(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 107111596 B
(45)授权公告日 2018.10.30

(21)申请号 201580002558.6

(22)申请日 2015.12.14

(65)同一申请的已公布的文献号
申请公布号 CN 107111596 A

(43)申请公布日 2017.08.29

(85)PCT国际申请进入国家阶段日
2016.05.11

(86)PCT国际申请的申请数据
PCT/CN2015/097267 2015.12.14

(87)PCT国际申请的公布数据
WO2017/100978 ZH 2017.06.22

(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 蓝文海 张巍 王学友 张粤

(51)Int.Cl.
G06F 15/177(2006.01)
H04L 29/08(2006.01)

(56)对比文件
CN 104536834 A,2015.04.22,
CN 104461707 A,2015.03.25,
CN 104063355 A,2014.09.24,
US 2006242644 A1,2006.10.26,
US 2005039049 A1,2005.02.17,

审查员 李晨

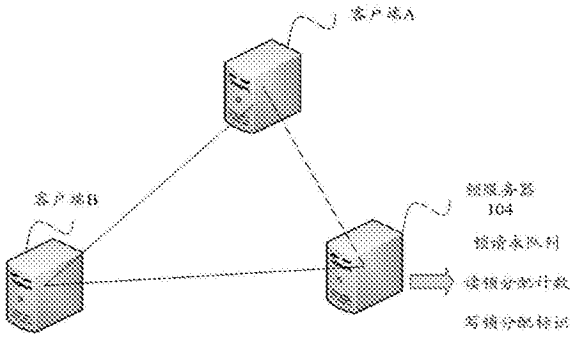
权利要求书3页 说明书24页 附图4页

(54)发明名称

一种集群中锁管理的方法、锁服务器及客户端

(57)摘要

本发明提供了一种集群中锁管理的方案，集群中包括客户端和锁服务器，锁服务器包含接口卡和内存，内存中存储读锁请求队列，记录处于读锁等待或获得读锁的客户端的标识，内存中还存储写锁请求队列记录处于写锁等待或获得写锁的客户端的标识，内存中还包含读锁分配计数和写锁分配标识，前者用于记录分配的读锁，后者用于表示是否已经分配写锁，客户端与锁服务器的接口卡通过RDMA协议通信进行锁操作。



CN 107111596 B

1.一种集群中锁管理方法，所述集群包含多个客户端和锁服务器，所述锁服务器包含接口卡和内存，其特征在于，所述内存中存储有写锁请求队列、写锁分配标识和读锁分配计数；其中，所述写锁请求队列用于记录所述集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，所述写锁分配标识用于表示是否分配针对所述共享资源的写锁，所述读锁分配计数用于表示针对所述共享资源分配读锁的数量；所述方法包括：

所述接口卡接收所述多个客户端中的第一客户端发送的第一写锁请求；所述第一写锁请求携带第一客户端的标识和写锁请求队列队首的内存地址；

所述接口卡将所述第一客户端的标识写入所述写锁请求队列队首的内存地址，所述接口卡查询所述读锁分配计数确定是否已经分配读锁，当没有分配所述读锁，所述接口卡将所述写锁分配标识置为已经分配写锁状态；所述接口卡向所述第一客户端发送消息，表明第一客户端获得所述写锁；

当已经分配所述读锁，所述接口卡向所述第一客户端发送第二消息，所述第二消息表明所述第一客户端处于写锁等待状态；

当所述第一客户端获得所述写锁，所述接口卡接收所述第一客户端发送的写锁释放请求或者当所述第一客户端获得所述写锁后发生故障，所述接口卡将所述写锁分配标识置为未分配写锁状态；

