字信号处理器 (Digital Signal Processor, DSP)、专用集成电路 (Application Specific Integrated Circuit, ASIC)、现场可编程门阵列 (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。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。另外,处理器42可以由集成电路芯片共同实现。

[0078] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标

备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽可能降低数据丢失对于用户侧业务影响。

[0079] 在一个实施场景中,数据备份系统,包括数据备份设备和多个集群,数据备份设备分别与多个集群通信连接,数据备份设备用于执行上述任一数据备份方法实施例中的步骤,以在多个集群之间实现数据备份。

[0080] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽可能降低数据丢失对于用户侧业务影响。

[0081] 请参阅图7,图7是本申请计算机可读存储介质一实施例的框架示意图。计算机可读存储介质50存储有能够被处理器运行的程序指令51,程序指令51用于实现上述任一数据备份方法实施例中的步骤。

[0082] 上述方案,响应于备份指令,对第一集群进行数据检测,得到数据检测结果,数据检测结果包括第一集群中是否存入目标数据,再基于数据检测结果,在多种预设备份机制中选择得到目标备份机制,且多种预设备份机制至少包括快照机制、日志机制,再基于目标备份机制,对目标数据执行备份操作,以将目标数据备份至第二集群。一方面第一集群与第二集群之间并无业务关联,即两者之间相互独立,任一集群出现故障并不会影响另一集群正常工作,另一方面基于数据检测结果在多种预设备份机制中选择得到目标备份机制,再将目标数据进行备份,多种备份机制的灵活选择有助于减少数据丢失的情况。故此,能够尽可能降低数据丢失对于用户侧业务影响。

[0083] 在一些实施例中,本公开实施例提供的装置具有的功能或包含的模块可以用于执行上文方法实施例描述的方法,其具体实现可以参照上文方法实施例的描述,为了简洁,这里不再赘述。

[0084] 上文对各个实施例的描述倾向于强调各个实施例之间的不同之处,其相同或相似之处可以互相参考,为了简洁,本文不再赘述。

[0085] 在本申请所提供的几个实施例中,应该理解到,所揭露的方法和装置,可以通过其它的方式实现。例如,以上所描述的装置实施方式仅仅是示意性的,例如,模块或单元的划分,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。

[0086] 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方式方案的目的。

[0087] 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。

[0088] 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,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)或处理器(processor)执行本申请各个实施方式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。

