



(12)发明专利申请

(10)申请公布号 CN 107122354 A

(43)申请公布日 2017. 09. 01

(21)申请号 201610101424.6

(22)申请日 2016.02.24

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 文继军 熊亮春 杨春铨

(74)专利代理机构 北京三高永信知识产权代理有限公司 11138

代理人 罗振安

(51)Int.Cl.

G06F 17/30(2006.01)

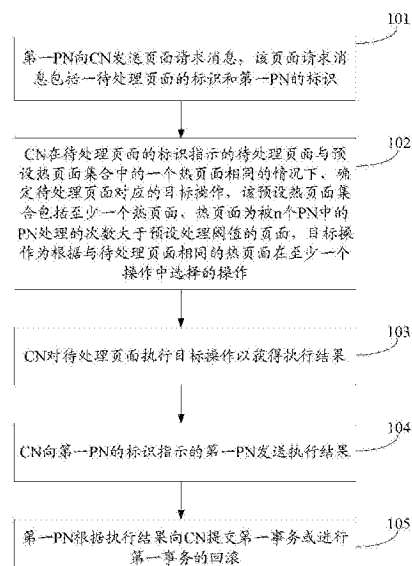
权利要求书6页 说明书22页 附图12页

(54)发明名称

事务执行方法、装置及系统

(57)摘要

本发明公开了一种事务执行方法、装置及系统,属于计算机领域。该方法用于OLTP集群数据库的CN,该方法包括:接收第一PN发送的页面请求消息,第一PN配置有第一事务,第一事务包括至少一个操作,页面请求消息包括一待处理页面的标识和第一PN的标识;在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作,预设热页面集合包括至少一个热页面;对待处理页面执行目标操作以获得执行结果;向第一PN的标识指示的第一PN发送执行结果,解决了事务执行能力较低,OLTP集群数据库的性能较低的问题,提高了事务执行能力和OLTP集群数据库的性能,用于执行事务。



1. 一种事务执行方法,其特征在于,用于联机事务处理OLTP集群数据库的协调节点CN,所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN,所述n大于或等于1,所述方法包括:

接收第一PN发送的页面请求消息,所述第一PN为所述n个PN中的任一PN,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;

在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定所述待处理页面对应的目标操作,所述预设热页面集合包括至少一个热页面,所述热页面为被所述n个PN中的PN处理的次数大于预设处理阈值的页面,所述目标操作为根据与所述待处理页面相同的热页面在所述至少一个操作中选择的操作;

对所述待处理页面执行所述目标操作以获得执行结果;

向所述第一PN的标识指示的第一PN发送所述执行结果,以便于所述第一PN根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

2. 根据权利要求1所述的方法,其特征在于,所述在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定所述待处理页面对应的目标操作,包括:

在所述待处理页面的标识指示的待处理页面与所述预设热页面集合中的一个热页面相同的情况下,检测所述第一事务的操作是否能由所述第一PN和所述CN拆分执行;

在所述第一事务的操作能由所述第一PN和所述CN拆分执行的情况下,确定所述待处理页面对应的目标操作。

3. 根据权利要求1所述的方法,其特征在于,所述对所述待处理页面执行所述目标操作以获得执行结果,包括:

将所述目标操作加入等待队列中,所述等待队列包括待对所述待处理页面执行的操作;

确定所述第一事务的优先级;

根据预设的操作与事务标识的对应关系以及所述第一事务的优先级,确定从所述等待队列中获取所述操作的获取顺序;

对所述待处理页面执行所述目标操作,所述目标操作是根据所述获取顺序从所述等待队列中获取的;

将执行后的所述目标操作加入完成队列中,所述完成队列包括已对所述待处理页面执行完成的操作。

4. 根据权利要求3所述的方法,其特征在于,所述待处理页面包括至少两行记录,所述方法还包括:

在所述对所述待处理页面执行所述目标操作之后,对所述待处理页面执行所述等待队列中的至少一个其他操作;

其中,所述其他操作与所述目标操作用于处理所述待处理页面的不同行的记录,且所述其他操作对应的事务的优先级低于所述目标操作对应的第一事务的优先级。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

从所述n个PN上获取被所述n个PN中的PN处理的次数大于所述预设处理阈值的m个页

面,所述 m 大于或等于1;

建立所述预设热页面集合,所述预设热页面集合包括所述 m 个页面。

6.根据权利要求1所述的方法,其特征在于,所述页面请求消息还包括:所述第一事务的标识,所述确定所述待处理页面对应的目标操作,包括:

根据预设的热页面与操作的对应关系,确定所述待处理页面对应的至少一个操作,所述预设的热页面与操作的对应关系记录有所述预设热页面集合中的各个热页面对应的操作;

根据预设的操作与事务标识的对应关系确定所述第一事务的标识对应的至少一个操作;

将所述待处理页面对应的至少一个操作与所述第一事务的标识对应的至少一个操作中相同的操作确定为所述目标操作。

7.根据权利要求1所述的方法,其特征在于,在所述向所述第一PN的标识指示的第一PN发送所述执行结果之后,所述方法还包括:

接收所述第一PN发送的事务处理消息,所述事务处理消息为事务提交消息或事务回滚消息,所述事务提交消息用于指示所述第一事务需要向所述CN提交,所述事务回滚消息用于指示所述第一事务需要进行回滚,所述事务处理消息是所述第一PN根据所述执行结果生成的;

执行所述事务处理消息指示的动作。

8.根据权利要求5所述的方法,其特征在于,所述方法还包括:

每隔预设时间段检测所述 n 个PN上的第一页面被所述 n 个PN中的PN处理的次数是否大于所述预设处理阈值,所述第一页面为所述 n 个PN上的任一页面;

在所述第一页面被所述 n 个PN中的PN处理的次数大于所述预设处理阈值的情况下,将所述第一页面存储在所述预设热页面集合中。

9.根据权利要求5所述的方法,其特征在于,所述方法还包括:

每隔预设时间段检测所述预设热页面集合中的第二页面被所述CN处理的次数是否大于所述预设处理阈值,所述第二页面为所述预设热页面集合中的任一页面;

在所述第二页面被所述CN处理的次数不大于所述预设处理阈值的情况下,将所述第二页面从所述预设热页面集合中移除。

10.根据权利要求2所述的方法,其特征在于,所述方法还包括:

在所述第一事务的操作不能由所述第一PN和所述CN拆分执行的情况下,向所述第一PN发送所述待处理页面,以便于所述第一PN对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

11.一种事务执行方法,其特征在于,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括 n 个PN和一个协调节点CN,所述 n 大于或等于1,所述第一PN为所述 n 个PN中的任一PN,所述方法包括:

向所述CN发送页面请求消息,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;

接收所述CN发送的所述待处理页面的执行结果,所述执行结果是在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对所述待处理页

面执行所述目标操作后获得的；

根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

12. 根据权利要求11所述的方法，其特征在于，所述根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚，包括：

根据所述执行结果生成事务处理消息，所述事务处理消息为事务提交消息或事务回滚消息，所述事务提交消息用于指示所述第一事务需要向所述CN提交，所述事务回滚消息用于指示所述第一事务需要进行回滚；

向所述CN发送所述事务处理消息，以便于所述CN执行所述事务处理消息指示的动作。

13. 根据权利要求11所述的方法，其特征在于，所述方法还包括：

接收所述CN发送的所述待处理页面，所述待处理页面的标识指示的待处理页面与所述CN中预设热页面集合中的一个热页面相同，且所述第一事务的操作不能由所述第一PN和所述CN拆分执行；

对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

14. 一种事务执行装置，其特征在于，用于联机事务处理OLTP集群数据库的协调节点CN，所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN，所述n大于或等于1，所述装置包括：

接收单元，用于接收第一PN发送的页面请求消息，所述第一PN为所述n个PN中的任一PN，所述第一PN配置有第一事务，所述第一事务包括至少一个操作，所述页面请求消息包括一待处理页面的标识和所述第一PN的标识；

确定单元，用于在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下，确定所述待处理页面对应的目标操作，所述预设热页面集合包括至少一个热页面，所述热页面为被所述n个PN中的PN处理的次数大于预设处理阈值的页面，所述目标操作为根据与所述待处理页面相同的热页面在所述至少一个操作中选择的操作；

执行单元，用于对所述待处理页面执行所述目标操作以获得执行结果；

发送单元，用于向所述第一PN的标识指示的第一PN发送所述待处理页面的执行结果，以便于所述第一PN根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

15. 根据权利要求14所述的装置，其特征在于，所述确定单元，具体用于：

在所述待处理页面的标识指示的待处理页面与所述预设热页面集合中的一个热页面相同的情况下，检测所述第一事务的操作是否能由所述第一PN和所述CN拆分执行；

在所述第一事务的操作能由所述第一PN和所述CN拆分执行的情况下，确定所述待处理页面对应的目标操作。

16. 根据权利要求14所述的装置，其特征在于，所述执行单元，具体用于：

将所述目标操作加入所述等待队列中，所述等待队列包括用于执行所述待处理页面的操作；

确定所述第一事务的优先级；

根据预设的操作与事务标识的对应关系以及所述第一事务的优先级，确定从所述等待队列中获取所述操作的获取顺序；

对所述待处理页面执行所述目标操作，所述目标操作是根据所述获取顺序从所述等待

队列中获取的；

将执行后的所述目标操作加入完成队列中，所述完成队列包括已执行完所述待处理页面的操作。

17. 根据权利要求16所述的装置，其特征在于，所述待处理页面包括至少两行记录，所述装置还包括：

第一选取单元，用于在所述对所述待处理页面执行所述目标操作之后，对所述待处理页面执行所述等待对列中的至少一个其他操作；

其中，所述其他操作与所述目标操作用于处理所述待处理页面的不同行的记录，且所述其他操作对应的事务的优先级低于所述目标操作对应的第一事务的优先级。

18. 根据权利要求14所述的装置，其特征在于，所述装置还包括：

第二选取单元，用于从所述n个PN上获取被所述n个PN中的PN处理的次数大于所述预设处理阈值的m个页面，所述m大于或等于1；

建立单元，用于建立所述预设热页面集合，所述预设热页面集合包括所述m个页面。

19. 根据权利要求14所述的装置，其特征在于，所述页面请求消息还包括：所述第一事务的标识，所述确定单元，具体用于：

根据预设的热页面与操作的对应关系，确定所述待处理页面对应的至少一个操作，所述预设的热页面与操作的对应关系记录有所述预设热页面集合中的各个热页面对应的操作；

根据预设的操作与事务标识的对应关系确定所述第一事务的标识对应的至少一个操作；

将所述待处理页面对应的至少一个操作与所述第一事务的标识对应的至少一个操作中相同的操作确定为所述目标操作。

20. 根据权利要求14所述的装置，其特征在于，

所述接收单元，还用于接收所述第一PN发送的事务处理消息，所述事务处理消息为事务提交消息或事务回滚消息，所述事务提交消息用于指示所述第一事务需要向所述CN提交，所述事务回滚消息用于指示所述第一事务需要进行回滚，所述事务处理消息是所述第一PN根据所述执行结果生成的；

所述执行单元，还用于执行所述事务处理消息指示的动作。

21. 根据权利要求18所述的装置，其特征在于，所述装置还包括：

第一检测单元，用于每隔预设时间段检测所述n个PN上的第一页面被所述n个PN中的PN处理的次数是否大于所述预设处理阈值，所述第一页面为所述n个PN上的任一页面；

存储单元，用于在所述第一页面被所述n个PN中的PN处理的次数大于所述预设处理阈值的情况下，将所述第一页面存储在所述预设热页面集合中。

22. 根据权利要求18所述的装置，其特征在于，所述装置还包括：

第二检测单元，用于每隔预设时间段检测所述预设热页面集合中的第二页面被所述CN处理的次数是否大于所述预设处理阈值，所述第二页面为所述预设热页面集合中的任一页面；

移除单元，用于在所述第二页面被所述CN处理的次数不大于所述预设处理阈值的情况下，将所述第二页面从所述预设热页面集合中移除。

23. 根据权利要求15所述的装置,其特征在于,

所述发送单元,还用于在所述第一事务的操作不能由所述第一PN和所述CN拆分执行的情况下,向所述第一PN发送所述待处理页面,以便于所述第一PN对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

24. 一种事务执行装置,其特征在于,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括n个PN和一个协调节点CN,所述n大于或等于1,所述第一PN为所述n个PN中的任一PN,所述装置包括:

发送单元,用于向所述CN发送页面请求消息,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;

接收单元,用于接收所述CN发送的所述待处理页面的执行结果,所述执行结果是在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对所述待处理页面执行所述目标操作后获得的;

处理单元,用于根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

25. 根据权利要求24所述的装置,其特征在于,所述处理单元,具体用于:

根据所述执行结果生成事务处理消息,所述事务处理消息为事务提交消息或事务回滚消息,所述事务提交消息用于指示所述第一事务需要向所述CN提交,所述事务回滚消息用于指示所述第一事务需要进行回滚;

向所述CN发送所述事务处理消息,以便于所述CN执行所述事务处理消息指示的动作。

26. 根据权利要求24所述的装置,其特征在于,

所述接收单元,还用于接收所述CN发送的所述待处理页面,所述待处理页面的标识指示的待处理页面与所述CN中预设热页面集合中的一个热页面相同,且所述第一事务的操作不能由所述第一PN和所述CN拆分执行;

所述处理单元,还用于对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

27. 一种事务执行装置,其特征在于,用于联机事务处理OLTP集群数据库的协调节点CN,所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN,所述n大于或等于1,所述装置包括:处理器、存储器、网络接口和总线;所述总线用于连接所述处理器、所述存储器和所述网络接口,所述处理器用于执行所述存储器中存储的程序,所述程序包括权利要求1至10任一所述的事务执行方法。

28. 一种事务执行装置,其特征在于,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括n个PN和一个协调节点CN,所述n大于或等于1,所述第一PN为所述n个PN中的任一PN,所述装置包括:处理器、存储器、网络接口和总线;所述总线用于连接所述处理器、所述存储器和所述网络接口,所述处理器用于执行所述存储器中存储的程序,所述程序包括权利要求11至13任一所述的事务执行方法。

29. 一种事务执行系统,其特征在于,包括协调节点CN和第一数据处理节点PN,

所述CN包括权利要求14至23任一所述的事务执行装置;

