



(12)发明专利申请

(10)申请公布号 CN 111352959 A  
(43)申请公布日 2020.06.30

(21)申请号 202010131442.5

(22)申请日 2020.02.28

(71)申请人 中国工商银行股份有限公司  
地址 100140 北京市西城区复兴门内大街  
55号

(72)发明人 武文斌 王辉 刘静 郑宝城

(74)专利代理机构 北京三友知识产权代理有限公司 11127  
代理人 赵平 周永君

(51) Int. Cl.  
G06F 16/2455(2019.01)  
G06F 16/23(2019.01)  
G06F 16/27(2019.01)

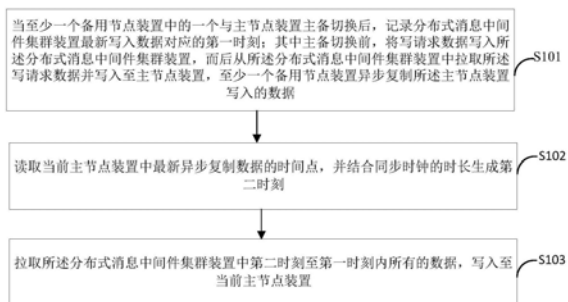
权利要求书4页 说明书16页 附图5页

(54)发明名称

数据同步补救、存储方法及集群装置

(57)摘要

申请提供一种数据同步补救、存储方法及集群装置,当需要主备切换时,本发明中在保证具有存储节点装置高可用能力的前提下,通过分布式消息中间件集群装置记录主节点装置所有写请求数据包,并在备节点装置中记录同步主节点装置数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件集群装置可以根据时间戳来查找对应的数据包的特性,结合时间戳回退,找到主备节点装置间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点装置中,从而避免了主备切换后的数据丢失。



1. 一种数据同步补救方法,应用于代理服务器集群装置,其特征在于,包括:

当至少一个备用节点装置中的一个与主节点装置主备切换后,记录分布式消息中间件集群装置最新写入数据对应的第一时刻;其中主备切换前,将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;

拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

2. 根据权利要求1所述的数据同步补救方法,其特征在于,所述第一时刻和第二时刻以时间戳或偏移量的形式被记录,在记录分布式消息中间件集群装置最新写入数据对应的第一时刻之后,所述数据同步补救方法还包括:

将所述第一时刻对应的时间戳或偏移量发送到分布式协调装置;

相对应地,在生成第二时刻后,所述数据同步补救方法还包括:

将所述第二时刻对应的时间戳或偏移量发送到分布式协调装置。

3. 根据权利要求2所述的数据同步补救方法,其特征在于,当至少一个备用节点装置中的一个与主节点装置主备切换后,所述当前主节点装置被监控系统集群装置配置为可写不可读状态;

当从所述分布式消息中间件集群装置拉取到第一时刻的数据后,所述数据同步补救方法还包括:

向所述分布式协调装置反馈补数完成事件,以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

4. 根据权利要求1所述的数据同步补救方法,其特征在于,所述读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻,包括:

设定一大于所述同步时钟的预设时长;

以所述当前主节点装置中最新异步复制数据的时间点向前回退所述预设时长,得到所述第二时刻。

5. 一种数据同步补救方法,应用于监控系统集群装置,其特征在于,包括:

检测主节点装置是否宕机;

若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;其中,主备切换前,将数据写入分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻;代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入

至当前主节点装置；

所述数据同步补救方法，还包括：

当当前主节点装置数据写入完成后，修改当前主节点装置为可读可写状态。

6. 一种数据同步补救方法，应用于分布式消息中间件集群装置，其特征在于，包括：

主备切换前，写入写请求数据，而后按序写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；

主备切换后，响应于代理服务器集群装置请求反馈最新写入数据对应的第一时刻；以使所述代理服务器集群装置记录所述第一时刻；

响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置；以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置；其中，

所述第二时刻根据当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成。

7. 根据权利要求6所述的数据同步补救方法，其特征在于，所述第一时刻和所述第二时刻以时间戳或偏移量的形式被记录。

8. 一种数据存储方法，应用于代理服务器集群装置，其特征在于，在包括：

接收写请求数据，并将所述写请求数据发送至分布式消息中间件集群装置；

重新读取所述分布式消息中间件集群装置中所述写请求数据的数据，写入至主节点装置；其中，至少一个备用节点装置异步复制所述主节点装置写入的数据。

9. 一种数据存储方法，应用于分布式消息中间件集群装置，其特征在于，在包括：

写入代理服务器集群装置发送的写请求数据；

根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置，以使所述代理服务器集群装置将所述写请求数据写入至主节点装置，其中，至少一个备用节点装置异步复制所述主节点装置写入的数据。

10. 一种代理服务器集群装置，其特征在于，包括：

第一时刻记录模块，当至少一个备用节点装置中的一个与主节点装置主备切换后，记录分布式消息中间件集群装置最新写入数据对应的第一时刻；其中主备切换前，将写请求数据写入所述分布式消息中间件集群装置，而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；

第二时刻生成模块，读取当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成第二时刻；

数据拉取模块，拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据，写入至当前主节点装置。

11. 根据权利要求10所述的代理服务器集群装置，其特征在于，所述第一时刻和第二时刻以时间戳或偏移量的形式被记录，所述代理服务器集群装置还包括：

第一时刻发送模块，将所述第一时刻对应的时间戳或偏移量发送到分布式协调装置；

第二时刻发送模块，将所述第二时刻对应的时间戳或偏移量发送到分布式协调装置。

12. 根据权利要求11所述的代理服务器集群装置，其特征在于，当至少一个备用节点装

置中的一个与主节点装置主备切换后,所述当前主节点装置被监控系统集群装置配置为可写不可读状态;

所述代理服务器集群装置还包括:

补数完成反馈模块,当从所述分布式消息中间件集群装置拉取到第一时刻的数据后,向所述分布式协调装置反馈补数完成事件,以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

13. 根据权利要求10所述的代理服务器集群装置,其特征在于,还包括:

写请求数据收发模块,接收写请求数据,并将所述写请求数据发送至分布式消息中间件集群装置;

写请求数据重新读取模块,在重新读取所述分布式消息中间件集群装置中所述写请求数据的数据,写入至主节点装置;其中,至少一个备用节点装置异步复制所述主节点装置写入的数据。

14. 一种监控系统集群装置,其特征在于,包括:

宕机检测模块,检测主节点装置是否宕机;

主备关系断开模块,若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

路由信息修改模块,修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;其中,

主备切换前,将数据写入分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;并且,

当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻;代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置;

所述监控系统集群装置,还包括:

当前主节点装置状态修改模块,当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

15. 一种分布式消息中间件集群装置,其特征在于,包括:

写请求数据写入模块,主备切换前,写入写请求数据,而后按序写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

第一时刻反馈模块,主备切换后,响应于代理服务器集群装置的请求反馈最新写入数据对应的第一时刻;以使所述代理服务器集群装置记录所述第一时刻;

数据拉取模块,响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置;以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置;其中,

所述第二时刻根据当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成。

16. 根据权利要求15所述的分布式消息中间件集群装置,其特征在于,所述第一时刻和所述第二时刻以时间戳或偏移量的形式被记录。

17. 根据权利要求15所述的分布式消息中间件集群装置,其特征在于,还包括:

读取请求反馈模块,根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置,以使所述代理服务器集群装置将所述写请求数据写入至主节点装置。

18. 一种分布式数据存储系统,其特征在于,包括:

代理服务器集群装置、分布式消息中间件集群装置、主节点装置以及至少一个备用节点装置;

主备切换前,代理服务器集群装置将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻,然后读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻,最后进一步拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

19. 根据权利要求18所述的分布式数据存储系统,其特征在于,还包括:

监控系统集群装置,所述监控系统集群装置包括:

宕机检测模块,检测主节点装置是否宕机;

主备关系断开模块,若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

路由信息修改模块,修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;