所述接口卡将所述第一客户端的标识从所述写锁请求队列队首的内存地址中删除；其中，所述写锁释放请求携带所述写锁请求队列队首的内存地址。

2.根据权利要求1所述的方法，其特征在于，所述方法还包括：

所述接口卡接收所述多个客户端中的第二客户端发送的第二写锁请求；所述第二写锁请求携带所述第二客户端的标识和所述写锁请求队列的第一内存地址；所述第一内存地址为从所述队首到队尾的顺序中，除所述写锁请求队列队首内存地址外可写入所述第二客户端的标识的内存地址；

所述接口卡将所述第二客户端的标识写入写锁请求队列的所述第一内存地址；

所述接口卡向所述第二客户端发送消息表明所述第二客户端处于写锁等待状态。

3.根据权利要求1所述的方法，其特征在于，所述锁服务器的内存还包括读锁请求队列，当所述读锁分配计数表明已经分配所述读锁，所述读锁请求队列用于记录所述集群中针对所述共享资源获得读锁的第三客户端的标识。

4.根据权利要求1所述的方法，其特征在于，所述第一写锁请求还携带锁租约；所述锁租约用于表明写锁的有效期。

5.根据权利要求2所述的方法，其特征在于，还包括：

所述接口卡查询所述写锁请求队列确定处于写锁等待状态的所述第二客户端；

所述接口卡向所述第二客户端发送通知，通知所述第二客户端向所述接口卡申请写锁。

6.根据权利要求1所述的方法，其特征在于，还包括：

所述接口卡向所述集群中除所述第一客户端外的客户端发送通知，所述通知表明已经释放了所述第一客户端持有的写锁。

7.一种集群中锁管理方法，所述集群包含多个客户端和锁服务器，所述锁服务器包含接口卡和内存，其特征在于，所述内存中存储有读锁请求队列和写锁分配标识；其中，所述

读锁请求队列用于记录所述集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，所述写锁分配标识用于表示是否分配针对所述共享资源的写锁；所述方法包括：

所述接口卡接收所述多个客户端中的第一客户端发送的第一读锁请求；所述第一读锁请求携带所述第一客户端的标识和所述读锁请求队列的第一内存地址；

所述接口卡将所述第一客户端的标识写入所述读锁请求队列的所述第一内存地址；

所述接口卡向所述第一客户端发送第一消息，所述第一消息携带所述写锁分配标识；

当所述写锁分配标识表示未分配所述写锁，所述第一客户端获得所述读锁时，所述接口卡根据为所述第一客户端分配的读锁数量增加所述内存中的读锁分配计数；

所述接口卡接收所述第一客户端发送的读锁释放请求，或者，当所述第一客户端获得所述读锁后发生故障，所述接口卡根据为所述第一客户端分配的所述读锁数量减少所述内存中的所述读锁分配计数；

所述接口卡将所述读锁请求队列中所述第一内存地址中的第一客户端的标识删除；其中，所述读锁请释放求携带所述读锁请求队列中存储所述第一客户端的标识的所述第一内存地址。

8.根据权利要求7所述的方法，其特征在于，当所述写锁分配标识表明已经分配了写锁，则所述第一消息表明所述第一客户端处于读锁等待状态。

9.根据权利要求7所述的方法，其特征在于，所述多个客户端中还包括第二客户端，所述内存中还存储写锁请求队列，所述写锁请求队列用于记录所述集群中针对所述共享资源处于写锁等待状态和获得写锁的客户端的标识；当所述写锁分配标识表明已经分配了写锁，所述写锁请求队列队首的内存地址中存储所述第二客户端的标识。

10.根据权利要求7所述的方法，其特征在于，所述多个客户端中还包括第三客户端，所述读锁请求队列队首的内存地址存储所述第三客户端的标识；

则所述第一读锁请求携带所述读锁请求队列的所述第一内存地址为从所述读锁请求队列队首到队尾的顺序中，可写入所述第一客户端的标识的内存地址。

11.根据权利要求7所述的方法，其特征在于，所述第一读锁请求还携带锁租约；所述锁租约用于表明读锁的有效期。

12.一种锁服务器，所述锁服务器应用于集群中，所述集群还包括多个客户端，其特征在于，所述锁服务器包括接口卡和内存，所述内存用于存储有写锁请求队列、写锁分配标识和读锁分配计数；其中，所述写锁请求队列用于记录所述集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，所述写锁分配标识用于表示是否分配针对所述共享资源的写锁，所述读锁分配计数用于表示针对所述共享资源分配读锁的数量；

