



(12)发明专利

(10)授权公告号 CN 104462225 B

(45)授权公告日 2018.01.12

(21)申请号 201410635413.7

(22)申请日 2014.11.12

(65)同一申请的已公布的文献号
申请公布号 CN 104462225 A

(43)申请公布日 2015.03.25

(73)专利权人 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 肖宇雷 陈志远 叶涛

(74)专利代理机构 深圳市深佳知识产权代理事务所(普通合伙) 44285
代理人 王仲凯

(51)Int.Cl.
G06F 17/30(2006.01)

(56)对比文件

CN 1595905 A, 2005.03.16,
CN 103237313 A, 2013.08.07,
US 8818942 B2, 2014.08.26,
US 2012017037 A1, 2012.01.19,

审查员 张倩

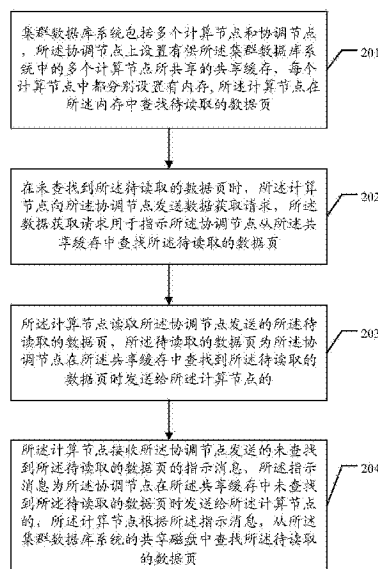
权利要求书3页 说明书14页 附图7页

(54)发明名称

一种数据读取的方法、装置及系统

(57)摘要

本发明公开了一种数据读取的方法,所述集群数据库系统包括计算节点和协调节点,协调节点上设置有多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:计算节点在内存中查找待读取的数据页,在未查找到待读取的数据页时,向协调节点发送数据获取请求,所述数据获取请求用于指示协调节点从所述共享缓存中查找所述待读取的数据页,计算节点读取协调节点发送的所述待读取的数据页,所述待读取的数据页为协调节点在所述共享缓存中查找到所述待读取的数据页时发送给计算节点的。本发明实施例提供的数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。



1. 一种数据读取的方法,其特征在于,所述方法应用于集群数据库系统中的计算节点,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:

所述计算节点在所述内存中查找待读取的数据页;

在未查找到所述待读取的数据页时,所述计算节点向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

所述计算节点读取所述协调节点发送的所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的;或者,

所述计算节点接收所述协调节点发送的未查找到所述待读取的数据页的指示消息,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的,所述计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

2. 根据权利要求1所述的方法,其特征在于,所述计算节点向所述协调节点发送数据获取请求之前,所述方法还包括:

所述计算节点根据淘汰算法,从内存中确定出淘汰页;

当所述淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;

当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

3. 一种数据读取的方法,其特征在于,所述方法应用于集群数据库系统中的协调节点,所述集群数据库系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:

所述协调节点接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的;

所述协调节点根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;

所述协调节点在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页;

所述协调节点在未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

4. 根据权利要求3所述的方法,其特征在于,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,所述从所述共享缓存中查找所述待读取的数据页,包括:

从设置在所述协调节点上的内存中查找是否有所述待读取的数据页;

当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

5. 根据权利要求4所述的方法,其特征在于,所述从所述SSD中查找对应的待读取的数据页,包括:

从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息;

当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

6. 根据权利要求5所述的方法,其特征在于,所述协调节点接收数据获取请求之前,所述方法还包括:

所述协调节点接收所述计算节点发送的淘汰页;

将所述淘汰页存放到所述SSD中,并将所述淘汰页的标识信息添加到所述SSD哈希表中。

7. 一种应用于集群数据库系统中的计算节点,其特征在于,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述计算节点包括:

查找单元,用于在所述内存中查找待读取的数据页;

发送单元,用于所述查找单元在内存中未查找到待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

接收单元,用于接收所述协调节点发送的所述待读取的数据页或未查找到所述待读取的数据页的指示消息,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的;

读取单元,用于读取所述待读取的数据页;

所述查找单元,还用于根据所述接收单元接收的所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

8. 根据权利要求7所述的计算节点,其特征在于,所述计算节点还包括:确定单元,

所述确定单元,用于根据淘汰算法,从内存中确定出淘汰页;

所述发送单元,还用于当所述确定单元确定的淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

9. 一种应用于集群数据库系统中的协调节点,其特征在于,所述集群数据系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述协调节点包括:

接收单元,用于接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的;

查找单元,用于根据所述接收单元接收的所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;

发送单元,用于在所述查找单元查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页,在所述查找单元未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的

共享磁盘中查找所述待读取的数据页。

10. 根据权利要求9所述的协调节点,其特征在于,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,

所述查找单元,具体用于从设置在所述协调节点上的内存中查找是否有所述待读取的数据页,当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

11. 根据权利要求10所述的协调节点,其特征在于,

所述查找单元,具体用于从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息,当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

12. 根据权利要求11所述的协调节点,其特征在于,所述协调节点还包括:存放单元和添加单元,

所述接收单元,还用于接收所述计算节点发送的淘汰页;

所述存放单元,用于将所述接收单元接收的所述淘汰页存放到所述SSD中;

所述添加单元,用于并将所述存放单元存放到所述SSD中的淘汰页的标识信息添加到所述SSD哈希表中。

13. 一种集群数据库系统,其特征在于,包括:协调节点和多个计算节点,所述协调节点上设置有供所述集群数据库系统中所述多个计算节点所共享的共享缓存,所述多个计算节点中的每个计算节点中都分别设置有内存;

所述每个计算节点,用于在未查找到待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

所述协调节点,用于根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页,在查找到所述待读取的数据页时,向发送所述数据获取请求的所述每个计算节点发送所述待读取的数据页;

所述每个计算节点,还用于接收所述协调节点发送的所述待读取的数据页或未查找到所述待读取的数据页的指示消息,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述每个计算节点的,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送的。

一种数据读取的方法、装置及系统

技术领域

[0001] 本发明涉及计算机技术领域,具体涉及一种业务数据处理的方法、装置及系统。

背景技术

[0002] 在集群数据库系统中数据读取的过程是:计算节点先从自身内存中查找数据,如果自身内存中不存在查找的数据,则从自身固态硬盘(Solid State Drive,SSD)中查找,如果自身SSD中也不存在查找的数据,则从再从磁盘中查找。SSD是数据库系统中的扩展缓存,存放内存中淘汰出来的干净页。在读取数据的时候,如果内存中不命中,先尝试在SSD扩展缓存中命中。因为SSD的容量一般是内存容量的2-8倍,可以缓存大部分的热数据。所以,大部分的不在内存中命中的热数据可以快速在SSD中命中。

[0003] 现有技术的数据库系统中每个计算节点都要配一块SSD卡。如果集群中有10个计算节点,则共需要配10块SSD卡,导致数据库系统中要配多块SSD卡,而且单块SSD只为单个计算节点做读缓存无法充分利用SSD的容量,导致SSD的利用率低下。