当前主节点装置状态修改模块,当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

20. 根据权利要求19所述的分布式数据存储系统,其特征在于,还包括:

分布式协调装置,接收所述代理服务器集群装置反馈的补数完成事件,并发送至所述监控系统集群装置,以告知所述监控系统集群装置当前主节点装置数据写入完成。

21. 一种电子设备,包括存储器、处理器及存储在存储器上并可在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程序时实现权利要求1至7任一项所述的数据同步补救方法,或者权利要求8-9任一项所述的数据存储方法。

22. 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,该计算机程序被处理器执行时实现权利要求1至7任一项所述的数据同步补救方法,或者权利要求8-9任一项所述的数据存储方法。

## 数据同步补救、存储方法及集群装置

### 技术领域

[0001] 本申请涉及数据同步补救技术领域,具体涉及数据同步补救、存储方法及集群装置。

### 背景技术

[0002] 很多追求高性能的分布式存储系统的主备数据同步机制为异步复制,所以当主节点装置出现故障并触发主备切换时,会有部分数据丢失的风险。

### 发明内容

[0003] 针对现有技术中的问题,本申请提供一种数据同步补救、存储方法及集群装置,能够在不影响整个系统对外服务的高性能基础上,保证了数据零丢失且不中断对外服务。

[0004] 为解决上述技术问题,本申请提供以下技术方案:

[0005] 第一方面,一种数据同步补救方法,应用于代理服务器集群装置,包括:

[0006] 当至少一个备用节点装置中的一个与主节点装置主备切换后,记录分布式消息中间件集群装置最新写入数据对应的第一时刻;其中主备切换前,将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0007] 读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;

[0008] 拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

[0009] 优选的实施例中,所述第一时刻和第二时刻以时间戳或偏移量的形式被记录,在记录分布式消息中间件集群装置最新写入数据对应的第一时刻之后,所述数据同步补救方法还包括:

[0010] 将所述第一时刻对应的的时间戳或偏移量发送到分布式协调装置;

[0011] 相对应地,在生成第二时刻后,所述数据同步补救方法还包括:

[0012] 将所述第二时刻对应的的时间戳或偏移量发送到分布式协调装置。

[0013] 优选的实施例中,当至少一个备用节点装置中的一个与主节点装置主备切换后,所述当前主节点装置被监控系统集群装置配置为可写不可读状态;

[0014] 当从所述分布式消息中间件集群装置拉取到第一时刻的数据后,所述数据同步补救方法还包括:

[0015] 向所述分布式协调装置反馈补数完成事件,以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

[0016] 优选的实施例中,所述主节点装置与备用节点装置之间异步复制产生的时差为3s。

[0017] 第二方面,一种数据同步补救方法,应用于监控系统集群装置,包括:

- [0018] 检测主节点装置是否宕机；
- [0019] 若是，断开至少一个备用节点装置与所述主节点装置的主备关系，并将其中一个备用节点装置置为可写不可读状态；
- [0020] 修改对应的路由信息，将所述其中一个备用节点装置的身份变更为主节点装置；其中，主备切换前，将数据写入所述分布式消息中间件集群装置，而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；
- [0021] 当至少一个备用节点装置中的一个与主节点装置主备切换后，代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻；代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成第二时刻；而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据，写入至当前主节点装置；
- [0022] 所述数据同步补救方法，还包括：
- [0023] 当当前主节点装置数据写入完成后，修改当前主节点装置为可读可写状态。
- [0024] 第三方面，一种数据同步补救方法，应用于分布式消息中间件集群装置，包括：
- [0025] 主备切换前，写入写请求数据，而后按序写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；
- [0026] 主备切换后，响应于代理服务器集群装置请求反馈最新写入数据对应的第一时刻；以使所述代理服务器集群装置记录所述第一时刻；
- [0027] 响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置；以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置；其中，
- [0028] 所述第二时刻根据当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成。
- [0029] 优选的实施例中，所述第一时刻和所述第二时刻以时间戳或偏移量的形式被记录。
- [0030] 第三方面，一种数据存储方法，应用于代理服务器集群装置，在包括：
- [0031] 接收写请求数据，并将所述写请求数据发送至分布式消息中间件集群装置；
- [0032] 重新读取所述分布式消息中间件集群装置中所述写请求数据的数据，写入至主节点装置；其中，至少一个备用节点装置异步复制所述主节点装置写入的数据。
- [0033] 优选的实施例中，还包括：
- [0034] 记录预设的主节点装置与备用节点装置之间异步复制产生的时差。
- [0035] 第四方面，一种数据存储方法，应用于分布式消息中间件集群装置，在包括：
- [0036] 写入代理服务器集群装置发送的写请求数据；
- [0037] 根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置，以使所述代理服务器集群装置将所述写请求数据写入至主节点装置，其中，至少一个备用节点装置异步复制所述主节点装置写入的数据。
- [0038] 第五方面，一种数据存储方法，应用于主节点装置，在包括：
- [0039] 接收代理服务器集群装置发送的写请求数据；其中所述写请求数据通过所述代理

服务器集群装置读取分布式消息中间件集群装置得到；

[0040] 向至少一个备用节点装置异步写入所述写请求数据。

[0041] 第六方面，一种数据存储方法，应用于备用节点装置，包括：

[0042] 接收主节点装置异步写入的写请求数据；其中，代理服务器集群装置将所述写请求数据发送至所述主节点装置；所述写请求数据通过所述代理服务器集群装置读取分布式消息中间件集群装置得到。

[0043] 第七方面，一种代理服务器集群装置，包括：

[0044] 第一时刻记录模块，当至少一个备用节点装置中的一个与主节点装置主备切换后，记录分布式消息中间件集群装置最新写入数据对应的第一时刻；其中主备切换前，将写请求数据写入所述分布式消息中间件集群装置，而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；

[0045] 第二时刻生成模块，读取当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成第二时刻；

[0046] 数据拉取模块，拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据，写入至当前主节点装置。

[0047] 优选的实施例中，所述第一时刻和第二时刻以时间戳或偏移量的形式被记录，所述代理服务器集群装置还包括：

[0048] 第一时刻发送模块，将所述第一时刻对应的的时间戳或偏移量发送到分布式协调装置；

[0049] 第二时刻发送模块，将所述第二时刻对应的的时间戳或偏移量发送到分布式协调装置。

[0050] 优选的实施例中，当至少一个备用节点装置中的一个与主节点装置主备切换后，所述当前主节点装置被监控系统集群装置配置为可写不可读状态；

[0051] 所述代理服务器集群装置还包括：

[0052] 当从所述分布式消息中间件集群装置拉取到第一时刻的数据后，向所述分布式协调装置反馈补数完成事件，以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

[0053] 优选的实施例中，所述主节点装置与备用节点装置之间异步复制产生的时差为3s。

[0054] 优选的实施例中，还包括：

[0055] 写请求数据收发模块，接收写请求数据，并将所述写请求数据发送至分布式消息中间件集群装置；

[0056] 写请求数据重新读取模块，在重新读取所述分布式消息中间件集群装置中所述写请求数据的数据，写入至主节点装置；其中，至少一个备用节点装置异步复制所述主节点装置写入的数据。

[0057] 优选的实施例中，还包括：

[0058] 时差记录模块，记录预设的主节点装置与备用节点装置之间异步复制产生的时差。



[0059] 第八方面,一种监控系统集群装置,包括:

[0060] 宕机检测模块,检测主节点装置是否宕机;

[0061] 主备关系断开模块,若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

[0062] 路由信息修改模块,修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;其中,

[0063] 主备切换前,将数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;并且,

[0064] 当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻;代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置;

[0065] 所述监控系统集群装置,还包括:

[0066] 当前主节点装置状态修改模块,当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

[0067] 第九方面,一种分布式消息中间件集群装置,包括:

[0068] 写请求数据写入模块,主备切换前,写入写请求数据,而后按序写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0069] 第一时刻反馈模块,主备切换后,响应于代理服务器集群装置请求反馈最新写入数据对应的第一时刻;以使所述代理服务器集群装置记录所述第一时刻;