所述第一PN包括权利要求24至26任一所述的事务执行装置。

30. 一种事务执行系统,其特征在于,包括协调节点CN和第一数据处理节点PN,

所述CN包括权利要求27所述的事务执行装置；
所述第一PN包括权利要求28所述的事务执行装置。

事务执行方法、装置及系统

技术领域

[0001] 本发明涉及计算机领域,特别涉及一种事务执行方法、装置及系统。

背景技术

[0002] 联机事务处理(英文:Online Transaction Processing;简称:OLTP)集群数据库用于执行如银行、航空等行业中大量的相对较小的事务(相对较小的事务指的是执行时间较短,无需进行复杂计算的事务)。其中,事务是一个操作序列,该操作序列中的所有操作存在全部执行,或者全部不执行两种执行情况。当事务中的所有操作都执行成功,OLTP集群数据库进行事务提交(即保存所有操作进行的修改)以保存执行结果;当事务中有一个操作未执行成功,OLTP集群数据库进行事务回滚以撤销之前所有执行成功的操作进行的修改。衡量OLTP集群数据库的集群能力的一个重要指标为事务执行能力,因此,需要一种事务执行方法提高事务执行能力。OLTP集群数据库包括协调节点(英文:coordinator;简称:CN)和至少一个数据处理节点(英文:Processing Node;简称:PN),其中,协调节点用于统一管理与协调整个数据库中各数据处理节点对数据访问的一致性,正确性等问题。

[0003] 现有技术中有一种事务执行方法,通过该方法,OLTP集群数据库(如高斯集群数据库)中的CN在接收到某一PN发送的页面请求消息后,检测到请求的页面被另一PN所持有,CN会从另一PN处获取相应页面,并将获取到的相应页面发送至请求相应页面的PN,接着,该PN对相应页面执行相应操作,如根据相应操作对相应页面中的某一行记录进行修改或删除等,最后,该PN根据执行结果确定是否向CN提交事务。

[0004] 上述方法在高并发激烈冲突(即多个事务并行执行)的场景下,页面转移的开销较大,事务执行的效率较低,因此,事务执行能力较低,OLTP集群数据库的性能较低。

发明内容

[0005] 为了解决现有技术中事务执行能力较低,OLTP集群数据库的性能较低的问题,本发明实施例提供了一种事务执行方法、装置及系统。所述技术方案如下:

[0006] 第一方面,提供了一种事务执行方法,用于联机事务处理OLTP集群数据库的协调节点CN,所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN,所述n大于或等于1,所述方法包括:接收第一PN发送的页面请求消息,所述第一PN为所述n个PN中的任一PN,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定所述待处理页面对应的目标操作,所述预设热页面集合包括至少一个热页面,所述热页面为被所述n个PN中的PN处理的次数大于预设处理阈值的页面,所述目标操作为根据与所述待处理页面相同的热页面在所述至少一个操作中选择的操作;对所述待处理页面执行所述目标操作以获得执行结果;向所述第一PN的标识指示的第一PN发送所述执行结果,以便于所述第一PN根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

[0007] CN能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能。

[0008] 结合第一方面,在第一方面的第一种可能的实施方式中,所述在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定所述待处理页面对应的目标操作,包括:在所述待处理页面的标识指示的待处理页面与所述预设热页面集合中的一个热页面相同的情况下,检测所述第一事务的操作是否能由所述第一PN和所述CN拆分执行;在所述第一事务的操作能由所述第一PN和所述CN拆分执行的情况下,确定所述待处理页面对应的目标操作。

[0009] 在第一事务的操作能由第一PN和CN拆分执行的情况下,CN确定待处理页面对应的目标操作,再对待处理页面执行目标操作,从而保证第一事务被准确执行。第一事务的操作能由第一PN和CN拆分执行,也就是说,第一事务包括的多个操作中的部分操作能由第一PN执行,且该多个操作中的其余操作能由CN执行。

[0010] 结合第一方面,在第一方面的第二种可能的实施方式中,所述对所述待处理页面执行所述目标操作以获得执行结果,包括:将所述目标操作加入等待队列中,所述等待队列包括待对所述待处理页面执行的操作;确定所述第一事务的优先级;根据预设的操作与事务标识的对应关系以及所述第一事务的优先级,确定从所述等待队列中获取所述操作的获取顺序;对所述待处理页面执行所述目标操作,所述目标操作是根据所述获取顺序从所述等待队列中获取的;将执行后的所述目标操作加入完成队列中,所述完成队列包括已对所述待处理页面执行完成的操作。

[0011] CN能够根据事务的优先级确定从等待队列中获取操作的获取顺序,并按照该获取顺序从等待队列中获取目标操作对待处理页面执行目标操作,从而保证各事务被有效执行。CN在建立操作与事务标识的对应关系时,可以给每个事务分配一个优先级,该优先级可以用数字来表示,数字越大,表明事务的优先级就越高。此外,CN除了可以根据事务的优先级确定从等待队列中获取操作的获取顺序,进而确定操作执行的先后顺序外,还可以根据其他优先规则确定从等待队列中获取操作的获取顺序,进而确定操作执行的先后顺序,完成对待处理页面的处理,如操作过程的复杂程度等。

[0012] 结合第一方面的第二种可能的实施方式,在第一方面的第三种可能的实施方式中,所述待处理页面包括至少两行记录,所述方法还包括:在所述对所述待处理页面执行所述目标操作之后,对所述待处理页面执行所述等待队列中的至少一个其他操作;其中,所述其他操作与所述目标操作用于处理所述待处理页面的不同行的记录,且所述其他操作对应的事务的优先级低于所述目标操作对应的第一事务的优先级。

[0013] 在不同事务的操作处理待处理页面的不同行记录的情况下,可以在对待处理页面执行前一事务的操作后,对待处理页面执行后一事务(后一事务的优先级低于前一事务的优先级)的操作,无需等待前一事务的提交,提高了事务执行能力和OLTP集群数据库的性能。

[0014] 结合第一方面,在第一方面的第四种可能的实施方式中,所述方法还包括:从所述n个PN上获取被所述n个PN中的PN处理的次数大于所述预设处理阈值的m个页面,所述m大于

或等于1;建立所述预设热页面集合,所述预设热页面集合包括所述m个页面。

[0015] 可以统计页面被PN选中处理的次数,当页面被PN处理的次数大于预设处理阈值时,可以将该页面作为热页面;还可以通过时间窗来统计一个时间段内页面被PN选中处理的次数,当页面被PN处理的次数大于预设处理阈值时,可以将该页面作为热页面。可选的,CN在建立预设热页面集合时,还可以建立热页面与操作的对应关系。CN建立预设热页面集合,以便于在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作。

[0016] 结合第一方面,在第一方面的第五种可能的实施方式中,所述页面请求消息还包括:所述第一事务的标识,所述确定所述待处理页面对应的目标操作,包括:根据预设的热页面与操作的对应关系,确定所述待处理页面对应的至少一个操作,所述预设的热页面与操作的对应关系记录有所述预设热页面集合中的各个热页面对应的操作;根据预设的操作与事务标识的对应关系确定所述第一事务的标识对应的至少一个操作;将所述待处理页面对应的至少一个操作与所述第一事务的标识对应的至少一个操作中相同的操作确定为所述目标操作。

[0017] CN根据待处理页面对应的至少一个操作与第一事务的标识对应的至少一个操作确定目标操作,进而对待处理页面执行目标操作。

[0018] 结合第一方面,在第一方面的第六种可能的实施方式中,在所述向所述第一PN的标识指示的第一PN发送所述执行结果之后,所述方法还包括:接收所述第一PN发送的事务处理消息,所述事务处理消息为事务提交消息或事务回滚消息,所述事务提交消息用于指示所述第一事务需要向所述CN提交,所述事务回滚消息用于指示所述第一事务需要进行回滚,所述事务处理消息是所述第一PN根据所述执行结果生成的;执行所述事务处理消息指示的动作。

[0019] PN根据待处理页面的执行结果生成事务处理消息,再将该事务处理消息发送至CN,以使CN执行该事务处理消息指示的动作,完成事务的提交或回滚。

[0020] 结合第一方面的第四种可能的实施方式,在第一方面的第七种可能的实施方式中,所述方法还包括:每隔预设时间段检测所述n个PN上的第一页面被所述n个PN中的PN处理的次数是否大于所述预设处理阈值,所述第一页面为所述n个PN上的任一页面;在所述第一页面被所述n个PN中的PN处理的次数大于所述预设处理阈值的情况下,将所述第一页面存储在所述预设热页面集合中。

[0021] CN能够定期对新的热页面进行判定,及时将PN上产生的新的热页面存储在预设热页面集合中,以便于更多的热页面由CN处理,减小了页面转移的开销,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能。

[0022] 结合第一方面的第四种可能的实施方式,在第一方面的第八种可能的实施方式中,所述方法还包括:每隔预设时间段检测所述预设热页面集合中的第二页面被所述CN处理的次数是否大于所述预设处理阈值,所述第二页面为所述预设热页面集合中的任一页面;在所述第二页面被所述CN处理的次数不大于所述预设处理阈值的情况下,将所述第二页面从所述预设热页面集合中移除。

[0023] CN能够定期将预设热页面集合中的非热页面进行移除,提高了事务执行能力和OLTP集群数据库的性能。

[0024] 结合第一方面的第一种可能的实施方式,在第一方面的第九种可能的实施方式中,在所述第一事务的操作不能由所述第一PN和所述CN拆分执行的情况下,向所述第一PN发送所述待处理页面,以便于所述第一PN对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

[0025] 在第一事务的操作不能由第一PN和CN拆分执行的情况下,CN向第一PN发送待处理页面,由第一PN对待处理页面执行待处理页面对应的操作,从而保证第一事务被准确执行。

[0026] 可选的,在事务处理消息为所述事务提交消息的情况下,所述方法还包括:释放所述目标操作占有的资源。事务提交后,CN释放目标操作占有的资源,节省资源。

[0027] 第二方面,提供了一种事务执行方法,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括n个PN和一个协调节点CN,所述n大于或等于1,所述第一PN为所述n个PN中的任一PN,所述方法包括:向所述CN发送页面请求消息,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;接收所述CN发送的所述待处理页面的执行结果,所述执行结果是在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对所述待处理页面执行所述目标操作后获得的;根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

[0028] 第一PN向CN发送页面请求消息,CN在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得第一PN上的第一事务能够由第一PN和CN拆分执行,提高了第一PN执行事务的效率,因此,提高了事务执行能力和OLTP集群数据库的性能。

[0029] 结合第二方面,在第二方面的第一种可能的实施方式中,所述根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚,包括:根据所述执行结果生成事务处理消息,所述事务处理消息为事务提交消息或事务回滚消息,所述事务提交消息用于指示所述第一事务需要向所述CN提交,所述事务回滚消息用于指示所述第一事务需要进行回滚;向所述CN发送所述事务处理消息,以便于所述CN执行所述事务处理消息指示的动作。

[0030] 第一PN可以事先建立一个热页面处理结构,该热页面处理结构用于存储后续CN发送的待处理页面的执行结果,第一PN能够根据热页面处理结构中的执行结果生成事务处理消息,并将该事务处理消息发送至CN,以使CN执行事务处理消息指示的动作,完成事务的执行。第一PN根据热页面处理结构中的执行结果生成事务处理消息时,可以分三种情况:1、在第一PN根据热页面处理结构中的执行结果,确定第一事务对应的所有由CN处理的待处理页面都处理成功的情况下,第一PN生成事务提交消息;2、在第一PN根据热页面处理结构中的执行结果,确定存在由CN处理的待处理页面处理失败的情况下,第一PN生成事务回滚消息;3、在第一PN在第一预设时间段内未全部接收到第一事务对应的所有由CN处理的待处理页面的执行结果的情况下,可以检测第二预设时间段内是否接收到由CN处理的其余待处理页面的执行结果,在第一PN在第二预设时间段内接收到由CN处理的其余待处理页面的执行结果的情况下,第一PN再根据执行结果生成事务处理消息,在第一PN在第二预设时间段内仍未接收到由CN处理的其余待处理页面的执行结果的情况下,第一PN生成事务回滚消息。

[0031] 结合第二方面,在第二方面的第二种可能的实施方式中,所述方法还包括:接收所述CN发送的所述待处理页面,所述待处理页面的标识指示的待处理页面与所述CN中预设热

页面集合中的一个热页面相同,且所述第一事务的操作不能由所述第一PN和所述CN拆分执行;对所述待处理页面执行所述待处理页面对应的操作以获得处理结果。

[0032] 在第一事务的操作不能由第一PN和CN拆分执行的情况下,CN向第一PN发送待处理页面,由第一PN对待处理页面执行待处理页面对应的操作,从而保证第一事务被准确执行。

[0033] 第三方面,提供了一种事务执行装置,用于联机事务处理OLTP集群数据库的协调节点CN,所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN,所述n大于或等于1,所述装置包括:接收单元,用于接收第一PN发送的页面请求消息,所述第一PN为所述n个PN中的任一PN,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;确定单元,用于在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定所述待处理页面对应的目标操作,所述预设热页面集合包括至少一个热页面,所述热页面为被所述n个PN中的PN处理的次数大于预设处理阈值的页面,所述目标操作为根据与所述待处理页面相同的热页面在所述至少一个操作中选择的操作;执行单元,用于对所述待处理页面执行所述目标操作以获得执行结果;发送单元,用于向所述第一PN的标识指示的第一PN发送所述待处理页面的执行结果,以便于所述第一PN根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

[0034] 可选的,所述装置还包括:释放单元,用于释放所述目标操作占有的资源。

[0035] 第三方面描述的装置和单元的具体工作过程,可以参考第一方面对应的实施方式。

[0036] 第四方面,提供了一种事务执行装置,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括n个PN和一个协调节点CN,所述n大于或等于1,所述第一PN为所述n个PN中的任一PN,所述装置包括:发送单元,用于向所述CN发送页面请求消息,所述第一PN配置有第一事务,所述第一事务包括至少一个操作,所述页面请求消息包括一待处理页面的标识和所述第一PN的标识;接收单元,用于接收所述CN发送的所述待处理页面的执行结果,所述执行结果是在所述待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对所述待处理页面执行所述目标操作后获得的;处理单元,用于根据所述执行结果向所述CN提交所述第一事务或进行所述第一事务的回滚。

[0037] 第四方面描述的装置和单元的具体工作过程,可以参考第二方面对应的实施方式。

[0038] 第五方面,提供了一种事务执行装置,用于联机事务处理OLTP集群数据库的协调节点CN,所述OLTP集群数据库包括n个数据处理节点PN和一个所述CN,所述n大于或等于1,所述装置包括:处理器、存储器、网络接口和总线;所述总线用于连接所述处理器、所述存储器和所述网络接口,所述处理器用于执行所述存储器中存储的程序,所述程序包括第一方面所述的事务执行方法。

[0039] 第六方面,提供了一种事务执行装置,用于联机事务处理OLTP集群数据库的第一数据处理节点PN,所述OLTP集群数据库包括n个PN和一个协调节点CN,所述n大于或等于1,所述第一PN为所述n个PN中的任一PN,所述装置包括:处理器、存储器、网络接口和总线;所述总线用于连接所述处理器、所述存储器和所述网络接口,所述处理器用于执行所述存储