发明内容

[0004] 本发明实施例提供的一种数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。本发明实施例还提供了相应的装置及系统。

[0005] 本发明第一方面提供一种数据读取的方法,所述方法应用于集群数据库系统中的计算节点,所述集群数据系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:

[0006] 所述计算节点在所述内存中查找待读取的数据页;

[0007] 在未查找到所述待读取的数据页时,所述计算节点向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

[0008] 所述计算节点读取所述协调节点发送的所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的;

[0009] 所述计算节点接收所述协调节点发送的未查找到所述待读取的数据页的指示消息,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的,所述计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0010] 结合第一方面,在第一种可能的实现方式中,所述计算节点所述向所述协调节点发送数据获取请求之前,所述方法还包括:

[0011] 所述计算节点根据淘汰算法,从内存中确定出淘汰页;

[0012] 当所述淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;

[0013] 当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

[0014] 本发明第二方面提供一种数据读取的方法,所述方法应用于集群数据库系统中的协调节点,所述集群数据库系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:

[0015] 所述协调节点接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的;

[0016] 所述协调节点根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;

[0017] 所述协调节点在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页;

[0018] 所述协调节点在未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0019] 结合第二方面,在第一种可能的实现方式中,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,所述从所述共享缓存中查找所述待读取的数据页,包括:

[0020] 从设置在所述协调节点上的内存中查找是否有所述待读取的数据页;

[0021] 当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

[0022] 结合第二方面第一种可能的实现方式,在第二种可能的实现方式中,所述从所述SSD中查找对应的待读取的数据页,包括:

[0023] 从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息;

[0024] 当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0025] 结合第二方面第二种可能的实现方式,在第三种可能的实现方式中,所述协调节点接收数据获取请求之前,所述方法还包括:

[0026] 所述协调节点接收所述计算节点发送的淘汰页;

[0027] 将所述淘汰页存放到所述SSD中,并将所述淘汰页的标识信息添加到所述SSD哈希表中。

[0028] 本发明第三方面提供一种应用于集群数据库系统中的计算节点,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述计算节点包括:

[0029] 查找单元,用于在所述内存中查找待读取的数据页;

[0030] 发送单元,用于所述查找单元在内存中未查找到待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找

所述待读取的数据页；

[0031] 接收单元,用于接收所述协调节点发送的所述待读取的数据页或未查找到所述待读取的数据页的指示消息,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的；

[0032] 读取单元,用于读取所述待读取的数据页；

[0033] 所述查找单元,还用于根据所述接收单元接收的所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0034] 结合第三方面,在第一种可能的实现方式中,所述计算节点还包括:确定单元,

[0035] 所述确定单元,用于根据淘汰算法,从内存中确定出淘汰页；

[0036] 所述发送单元,还用于当所述确定单元确定的淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

[0037] 本发明第四方面提供一种应用于集群数据库系统中的协调节点,所述集群数据系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述协调节点包括:

[0038] 接收单元,用于接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的；

[0039] 查找单元,用于根据所述接收单元接收的所述数据获取请求,从所述共享缓存中查找所述待读取的数据页；

[0040] 发送单元,用于在所述查找单元查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页,在所述查找单元未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0041] 结合第四方面,在第一种可能的实现方式中,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,

[0042] 所述查找单元,具体用于从设置在所述协调节点上的内存中查找是否有所述待读取的数据页,当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

[0043] 结合第四方面第一种可能的实现方式,在第二种可能的实现方式中,

[0044] 所述查找单元,具体用于从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息,当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0045] 结合第四方面第二种可能的实现方式,在第三种可能的实现方式中,所述协调节点还包括:存放单元和添加单元,

[0046] 所述接收单元,还用于接收所述计算节点发送的淘汰页；

[0047] 所述存放单元,用于将所述接收单元接收的所述淘汰页存放到所述SSD中；

[0048] 所述添加单元,用于并将所述存放单元存放到所述SSD中的淘汰页的标识信息添

加到所述SSD哈希表中。

[0049] 本发明第五方面提供一种集群数据库系统,包括:多个计算节点和协调节点,所述协调节点上设置有供所述集群数据库系统中所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存;

[0050] 所述计算节点,用于在所述内存中查找待读取的数据页,在未查找到所述待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

[0051] 所述协调节点,用于根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页,在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页;

[0052] 所述计算节点,用于读取所述协调节点发送的所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的。

[0053] 本发明实施例采用在集群数据库系统的协调节点上设置可供多个计算节点共享的共享缓存,计算节点在所述内存中查找待读取的数据页,计算节点在内存中未查找到待读取的数据页时,向协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页,在协调节点查找到所述待读取的数据页时,将所述待读取的数据页发送给所述计算节点,当协调节点未查找到所述待读取的数据页时,发送未查找到所述待读取的数据页的指示消息给所述计算节点,计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

附图说明

[0054] 为了更清楚地说明本发明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。

[0055] 图1是本发明实施例中集群数据库系统的一实施例示意图;

[0056] 图2是本发明实施例中集群数据库系统的另一实施例示意图;

[0057] 图3是本发明实施例中数据读取的方法的一实施例示意图;

[0058] 图4是本发明实施例中数据读取的方法的另一实施例示意图;

[0059] 图5是本发明实施例中计算节点的一实施例示意图;

[0060] 图6是本发明实施例中计算节点的另一实施例示意图;

[0061] 图7是本发明实施例中协调节点的一实施例示意图;

[0062] 图8是本发明实施例中协调节点的另一实施例示意图;

[0063] 图9是本发明实施例中计算节点的另一实施例示意图;

[0064] 图10是本发明实施例中协调节点的另一实施例示意图；

[0065] 图11是本发明实施例中集群数据库系统的一实施例示意图。

具体实施方式

[0066] 本发明实施例提供一种数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。本发明实施例还提供了相应的装置及系统以下分别进行详细说明。

[0067] 为了使本技术领域的人员更好地理解本发明方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分的实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都应当属于本发明保护的范围。

[0068] 本发明实施例中计算节点和协调节点都可以为计算机、服务器等通信设备。

[0069] 参阅图1,本发明实施例提供的集群数据库系统包括多个计算节点,如图1中的计算节点1、计算节点2...计算节点K,K为大于2的整数。多个计算节点共同使用共享磁盘,共享磁盘作为多个计算节点的共享存储设备。多个计算节点通过网络通道与协调节点通信,图1所示的协调节点包括主协调节点和备协调节点两个,主协调节点正常时,备协调节点处于空闲状态,当主协调节点故障时,备协调节点接替主协调节点的工作。主协调节点和备协调节点上都设置有内存和固态硬盘(Solid State Drive,SSD),SSD也被叫做闪存、FLASH等。每个计算节点上都有各自的内存,每个计算节点可以通过网络通道与主协调节点通信。在本发明实施例中,内存的读取速度最快、SSD的读取速度居中,共享磁盘的读取数据最慢。