[0070] 数据拉取模块,响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置;以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置;其中,

[0071] 所述第二时刻根据当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成。

[0072] 优选的实施例中,所述第一时刻和所述第二时刻以时间戳或偏移量的形式被记录。

[0073] 优选的实施例中,还包括:

[0074] 读取请求反馈模块,根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置,以使所述代理服务器集群装置将所述写请求数据写入至主节点装置。

[0075] 第十方面,一种主节点装置,包括:

[0076] 写请求数据接收模块,接收代理服务器集群装置发送的写请求数据;其中所述写请求数据通过所述代理服务器集群装置读取分布式消息中间件集群装置得到;

[0077] 异步写入模块,在向至少一个备用节点装置异步写入所述写请求数据。

[0078] 第十一方面,一种备用节点装置,包括:

[0079] 写请求数据异步写入模块,接收主节点装置异步写入的写请求数据;其中,代理服

服务器集群装置将所述写请求数据发送至所述主节点装置；所述写请求数据通过所述代理服务器集群装置读取分布式消息中间件集群装置得到。

[0080] 第十二方面，一种分布式数据存储系统，包括：

[0081] 代理服务器集群装置、分布式消息中间件集群装置、主节点装置以及至少一个备用节点装置；

[0082] 主备切换前，代理服务器集群装置将写请求数据写入所述分布式消息中间件集群装置，而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置，至少一个备用节点装置异步复制所述主节点装置写入的数据；

[0083] 当至少一个备用节点装置中的一个与主节点装置主备切换后，代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻，然后读取当前主节点装置中最新异步复制数据的时间点，并结合同步时钟的时长生成第二时刻，最后进一步拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据，写入至当前主节点装置。

[0084] 在某些实施例中，还包括：

[0085] 监控系统集群装置，所述监控系统集群装置包括：

[0086] 宕机检测模块，检测主节点装置是否宕机；

[0087] 主备关系断开模块，若是，断开至少一个备用节点装置与所述主节点装置的主备关系，并将其中一个备用节点装置置为可写不可读状态；

[0088] 路由信息修改模块，修改对应的路由信息，将所述其中一个备用节点装置的身份变更为主节点装置；

[0089] 当前主节点装置状态修改模块，当当前主节点装置数据写入完成后，修改当前主节点装置为可读可写状态。

[0090] 在某些实施例中，还包括：

[0091] 分布式协调装置，接收所述代理服务器集群装置反馈的补数完成事件，并发送至所述监控系统集群装置，以告知所述监控系统集群装置当前主节点装置数据写入完成。

[0092] 第十三方面，本申请提供一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时实现所述的数据补救或者数据存储方法。

[0093] 第十四方面，本申请提供一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现所述的数据补救或者数据存储方法。

[0094] 由上述技术方案可知，本申请提供的数据同步补救、存储方法及集群装置，在数据同步时，写入的数据首先通过代理服务器集群装置传输到分布式消息中间件集群装置，而后通过读取分布式消息中间件集群装置中的数据，并写入到主备节点装置，完成数据同步，当需要主备切换时，本发明中在保证具有存储节点装置高可用能力的前提下，通过分布式消息中间件集群装置记录主节点装置所有写请求数据包，并在备节点装置中记录同步主节点装置数据时最后一条数据完成同步的时间戳，当出现主备切换时，利用了分布式消息中间件集群装置可以根据时间戳来查找对应的数据包的特性，结合时间戳回退，找到主备节点装置间的差异数据，并将这些数据通过补数机制写入主备切换后新的主节点装置中，从而避免了主备切换后的数据丢失。

## 附图说明

[0095] 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。

[0096] 图1为本申请分布式数据存储系统的结构示意图。

[0097] 图2为本申请实施例中一种数据同步补救方法的流程示意图之一。

[0098] 图3为本申请实施例中一种数据同步补救方法的流程示意图之二。

[0099] 图4为本申请实施例中一种数据同步补救方法的流程示意图之三。

[0100] 图5为本申请实施例中一种数据存储方法的流程示意图之一。

[0101] 图6为本申请实施例中一种数据存储方法的流程示意图之二。

[0102] 图7为本申请应用实例中代理服务器集群装置的结构示意图。

[0103] 图8为本申请应用实例中监控系统集群装置的结构示意图。

[0104] 图9为本申请应用实例中分布式消息中间件集群装置的结构示意图。

[0105] 图10为本申请实施例中的电子设备的结构示意图。

## 具体实施方式

[0106] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整的描述,显然,所描述的实施例是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0107] 考虑到很多追求高性能的分布式存储系统的主备数据同步机制为异步复制,所以当主节点装置出现故障并触发主备切换时,会有部分数据丢失的风险。本申请提供一种数据补救或者数据存储方法、服务器、装置、分布式消息中间件集群装置、电子设备和计算机可读存储介质,在数据同步时,写入的数据首先通过代理服务器集群装置传输到分布式消息中间件集群装置,而后通过读取分布式消息中间件集群装置中的数据,并写入到主备节点装置,完成数据同步,当需要主备切换时,本发明中在保证具有存储节点装置高可用能力的前提下,通过分布式消息中间件集群装置记录主节点装置所有写请求数据包,并在备节点装置中记录同步主节点装置数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件集群装置可以根据时间戳来查找对应的数据包的特性,找到主备节点装置间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点装置中,从而避免了主备切换后的数据丢失。

[0108] 在本申请的一个或多个实施例中,如图1所示,分布式存储系统具体包括:应用服务器集群装置1、代理服务器集群装置2、分布式消息中间件集群装置4、分布式协调装置3、监控系统集群装置9、主节点装置以及至少一个备用节点装置(图1中的5和7是主节点装置的其中两个示例,6和8是备用节点装置的其中两个示例),其中主节点装置和备用节点装置属于数据存储集群。

[0109] 代理服务器集群装置2、分布式消息中间件集群装置4、主节点装置以及至少一个备用节点装置等可以是集群装置,整个系统属于分布式存储系统。

[0110] 在本申请的一个或多个实施例中,如图2所示的一种数据同步补救方法,该方法由代理服务器集群装置执行,具体如下:

[0111] S101:当至少一个备用节点装置中的一个与主节点装置主备切换后,记录分布式消息中间件集群装置最新写入数据对应的第一时刻;其中主备切换前,将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0112] S102:读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;

[0113] S103:拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

[0114] 可以理解的是,各个服务器之间可以使用任何合适的网络协议进行通信,包括在本申请提交日尚未开发出的网络协议。所述网络协议例如可以包括TCP/IP协议、UDP/IP协议、HTTP协议、HTTPS协议等。当然,所述网络协议例如还可以包括在上述协议之上使用的RPC协议(Remote Procedure Call Protocol,远程过程调用协议)、REST协议(Representational State Transfer,表述性状态转移协议)等。

[0115] 由于主备节点间的数据同步机制为异步复制,所以存储主节点宕机时会出现部分数据还未同步至存储备节点,因此读取当前主节点装置中最新异步复制数据的时间点可以知晓当前同步的数据。

[0116] 在本申请的一个或多个实施例中,在数据存储/写入时,首先将数据写入到分布式消息中间件集群装置,然后从分布式消息中间件集群装置中调取数据到主节点装置中,以此使分布式消息中间件集群装置中存储有主节点装置中的所有数据。

[0117] 分布式消息中间件集群装置以将数据以消息队列的形式进行排序,并按序进行消费,借助于分布式消息中间件集群装置,数据本身按照其写入数据的时间戳或者偏移量进行记录,从而可以锚定数据处于序列中的位置。

[0118] 在本申请的一个或多个实施例中,步骤S102具体包括:

[0119] S210:设定一大于所述同步时钟的预设时长;