所述接口卡用于接收所述多个客户端中的第一客户端发送的第一写锁请求；所述第一写锁请求携带第一客户端的标识和写锁请求队列队首的内存地址；

所述接口卡还用于将所述第一客户端的标识写入所述写锁请求队列队首的内存地址，查询所述读锁分配计数确定是否已经分配读锁，当没有分配所述读锁，将所述写锁分配标识置为已经分配写锁状态，向所述第一客户端发送消息，表明第一客户端获得所述写锁；

当已经分配所述读锁，向所述第一客户端发送第二消息，所述第二消息表明所述第一客户端处于写锁等待状态；

所述接口卡还用于当所述第一客户端获得所述写锁后，接收所述第一客户端发送的写锁释放请求，或者当所述第一客户端获得所述写锁后发生故障，将所述写锁分配标识置为未分配写锁状态，将所述第一客户端的标识从所述写锁请求队列队首的内存地址中删除；其中，所述写锁释放请求携带所述写锁请求队列队首的内存地址。

13.根据权利要求12所述的锁服务器，其特征在于，所述接口卡还用于接收所述多个客户端中的第二客户端发送的第二写锁请求，所述第二写锁请求携带所述第二客户端的标识和所述写锁请求队列的第一内存地址；所述第一内存地址为从所述队首到队尾的顺序中，除所述写锁请求队列队首内存地址外可写入所述第二客户端的标识的内存地址；

所述接口卡将所述第二客户端的标识写入写锁请求队列的所述第一内存地址；

所述接口卡向所述第二客户端发送消息表明所述第二客户端处于写锁等待状态。

14.根据权利要求12所述的锁服务器，其特征在于，所述内存还用于存储读锁请求队列，当所述读锁分配计数表明已经分配所述读锁，所述读锁请求队列用于记录所述集群中针对所述共享资源获得读锁的第三客户端的标识。

15.一种锁服务器，所述锁服务器应用于集群中，所述集群还包括多个客户端，其特征在于，所述锁服务器包括接口卡和内存，所述内存用于存储有读锁请求队列和写锁分配标识；其中，所述读锁请求队列用于记录所述集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，所述写锁分配标识用于表示是否分配针对所述共享资源的写锁；

所述接口卡用于接收所述多个客户端中的第一客户端发送的第一读锁请求，所述第一读锁请求携带第一客户端的标识和读锁请求队列的第一内存地址，将所述第一客户端的标识写入所述读锁请求队列的所述第一内存地址，向所述第一客户端发送第一消息，所述第一消息携带所述写锁分配标识；

所述接口卡还用于，当所述写锁分配标识表示未分配所述写锁，所述第一客户端获得所述读锁时，根据为所述第一客户端分配的读锁数量增加所述内存中的读锁分配计数，接收所述第一客户端发送的读锁释放请求，或者当所述第一客户端获得所述读锁后发生故障，根据为所述第一客户端分配的所述读锁数量减少所述内存中的所述读锁分配计数，将所述读锁请求队列中所述第一内存地址中的第一客户端的标识删除；其中，所述读锁请释放求携带所述读锁请求队列中存储所述第一客户端的标识的所述第一内存地址，。

16.根据权利要求15所述的锁服务器，其特征在于，所述多个客户端中还包括第二客户端，所述内存还用于存储写锁请求队列，所述写锁请求队列用于记录所述集群中针对所述共享资源处于写锁等待状态和获得写锁的客户端的标识；当所述写锁分配标识表明已经分配了写锁，所述写锁请求队列队首的内存地址中存储所述第二客户端的标识。