器中存储的程序,所述程序包括第二方面所述的事务执行方法。

[0040] 第七方面,提供了一种事务执行系统,包括协调节点CN和第一数据处理节点PN,所述CN包括第三方面所述的事务执行装置;所述第一PN包括第四方面所述的事务执行装置。

[0041] 第八方面,提供了一种事务执行系统,包括协调节点CN和第一数据处理节点PN,所述CN包括第五方面所述的事务执行装置;所述第一PN包括第六方面所述的事务执行装置。

[0042] 本发明提供的技术方案的有益效果是:

[0043] 本发明提供的事务执行方法、装置及系统,由于CN接收PN发送的页面请求消息后,能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能。

附图说明

[0044] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,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。

[0045] 图1-1是本发明各个实施例所涉及的OLTP集群数据库的结构示意图;

[0046] 图1-2是本发明实施例提供的一种事务执行方法的流程图;

[0047] 图2-1是本发明实施例提供的一种事务执行方法的流程图;

[0048] 图2-2是图2-1所示实施例中CN确定待处理页面对应的目标操作的流程图;

[0049] 图2-3是图2-1所示实施例中CN对待处理页面执行目标操作以获得执行结果的流程图;

[0050] 图2-4示出了现有技术中事务执行方法所涉及的一种实施环境的示意图;

[0051] 图2-5是现有技术中事务执行方法的流程图;

[0052] 图2-6示出了现有技术中事务执行方法所涉及的另一种实施环境的示意图;

[0053] 图2-7示出了本发明实施例提供的事务执行方法所涉及的实施环境的示意图;

[0054] 图2-8是在图2-7所示实施环境中进行事务执行所对应的示意图;

[0055] 图3-1是本发明实施例提供的一种事务执行装置的结构示意图;

[0056] 图3-2是本发明实施例提供的另一种事务执行装置的结构示意图;

[0057] 图4是本发明实施例提供的又一种事务执行装置的结构示意图;

[0058] 图5是本发明实施例提供的一种事务执行装置的结构示意图;

[0059] 图6是本发明实施例提供的一种事务执行装置的结构示意图。

[0060] 通过上述附图,已示出本发明明确的实施例,后文中将有更详细的描述。这些附图和文字描述并不是为了通过任何方式限制本发明构思的范围,而是通过参考特定实施例为本领域技术人员说明本发明的概念。

具体实施方式

[0061] 为使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明实施方

式作进一步地详细描述。

[0062] 图1-1是本发明各个实施例所涉及的OLTP集群数据库的结构示意图,该OLTP集群数据库可以包括 n 个PN001和一个CN002, n 大于或等于1。图1-1中,PN001的个数为2。其中,CN002用于统一管理与协调整个数据库中各PN001对数据访问的一致性,正确性等问题。 n 个PN001共享同一个存储单元003。每个PN001可以配置有至少一个事务,每个事务包括至少一个操作,当需要执行第一PN001(第一PN为 n 个PN中的任一PN)上的第一事务T1(第一事务为该至少一个事务中的任一事务)时,如图1-2所示,该执行过程可以包括:

[0063] 步骤101、第一PN向CN发送页面请求消息,该页面请求消息包括一待处理页面的标识和第一PN的标识。

[0064] 第一PN001向CN002发送页面请求消息,该页面请求消息包括一待处理页面的标识和第一PN001的标识。

[0065] 步骤102、CN在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作。该预设热页面集合包括至少一个热页面,热页面为被 n 个PN中的PN处理的次数大于预设处理阈值的页面,目标操作为根据与待处理页面相同的热页面在至少一个操作中选择的操作。

[0066] CN002在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作。预设热页面集合包括的热页面可以为被 n 个PN中的一个PN处理的次数大于预设处理阈值的页面,也可以为被 n 个PN中的至少两个PN处理的次数大于预设处理阈值的页面,该次数是统计的关于页面的所有处理的次数。

[0067] 步骤103、CN对待处理页面执行目标操作以获得执行结果。

[0068] CN002对待处理页面执行目标操作以获得执行结果。

[0069] 步骤104、CN向第一PN的标识指示的第一PN发送执行结果。

[0070] CN002向第一PN的标识指示的第一PN001发送执行结果。

[0071] 步骤105、第一PN根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0072] 第一PN001根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0073] 图2-1是本发明实施例提供的一种事务执行方法的流程图,本实施例以该事务执行方法应用于图1-1所示OLTP集群数据库来举例说明,该事务执行方法可以包括:

[0074] 步骤501、第一PN向CN发送事务执行请求消息。

[0075] 在一种可能的实现方式中,第一PN向CN发送事务执行消息,其中,第一PN为 n 个PN中的任一PN。第一PN配置有第一事务,该事务执行请求消息是由第一事务生成的。第一事务包括至少一个操作。示例的,第一事务包括五个操作,这五个操作存在全部执行,或者全部不执行两种执行情况。当五个操作都执行成功,OLTP集群数据库进行事务提交以保存所有操作的执行结果;当五个操作中有一个操作未执行成功,OLTP集群数据库进行事务回滚以撤销之前所有执行成功的操作进行的修改。步骤501的具体过程可以参考现有技术进行说明。

[0076] 步骤502、CN根据事务执行请求消息向第一PN发送事务信息消息。

[0077] 在一种可能的实现方式中,在第一PN向CN发送事务执行消息后,CN根据事务执行请求消息向第一PN发送事务信息消息。该事务信息消息包括第一事务的标识。事务的标识用于区分不同的事务。步骤502的具体过程可以参考现有技术进行说明。

[0078] 步骤503、CN从n个PN上获取被n个PN中的PN处理的次数大于预设处理阈值的m个页面,该m大于或等于1。

[0079] 在一种可能的实现方式中,如图1-1所示,CN002可以先从n个PN001上获取n个PN中的PN处理的次数大于预设处理阈值的m个页面,以便于CN002建立包括m个页面的预设热页面集合。需要说明的是,可以通过统计页面被PN选中处理的次数确定热页面的,具体的,当页面被PN处理的次数大于预设处理阈值时,则将该页面作为热页面。或者,可以通过时间窗来统计一个时间段内页面被PN选中处理的次数,当页面被PN处理的次数大于预设处理阈值时,将该页面作为热页面。本发明实施例对确定PN上的页面为热页面的方式不做限定。需要说明的是,步骤503也可以在步骤501之前执行。

[0080] 步骤504、CN建立预设热页面集合。

[0081] 在一种可能的实现方式中,CN建立预设热页面集合,该预设热页面集合包括m个页面。假设步骤503中,CN从n个PN上获取到被n个PN中的PN处理的次数大于预设处理阈值的三个页面,这个三个页面分别为y1,y2和y3,那么CN将这三个页面确定为热页面,CN建立预设热页面集合,该预设热页面集合即包括:y1,y2和y3。

[0082] 需要说明的是,CN在根据m个页面建立预设热页面集合时,还可以建立热页面与操作的对应关系,热页面与操作的对应关系记录有预设热页面集合中的各个热页面对应的操作。示例的,热页面与操作的对应关系可以如表1所示。表1中,热页面y1对应的操作为p11,p12和p13,热页面y2对应的操作为p21,p22,p23和p24,热页面y3对应的操作为p31和p32。步骤503和步骤504也可以在步骤501之前执行。

[0083] 表1

[0084]

热页面	操作
y1	p11,p12,p13
y2	p21,p22,p23,p24
y3	p31,p32

[0085] 步骤505、第一PN向CN发送页面请求消息。

[0086] 第一PN可以根据事务信息消息向CN发送页面请求消息。该页面请求消息包括一待处理页面的标识和第一PN的标识。假设该待处理页面为表1中的y1,第一PN为N1,则待处理页面的标识可为y1,第一PN的标识可以为N1。第一PN向CN发送的页面请求消息可以包括待处理页面的标识y1和第一PN的标识N1。步骤505的具体过程可以参考现有技术进行说明。

[0087] 步骤506、CN在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,检测第一事务的操作是否能由第一PN和CN拆分执行。

[0088] 在一种可能的实现方式中,CN在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,检测第一事务的操作是否能由第一PN和CN拆分执行,也就是说,第一事务包括的多个操作中的部分操作是否能由第一PN执行,同时该多个操作中的其余操作是否能由CN执行。如果第一事务的操作能由第一PN和CN拆分执行,说明待处理页面适合CN执行。如果第一事务的操作不能由PN和CN拆分执行,说明待处理页面不适合CN执行。在第一事务的操作能由第一PN和CN拆分执行的情况下,执行步骤507;在第一事务的操作不能由第一PN和CN拆分执行的情况下,执行步骤513。可选的,CN可以根据用户提

交的查询语句中的函数来检测第一事务的操作是否能由第一PN和CN拆分执行。

[0089] 示例的,假设第一事务包括两个操作:p11和p8,预设热页面集合包括三个热页面:y1,y2和y3,如表1所示,热页面y1对应的操作为p11,p12和p13,热页面y2对应的操作为p21,p22,p23和p24,热页面y3对应的操作为p31和p32,如果第一PN向CN发送的页面请求消息包括的待处理页面的标识为y1,该待处理页面的标识指示的待处理页面为y1,待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同,CN可以根据用户提交的查询语句中的函数检测p11是否能由CN执行,p8是否能由第一PN执行,示例的,在该函数为随机函数的情况下,可以确定第一事务的操作不能由第一PN和CN拆分执行。

[0090] 在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面不相同的情况下,执行步骤515。

[0091] 步骤507、CN在第一事务的操作能由第一PN和CN拆分执行的情况下,确定待处理页面对应的目标操作。

[0092] 在一种可能的实现方式中,CN在第一事务的操作能由第一PN和CN拆分执行的情况下,确定待处理页面对应的目标操作。目标操作为根据与待处理页面相同的热页面在至少一个操作中选择的操作。页面请求消息还可以包括:第一事务的标识,可选的,如图2-2所示,步骤507可以包括:

[0093] 步骤5071、根据预设的热页面与操作的对应关系,确定待处理页面对应的至少一个操作。

[0094] 在一种可能的实现方式中,可以根据步骤504中建立的热页面与操作的对应关系,确定待处理页面对应的至少一个操作。示例的,以表1为例,待处理页面为y1,那么根据表1可以得到,待处理页面y1对应三个操作,这三个操作分别为p11,p12和p13。

[0095] 步骤5072、根据预设的操作与事务标识的对应关系确定第一事务的标识对应的至少一个操作。

[0096] 操作与事务标识的对应关系是CN预先建立的。示例的,操作与事务标识的对应关系可以如表2所示。表2中,第一事务的标识T1对应三个操作:p11、p23和p8,第二事务的标识T2对应两个操作:p12和p32,第三事务的标识T3对应三个操作:p41,p51和p61。

[0097] 表2

[0098]

事务标识	操作
T1	p11,p23,p8
T2	p12,p32
T3	p41,p51,p61

[0099] 步骤5073、将待处理页面对应的至少一个操作与第一事务的标识对应的至少一个操作中相同的操作确定为目标操作。

[0100] 以表1和表2为例,待处理页面y1对应三个操作:p11,p12和p13,第一事务的标识T1对应三个操作:p11、p23和p8,所以可以求取待处理页面y1对应的三个操作和第一事务的标识T1对应的三个操作的交集,并将该交集集中的操作作为目标操作,即将p11确定为目标操作。

[0101] 步骤508、CN对待处理页面执行目标操作以获得执行结果。

[0102] 可选的,如图2-3所示,步骤508可以包括:

[0103] 步骤5081、将目标操作加入等待队列中,该等待队列包括待对待处理页面执行的操作。

[0104] 以步骤5073中的目标操作p11为例,CN对待处理页面执行目标操作时,可以先将目标操作p11加入等待队列中。

[0105] 步骤5082、确定第一事务的优先级。

[0106] CN在建立操作与事务标识的对应关系时,可以给每个事务分配一个优先级,该优先级可以用数字来表示,数字越大,表明事务的优先级就越高。进一步的,操作与事务标识的对应关系还可以如表3所示。表3中,第一事务T1的优先级高于第二事务T2的优先级,且低于第三事务T3的优先级。

[0107] 表3

[0108]

事务标识	操作	事务优先级
T1	p11,p23,p8	2
T2	p12,p32	1
T3	p41,p51,p61	3

[0109] 步骤5083、根据预设的操作与事务标识的对应关系以及第一事务的优先级,确定从等待队列中获取操作的获取顺序。

[0110] 可选的,CN可以根据预设的操作与事务标识的对应关系、第一事务的优先级、等待队列中其余操作对应的事务的优先级,确定从等待队列中获取操作的获取顺序。以表3为例,假设等待队列中包括的操作有目标操作p11和另一操作p12,第一事务T1的优先级为2,第二事务T2的优先级为1,第一事务T1的优先级高于第二事务T2的优先级,则可以确定,先从等待队列中获取目标操作p11,再获取另一操作p12;假设第一事务T1的优先级低于第二事务T2的优先级,则可以确定从等待队列中获取了另一操作p12后,再从等待队列中获取目标操作p11。需要说明的是,CN除了可以根据事务的优先级确定从等待队列中获取操作的获取顺序,进而确定操作执行的先后顺序外,还可以根据其他优先规则确定从等待队列中获取操作的获取顺序,进而确定操作执行的先后顺序,完成对待处理页面的处理,如操作过程的复杂程度等,本发明实施例对此不做限定。

[0111] 步骤5084、对待处理页面执行目标操作,该目标操作是根据获取顺序从等待队列中获取的。

[0112] 可选的,CN对待处理页面执行目标操作,该目标操作是根据步骤5083中确定的获取顺序从等待队列中获取的。

[0113] 步骤5085、将执行后的目标操作加入完成队列中,该完成队列包括已对待处理页面执行完成的操作。

[0114] 可选的,在对待处理页面执行目标操作后,CN可以将执行后的目标操作加入完成队列中。

[0115] 在一种可能的实现方式中,待处理页面可以包括至少两行记录,该记录为待处理页面的数据元组。示例的,该待处理页面为一个关于公司5个员工的个人信息(如姓名、年龄、性别、入职时间等)的页面,每一行信息(即记录)对应一个员工的个人信息。如果有两个

操作等待处理该待处理页面,且这两个操作对应的事务不同,在这两个操作处理待处理页面的不同行记录的情况下,可以对待处理页面执行前一事务的操作后,直接对待处理页面执行后一事务的操作,无需等待前一事务的提交,提高了事务执行能力和OLTP集群数据库的性能。本发明实施例提供的事务执行方法中,CN能够判断不同事务是否处理待处理页面的同一行记录,当不同事务处理待处理页面的不同行记录时,可以对待处理页面执行前一事务的操作后,直接对待处理页面执行后一事务的操作。当该不同事务被配置在同一PN上时,该过程相当于对同一PN上配置的不同事务进行合并处理。