[0120] S220:以所述当前主节点装置中最新异步复制数据的时间点向前回退所述预设时长,得到所述第二时刻。

[0121] 根据生产环境的时钟同步机制,所有设备都会在例如2秒内进行一次时钟源同步,保证设备之间的时钟差异一定在2秒以内,该实施例中通过设定一个大于同步时钟的预设时长,例如3秒,通过时间戳向前移动3秒来保证数据的0丢失。

[0122] 在本申请的一个或多个实施例中,第一时刻和第二时刻以时间戳或偏移量的形式被记录(分布式消息中间件集群装置的特性),在记录分布式消息中间件集群装置最新写入数据对应的第一时刻之后,所述数据同步补救方法还包括:

[0123] 将所述第一时刻对应的时间戳或偏移量发送到分布式协调装置3;

[0124] 相对应地,在生成第二时刻后,所述数据同步补救方法还包括:

[0125] 将所述第二时刻对应的时间戳或偏移量发送到分布式协调装置3。

[0126] 分布式协调装置主要起到反馈等辅助功能,具体的,分布式协调装置可以反馈补

数完成事件等。该实施例中,当至少一个备用节点装置中的一个与主节点装置主备切换后,所述当前主节点装置被监控系统集群装置配置为可写不可读状态;

[0127] 当从所述分布式消息中间件集群装置拉取到第一时刻的数据后,所述数据同步补救方法还包括:

[0128] 向所述分布式协调装置反馈补数完成事件,以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

[0129] 举例来说,请继续结合图1所示,代理服务器集群装置2收到主备切换事件后,记录当前消费分布式消息中间件集群装置4的最新lastOffset,并将此lastOffset写入分布式协调装置3中。代理服务器集群装置2获取数据存储集群分片主节点装置6在作为备节点时同步主节点的最后一条数据的时间戳,然后减3秒,以此时间戳从分布式消息中间件装置4中找到对应的数据startOffset,并写入分布式协调器集群装3中。代理服务器集群装置2从分布式消息中间件集群装置4中的startOffset开始拉取数据,并将数据写入数据存储集群分片主节点装置6中;代理服务器集群装置2从分布式消息中间件集群装置4消费至lastOffset时,向分布式协调装置3反馈补数完成。

[0130] 原有分布式数据存储的高可用机制下,由于主备数据为异步复制,从而导致当主节点故障触发主备切换时,备节点数据将存在与主节点不一致的可能,从而导致应用感知丢数。本发明中:在保证具有存储节点高可用能力的前提下,通过分布式消息中间件记录主节点所有写请求数据包,并在备节点中记录同步主节点数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件可以根据时间戳来查找对应的数据包的特性,找到主备节点间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点中,从而避免了主备切换后的数据丢失。

[0131] 为了能够使得上述实施例进行,本申请提供一种监控系统集群装置实施的数据同步补救方法的实施例,进而可以判断出主节点装置宕机,进行主备切换操作。参见图3,所述数据同步补救方法具体包含有如下内容:

[0132] S201:检测主节点装置是否宕机;

[0133] S202:若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

[0134] S203:修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;其中,主备切换前,将数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0135] 当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻;代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置;

[0136] 所述数据同步补救方法,还包括:

[0137] S204:当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

[0138] 举例而言,监控系统集群装置9检测到数据存储集群分片主节点装置5失效时,触

发主备切换流程;断开数据存储集群分片主节点装置5与数据存储集群分片备节点装置6的主备关系;而后将数据存储集群分片备节点装置6置为可写不可读;之后修改分布式协调装置3中路由信息,将数据存储集群分片备节点装置6变更为主节点;最后,监控系统集群获取到补数完成事件后,将数据存储集群分片主节点装置6变更为可读可写。

[0139] 图4示出了本申请实施例中分布式消息中间件集群装置执行的数据同步补救方法,具体如下:

[0140] S301:主备切换前,写入写请求数据,而后按序写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0141] S302:主备切换后,响应于代理服务器集群装置请求反馈最新写入数据对应的第一时刻;以使所述代理服务器集群装置记录所述第一时刻;

[0142] S303:响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置;以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置;其中,

[0143] 所述第二时刻根据当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成。

[0144] 可以理解,在监控系统集群装置进行主备切换操作前的数据写入过程中,分布式消息中间件集群装置首先获得代理服务器集群装置转发的写入数据,然后调取分布式消息中间件集群装置中的写入数据,写入到主节点装置中,同时结合在数据补救过程中,向前时移的时长大于一个同步时钟的时长,这样可以避免在写入过程中主节点宕机的情况。现有技术中如若主节点宕机,则当前的数据无法写入至主节点,导致主节点的接收数据本身并非最新数据,本实施例中分布式消息中间件集群装置存储的是最新的数据,即使主节点已宕机,不影响最终恢复的数据,由于时移的时长大于一个同步时钟的时长,因此当前同步时钟的数据必然被恢复,当检测到当前同步时钟内主节点已宕机时,不影响写入数据的写入和后续的恢复。

[0145] 在本申请的一个或多个实施例中,第一时刻和第二时刻以时间戳或偏移量的形式被记录(分布式消息中间件集群装置的特性)。

[0146] 对应地,由于本申请的核心构思在于数据补救,为了保证数据补救的进行,需要将写请求数据首先写入到分布式消息中间件集群装置,基于此,如图5所示,本申请还提供一种数据存储方法,应用于代理服务器集群装置,在包括:

[0147] S401:接收写请求数据,并将所述写请求数据发送至分布式消息中间件集群装置;

[0148] S402:重新读取所述分布式消息中间件集群装置中所述写请求数据的数据,写入至主节点装置;其中,至少一个备用节点装置异步复制所述主节点装置写入的数据。

[0149] 对应地,如图6所示,本申请还提供一种数据存储方法,应用于分布式消息中间件集群装置,在包括:

[0150] S501:写入代理服务器集群装置发送的写请求数据;

[0151] S202:根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置,以使所述代理服务器集群装置将所述写请求数据写入至主节点装置,其中,至少一个备用节点装置异步复制所述主节点装置写入的数据。

[0152] 下面结合图1示出若干具体场景

[0153] 写模式

[0154] 1) 应用服务器集群装置1向代理服务器集群装置2发送写请求数据包;

[0155] 2) 代理服务器集群装置2中将所收到的写请求数据包发送至分布式消息中间件集群装置4;

[0156] 3) 代理服务器集群装置2读取分布式消息中间件集群装置4将写请求数据包,并将数据包写入数据存储服务器集群分片装置5或7;

[0157] 4) 代理服务器集群装置2的收到写入分布式消息中间件集群装置4和写入数据存储服务器集群分片装置5或7的两个确认ACK后,向应用服务器集群装置1返回写请求成功应答。

[0158] 读模式

[0159] 1) 应用服务器集群装置1向代理服务器集群装置2发送读请求;

[0160] 2) 代理服务器集群装置2根据路由算法将读请求发送至数据存储服务器集群装置5或7;

[0161] 3) 代理服务器集群装置2将从数据存储服务器集群装置5或7返回的应答数据包返回给应用服务器集群装置1。

[0162] 主备切换及补数流程

[0163] 1. 监控系统集群装置9检测到数据存储集群分片主节点装置5失效时,触发主备切换流程;

[0164] 2. 监控系统集群装置9断开数据存储集群分片主节点装置5与数据存储集群分片备节点装置6的主备关系;

[0165] 3. 监控系统集群装置9将数据存储集群分片备节点装置6置为可写不可读;

[0166] 4. 监控系统集群装置9修改分布式协调装置3中路由信息,将数据存储集群分片备节点装置6变更为主节点;

[0167] 5. 代理服务器集群装置2收到主备切换事件后,记录当前消费分布式消息中间件集群装置4的最新lastOffset,并将此lastOffset写入分布式协调装置3中;

[0168] 6. 代理服务器集群装置2获取数据存储集群分片主节点装置6在作为备节点时同步主节点的最后一条数据的时间戳,然后减3秒,以此时间戳从分布式消息中间件装置4中找到对应的数据startOffset,并写入分布式协调器集群装置3中;