[0070] 协调节点负责管理和维护多个计算节点的资源信息。在集群环境下,多个计算节点同时访问共享资源,可能会造成共享资源的并发访问,这种情况下,协调节点就需要对共享资源实现统一的管理。

[0071] 计算节点负责执行具体的事务,可以直接访问磁盘等共享存储设备。

[0072] 多个计算节点共享协调节点中设置的内存和SSD,为了满足计算节点对协调节点中SSD的使用,可以在协调节点中设置全局buffer接口装置,在每个计算节点中设置SSD读缓存管理装置,如图2所示。

[0073] 全局Buffer接口装置:主要提供接口供计算节点使用协调节点上面的SSD等公共资源。

[0074] SSD读缓存管理装置:位于内存与文件系统之间,将SSD作为内存与文件之间的缓存进行管理。在计算节点向协调节点请求SSD中的数据时,SSD读缓存管理装置与全局Buffer接口进行交互,从而获得SSD中的数据页。

[0075] 本发明实施例中数据读取的过程以计算节点1为例,可以参阅图2中的S100、S102、S104、106和108进行理解:

[0076] 本发明实施例中协调节点的共享缓存包括设置在协调节点中的内存和SSD。

[0077] 计算节点1中内存缺页时,根据淘汰算法,从内存中确定出淘汰页。当所述淘汰页为脏页时,执行步骤S100,当所述淘汰页为脏页时,执行步骤S102。

[0078] 淘汰算法可以有多种,例如:先进先出(First Input First Output,FIFO)算法、

最近最少使用(Least Recently Used,LRU)算法等。

[0079] S100、计算节点1将所述淘汰页发送到所述多个计算节点的共享磁盘进行存储。

[0080] S102、计算节点1将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述SSD中。

[0081] 协调节点上会维护一个SSD哈希表,所述SSD哈希表记录所述SSD中存储的数据页的标识信息,当SSD中新增淘汰页时,SSD哈希表也会添加该淘汰页的标识信息。

[0082] S104、计算节点1在所述内存中查找待读取的数据页,在自身内存中未查找到待读取的数据页时,向协调节点发送数据获取请求。

[0083] 计算节点1在所述内存中查找待读取的数据页的动作可能是由用户向所述计算节点1中执行写操作触发的,也可能是计算节点1在执行业务程序所触发的。

[0084] S106、所述协调节点根据所述数据获取请求先从协调节点内存中查找对应的待读取的数据页,协调节点的内存中没有该待读取的数据页时,从所述SSD中查找对应的待读取的数据页,当查找到所述待读取的数据页时,向计算节点1发送所述待读取的数据页,当未查找到所述待读取的数据页时,向计算节点1发送未查找到所述待读取的数据页的指示消息。

[0085] 计算节点1接收到所述待读取的数据页时,读取所述待读取的数据页。

[0086] 计算节点1接收到未查找到所述待读取的数据页的指示消息时,执行步骤S108。

[0087] 所述协调节点根据所述数据获取请求,从所述SSD中查找对应的待读取的数据页,可以包括:

[0088] 所述协调节点从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息;

[0089] 当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0090] S108、计算节点1从磁盘查找所述待读取的数据页。

[0091] 与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0092] 参阅图3,本发明实施例提供的数据读取的方法的一实施例,所述方法应用于集群数据库系统中的计算节点,所述集群数据系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法可以包括:

[0093] 201、所述计算节点在所述内存中查找待读取的数据页。

[0094] 202、在未查找到所述待读取的数据页时,所述计算节点向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页。

[0095] 203、所述计算节点读取所述协调节点发送的所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的。

[0096] 204、所述计算节点接收所述协调节点发送的未查找到所述待读取的数据页的指示消息,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的,所述计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0097] 本发明实施例采用在集群数据库系统的协调节点上设置可供多个计算节点共享的共享缓存,计算节点在所述内存中查找待读取的数据页,计算节点在内存中未查找到待读取的数据页时,向协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页,在协调节点查找到所述待读取的数据页时,将所述待读取的数据页发送给所述计算节点,当协调节点未查找到所述待读取的数据页时,发送未查找到所述待读取的数据页的指示消息,计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页,并发送给计算节点。与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0098] 可选地,在上述图3对应的实施例的基础上,本发明实施例提供的数据读取的方法的另一实施例中,

[0099] 所述计算节点所述向所述协调节点发送数据获取请求之前,所述方法还包括:

[0100] 所述计算节点根据淘汰算法,从内存中确定出淘汰页;

[0101] 当所述淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;

[0102] 当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

[0103] 本发明图3所示的实施例及相应的可选实施例提供的数据读取的过程,可以参阅图1和图2部分的描述进行理解,本处不做过多赘述。

[0104] 参阅图4,本发明实施例提供的数据读取的方法的一实施例,所述方法应用于集群数据库系统中的协调节点,所述集群数据库系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:

[0105] 301、所述协调节点接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的。

[0106] 302、所述协调节点根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页。

[0107] 303、所述协调节点在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页。

[0108] 304、所述协调节点在未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0109] 本发明实施例提供的数据读取的方法,应用于集群数据库系统中的协调节点,所

述集群数据库系统还包括多个计算节点,所述协调节点上设置有所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述方法包括:所述协调节点接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送给所述协调节点的;所述协调节点根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;所述协调节点在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页,所述协调节点在未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的数据读取的方法,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0110] 可选地,在上述图4对应的实施例的基础上,本发明实施例提供的数据读取的方法的另一实施例中,

[0111] 所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,所述从所述共享缓存中查找所述待读取的数据页,包括:

[0112] 从设置在所述协调节点上的内存中查找是否有所述待读取的数据页;

[0113] 当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

[0114] 可选地,在上一个可选实施例的基础上,本发明实施例提供的数据读取的方法的另一实施例中,所述从所述SSD中查找对应的待读取的数据页,可以包括:

[0115] 所述协调节点从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息;

[0116] 当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0117] 可选地,在上述图4对应的上一可选实施例的基础上,本发明实施例提供的数据读取的方法的另一实施例中,所述协调节点接收每个计算节点发送的数据获取请求之前,所述方法还可以包括:

[0118] 所述协调节点接收所述计算节点发送的淘汰页;

[0119] 将所述淘汰页存放到所述SSD中,并将所述淘汰页的标识信息添加到所述SSD哈希表中。

[0120] 本发明图4所示的实施例及相应的可选实施例提供的数据读取的过程,可以参阅图1和图2部分的描述进行理解,本处不做过多赘述。

[0121] 参阅图5,本发明实施例提供的应用于集群数据库系统中的计算节点的一实施例,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述计算节点40包括:

[0122] 查找单元401,用于在所述内存中查找待读取的数据页;

[0123] 发送单元402,用于所述查找单元401在内存中未查找到待读取的数据页时,向所

述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

[0124] 接收单元403,用于接收所述协调节点发送的所述待读取的数据页或未查找到所述待读取的数据页的指示消息,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的;