[0116] 因此,该事务执行方法还可以包括:CN在对待处理页面执行目标操作之后,对待处理页面执行等待对列中的至少一个其他操作。其中,其他操作与目标操作用于处理待处理页面的不同行的记录,且其他操作对应的事务的优先级低于目标操作对应的第一事务的优先级。

[0117] 示例的,待处理页面上的第一行记录对应的是员工甲的个人信息,第二行记录对应的是员工乙的个人信息,假设等待队列中有两个操作:p11和p12,操作p11对应第一事务T1,操作p12对应第二事务T2,第二事务T2的优先级低于第一事务T1的优先级。操作p11为对待处理待处理页面的第一行记录执行的操作,即操作p11为待对待处理页面的员工甲的个人信息执行的操作,操作p12为待对待处理页面的第二行记录执行的操作,即操作p12为待对待处理页面的员工乙的个人信息执行的操作,那么CN在对员工甲的个人信息执行操作p11之后,可以直接从等待队列中获取操作p12,以对员工乙的个人信息执行操作p12,而无需等待第一事务T1的提交。

[0118] 步骤509、CN向第一PN的标识指示的第一PN发送执行结果。

[0119] 在CN对待处理页面执行目标操作以获得执行结果后,CN向第一PN的标识指示的第一PN发送执行结果,以便于第一PN根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0120] 步骤510、第一PN根据执行结果生成事务处理消息。

[0121] 在一种可能的实现方式中,第一PN根据执行结果生成事务处理消息。该事务处理消息为事务提交消息或事务回滚消息,其中,事务提交消息用于指示第一事务需要向CN提交,事务回滚消息用于指示第一事务需要进行回滚。第一PN可以事先建立一个热页面处理结构,该热页面处理结构用于存储后续CN发送的待处理页面的执行结果。第一PN接收到CN发送的执行结果后,可将执行结果存储在热页面处理结构中,第一PN根据热页面处理结构中的执行结果生成事务处理消息。此外,在步骤506中,在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,CN还可以给第一PN发送一热页面通知消息,该热页面通知消息用于指示待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同,第一PN可以根据该页面通知消息建立待处理页面对应的热页面处理结构。

[0122] 在一种可能的实现方式中,第一PN根据热页面处理结构中的执行结果生成事务处理消息时,可以分三种情况:1、在第一PN根据热页面处理结构中的执行结果,确定第一事务对应的所有由CN处理的待处理页面都处理成功的情况下,第一PN生成事务提交消息,该事务提交消息用于指示第一事务需要向CN提交;2、在第一PN根据热页面处理结构中的执行结果,确定存在由CN处理的待处理页面处理失败的情况下,第一PN生成事务回滚消息,该事务

回滚消息用于指示第一事务需要进行回滚;3、在第一PN在第一预设时间段内未全部接收到第一事务对应的所有由CN处理的待处理页面的执行结果的情况下,可以检测第二预设时间段内是否接收到由CN处理的其余待处理页面的执行结果,在第一PN在第二预设时间段内接收到由CN处理的其余待处理页面的执行结果的情况下,第一PN再根据执行结果生成事务处理消息,在第一PN在第二预设时间段内仍未接收到由CN处理的其余待处理页面的执行结果的情况下,第一PN生成事务回滚消息。

[0123] 以表1和表2为例,表1中,热页面y1对应的操作为p11,p12和p13,热页面y2对应的操作为p21,p22,p23和p24,表2中,第一事务的标识T1对应三个操作为p11,p23和p8。第一PN向CN发送的页面请求消息中包括一待处理页面的标识y1,该待处理页面的标识y1指示的待处理页面y1与预设热页面集合中的一个热页面y1相同,在第一事务的三个操作能由第一PN和CN拆分执行的情况下(即p8能由第一PN执行,p11和p23能由CN执行),CN对待处理页面y1执行目标操作p11以获得执行结果,并向第一PN的标识指示的第一PN发送执行结果;为了完成第一事务的执行,第一PN继续向CN发送的页面请求消息中包括一待处理页面的标识y2,该待处理页面的标识y2指示的待处理页面y2与预设热页面集合中的一个热页面y2相同,在第一事务的三个操作能由第一PN和CN拆分执行的情况下,CN对待处理页面y2执行目标操作p23以获得执行结果,并向第一PN的标识指示的第一PN发送执行结果。如果是第一种情况,即第一PN根据两个执行结果,确定第一事务对应的所有由CN处理的待处理页面都处理成功,那么第一PN生成事务提交消息;如果是第二种情况,如第一PN根据两个执行结果,确定存在由CN处理的待处理页面处理失败,那么第一PN生成事务回滚消息;如果是第三种情况,如第一PN在第一预设时间段内接收到待处理页面y1的执行结果,未接收到待处理页面y2的执行结果,则可以检测第二预设时间段内是否接收到待处理页面y2的执行结果,在第一PN在第二预设时间段内接收到待处理页面y2的执行结果的情况下,第一PN再根据两个执行结果按照第一种情况和第二种情况所述的内容生成事务处理消息,在第一PN在第二预设时间段内仍未接收到待处理页面y2的执行结果的情况下,第一PN生成事务回滚消息。

[0124] 需要补充说明的是,第一PN向CN发送的页面请求消息可以包括一个待处理页面的标识,也可以包括多个待处理页面的标识,本发明实施例对此不做限定。

[0125] 还需要补充说明的是,CN可以一次向第一PN的标识指示的第一PN发送一个待处理页面的执行结果,也可以一次向第一PN的标识指示的第一PN发送至少两个待处理页面的执行结果,本发明实施例对此不作限定。

[0126] 步骤511、第一PN向CN发送事务处理消息。

[0127] 在一种可能的实现方式中,在第一PN根据执行结果生成事务处理消息后,可以将该事务处理消息发送至CN。当第一PN根据执行结果生成事务提交消息时,第一PN向CN发送该事务提交消息;当第一PN根据执行结果生成事务回滚消息时,第一PN向CN发送事务回滚消息。步骤511的具体过程可以参考现有技术进行说明。

[0128] 步骤512、CN执行事务处理消息指示的动作。

[0129] 在一种可能的实现方式中,当第一PN向CN发送事务提交消息时,CN对第一事务对应的所有由CN处理的待处理页面的修改进行保存,并向第一PN发送提交确认消息和待处理页面处理的日志;当第一PN向CN发送事务回滚消息时,CN对第一事务对应的所有由CN处理的待处理页面的修改进行撤销,并向第一PN发送回滚确认消息。

[0130] 需要说明的是,当事务处理消息为事务提交消息时,CN执行完事务提交消息指示的动作后,还可以释放目标操作占有的资源,示例的,该资源可以为目标操作占用的存储空间。此外,CN还可以释放第一事务的相关信息占有的资源,该相关信息可以包括第一事务的标识和第一事务的优先级等信息。CN释放第一事务的相关信息占有的资源的具体过程可以参考现有技术进行说明。

[0131] 步骤513、在第一事务的操作不能由第一PN和CN拆分执行的情况下,CN向第一PN发送待处理页面。执行步骤514。

[0132] 在一种可能的实现方式中,CN在第一事务的操作不能由第一PN和CN拆分执行的情况下,向第一PN发送待处理页面,以便于第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。CN在第一事务的操作不能由第一PN和CN拆分执行的情况下,临时解锁相应的热页面,选择合适的时机将该热页面(即待处理页面)发送至第一PN,以便于第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。其中,第一PN对待处理页面执行待处理页面对应的操作的具体过程可以参考现有技术。

[0133] 需要补充说明的是,在第一PN对待处理页面执行待处理页面对应的操作后,CN可以重新判断该待处理页面被PN处理的次数是否大于预设处理阈值,当该待处理页面被PN处理的次数大于预设处理阈值时,CN确定该待处理页面为热页面,CN获取该待处理页面,并将该待处理页面存储在预设热页面集合中。当该待处理页面被PN处理的次数不大于预设处理阈值时,后续对该待处理页面的处理继续采用现有技术进行。

[0134] 步骤514、第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。

[0135] 在一种可能的实现方式中,CN在第一事务的操作不能由第一PN和CN拆分执行的情况下,向第一PN发送待处理页面,第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。假设预设热页面集合包括三个热页面:y1,y2和y3,热页面y1对应的操作为p11,p12和p13,第一事务的标识T1对应三个操作:p11和p8,待处理页面为y1,CN按照步骤506检测到第一事务的操作不能由第一PN和CN拆分执行,即待处理页面y1不适合CN执行,CN将待处理页面为y1发送至第一PN,第一PN对待处理页面y1执行待处理页面y1对应的操作(即p11)以获得处理结果,并根据该处理结果生成事务处理消息。步骤514的具体过程可以参考现有技术进行说明。

[0136] 步骤515、在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面不相同的情况下,CN向第二PN发送页面请求消息,以便于第二PN将待处理页面通过CN转发至第一PN。

[0137] 在一种可能的实现方式中,CN向第二PN发送页面请求消息,该第二PN为待处理页面的持有者。具体的,CN向第二PN发送页面请求消息;第二PN根据该页面请求消息向CN发送待处理页面;CN再将该待处理页面发送至第一PN;第一PN接收到该待处理页面后,对该待处理页面执行该待处理页面对应的操作以获得处理结果,并根据该处理结果生成事务处理消息。步骤515的具体过程可以参考现有技术进行说明。

[0138] 需要说明的是,为了进一步提高事务执行能力和OLTP集群数据库的性能,可以对预设热页面集合中的热页面进行定期维护。在一种可能的实现方式中,该事务执行方法还可以包括:

[0139] 每隔预设时间段检测n个PN上的第一页面被n个PN中的PN处理的次数是否大于预

设处理阈值,第一页面为n个PN上的任一页面;

[0140] 在第一页面被n个PN中的PN处理的次数大于预设处理阈值的情况下,将第一页面存储在预设热页面集合中。

[0141] 本发明实施例提供的事务执行方法,CN能够定期对新的热页面进行判定,及时将PN上产生的新的热页面存储在预设热页面集合中,以便于更多的热页面由CN处理,在CN端对不同PN上配置的事务进行统一处理,实现了类似单点处理的功能,充分利用了CN冗余的处理能力,减小了页面转移的开销,提高了事务执行的效率,进而提高了事务执行能力和OLTP集群数据库的性能。

[0142] 在一种可能的实现方式中,该事务执行方法还可以包括:

[0143] 每隔预设时间段检测预设热页面集合中的第二页面被CN处理的次数是否大于预设处理阈值,第二页面为预设热页面集合中的任一页面;

[0144] 在第二页面被CN处理的次数不大于预设处理阈值的情况下,将第二页面从预设热页面集合中移除。

[0145] 本发明实施例提供的事务执行方法,CN能够定期将预设热页面集合中的非热页面进行移除,提高了事务执行能力和OLTP集群数据库的性能。

[0146] 图2-4示出了现有技术中事务执行方法所涉及的实施环境的示意图,如图2-4所示,OLTP集群数据库中的各个PN(如PN01)在进行事务提交时,不会将通过处理后的最新页面发送至CN02。当CN02接收到PN03发送的页面请求消息时,检测到请求的待处理页面被PN01所持有,CN02会向PN01发送页面转移消息,PN01将待处理页面发送至CN02,CN02再将待处理页面发送至PN03。PN03在接收到待处理页面时,对待处理页面执行对应的操作,如对待处理页面中的某一行记录进行修改或删除等。最后,PN03根据该事务对应的所有页面被处理的结果确定是否进行事务提交。图2-5示出了现有技术中事务执行方法的流程图。如图2-5所示,该事务执行方法包括:

[0147] 步骤601、第一PN向CN发送事务执行请求消息。执行步骤602。

[0148] 第一PN为n个PN中的任一PN。第一PN配置有第一事务,该事务执行请求消息由第一事务生成。

[0149] 步骤602、CN根据事务执行请求消息向第一PN发送事务信息消息。执行步骤603。

[0150] 该事务信息消息包括第一事务的标识。

[0151] 步骤603、第一PN根据事务信息消息向CN发送页面请求消息。执行步骤604。

[0152] 该页面请求消息包括一待处理页面的标识和第一PN的标识。

[0153] 步骤604、CN根据页面请求消息判断待处理页面是否在CN上。当待处理页面不在CN上时,执行步骤605;当待处理页面在CN上时,执行步骤609。

[0154] 步骤605、CN设置第一PN为待处理页面的持有者。执行步骤606。

[0155] 当待处理页面不在CN上,即待处理页面在存储单元如磁阵上,也就是说,待处理页面从未被处理过,CN设置第一PN为待处理页面的持有者。

[0156] 步骤606、CN向第一PN发送第一指示消息。执行步骤607。

[0157] 该第一指示消息用于指示待处理页面在存储单元上。

[0158] 步骤607、第一PN根据第一指示消息从存储单元中读取待处理页面。执行步骤608。

[0159] 步骤608、第一PN对待处理页面执行待处理页面对应的操作。执行步骤610。

[0160] 步骤609、CN根据待处理页面的状态对待处理页面进行处理。执行步骤610。

[0161] 步骤609可以包括以下三方面内容：

[0162] 第一方面，当待处理页面处于空闲状态时，CN将第一PN设置为待处理页面的持有者，并将该待处理页面发送至第一PN。

[0163] 第二方面，当待处理页面的持有者为第一PN时，CN使第一PN将待处理页面对应的操作加入本地的等待队列中，待处理页面对应的操作等待获得待处理页面。

[0164] 第三方面，当待处理页面的持有者不为第一PN时，示例的，待处理页面的持有者为第二PN，第二PN为n个PN中除第一PN之外的PN。CN使第一PN将待处理页面对应的操作加入本地的等待队列中，CN向第一PN发送第二指示消息，该第二指示消息用于指示待处理页面的持有者为第二PN。同时，CN向第二PN发送页面请求消息，第二PN根据页面请求消息对待处理页面进行处理。当处理的结果满足页面转移条件时，第二PN向CN发送待处理页面，CN再将待处理页面发送至第一PN，第一PN接收到待处理页面后，对待处理页面执行对应的操作。

[0165] 按照步骤603至步骤609继续对第一事务对应的其余待处理页面进行处理。

[0166] 步骤610、第一PN根据第一事务对应的所有页面的处理结果生成事务处理消息。执行步骤611。

[0167] 该事务处理消息为事务提交消息或事务回滚消息。由于事务是一个操作序列，该操作序列是一个不可分割的工作单元，所以，在第一PN检测到第一事务对应的所有页面均被处理成功的情况下，第一PN生成事务提交消息；当第一PN检测到第一事务对应的所有页面中存在页面被处理失败时，第一PN生成事务回滚消息。