[0169] 7. 代理服务器集群装置2从分布式消息中间件集群装置4中的startOffset开始拉取数据,并将数据写入数据存储集群分片主节点装置6中;

[0170] 8. 代理服务器集群装置2从分布式消息中间件集群装置4消费至lastOffset时,向分布式协调装置3反馈补数完成;

[0171] 9. 监控系统集群获取到补数完成事件后,将数据存储集群分片主节点装置6变更为可读可写。

[0172] 可以理解,本发明中:在保证具有存储节点高可用能力的前提下,通过分布式消息中间件记录主节点所有写请求数据包,并在备节点中记录同步主节点数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件可以根据时间戳来查找对应的数据包的特性,找到主备节点间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点中,从而避免了主备切换后的数据丢失。

[0173] 为了保障分布式数据存储的高可用性,避免主备切换时的数据丢失,本申请进一步提供一种代理服务器集群装置,如图7所示,包括:

[0174] 第一时刻记录模块101,当至少一个备用节点装置中的一个与主节点装置主备切换后,记录分布式消息中间件集群装置最新写入数据对应的第一时刻;其中主备切换前,将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0175] 第二时刻生成模块102,读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;

[0176] 数据拉取模块103,拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

[0177] 在本申请的一个或多个实施例中,所述第一时刻和第二时刻以时间戳或偏移量的形式被记录,所述代理服务器集群装置还包括:

[0178] 第一时刻发送模块,将所述第一时刻对应的时间戳或偏移量发送到分布式协调装置;

[0179] 第二时刻发送模块,将所述第二时刻对应的时间戳或偏移量发送到分布式协调装置。

[0180] 在本申请的一个或多个实施例中,至少一个备用节点装置中的一个与主节点装置主备切换后,所述当前主节点装置被监控系统集群装置配置为可写不可读状态;

[0181] 所述代理服务器集群装置还包括:

[0182] 补数完成反馈模块,当从所述分布式消息中间件集群装置拉取到第一时刻的数据后,向所述分布式协调装置反馈补数完成事件,以使监控系统集群装置将所述当前主节点装置配置为可读可写状态。

[0183] 在本申请的一个或多个实施例中,所述代理服务器集群装置还包括:

[0184] 写请求数据收发模块,接收写请求数据,并将所述写请求数据发送至分布式消息中间件集群装置;

[0185] 写请求数据重新读取模块,在重新读取所述分布式消息中间件集群装置中所述写请求数据的数据,写入至主节点装置;其中,至少一个备用节点装置异步复制所述主节点装置写入的数据。

[0186] 为了保障分布式数据存储的高可用性,避免主备切换时的数据丢失,本申请进一步提供一种监控系统集群装置,如图8所示,包括:

[0187] 宕机检测模块201,检测主节点装置是否宕机;

[0188] 主备关系断开模块202,若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

[0189] 路由信息修改模块203,修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;其中,

[0190] 主备切换前,将数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;并且,



[0191] 当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置记录分布式消息中间件集群装置最新写入数据对应的第一时刻;代理服务器集群装置进一步读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;而后拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置;

[0192] 所述监控系统集群装置,还包括:

[0193] 当前主节点装置状态修改模块204,当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

[0194] 基于相同的发明构思,如图9所示,本申请进一步提供一种分布式消息中间件集群装置,包括:

[0195] 写请求数据写入模块301,主备切换前,写入写请求数据,而后按序写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0196] 第一时刻反馈模块302,主备切换后,响应于代理服务器集群装置请求反馈最新写入数据对应的第一时刻;以使所述代理服务器集群装置记录所述第一时刻;

[0197] 数据拉取模块303,响应于所述代理服务器集群装置的数据拉取请求拉取第二时刻至第一时刻内所有的数据至所述代理服务器集群装置;以通过所述代理服务器集群装置将该所有的数据写入至当前主节点装置;其中,

[0198] 所述第二时刻根据当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成。

[0199] 在本申请的一个或多个实施例中,所述第一时刻和所述第二时刻以时间戳或偏移量的形式被记录。

[0200] 在本申请的一个或多个实施例中,还包括:

[0201] 读取请求反馈模块,根据所述代理服务器集群装置发送的读取请求发送所述写请求数据至所述代理服务器集群装置,以使所述代理服务器集群装置将所述写请求数据写入至主节点装置。

[0202] 基于上述实施例可以知晓,本申请实施例还可以提供一种分布式数据存储系统,如图1所示,包括:

[0203] 代理服务器集群装置2、分布式消息中间件集群装置4、主节点装置(例如5和7)以及至少一个备用节点装置(例如6和8);

[0204] 主备切换前,代理服务器集群装置2将写请求数据写入所述分布式消息中间件集群装置4,而后从所述分布式消息中间件集群装置4中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数据;

[0205] 当至少一个备用节点装置中的一个与主节点装置主备切换后,代理服务器集群装置2记录分布式消息中间件集群装置4最新写入数据对应的第一时刻,然后读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻,最后进一步拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

[0206] 在本申请的一个或多个实施例中,还包括:

[0207] 监控系统集群装置9,所述监控系统集群装置9包括:

[0208] 宕机检测模块,检测主节点装置是否宕机;

[0209] 主备关系断开模块,若是,断开至少一个备用节点装置与所述主节点装置的主备关系,并将其中一个备用节点装置置为可写不可读状态;

[0210] 路由信息修改模块,修改对应的路由信息,将所述其中一个备用节点装置的身份变更为主节点装置;

[0211] 当前主节点装置状态修改模块,当当前主节点装置数据写入完成后,修改当前主节点装置为可读可写状态。

[0212] 在本申请的一个或多个实施例中,还包括:

[0213] 分布式协调装置3,接收所述代理服务器集群装置反馈的补数完成事件,并发送至所述监控系统集群装置,以告知所述监控系统集群装置当前主节点装置数据写入完成。

[0214] 可以理解的是,本发明提供的分布式数据存储系统在数据同步时,写入的数据首先通过代理服务器集群装置传输到分布式消息中间件集群装置,而后通过读取分布式消息中间件集群装置中的数据,并写入到主备节点装置,完成数据同步,当需要主备切换时,本发明中在保证具有存储节点装置高可用能力的前提下,通过分布式消息中间件集群装置记录主节点装置所有写请求数据包,并在备节点装置中记录同步主节点装置数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件集群装置可以根据时间戳来查找对应的数据包的特性,结合时间戳回退,找到主备节点装置间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点装置中,从而避免了主备切换后的数据丢失。

[0215] 从硬件层面来说,为了保障分布式数据存储的高可用性,避免主备切换时的数据丢失,本申请提供一种用于实现所述数据补救或者数据存储方法中的全部或部分内容的电子设备的实施例,所述电子设备具体包含有如下内容:

[0216] 处理器(processor)、存储器(memory)、通信接口(Communications Interface)和总线;其中,所述处理器、存储器、通信接口通过所述总线完成相互间的通信;所述通信接口用于实现服务器、装置、分布式消息中间件集群装置、各类数据库以及用户终端等相关设备之间的信息传输;该电子设备可以是台式计算机、平板电脑及移动终端等,本实施例不限于此。在本实施例中,该电子设备可以参照实施例中的数据补救或者数据存储方法的实施例,以及,数据补救或者数据存储装置的实施例进行实施,其内容被合并于此,重复之处不再赘述。

[0217] 图10为本申请实施例的电子设备的系统构成的示意框图。如图10所示,该电子设备9600可以包括中央处理器9100和存储器9140;存储器9140耦合到中央处理器9100。值得注意的是,该图10是示例性的;还可以使用其他类型的结构,来补充或代替该结构,以实现电信功能或其他功能。

[0218] 一实施例中,数据补救或者数据存储功能可以被集成到中央处理器9100中。例如,中央处理器9100可以被配置为进行如下控制:

[0219] S101:当至少一个备用节点装置中的一个与主节点装置主备切换后,记录分布式消息中间件集群装置最新写入数据对应的第一时刻;其中主备切换前,将写请求数据写入所述分布式消息中间件集群装置,而后从所述分布式消息中间件集群装置中拉取所述写请求数据并写入至主节点装置,至少一个备用节点装置异步复制所述主节点装置写入的数

据;

[0220] S102:读取当前主节点装置中最新异步复制数据的时间点,并结合同步时钟的时长生成第二时刻;

[0221] S103:拉取所述分布式消息中间件集群装置中第二时刻至第一时刻内所有的数据,写入至当前主节点装置。

[0222] 从上述描述可知,本申请的实施例提供的电子设备,在保证具有存储节点高可用能力的前提下,通过分布式消息中间件记录主节点所有写请求数据包,并在备节点中记录同步主节点数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件可以根据时间戳来查找对应的数据包的特性,找到主备节点间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点中,从而避免了主备切换后的数据丢失。

[0223] 在另一个实施方式中,数据补救或者数据存储装置可以与中央处理器9100分开配置,例如可以将数据补救或者数据存储配置为与中央处理器9100连接的芯片,通过中央处理器的控制来实现数据补救或者数据存储功能。

[0224] 如图10所示,该电子设备9600还可以包括:通信模块9110、输入单元9120、音频处理器9130、显示器9160、电源9170。值得注意的是,电子设备9600也并不是必须要包括图10中所示的所有部件;此外,电子设备9600还可以包括图10中没有示出的部件,可以参考现有技术。

[0225] 如图10所示,中央处理器9100有时也称为控制器或操作控件,可以包括微处理器或其他处理器装置和/或逻辑装置,该中央处理器9100接收输入并控制电子设备9600的各个部件的操作。

[0226] 其中,存储器9140,例如可以是缓存器、闪存、硬驱、可移动介质、易失性存储器、非易失性存储器或其它合适装置中的一种或更多种。可储存上述与失败有关的信息,此外还可储存执行有关信息的程序。并且中央处理器9100可执行该存储器9140存储的该程序,以实现信息存储或处理等。

[0227] 输入单元9120向中央处理器9100提供输入。该输入单元9120例如为按键或触摸输入装置。电源9170用于向电子设备9600提供电力。显示器9160用于进行图像和文字等显示对象的显示。该显示器例如可为LCD显示器,但并不限于此。

[0228] 该存储器9140可以是固态存储器,例如,只读存储器(ROM)、随机存取存储器(RAM)、SIM卡等。还可以是这样的存储器,其即使在断电时也保存信息,可被选择性地擦除且设有更多数据,该存储器的示例有时被称为EPROM等。存储器9140还可以是某种其它类型的装置。存储器9140包括缓冲存储器9141(有时被称为缓冲器)。存储器9140可以包括应用/功能存储部9142,该应用/功能存储部9142用于存储应用程序和功能程序或用于通过中央处理器9100执行电子设备9600的操作的流程。

[0229] 存储器9140还可以包括数据存储部9143,该数据存储部9143用于存储数据,例如联系人、数字数据、图片、声音和/或任何其他由电子设备使用的数据。存储器9140的驱动程序存储部9144可以包括电子设备的用于通信功能和/或用于执行电子设备的其他功能(如消息传送应用、通讯录应用等)的各种驱动程序。

[0230] 通信模块9110即为经由天线9111发送和接收信号的发送机/接收机9110。通信模

块(发送机/接收机)9110耦合到中央处理器9100,以提供输入信号和接收输出信号,这可以和常规移动通信终端的情况相同。

[0231] 基于不同的通信技术,在同一电子设备中,可以设置有多个通信模块9110,如蜂窝网络模块、蓝牙模块和/或无线局域网模块等。通信模块(发送机/接收机)9110还经由音频处理器9130耦合到扬声器9131和麦克风9132,以经由扬声器9131提供音频输出,并接收来自麦克风9132的音频输入,从而实现通常的电信功能。音频处理器9130可以包括任何合适的缓冲器、解码器、放大器等。另外,音频处理器9130还耦合到中央处理器9100,从而使得可以通过麦克风9132能够在本机上录音,且使得可以通过扬声器9131来播放本机上存储的声音。

[0232] 本申请的实施例还提供能够实现上述实施例中的执行主体可以为服务器的数据补救或者数据存储方法中全部步骤的一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,该计算机程序被处理器执行时实现上述实施例中的执行主体为服务器或客户端的数据补救或者数据存储方法的全部步骤。

[0233] 从上述描述可知,本申请的实施例提供的计算机可读存储介质,在保证具有存储节点高可用能力的前提下,通过分布式消息中间件记录主节点所有写请求数据包,并在备节点中记录同步主节点数据时最后一条数据完成同步的时间戳,当出现主备切换时,利用了分布式消息中间件可以根据时间戳来查找对应的数据包的特性,找到主备节点间的差异数据,并将这些数据通过补数机制写入主备切换后新的主节点中,从而避免了主备切换后的数据丢失。

[0234] 本领域内的技术人员应明白,本发明的实施例可提供为方法、装置、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0235] 本发明是参照根据本发明实施例的方法、设备(装置)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0236] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0237] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0238] 本发明中应用了具体实施例对本发明的原理及实施方式进行了阐述,以上实施例