17.根据权利要求15所述的锁服务器，其特征在于，所述多个客户端中还包括第三客户端，所述读锁请求队列队首的内存地址存储所述第三客户端的标识；

则所述第一读锁请求携带所述读锁请求队列的所述第一内存地址为从所述读锁请求队列队首到队尾的顺序中，可写入所述第一客户端的标识的内存地址。

一种集群中锁管理的方法、锁服务器及客户端

技术领域

[0001] 本发明涉及信息技术领域，特别涉及一种集群中锁管理的方法、服务器及客户端。

背景技术

[0002] 在集群中包含多个客户端，一个客户端为获得集群中的共享资源的处理权限，如写权限，或者读取权限，需要向集群中的锁服务器申请处理权限，通常称为申请锁。通常写权限也称为写锁，或排他锁，读取权限称为读锁或共享锁。

[0003] 现有技术客户端申请写锁的方式，如图1所示，集群中包括锁服务器、客户端A和客户端B。客户端A使用远程直接内存访问(Remote Direct Memory Access,RDMA)协议向锁服务器的内存地址写入客户端A的标识，如使用RDMA协议中的原子操作Compare-and-Swap向锁服务器申请写锁，即客户端A向锁服务器申请到写锁，即将写锁分配给客户端A。客户端B使用Compare-and-Swap向锁服务器申请写锁，锁服务器向客户端B返回客户端A的标识，客户端B根据客户端A的标识使用Compare-and-Swap向客户端A申请写锁，在客户端A释放写锁前，客户端B处于写锁等待状态。

[0004] 现有技术的实现方案中，客户端A故障，由于锁服务器不知道客户端B处于写锁等待状态，导致客户端B无法申请到写锁，始终处于写锁等待状态。由上述描述可知，现有技术中锁服务器只能获得当前一个客户端的锁的状态。

发明内容

[0005] 第一方面，本发明实施例提供了一种集群中锁管理方法，集群中包含多个客户端和锁服务器，锁服务器包含接口卡和内存，内存中存储有写锁请求队列、写锁分配标识和读锁分配计数；其中，写锁请求队列用于记录集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁，读锁分配计数用于表示针对共享资源分配读锁的数量；锁管理方法包括：接口卡接收多个客户端中的第一客户端发送的第一写锁请求，第一写锁请求携带第一客户端的标识和写锁请求队列队首的内存地址，将第一客户端的标识写入写锁请求队列队首的内存地址；接口卡查询读锁分配计数确定是否已经分配读锁，当没有分配读锁，将写锁分配标识置为已经分配写锁状态，向第一客户端发送消息，表明第一客户端获得写锁；

[0006] 当已经分配读锁，接口卡向第一客户端发送第二消息，第二消息表明第一客户端处于写锁等待状态。本发明实施例通过内存中的写锁请求队列、写锁分配标识和读锁分配计数，锁服务器可以确定客户端获得写锁和处于写锁等待。

[0007] 结合第一方面，在第一种可能的实现方式中，集群还包括第二客户端，接口卡接收多个客户端中的第二客户端发送的第二写锁请求；第二写锁请求携带第二客户端的标识和写锁请求队列的第一内存地址；第一内存地址为从队首到队尾的顺序中，除写锁请求队列队首内存地址外可写入第二客户端的标识的内存地址；接口卡将第二客户端的标识写入写锁请求队列的第一内存地址；接口卡向第二客户端发送消息表明第二客户端处于写等待状

态。因此，锁服务器可通过写锁请求队列确定处于写锁等待状态的客户端。可选的，第一内存地址为从队首到队尾的顺序中，除写锁请求队列队首内存地址外第一个可写入第二客户端的标识的内存地址

[0008] 结合第一方面，在第二种可能的实现方式中，锁服务器的内存包括读锁请求队列，当读锁分配计数表明已经分配读锁，读锁请求队列用于记录集群中针对共享资源获得读锁的第三客户端的标识。