[0125] 读取单元404,用于读取所述接收单元403接收的所述待读取的数据页。

[0126] 所述查找单元401,还用于根据所述接收单元403接收的所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0127] 本发明实施例中,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述计算节点40包括:查找单元401在所述内存中查找待读取的数据页;发送单元402所述查找单元401在内存中未查找到待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;接收单元403接收所述协调节点发送的所述待读取的数据页或未查找到所述待读取的数据页的指示消息,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的;读取单元404读取所述接收单元403接收的所述待读取的数据页,所述查找单元401还根据所述接收单元403接收的所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的计算节点,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0128] 可选地,在上述图5对应的实施例的基础上,参阅图6,本发明实施例提供的计算节点的另一实施例中,所述计算节点还包括:确定单元405,

[0129] 所述确定单元405,用于根据淘汰算法,从内存中确定出淘汰页;

[0130] 所述发送单元402,还用于当所述确定单元405确定的淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储;当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

[0131] 本发明实施例提供的计算节点的实施例可以参阅图1、图2和图3部分的方法实施例进行理解,本处不做过多赘述。

[0132] 参阅图7,本发明实施例提供的应用于集群数据库系统中的协调节点的一实施例中,所述集群数据系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述协调节点50包括:

[0133] 接收单元501,用于接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送的;

[0134] 查找单元502,用于根据所述接收单元501接收的所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;

[0135] 发送单元503,用于在所述查找单元502查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页,在所述查找单元502未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0136] 本发明实施例中,所述集群数据系统还包括多个计算节点,所述协调节点上设置有供所述集群数据库系统中的所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,所述协调节点50包括:接收单元501接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送的;查找单元502根据所述接收单元501接收的所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;发送单元503在所述查找单元502查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页,在所述查找单元502未查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。与现有技术中每个计算节点都要配一块SSD卡,单块SSD只为单个计算节点做读缓存,无法充分利用SSD的容量,导致SSD的利用率低下相比,本发明实施例提供的协调节点,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0137] 可选地,在上述图7对应的实施例的基础上,本发明实施例提供的协调节点50的另一实施例中,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,

[0138] 所述查找单元502,具体用于从设置在所述协调节点上的内存中查找是否有所述待读取的数据页,当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

[0139] 可选地,在上一可选实施例的基础上,本发明实施例提供的协调节点50的另一实施例中,

[0140] 所述查找单元502,具体用于从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息,当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0141] 可选地,在上一个可选实施例的基础上,本发明实施例提供的协调节点50的另一实施例中,参阅图8,所述协调节点还包括:存放单元504和添加单元505,

[0142] 所述接收单元501,还用于接收所述计算节点发送的淘汰页;

[0143] 所述存放单元504,用于将所述接收单元501接收的所述淘汰页存放到所述SSD中;

[0144] 所述添加单元505,用于并将所述存放单元504存放到所述SSD中的淘汰页的标识信息添加到所述SSD哈希表中。

[0145] 本发明实施例提供的协调节点的实施例可以参阅图1、图2和图4部分的方法实施例进行理解,本处不做过多赘述。

[0146] 图9是本发明实施例计算节点40的结构示意图。计算节点40可包括输入设备410、输出设备420、处理器430和存储器440。

[0147] 本发明实施例提供的计算节点40应用于集群数据库系统中的计算节点,所述集群

数据系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存,内存可以是存储器440的一部分。

[0148] 存储器440可以包括只读存储器和随机存取存储器,并向处理器430提供指令和数据。存储器440的一部分还可以包括非易失性随机存取存储器(NVRAM)。

[0149] 存储器440存储了如下的元素,可执行模块或者数据结构,或者它们的子集,或者它们的扩展集:

[0150] 操作指令:包括各种操作指令,用于实现各种操作。

[0151] 操作系统:包括各种系统程序,用于实现各种基础业务以及处理基于硬件的任务。

[0152] 在本发明实施例中,处理器430通过调用存储器440存储的操作指令(该操作指令可存储在操作系统中),执行如下操作:

[0153] 在所述内存中查找待读取的数据页;

[0154] 在未查找到所述待读取的数据页时,通过输出设备420向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

[0155] 读取所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的。

[0156] 通过输入设备410接收所述协调节点发送的未查找到所述待读取的数据页的指示消息,所述指示消息为所述协调节点在所述共享缓存中未查找到所述待读取的数据页时发送给所述计算节点的,所述计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0157] 本发明实施例提供的计算节点40,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0158] 处理器430控制计算节点40的操作,处理器430还可以称为CPU(Central Processing Unit,中央处理单元)。存储器440可以包括只读存储器和随机存取存储器,并向处理器430提供指令和数据。存储器440的一部分还可以包括非易失性随机存取存储器(NVRAM)。具体的应用中,计算节点40的各个组件通过总线系统450耦合在一起,其中总线系统450除包括数据总线之外,还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见,在图中将各种总线都标为总线系统450。

[0159] 上述本发明实施例揭示的方法可以应用于处理器430中,或者由处理器430实现。处理器430可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器430中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器430可以是通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(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。可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。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。结合本发明实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器440,处理器430读

取存储器440中的信息,结合其硬件完成上述方法的步骤。

[0160] 可选地,所述处理器430根据淘汰算法,从内存中确定出淘汰页;

[0161] 所述输出设备420用于当所述淘汰页为脏页时,将所述淘汰页发送到所述共享磁盘进行存储,当所述淘汰页为干净页时,将所述淘汰页发送到所述协调节点,以便所述协调节点将所述淘汰页存储到所述共享缓存中。

[0162] 图10是本发明实施例协调节点50的结构示意图。协调节点50可包括输入设备510、输出设备520、处理器530和存储设备540。

[0163] 本发明实施例提供的协调节点50应用于集群数据库系统中的计算节点,所述计算节点有多个,所述集群数据库系统还包括协调节点,所述协调节点上设置有供所述集群数据库系统中的多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存。

[0164] 存储设备540可以包括只读存储器和随机存取存储器,并向处理器530提供指令和数据。存储设备540的一部分还可以包括非易失性随机存取存储器(NVRAM)。

[0165] 存储设备540存储了如下的元素,可执行模块或者数据结构,或者它们的子集,或者它们的扩展集:

[0166] 操作指令:包括各种操作指令,用于实现各种操作。

[0167] 操作系统:包括各种系统程序,用于实现各种基础业务以及处理基于硬件的任务。

[0168] 在本发明实施例中,处理器530通过调用存储设备540存储的操作指令(该操作指令可存储在操作系统中),执行如下操作:

[0169] 通过输入设备510接收数据获取请求,所述数据获取请求为计算节点在内存中未查找到待读取的数据页时发送的;

[0170] 所述协调节点根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页;

[0171] 通过输出设备520在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页;

[0172] 在未查找到所述待读取的数据页时,通过所述输出设备520向发送所述数据获取请求的计算节点发送未查找到所述待读取的数据页的指示消息,以使所述发送所述数据获取请求的计算节点根据所述指示消息,从所述集群数据库系统的共享磁盘中查找所述待读取的数据页。