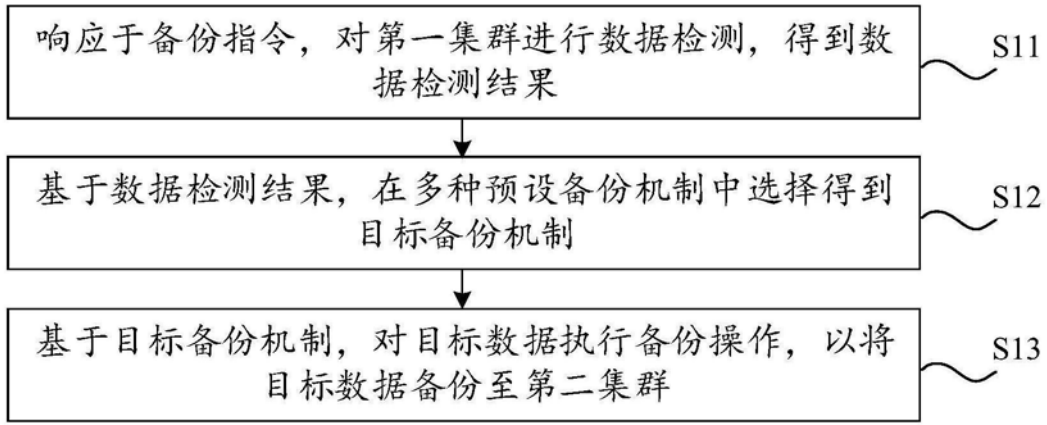


图1

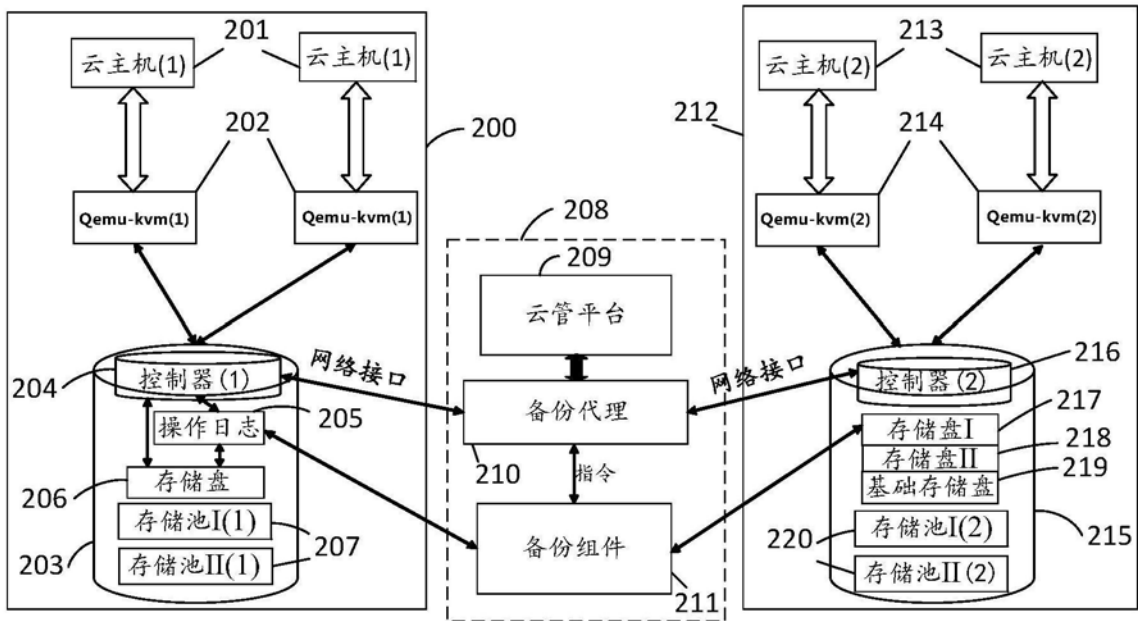


图2

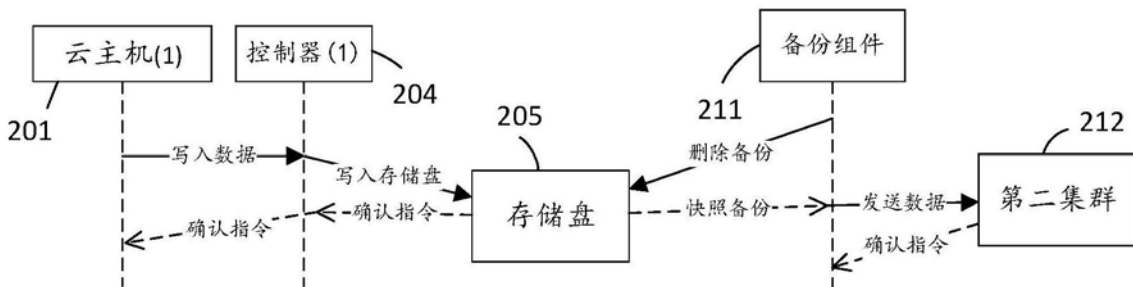


图3

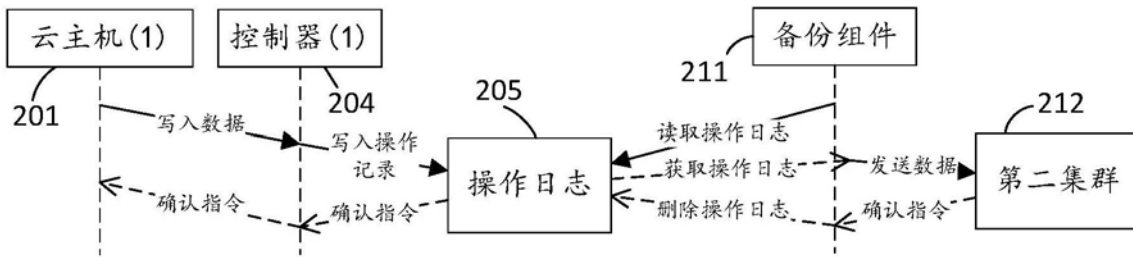


图4

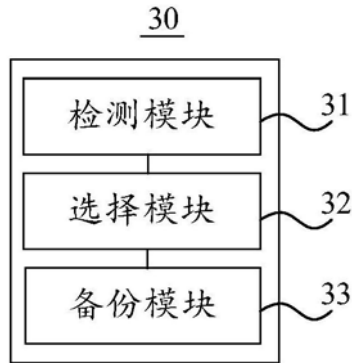


图5

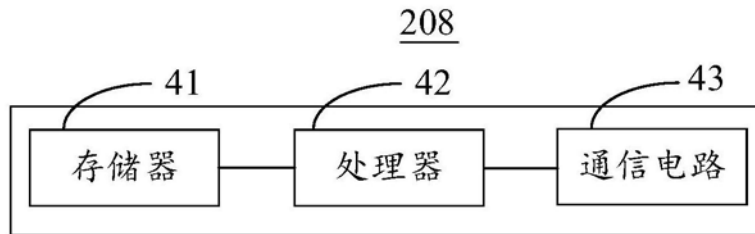


图6

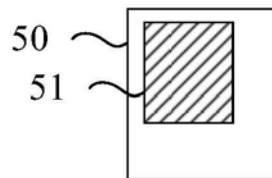


图7