[0168] 步骤611、第一PN向CN发送事务处理消息。执行步骤612。

[0169] 步骤612、CN执行事务处理消息指示的动作。执行步骤613。

[0170] 当第一PN向CN发送事务提交消息时，CN对第一事务对应的所有页面的修改进行保存；当第一PN向CN发送事务回滚消息时，CN对第一事务对应的所有页面的修改进行撤销。

[0171] 步骤613、CN向第一PN发送事务处理确定消息。

[0172] 事务处理确定消息为提交确认消息和回滚确认消息。CN对第一事务对应的所有页面的修改进行保存后，向第一PN发送提交确认消息；CN对第一事务对应的所有页面的修改进行撤销后，向第一PN发送回滚确认消息。

[0173] 需要说明的是，采用现有技术中的事务执行方法完成步骤609的第一方面的内容和第二方面的内容时，由于待处理页面可以直接从CN上获取，无需在PN和CN上进行转移，所以事务执行的时延较小，但是在完成步骤609的第三方面的内容时，由于待处理页面需要在PN和CN上进行转移，所以事务执行的时延会特别大，该时延可以达到页面不转移时的时延的300倍。现有技术中的事务执行方法是面向操作步骤的任务分解，缺少对数据运行时分布情况的考虑，这样一来，在高并发激烈冲突的场景下，页面争用会导致页面转移的开销急剧增加。现有技术中的事务执行方法实现的复杂度较低，在高并发激烈冲突的场景下，执行事务的效率较低，同时，CN的计算资源无法得到充分利用。

[0174] 而本发明实施例提供的事务执行方法，CN能够将热页面固定在CN上，当事务需要处理该热页面时，由CN来处理涉及到该热页面的部分操作；CN能够集中对等待处理热页面的事务进行合理调度，具体表现为：1)CN能够对同一个PN上配置的事务进行合并处理，即当CN判断同一PN上的不同事务处理同一待处理页面的不同行记录时，可以根据前一事务对待

处理页面进行处理后,继续根据后一事务对待处理页面进行处理。2)CN能够对不同PN上配置的事务进行统一处理,实现了类似单节点处理的功能。3)当事务需要向CN提交时,CN能够合理获取本地等待队列中的操作,如根据操作对应的事务的优先级获取等待队列中的操作;CN能够定期扫描预设热页面集合,将PN上的新的热页面存储在预设热页面集合中,将预设热页面集合中被CN处理的次数不大于预设处理阈值的页面移除,移除之后的页面按照现有技术进行处理;通过定期刷新PN的页面或者采用缓存机制完成读操作过程,由于读操作无需进行数据修改,所以为了实现读操作和写操作互不阻塞,CN可以将当前相应事务可见的数据页版本发送至作为读操作请求节点的PN;在事务执行过程中,CN能够将预设热页面集合中的热页面下发至配置有事务的PN,然后按照现有技术完成事务执行过程,待事务被提交后,该PN再将该热页面发送至CN,由CN对该热页面进行管理。

[0175] 本发明实施例提供的事务执行方法适用于高并发激烈冲突的场景,由于在高并发激烈冲突的场景下,页面争用会导致页面转移的开销急剧增加,而本发明实施例提供的事务执行方法由CN统一处理热页面,使得事务间的热页面的冲突得到了很好的统一处理,该方法充分利用了CN冗余的处理能力,即充分利用了CN的计算资源,提升了热页面的处理效率,减小了页面转移的开销,达到提升事务所需页面转移效率的效果。且由于CN对整个OLTP集群数据库系统的全局信息了解得比较充分,所以CN能够对各个事务进行高效的统筹调度,实现各种优化处理,通过改变事务执行的方式,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性。

[0176] 图2-6示出了现有技术中的事务执行方法所涉及的一种实施环境的示意图,图2-7示出了本发明实施例提供的事务执行方法所涉及的实施环境的示意图。

[0177] 图2-6所示的实施环境包括CN02和两个PN(即PN03和PN01),PN03和PN01共享一个存储单元04,PN03上存储有页面1,页面2和页面3等,PN01上存储有页面4,页面5和页面6等。其中,页面4被PN处理的次数大于预设处理阈值,即页面4为热页面。

[0178] 图2-7中所示的实施环境包括CN02和两个PN(即PN03和PN01),PN03和PN01共享一个存储单元04。假设PN03配置有第一事务T1和第二事务T2,PN03上存储有页面1,页面2和页面3等。假设PN01配置有第三事务T3、第四事务T4和第五事务T5,PN01上存储有页面5和页面6等。与图2-6不同的是,图2-7中的CN02将PN01上的热页面4存储在CN 02上。图2-7中CN02还存储有热页面与操作的对应关系,如图2-7所示,热页面4对应的操作为t1',t2',t3',t4'和t5',其中,t1'为从第一事务T1中剥离出来的对应于热页面4的操作,t2'为从第一事务T2中剥离出来的对应于热页面4的操作,t3'为从第一事务T3中剥离出来的对应于热页面4的操作,t4'为从第一事务T4中剥离出来的对应于热页面4的操作,t5'为从第一事务T5中剥离出来的对应于热页面4的操作。当第一事务T1向CN02发送热页面4对应的页面请求消息时,如果在图5-6所示的实施环境中执行事务,完成对热页面4的处理,则事务执行过程对应的页面转移示意图如图2-4所示,热页面4需要在CN02、PN01和PN03上进行转移;如果在图2-7所示的实施环境中执行事务,完成对热页面4的处理,则事务执行过程对应的示意图如图2-8所示,CN对热页面4(即待处理页面)执行目标操作t1'以获得执行结果。热页面4无需在CN和PN上进行转移。由图2-4和图2-8可知,本发明实施例提供的事务执行方法减小了高并发激烈冲突的场景下的页面转移的开销,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能。

[0179] 需要说明的是,本发明实施例提供的事务执行方法步骤的先后顺序可以进行适当调整,步骤也可以根据情况进行相应增减,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化的方法,都应涵盖在本发明的保护范围之内,因此不再赘述。

[0180] 综上所述,本发明实施例提供的事务执行方法,由于CN接收PN发送的页面请求消息后,能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提升事务所需页面转移效率,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性。

[0181] 下述为本发明装置实施例,可以用于执行本发明方法实施例。对于本发明装置实施例中未披露的细节,请参照本发明方法实施例。

[0182] 图3-1是本发明实施例提供的一种事务执行装置的结构示意图,该事务执行装置可以通过软件、硬件或者两者的结合实现成为OLTP集群数据库的CN的部分或者全部。OLTP集群数据库包括 n 个PN和一个CN, n 大于或等于1,该装置包括:

[0183] 接收单元310,用于接收第一PN发送的页面请求消息,第一PN为 n 个PN中的任一PN,第一PN配置有第一事务,第一事务包括至少一个操作,该页面请求消息包括一待处理页面的标识和第一PN的标识。

[0184] 确定单元320,用于在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作,该预设热页面集合包括至少一个热页面,热页面为被 n 个PN中的PN处理的次数大于预设处理阈值的页面,目标操作为根据与待处理页面相同的热页面在至少一个操作中选择的操作。

[0185] 执行单元330,用于对待处理页面执行目标操作以获得执行结果。

[0186] 发送单元340,用于向第一PN的标识指示的第一PN发送待处理页面的执行结果,以便于第一PN根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0187] 在一种可能的实施方式中,确定单元320,具体用于:在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,检测第一事务的操作是否能由第一PN和CN拆分执行;在第一事务的操作能由第一PN和CN拆分执行的情况下,确定待处理页面对应的目标操作。

[0188] 在一种可能的实施方式中,执行单元330,具体用于:将目标操作加入等待队列中,等待队列包括用于执行待处理页面的操作;确定第一事务的优先级;根据预设的操作与事务标识的对应关系以及第一事务的优先级,确定从等待队列中获取操作的获取顺序;对待处理页面执行目标操作,该目标操作是根据获取顺序从等待队列中获取的;将执行后的目标操作加入完成队列中,完成队列包括已执行完待处理页面的操作。

[0189] 在一种可能的实施方式中,待处理页面包括至少两行记录,如图3-2所示,该装置还包括:

[0190] 第一选取单元350,用于在对待处理页面执行目标操作之后,对待处理页面执行等待队列中的至少一个其他操作;其中,其他操作与目标操作用于处理待处理页面的不同行的记录,且其他操作对应的事务的优先级低于目标操作对应的第一事务的优先级。

[0191] 在一种可能的实施方式中,如图3-2所示,该装置还包括:

[0192] 第二选取单元360,用于从 n 个PN上获取被 n 个PN中的PN处理的次数大于预设处理阈值的 m 个页面, m 大于或等于1。

[0193] 建立单元370,用于建立预设热页面集合,该预设热页面集合包括 m 个页面。

[0194] 在一种可能的实施方式中,页面请求消息还包括:第一事务的标识,确定单元320,具体用于:根据预设的热页面与操作的对应关系,确定待处理页面对应的至少一个操作,预设的热页面与操作的对应关系记录有预设热页面集合中的各个热页面对应的操作;根据预设的操作与事务标识的对应关系确定第一事务的标识对应的至少一个操作;将待处理页面对应的至少一个操作与第一事务的标识对应的至少一个操作中相同的操作确定为目标操作。

[0195] 在一种可能的实施方式中,接收单元310,还用于接收第一PN发送的事务处理消息,该事务处理消息为事务提交消息或事务回滚消息,事务提交消息用于指示第一事务需要向CN提交,事务回滚消息用于指示第一事务需要进行回滚,该事务处理消息是第一PN根据执行结果生成的。

[0196] 执行单元330,还用于执行事务处理消息指示的动作。

[0197] 在一种可能的实施方式中,如图3-2所示,该装置还包括:

[0198] 第一检测单元380,用于每隔预设时间段检测 n 个PN上的第一页面被 n 个PN中的PN处理的次数是否大于预设处理阈值,第一页面为 n 个PN上的任一页面。

[0199] 存储单元390,用于在第一页面被 n 个PN中的PN处理的次数大于预设处理阈值的情况下,将第一页面存储在预设热页面集合中。

[0200] 在一种可能的实施方式中,如图3-2所示,该装置还包括:

[0201] 第二检测单元410,用于每隔预设时间段检测预设热页面集合中的第二页面被CN处理的次数是否大于预设处理阈值,第二页面为预设热页面集合中的任一页面。

[0202] 移除单元420,用于在第二页面被CN处理的次数不大于预设处理阈值的情况下,将第二页面从预设热页面集合中移除。

[0203] 在一种可能的实施方式中,发送单元340,还用于在第一事务的操作不能由第一PN和CN拆分执行的情况下,向第一PN发送待处理页面,以便于第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。

[0204] 综上所述,本发明实施例提供的事务执行装置,由于CN接收PN发送的页面请求消息后,能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提升事务所需页面转移效率,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性。

[0205] 图4是本发明实施例提供的一种事务执行装置的结构示意图,该事务执行装置可以通过软件、硬件或者两者的结合实现成为OLTP集群数据库的第一PN的部分或者全部。OLTP集群数据库包括 n 个PN和一个CN, n 大于或等于1,第一PN为 n 个PN中的任一PN。该事务执行装置可以包括:

[0206] 发送单元410,用于向CN发送页面请求消息,第一PN配置有第一事务,第一事务包

括至少一个操作,页面请求消息包括一待处理页面的标识和第一PN的标识。

[0207] 接收单元420,用于接收CN发送的待处理页面的执行结果,执行结果是在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作后获得的。

[0208] 处理单元430,用于根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0209] 在一种可能的实施方式中,处理单元430,具体用于:

[0210] 根据执行结果生成事务处理消息,该事务处理消息为事务提交消息或事务回滚消息,事务提交消息用于指示第一事务需要向CN提交,事务回滚消息用于指示第一事务需要进行回滚;向CN发送事务处理消息,以便于CN执行事务处理消息指示的动作。

[0211] 在一种可能的实施方式中,接收单元420,还用于接收CN发送的待处理页面,该待处理页面的标识指示的待处理页面与CN中预设热页面集合中的一个热页面相同,且第一事务的操作不能由第一PN和CN拆分执行。

[0212] 处理单元430,还用于对待处理页面执行待处理页面对应的操作以获得处理结果。

[0213] 综上所述,本发明实施例提供的事务执行装置,由于PN向CN发送页面请求消息,CN能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提升事务所需页面转移效率,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性。

[0214] 图5是本发明实施例提供的一种事务执行装置的结构示意图,用于OLTP集群数据库的CN,所述OLTP集群数据库包括n个PN和一个CN,所述n大于或等于1,该装置包括:处理器801、存储器802、网络接口803和总线804;总线804用于连接处理器801、存储器802和网络接口803,处理器801用于执行存储器802中存储的程序8021。上文中图3-1的接收单元310和发送单元340可以是由网络接口803来实现;图3-1的确定单元320和执行单元330,以及图3-2的第一选取单元350、第二选取单元360、建立单元370、第一检测单元380、存储单元390、第二检测单元410和移除单元420,可以是由处理器801执行存储器802中存储的程序8021来实现。

[0215] 存储器802可能包含随机存取存储器(英文:Random Access Memory;简称:RAM),也可能包含非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过网络接口803(可以是有线或者无线)实现CN与PN之间的通信连接。处理器801可以执行存储器802中存储的程序8021来实现OLTP集群数据库的CN,并使得该CN执行下述事务执行方法:

[0216] 接收第一PN发送的页面请求消息,第一PN为n个PN中的任一PN,第一PN配置有第一事务,第一事务包括至少一个操作,页面请求消息包括一待处理页面的标识和第一PN的标识;

[0217] 在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作,该预设热页面集合包括至少一个热页面,热页面为被n个PN中的PN处理的次数大于预设处理阈值的页面,目标操作为根据与待处理页面相同的热页面在至少一个操作中选择的操作;

[0218] 对待处理页面执行目标操作以获得执行结果;

[0219] 向第一PN的标识指示的第一PN发送执行结果,以便于第一PN根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0220] 可选的,在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,确定待处理页面对应的目标操作,包括:

[0221] 在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,检测第一事务的操作是否能由第一PN和CN拆分执行;

[0222] 在第一事务的操作能由第一PN和CN拆分执行的情况下,确定待处理页面对应的目标操作。

[0223] 可选的,对待处理页面执行目标操作以获得执行结果,包括:

[0224] 将目标操作加入等待队列中,该等待队列包括待对待处理页面执行的操作;

[0225] 确定第一事务的优先级;

[0226] 根据预设的操作与事务标识的对应关系以及第一事务的优先级,确定从等待队列中获取操作的获取顺序;

[0227] 对待处理页面执行目标操作,该目标操作是根据获取顺序从等待队列中获取的;

[0228] 将执行后的目标操作加入完成队列中,该完成队列包括已对待处理页面执行完成的操作。

[0229] 可选的,待处理页面包括至少两行记录,该方法还包括:

[0230] 在对待处理页面执行目标操作之后,对待处理页面执行等待对列中的至少一个其他操作;

[0231] 其中,其他操作与目标操作用于处理待处理页面的不同行的记录,且其他操作对应的事务的优先级低于目标操作对应的第一事务的优先级。

[0232] 可选的,该方法还包括:

[0233] 从n个PN上获取被n个PN中的PN处理的次数大于预设处理阈值的m个页面,m大于或等于1;

[0234] 建立预设热页面集合,该预设热页面集合包括m个页面。

[0235] 可选的,页面请求消息还包括:第一事务的标识,确定待处理页面对应的目标操作,包括:

[0236] 根据预设的热页面与操作的对应关系,确定待处理页面对应的至少一个操作,预设的热页面与操作的对应关系记录有预设热页面集合中的各个热页面对应的操作;

[0237] 根据预设的操作与事务标识的对应关系确定第一事务的标识对应的至少一个操作;

[0238] 将待处理页面对应的至少一个操作与第一事务的标识对应的至少一个操作中相同的操作确定为目标操作。

[0239] 可选的,在向第一PN的标识指示的第一PN发送执行结果之后,该方法还包括:

[0240] 接收第一PN发送的事务处理消息,该事务处理消息为事务提交消息或事务回滚消息,事务提交消息用于指示第一事务需要向CN提交,事务回滚消息用于指示第一事务需要进行回滚,事务处理消息是第一PN根据执行结果生成的;

[0241] 执行事务处理消息指示的动作。

[0242] 可选的,该方法还包括:

[0243] 每隔预设时间段检测n个PN上的第一页面被n个PN中的PN处理的次数是否大于预设处理阈值,第一页面为n个PN上的任一页面;

[0244] 在第一页面被n个PN中的PN处理的次数大于预设处理阈值的情况下,将第一页面存储在预设热页面集合中。

[0245] 可选的,该方法还包括:

[0246] 每隔预设时间段检测预设热页面集合中的第二页面被CN处理的次数是否大于预设处理阈值,第二页面为预设热页面集合中的任一页面;

[0247] 在第二页面被CN处理的次数不大于预设处理阈值的情况下,将第二页面从预设热页面集合中移除。

[0248] 可选的,该方法还包括:

[0249] 在第一事务的操作不能由第一PN和CN拆分执行的情况下,向第一PN发送待处理页面,以便于第一PN对待处理页面执行待处理页面对应的操作以获得处理结果。

[0250] 综上所述,本发明实施例提供的事务执行装置,由于CN接收PN发送的页面请求消息后,能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提升事务所需页面转移效率,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性的。

[0251] 图6是本发明实施例提供的一种事务执行装置的结构示意图,用于OLTP集群数据库的第一PN,OLTP集群数据库包括n个PN和一个CN,n大于或等于1,第一PN为n个PN中的任一PN,该装置包括:处理器901、存储器902、网络接口903和总线904;总线904用于连接处理器901、存储器902和网络接口903,处理器901用于执行存储器902中存储的程序9021。上文中图4的发送单元410和接收单元420可以是由网络接口903来实现;图4的处理单元430可以是由处理器901执行存储器902中存储的程序9021来实现。

[0252] 存储器902可能包含随机存取存储器,也可能包含非不稳定的存储器,例如至少一个磁盘存储器。通过网络接口903(可以是有线或者无线)实现PN与CN之间的通信连接。处理器901可以执行存储器902中存储的程序9021来实现OLTP集群数据库的第一PN,并使得该第一PN执行下述事务执行方法:

[0253] 向CN发送页面请求消息,第一PN配置有第一事务,第一事务包括至少一个操作,页面请求消息包括一待处理页面的标识和第一PN的标识;

[0254] 接收CN发送的待处理页面的执行结果,该执行结果是在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作后获得的;

[0255] 根据执行结果向CN提交第一事务或进行第一事务的回滚。

[0256] 可选的,根据执行结果向CN提交第一事务或进行第一事务的回滚,包括:

[0257] 根据执行结果生成事务处理消息,该事务处理消息为事务提交消息或事务回滚消息,事务提交消息用于指示第一事务需要向CN提交,事务回滚消息用于指示第一事务需要进行回滚;

[0258] 向CN发送事务处理消息,以便于CN执行事务处理消息指示的动作。

[0259] 可选的,该方法还包括:

[0260] 接收CN发送的待处理页面,待处理页面的标识指示的待处理页面与CN中预设热页面集合中的一个热页面相同,且第一事务的操作不能由第一PN和CN拆分执行;

[0261] 对待处理页面执行待处理页面对应的操作以获得处理结果。

[0262] 综上所述,本发明实施例提供的事务执行装置,由于PN向CN发送页面请求消息,CN能够在待处理页面的标识指示的待处理页面与预设热页面集合中的一个热页面相同的情况下,对待处理页面执行目标操作,使得待处理页面无需在PN和CN上进行频繁转移,减小了页面转移的开销,提升事务所需页面转移效率,提高了事务执行的效率,因此,提高了事务执行能力和OLTP集群数据库的性能,提高了整个OLTP集群数据库在高并发激烈冲突的场景下的扩展性和性能的稳定性。

[0263] 本发明实施例提供了一种事务执行系统,包括CN和第一PN。

[0264] 其中,CN包括图3-1或图3-2所示的事务执行装置;

[0265] 第一PN包括图4所示的事务执行装置。

[0266] 本发明实施例提供了另一种事务执行系统,包括CN和第一PN。

[0267] 其中,CN包括图5所示的事务执行装置;

[0268] 第一PN包括图6所示的事务执行装置。

[0269] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,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。

[0270] 所述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例方案的目的。

[0271] 另外,在本发明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包括,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,也可以采用硬件加软件功能单元的形式实现。