[0173] 本发明实施例提供的协调节点50,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0174] 处理器530控制协调节点50的操作,处理器530还可以称为CPU(Central Processing Unit,中央处理单元)。存储设备540可以包括只读存储器和随机存取存储器,并向处理器530提供指令和数据。存储设备540的一部分还可以包括非易失性随机存取存储器(NVRAM)。具体的应用中,协调节点50的各个组件通过总线系统550耦合在一起,其中总线系统550除包括数据总线之外,还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见,在图中将各种总线都标为总线系统550。

[0175] 上述本发明实施例揭示的方法可以应用于处理器530中,或者由处理器530实现。处理器530可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器530中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理

器530可以是通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现成可编程门阵列(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。可以实现或者执行本发明实施例中的公开的各方法、步骤及逻辑框图。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。结合本发明实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储设备540,处理器530读取存储设备540中的信息,结合其硬件完成上述方法的步骤。

[0176] 可选地,所述共享缓存包括设置在所述协调节点上的内存和固态硬盘SSD,所述处理器530具体用于从设置在所述协调节点上的内存中查找是否有所述待读取的数据页,当在所述协调节点内存中未查找到所述待读取的数据页时,从所述SSD中查找所述待读取的数据页。

[0177] 可选地,所述处理器530具体用于从SSD哈希表中查找是否有所述待读取的数据页的标识信息,所述SSD哈希表记录所述SSD中存储的数据页的标识信息,当所述SSD哈希表中有所述待读取的数据页的标识信息时,则从所述SSD中获取对应的待读取的数据页。

[0178] 可选地,所述输入设备510接收所述计算节点发送的淘汰页;

[0179] 所述处理器530将所述淘汰页存放到所述SSD中,并将所述淘汰页的标识信息添加到所述SSD哈希表中。

[0180] 参阅图11,本发明实施例提供的集群数据库系统的一实施例包括:多个计算节点40和协调节点50,所述协调节点50上设置有供所述集群数据库系统中所述多个计算节点所共享的共享缓存,每个计算节点中都分别设置有内存;

[0181] 所述计算节点40,在所述内存中查找待读取的数据页,在未查找到所述待读取的数据页时,向所述协调节点发送数据获取请求,所述数据获取请求用于指示所述协调节点从所述共享缓存中查找所述待读取的数据页;

[0182] 所述协调节点50,用于根据所述数据获取请求,从所述共享缓存中查找所述待读取的数据页,在查找到所述待读取的数据页时,向发送所述数据获取请求的计算节点发送所述待读取的数据页;

[0183] 所述计算节点40,用于读取所述协调节点发送的所述待读取的数据页,所述待读取的数据页为所述协调节点在所述共享缓存中查找到所述待读取的数据页时发送给所述计算节点的。

[0184] 本发明实施例提供的集群数据库系统,在保证数据读取速率的同时,可以减少集群数据库系统中SSD的数量,提高集群数据库系统中共享缓存的利用率。

[0185] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件(例如处理器)来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:ROM、RAM、磁盘或光盘等。

[0186] 以上对本发明实施例所提供的数据读取的方法、装置以及系统进行了详细介绍,本文中应用了具体个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。