的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

# 分布式数据存储系统

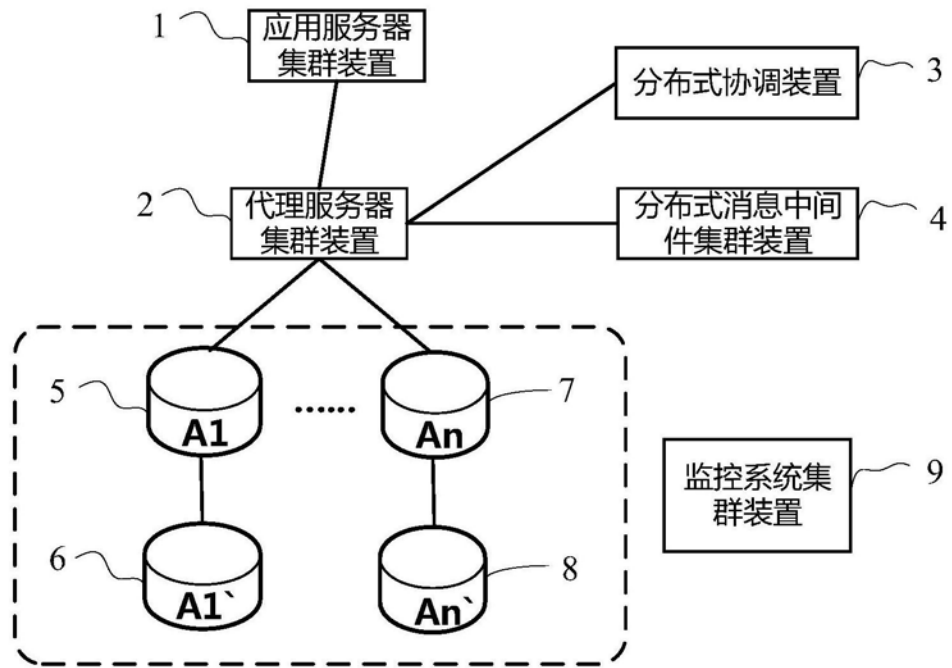


图1

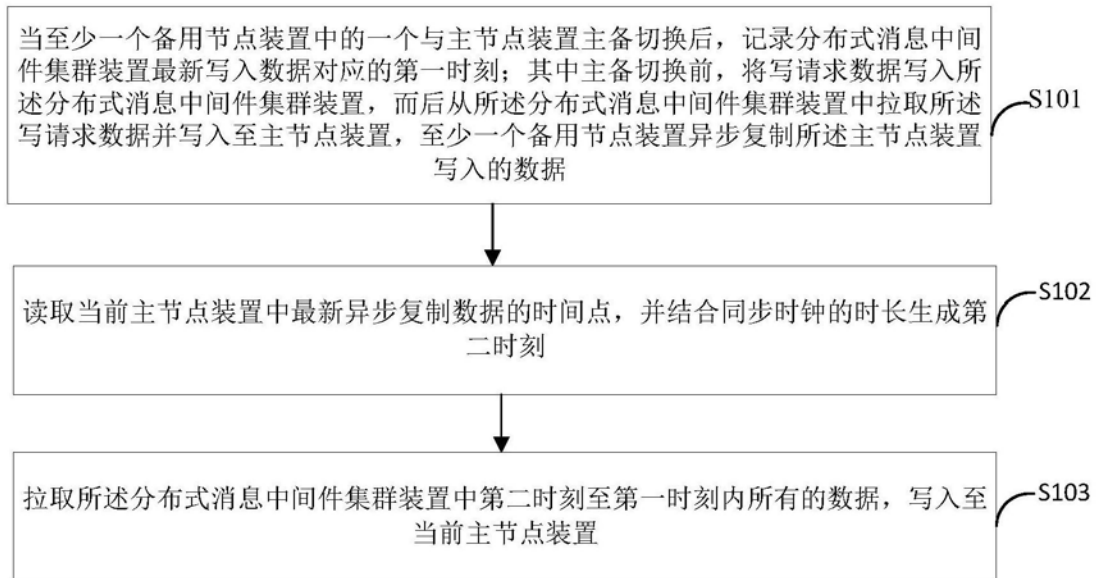


图2

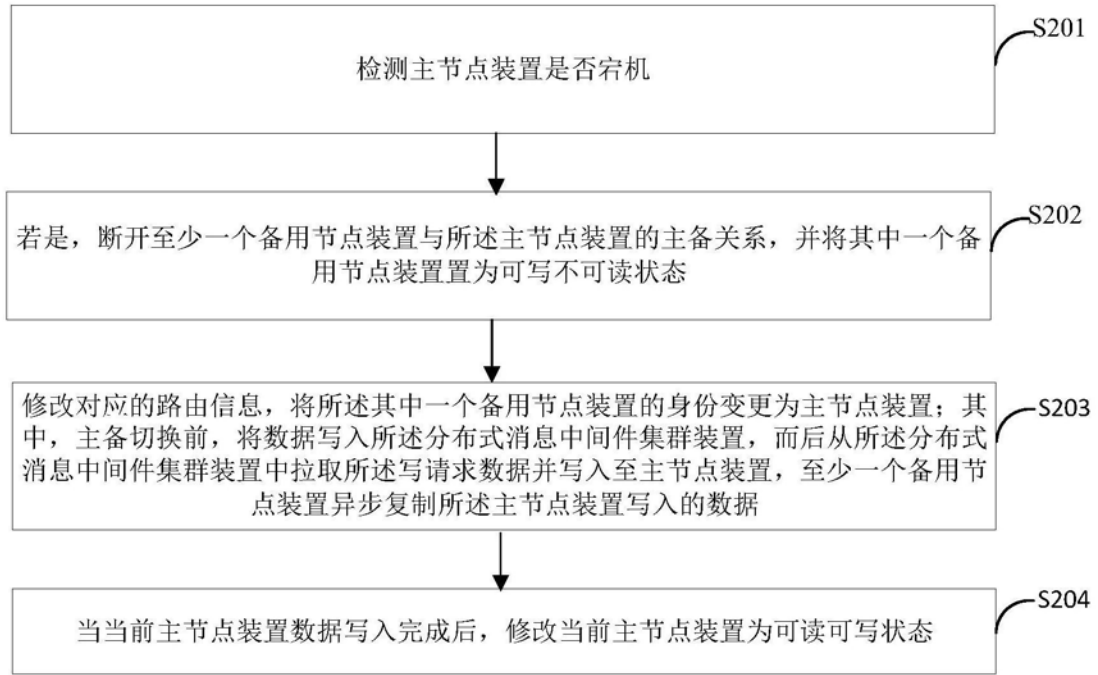


图3

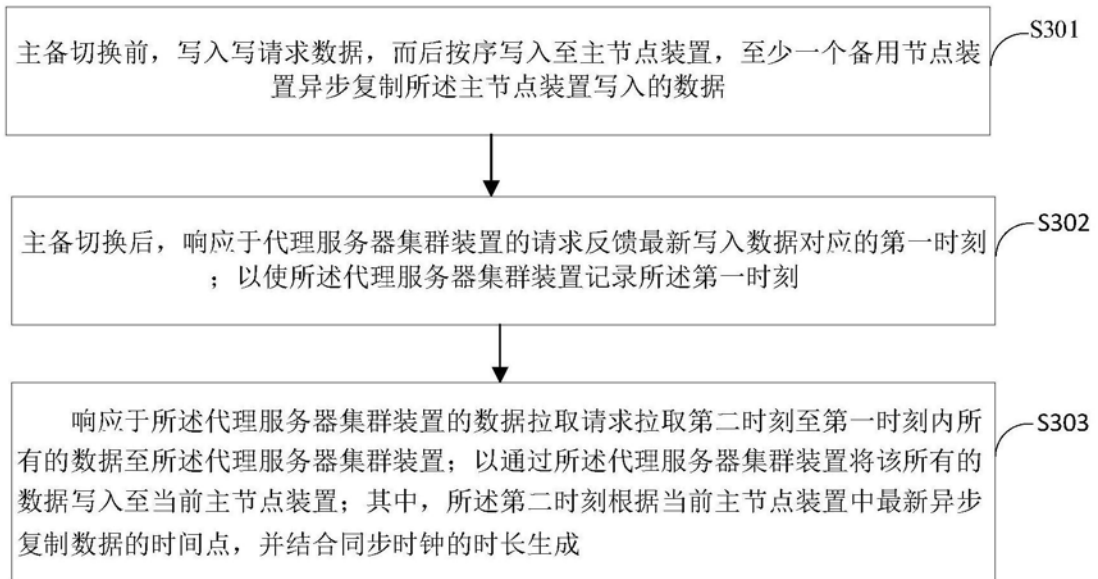


图4