[0272] 以上所述仅为本发明的较佳实施例,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

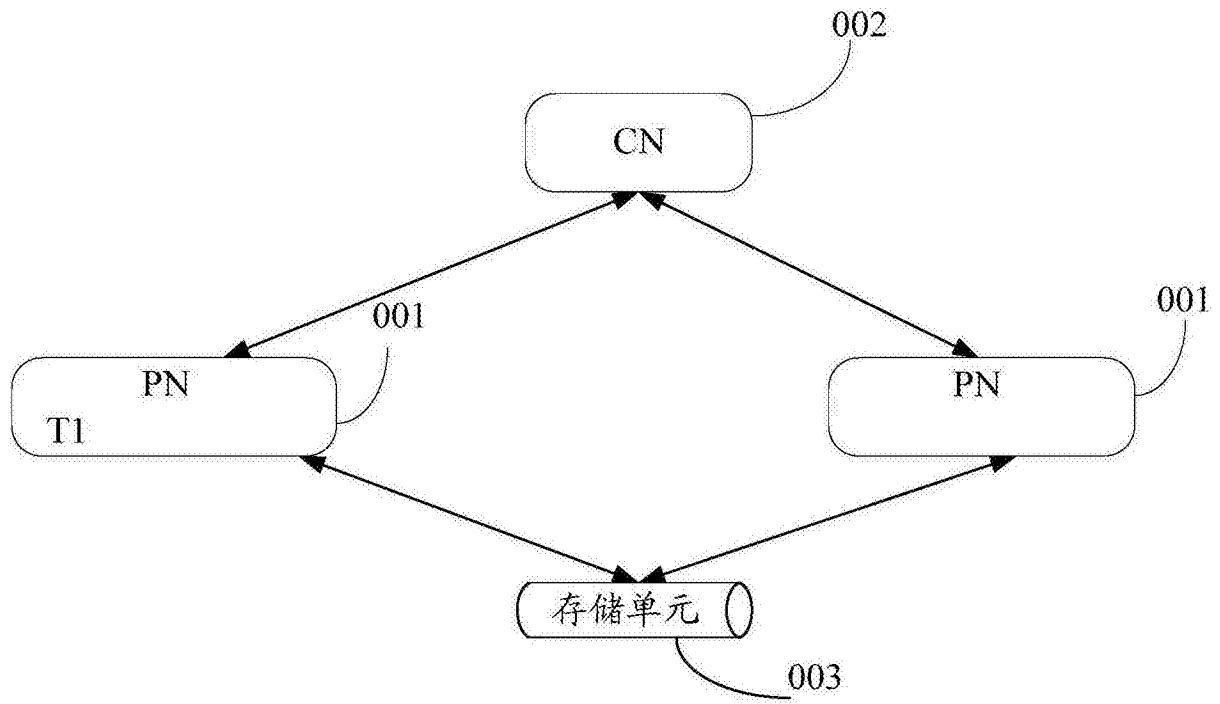


图1-1

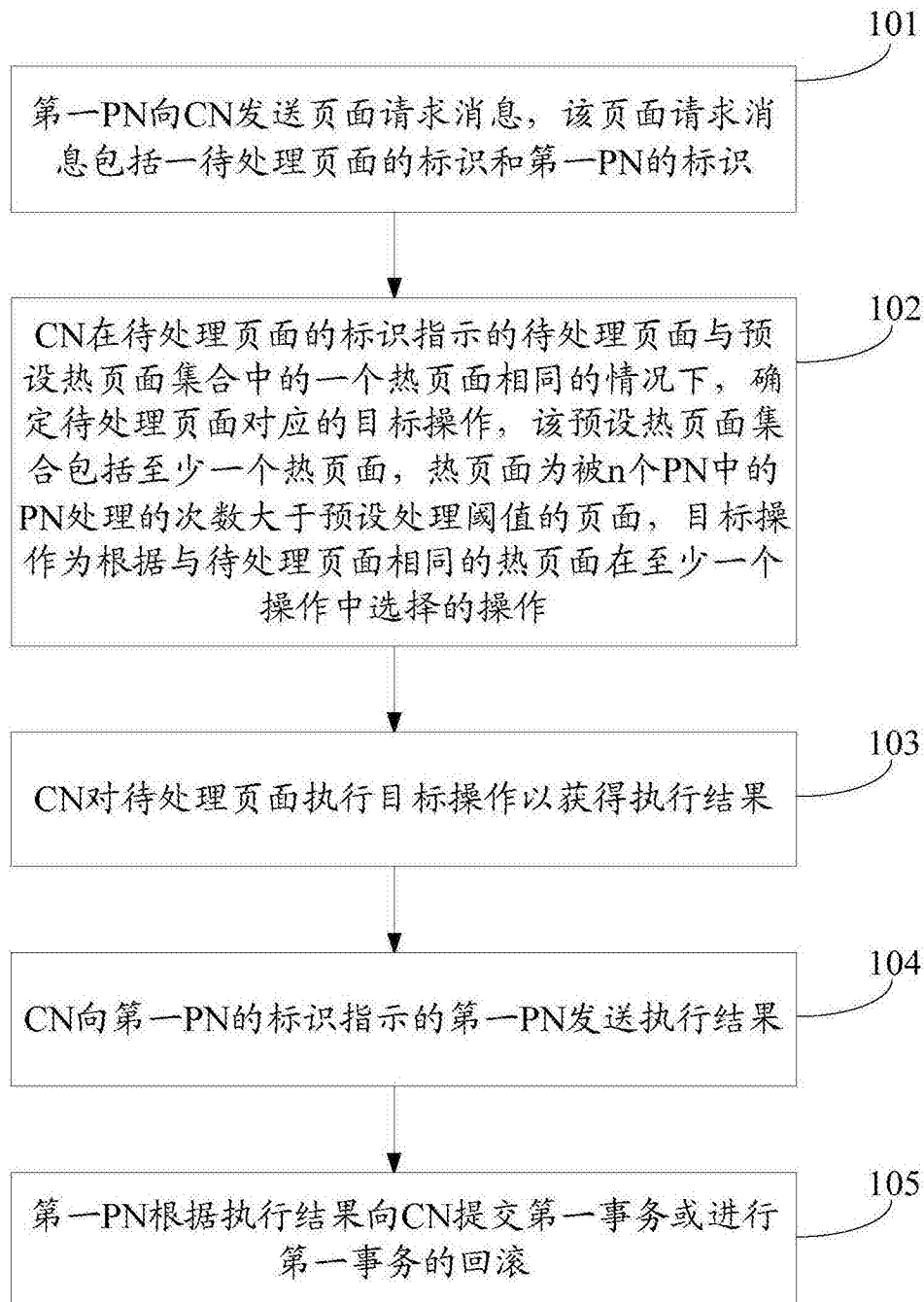


图1-2

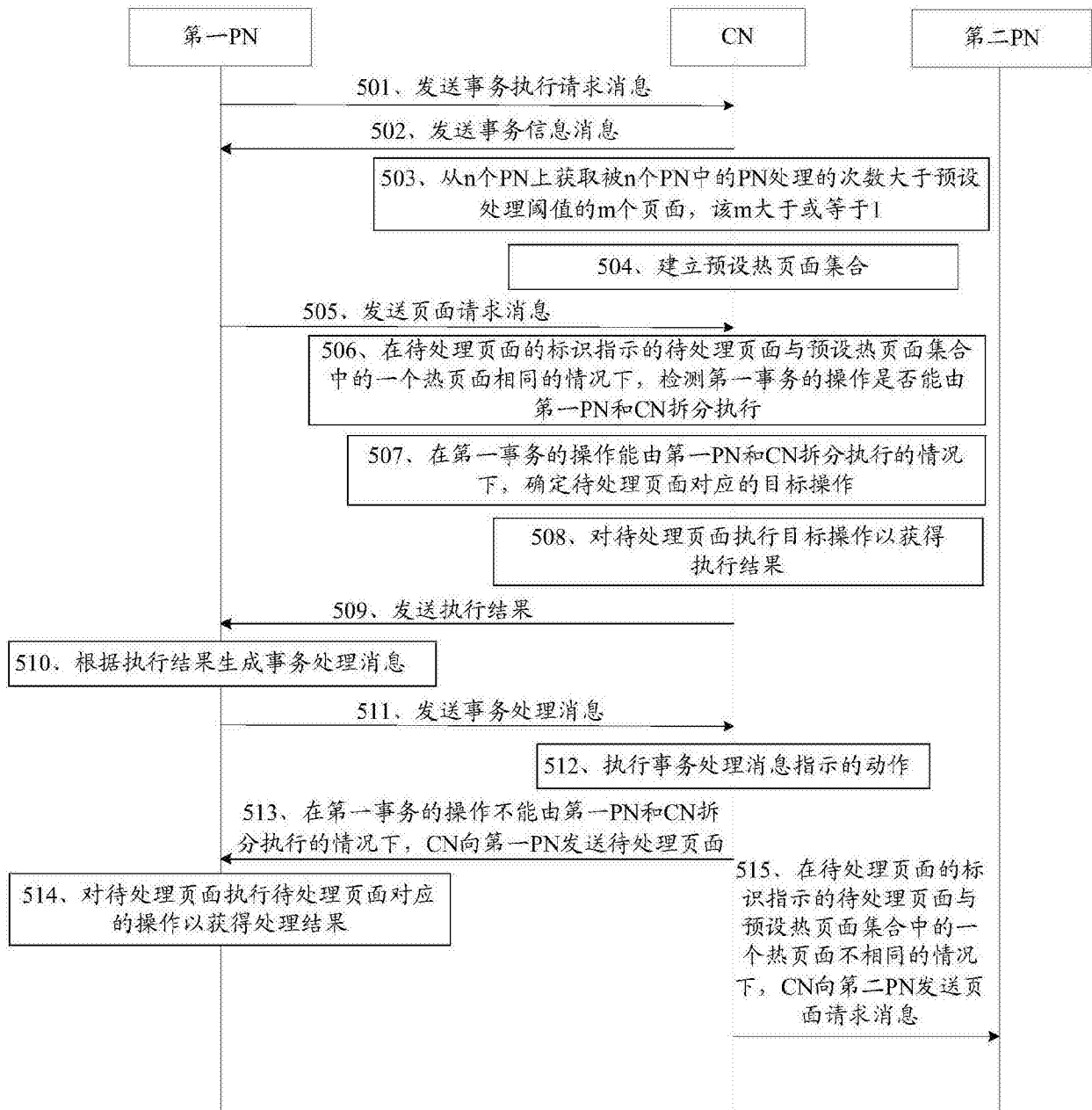


图2-1

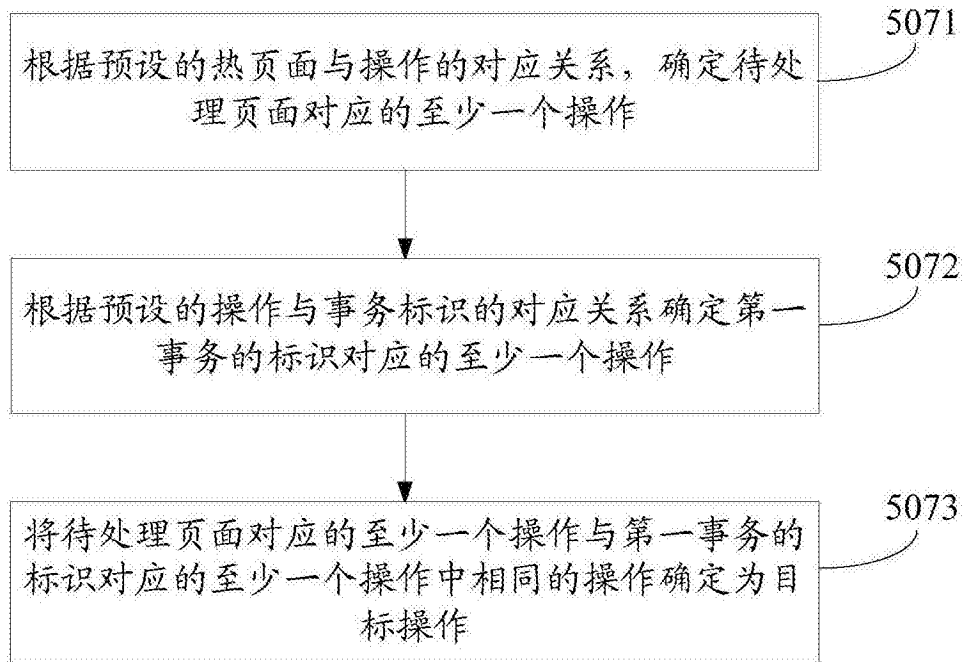


图2-2

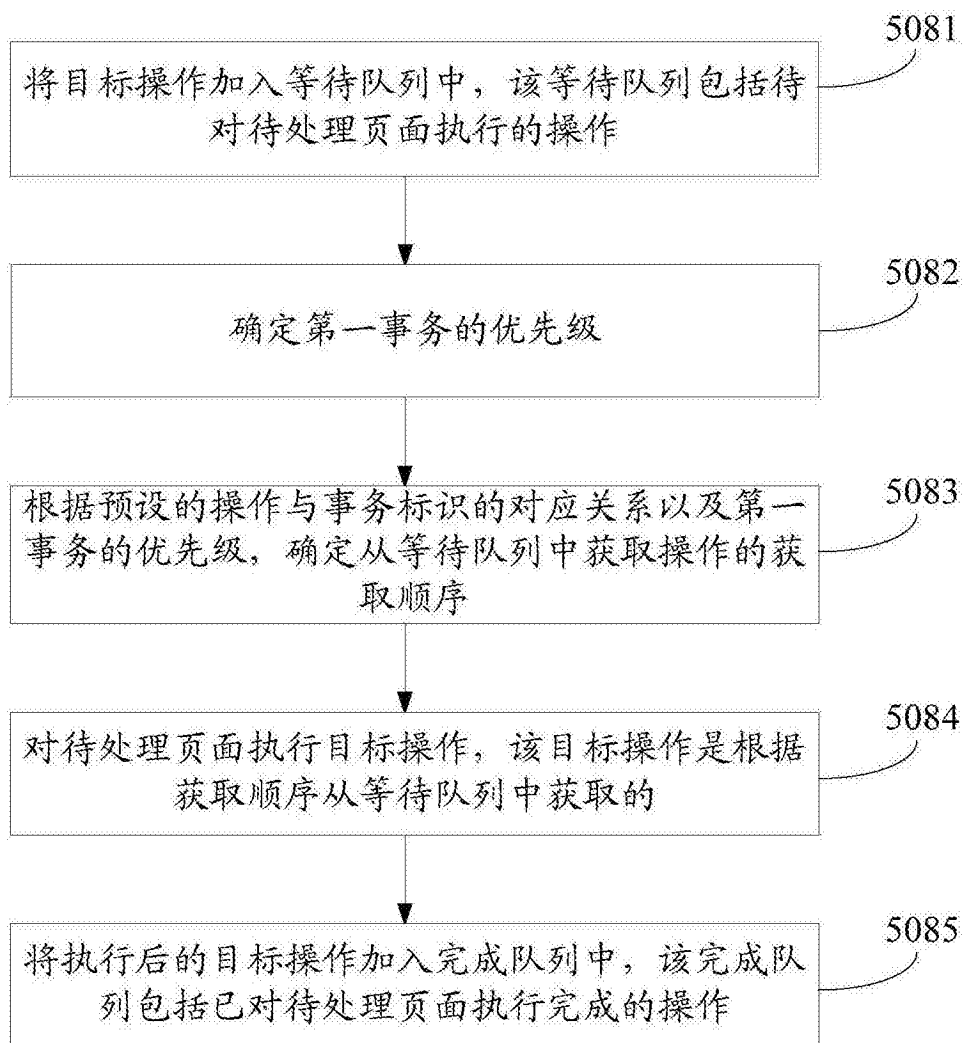


图2-3

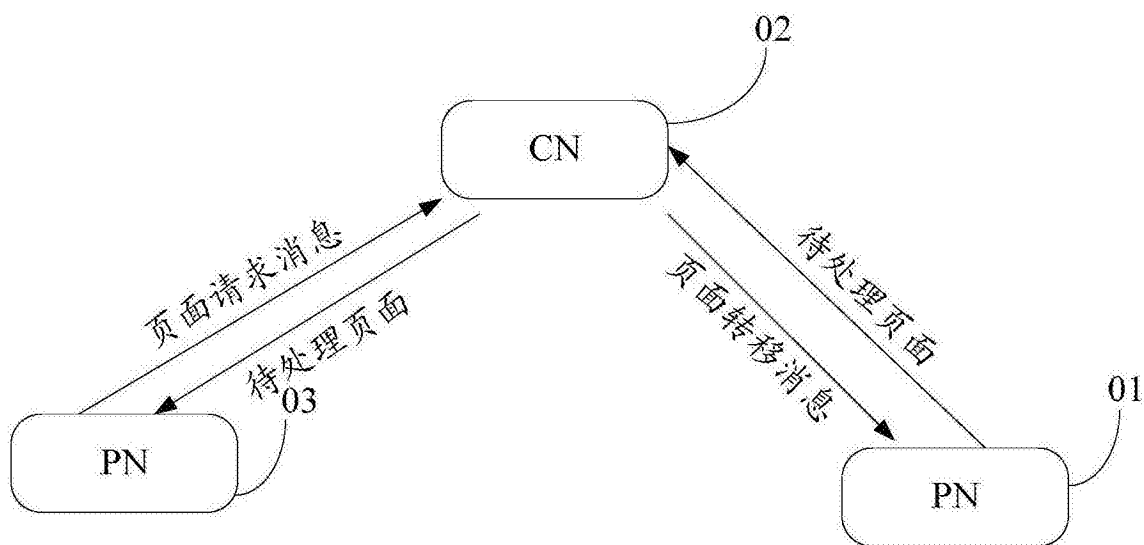


图2-4

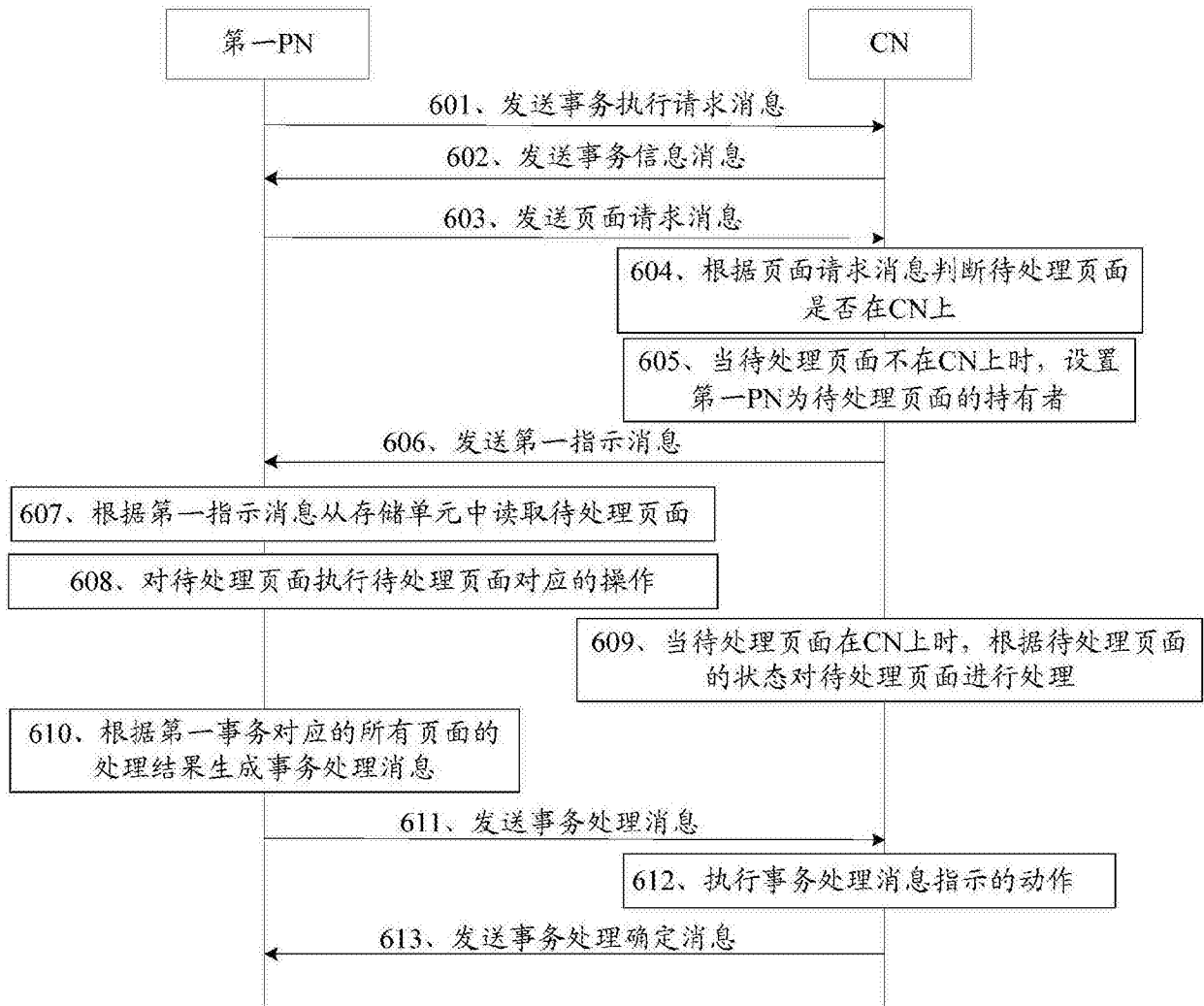