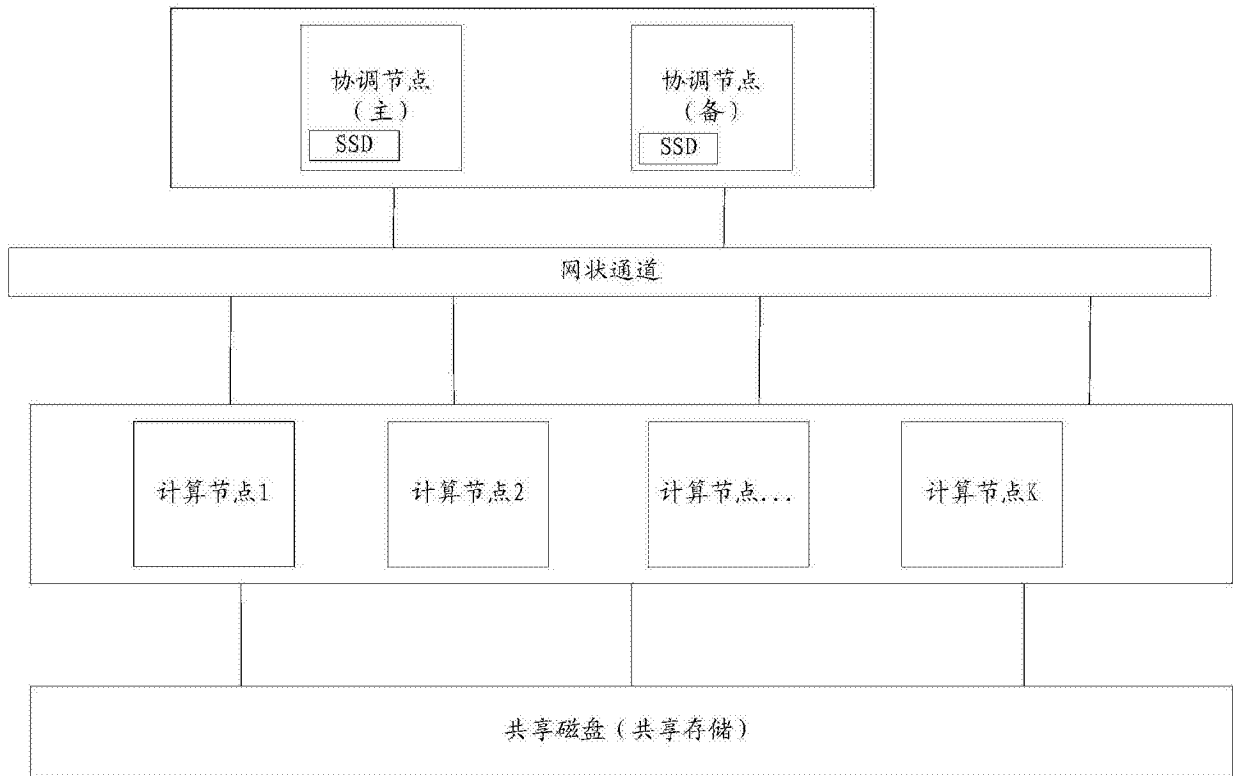


图1

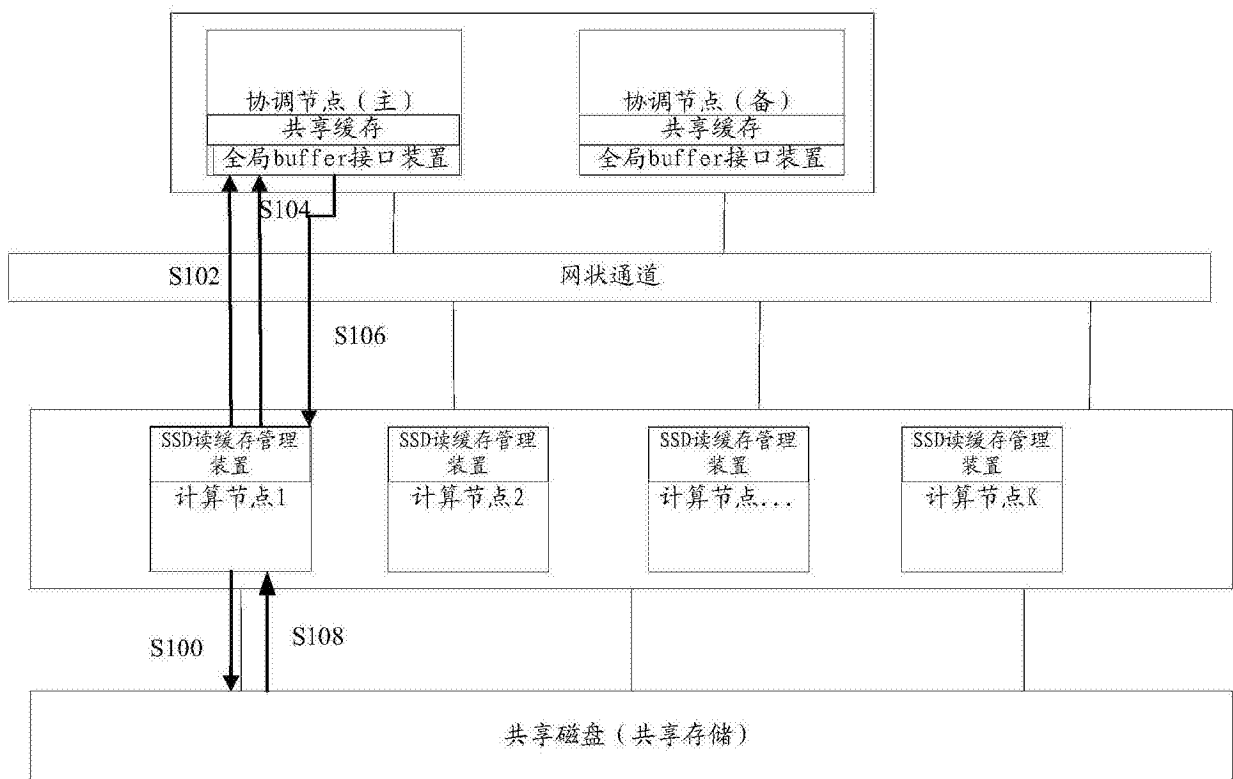


图2

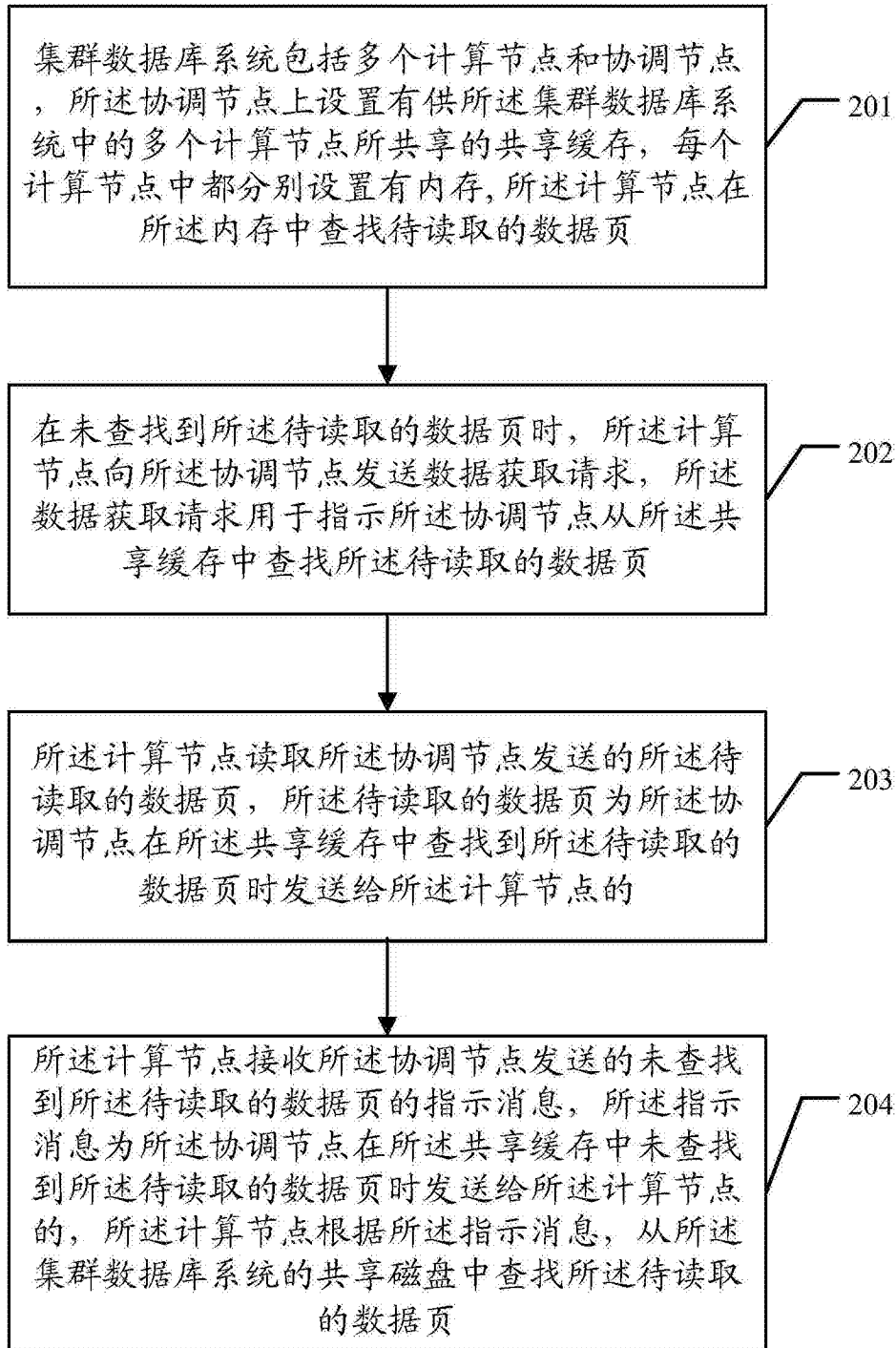


图3

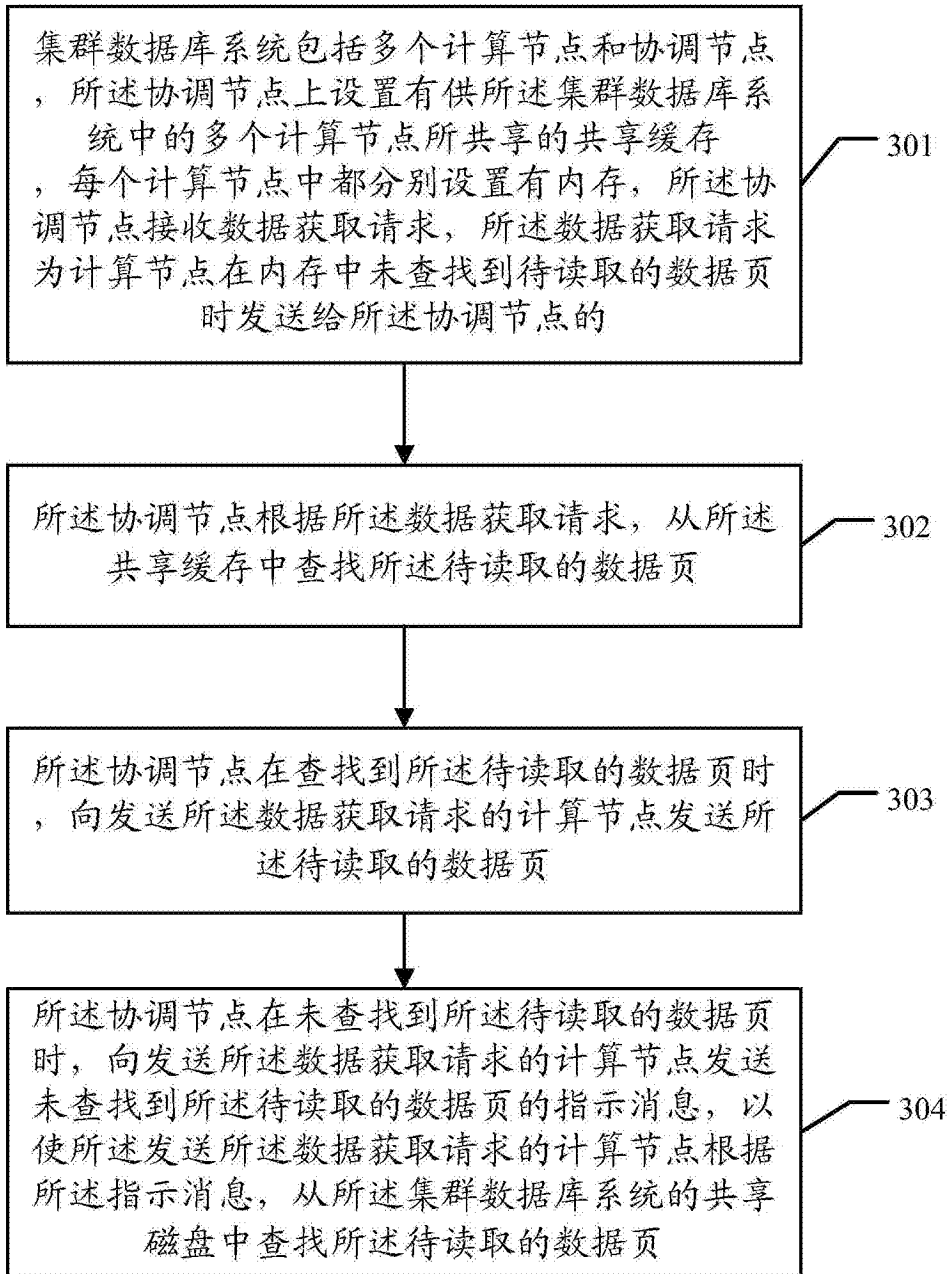