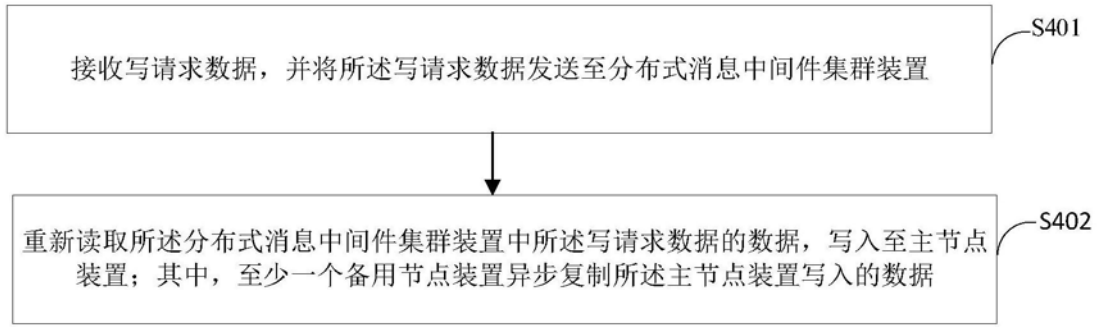


图5

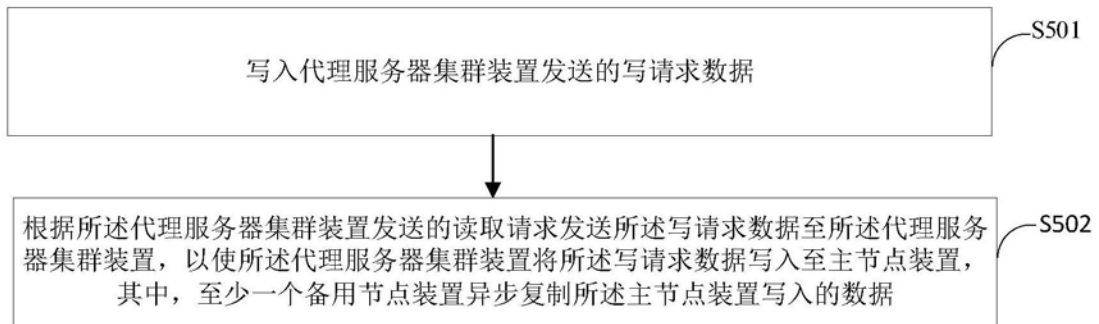


图6

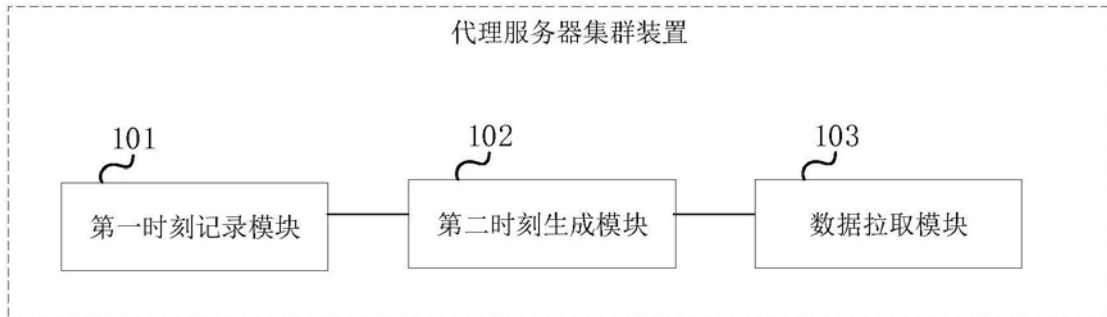


图7



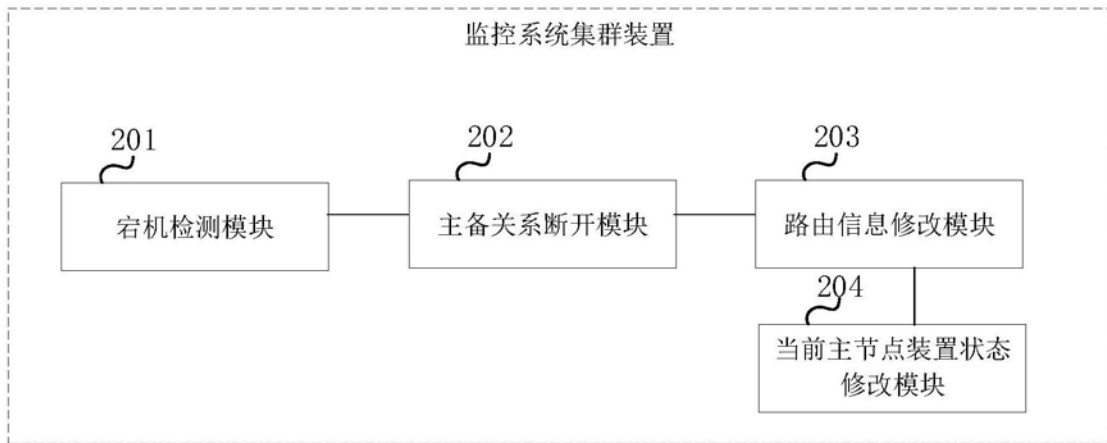


图8

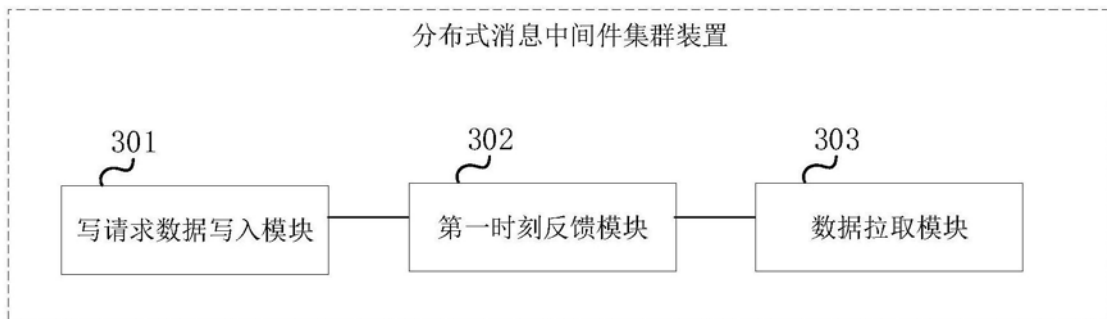


图9

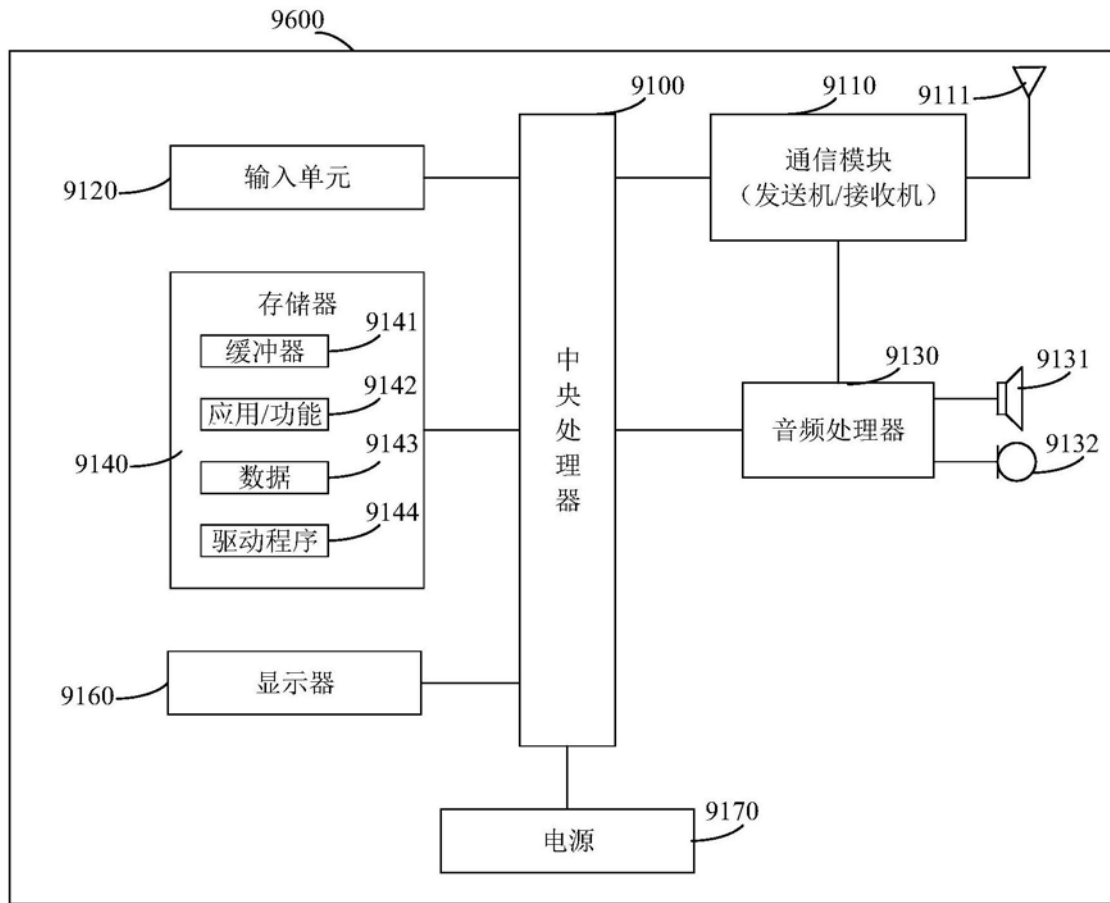


图10