图2-5

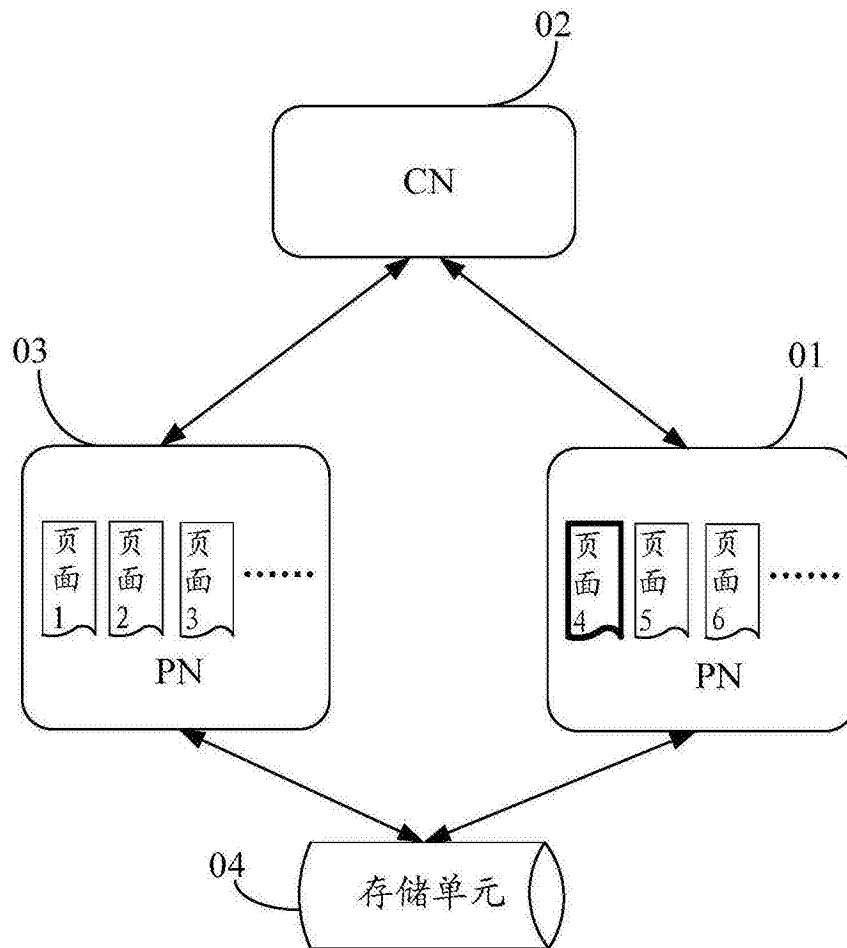


图2-6

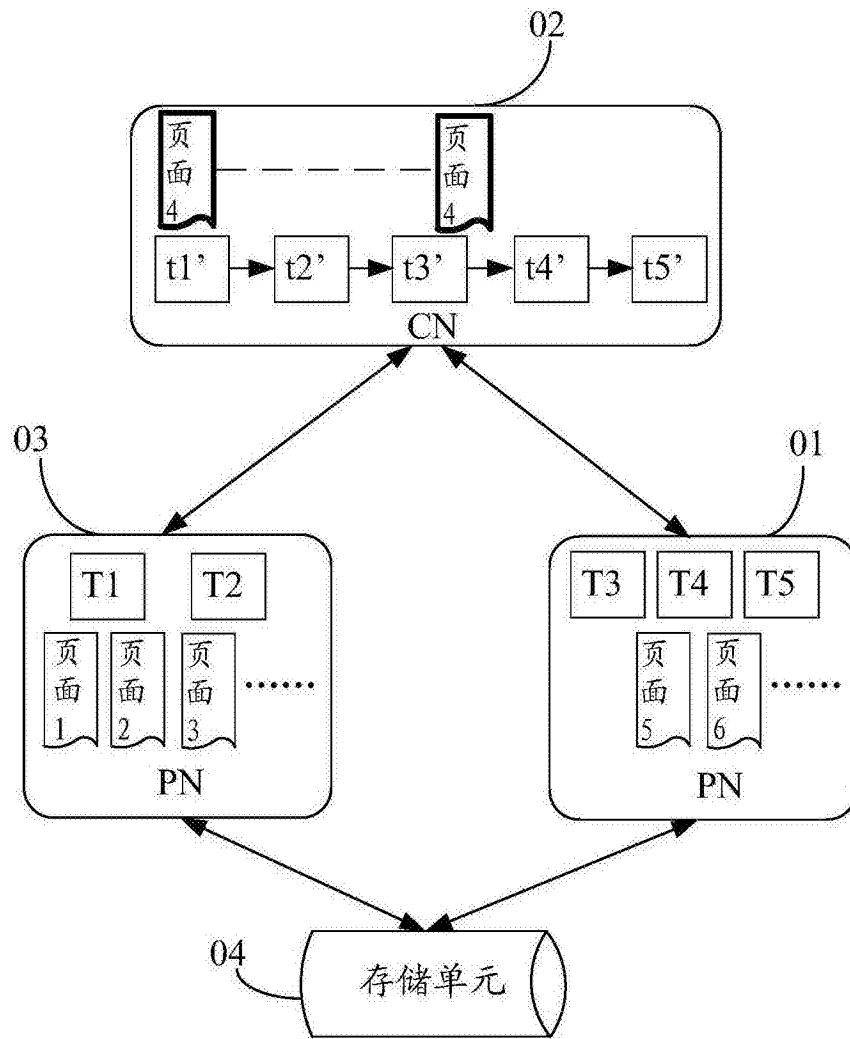


图2-7

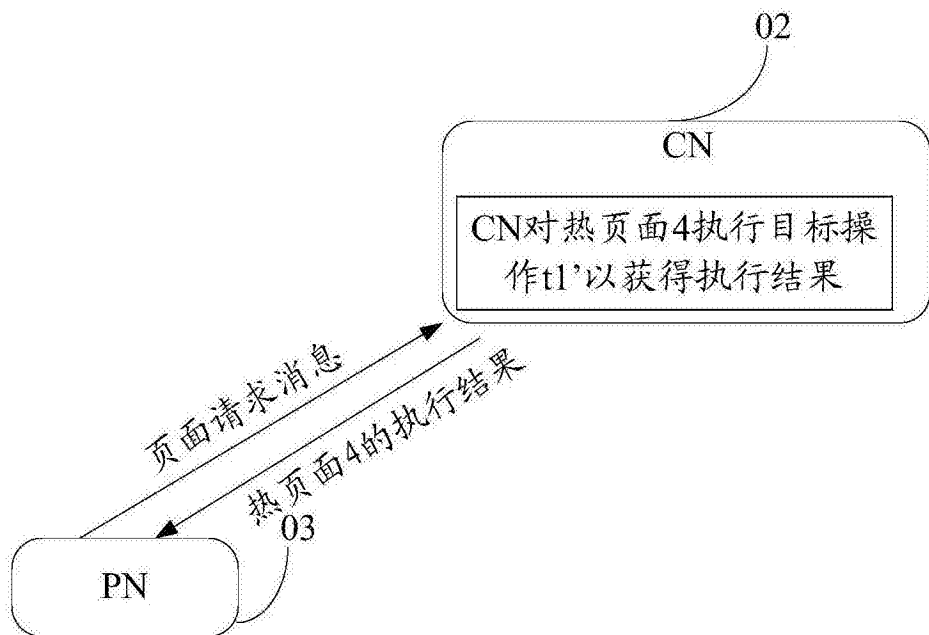


图2-8

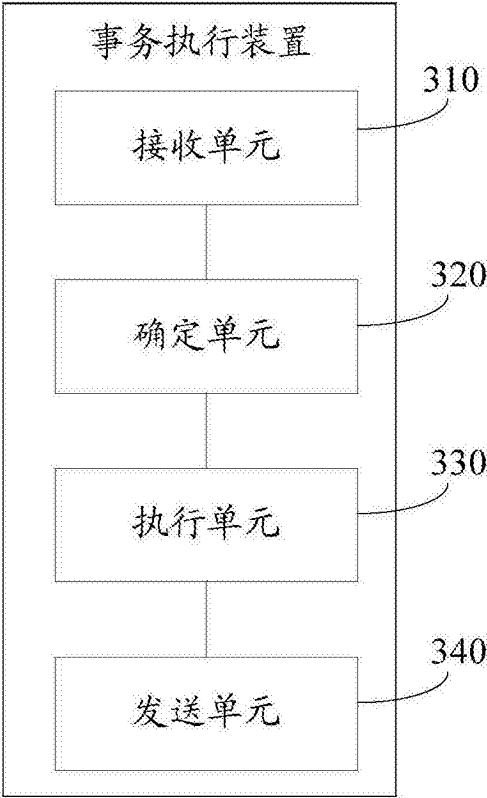


图3-1

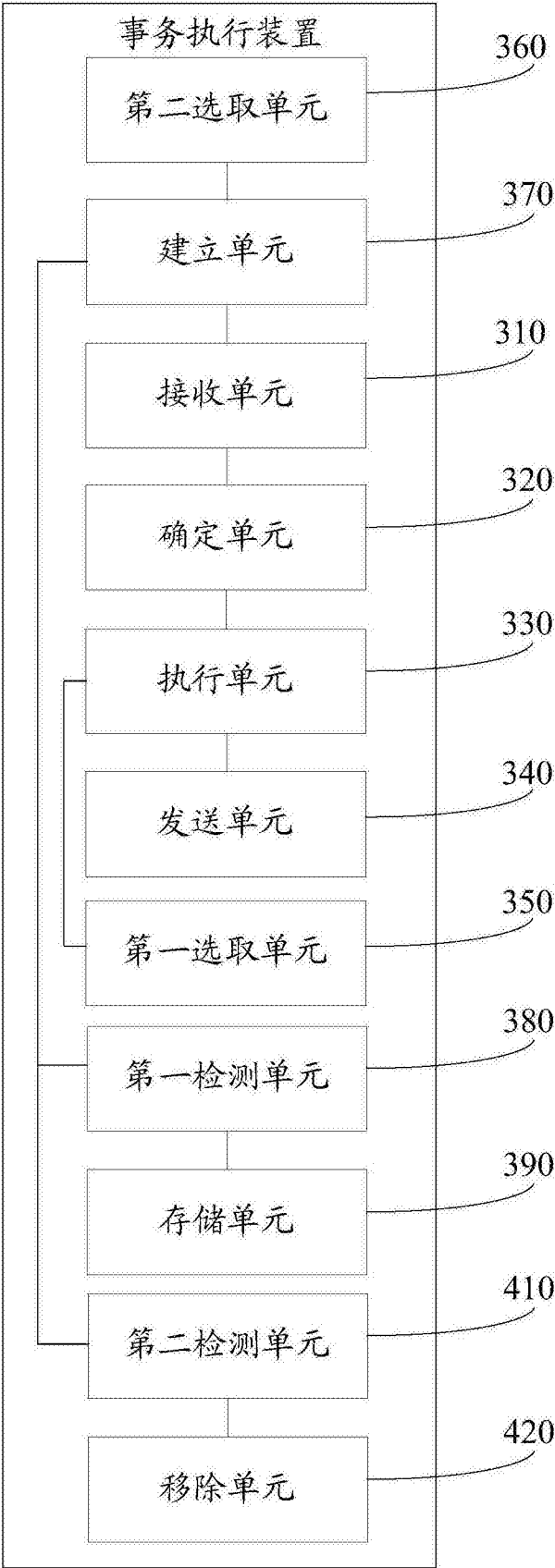


图3-2

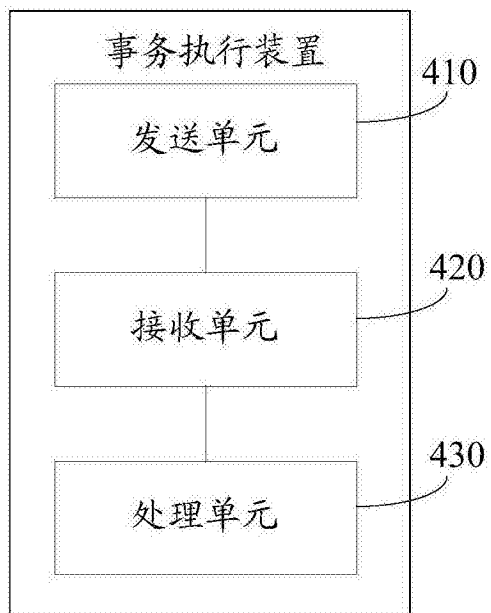


图4

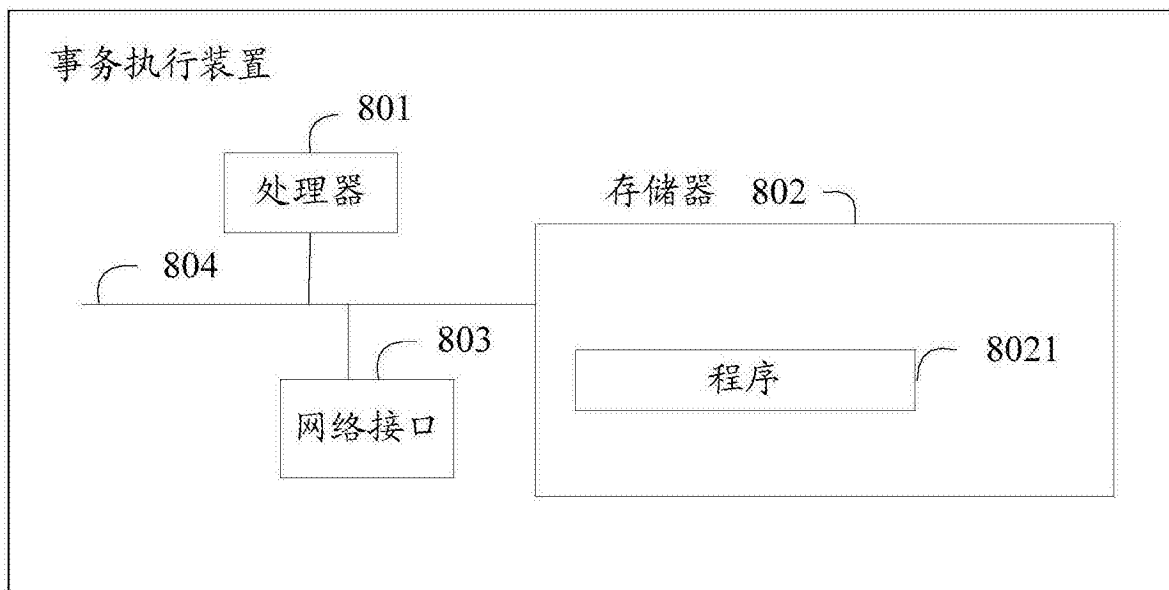


图5

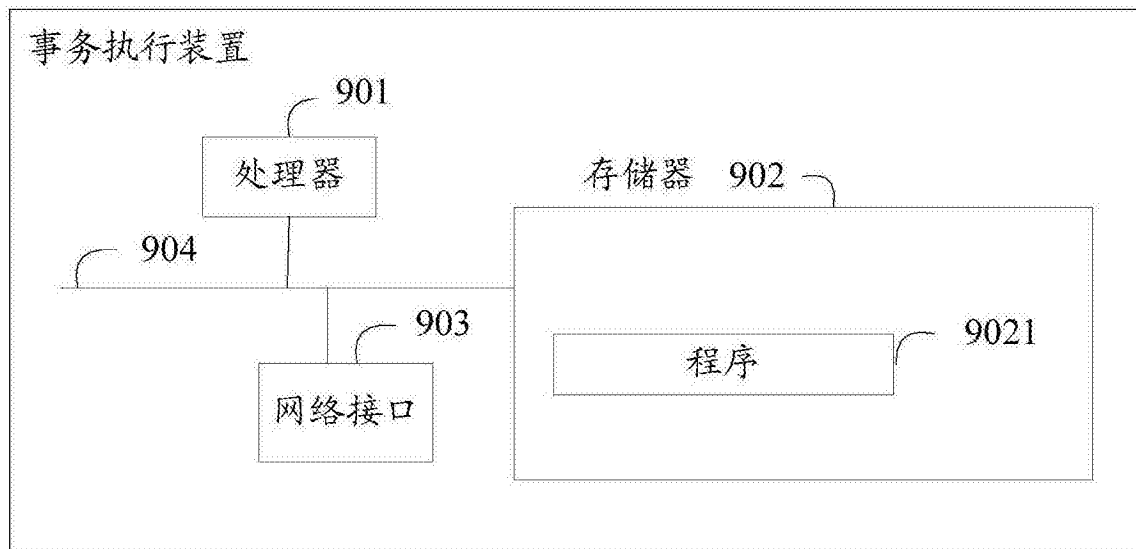


图6