图4

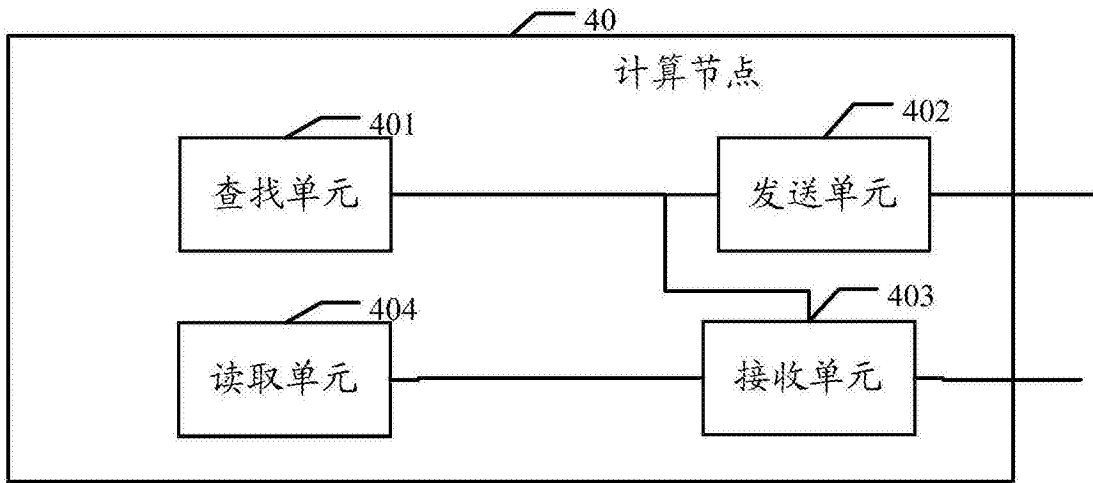


图5

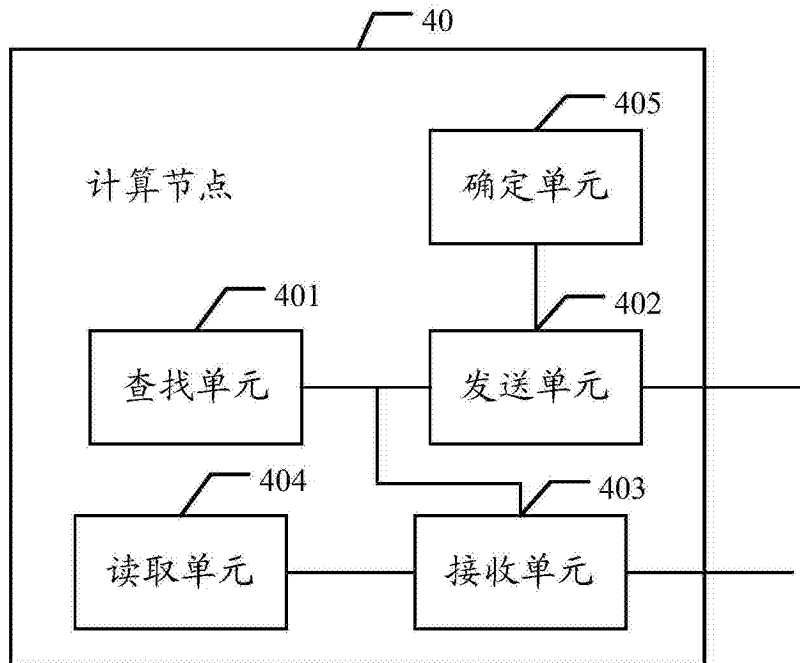


图6

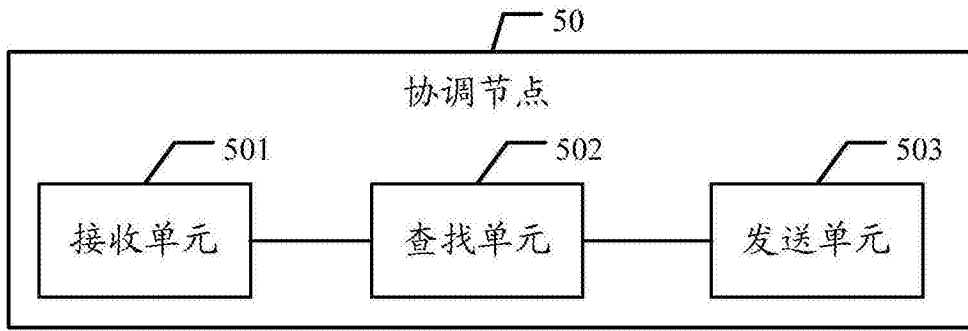


图7