[0009] 结合第一种可能的实现方式，在第三种可能的实现方式中，第一写锁请求还携带锁租约；锁租约用于表明写锁的有效期。通过锁租约，锁服务器可根据锁租约释放写锁。

[0010] 结合第一种可能的实现方式，在第四种可能的实现方式中，当第一客户端获得写锁后，接口卡接收第一客户端发送的写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；接口卡将写锁分配标识置为未分配写锁状态，将第一客户端的标识从写锁请求队列队首的内存地址中删除。第一客户端释放写锁，从而使集群中其他客户端可以获得写锁。

[0011] 结合第一种可能的实现方式，在第五种可能的实现方式中，当第一客户端获得写锁后发生故障，接口卡将写锁分配标识置为未分配写锁状态，将第一客户端的标识从写锁请求队列队首的内存地址中删除。可选地，接口卡接收集群中其他客户端发送的通知确定第一客户端故障。将故障的第一客户端的写锁释放，可以防止故障的第一客户端长期持有写锁。可选的，接口卡向集群中除第一客户端外的客户端发送通知，通知表明已经释放了第一客户端持有的写锁。

[0012] 结合第四种或第五种可能的实现方式，在第六种可能的实现方式中，还包括接口卡查询写锁请求队列确定处于写锁等待状态的第二客户端；接口卡向第二客户端发送通知，通知第二客户端向接口卡申请写锁。锁服务器通知处于写锁等待状态的客户端及时申请写锁。

[0013] 结合第五种可能的实现方式，在第七种可能的实现方式中，接口卡向集群中除第一客户端外的客户端发送通知，通知表明已经释放了第一客户端持有的写锁。

[0014] 结合第四种可能的实现方式，在第八种可能的实现方式中，还包括，接口卡接收第一客户端发送的写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首开始的下一个内存地址，接口卡查询写锁请求队列队首开始的下一个内存地址存储的客户端的标识，接口卡向第一客户端发送查询写锁请求查询响应消息，写锁请求查询响应消息中携带下一个内存地址存储的客户端的标识。

[0015] 结合第四种可能的实现方式，在第九种可能的实现方式中，还包括，接口卡查询写锁请求队列队首开始的下一个内存地址存储的客户端的标识，接口卡向下一个内存地址存储的客户端的标识对应的客户端发送通知，通知下一个内存地址存储的客户端的标识对应的客户端申请写锁。从而可根据处于写锁等待状态的客户端在写锁请求队列中从队首开始的顺序申请写锁。

[0016] 结合第四种可能的实现方式，在第十种可能的实现方式中，写锁请求队列中还存储处于写锁等待状态的客户端的写锁优先级，还包括，接口卡接收第一客户端发送的写锁请求查询消息，接口卡查询写锁请求队列中处于写锁等待状态的客户端的写锁优先级，接口卡向第一客户端发送写锁请求查询响应消息，查询响应消息中携带写锁请求队列中记录

的处于写锁等待状态的客户端的标识及写锁优先级。

[0017] 结合第四种可能的实现方式，在第十一种可能的实现方式中，写锁请求队列中还存储处于写锁等待状态的客户端的写锁优先级，还包括：接口卡查询写锁请求队列中记录的处于写锁等待状态的客户端的写锁优先级确定写锁优先级最高的客户端，接口卡向写锁优先级最高的客户端发送通知，通知写锁优先级最高的客户端申请写锁，从而使写锁优先级最高的客户端申请写锁。

[0018] 结合第四种可能的实现方式，在第十二种可能的实现方式中，还包括：接口卡查询写锁请求队列确定处于写锁等待状态的客户端，接口卡向处于写锁等待状态的客户端中的至少两个客户端发送通知，指示至少两个客户端向接口卡发送写锁请求。