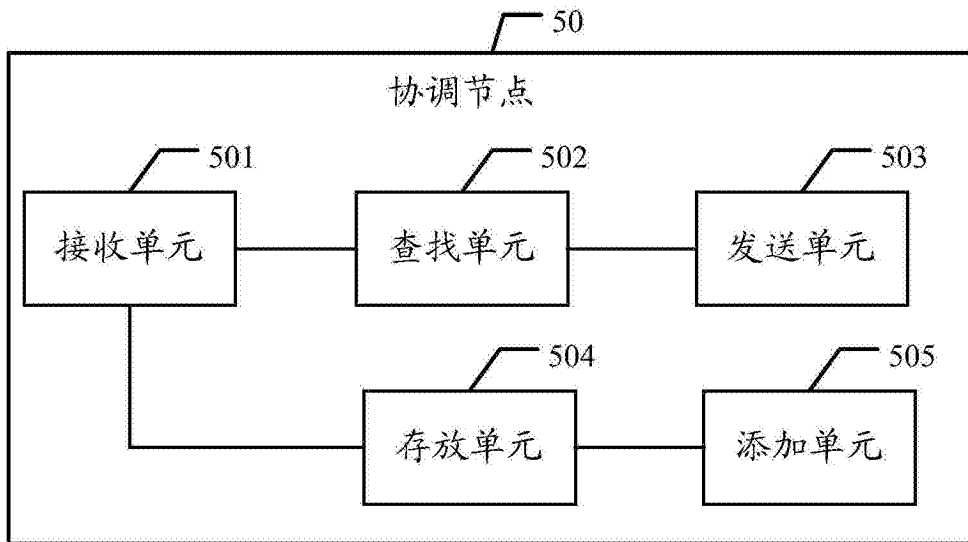


图8

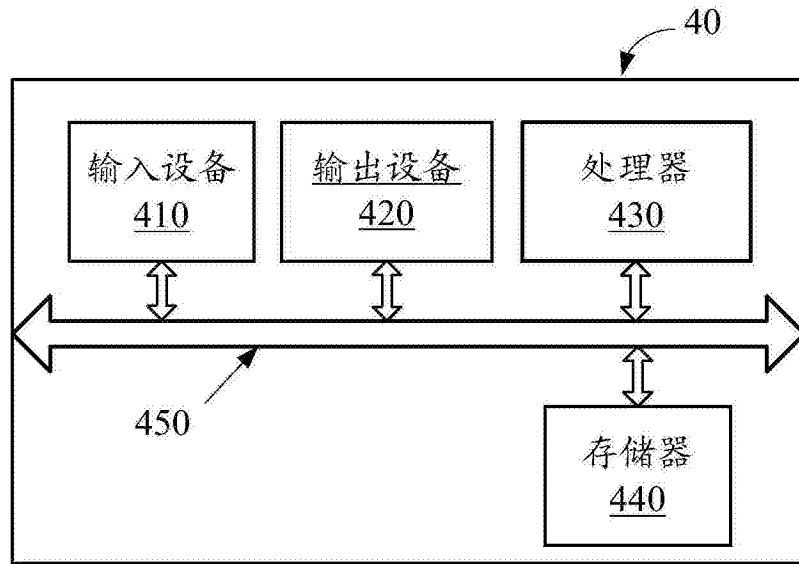


图9

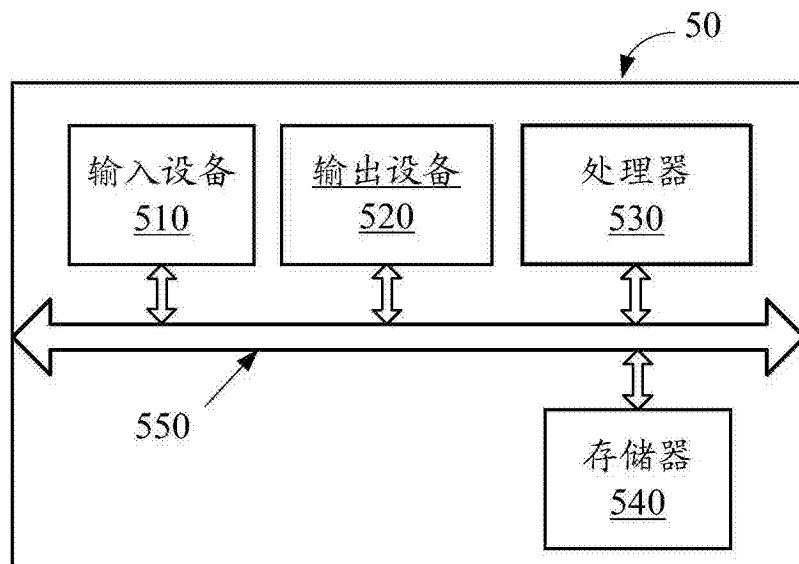


图10

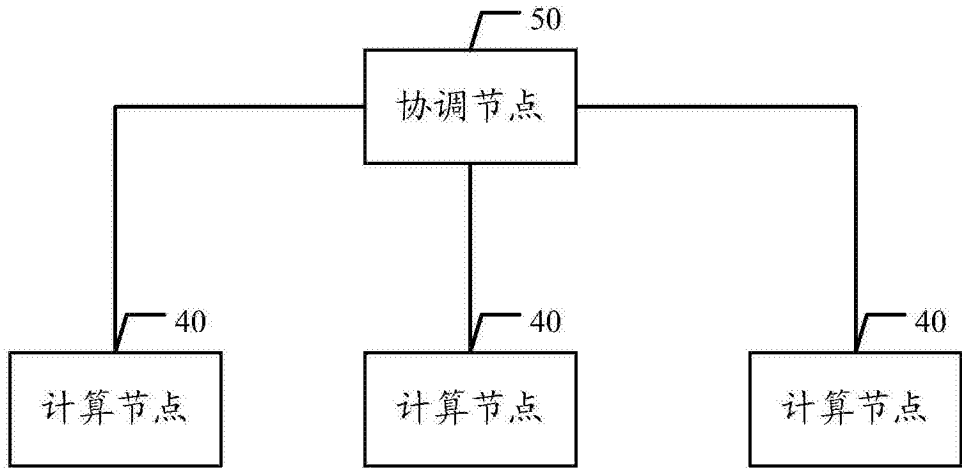


图11