[0019] 与第一方面相对应，第二方面，本发明实施例提供了一种集群中锁管理方法，集群包含多个客户端和锁服务器，锁服务器包含接口卡和内存，内存中存储有写锁请求队列、写锁分配标识和读锁分配计数；其中，写锁请求队列用于记录集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁，读锁分配计数用于表示针对共享资源分配读锁的数量；锁管理方法包括：

[0020] 多个客户端中的第一客户端向锁服务器的接口卡发送第一写锁请求；第一写锁请求携带第一客户端的标识和写锁请求队列队首的内存地址；

[0021] 第一客户端接收接口卡发送的第一消息，第一消息表明接口卡将第一客户端的标识写入写锁请求队列队首的内存地址以及表明第一客户端获得写锁或处于写锁等待状态；当第一消息表明接口卡将第一客户端的标识写入写锁请求队列队首的内存地址失败，第一客户端向接口卡发送第二写锁请求；第二写锁请求携带第一客户端的标识和写锁请求队列的第一内存地址；第一内存地址为从队首到队尾的顺序中，可写入第一客户端的标识的内存地址。可选的，第一内存地址为从队首到队尾的顺序中，第一个可写入第一客户端的标识的内存地址。

[0022] 结合第二方面第一种可能的实现方式，在第一种可能的实现方式中，第一写锁请求和第二写锁请求还携带锁租约；锁租约用于表明写锁的有效期。锁服务器可以根据锁租约释放第一客户端到期的写锁，防止第一客户端长期占用写锁。

[0023] 结合第二方面，在第二种可能的实现方式中，当第一客户端获得写锁后，第一客户端向接口卡发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址。

[0024] 结合第二种可能的实现方式，在第三种可能的实现方式中，还包括：第一客户端向接口卡发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首开始的下一个内存地址，接收接口卡发送的写锁请求查询响应消息，写锁请求查询响应消息中携带写锁请求队列队首开始的下一个内存地址存储的客户端的标识，第一客户端根据写锁请求查询响应消息携带的客户端的标识向写锁请求队列队首开始的下一个内存地址存储的客户端的标识对应的客户端发送通知指示申请写锁。

[0025] 结合第二种可能的实现方式，在第四种可能的实现方式中，还包括：第一客户端向接口卡发送写锁请求查询消息，接收接口卡发送的写锁请求查询响应消息，查询响应消息中携带写锁请求队列中记录的处于写锁等待状态的客户端的标识及写锁优先级，第一客户端根据写锁请求查询响应消息确定写锁优先级最高的客户端，第一客户端向写锁优先级最高的客户端发送通知，指示写锁优先级最高的客户端向接口卡发送写锁请求。

[0026] 第三方面，本发明实施例提供了另一种集群中锁管理方法，集群包含第一客户端和锁服务器，锁服务器包含接口卡和内存，内存中存储有读锁请求队列和写锁分配标识；其中，读锁请求队列用于记录集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁；锁管理方法包括：接口卡接收多个客户端中的第一客户端发送的第一读锁请求；第一读锁请求携带第一客户端的标识和读锁请求队列的第一内存地址；接口卡将第一客户端的标识写入读锁请求队列的第一内存地址，接口卡向第一客户端发送第一消息；第一消息携带写锁分配标识。通过读锁请求队列能够记录集群中客户端读锁请求状态，客户端处于读锁等待状态或获得读锁。

[0027] 结合第三方面，在第一种可能的实现方式中，当写锁分配标识表明已经分配了写锁，则第一消息表明第一客户端处于读锁等待状态。

[0028] 结合第三方面，在第二种可能的实现方式中，内存还存储有读锁分配计数，读锁分配计数用于表示针对共享资源分配读锁的数量；当写锁分配标识表明没有分配写锁，接口卡根据为第一客户端分配的读锁数量增加读锁分配计数，则第一消息表明第一客户端获得读锁。

[0029] 结合第三方面第一种可能的实现方式，在第三种可能的实现方式中，多个客户端中还包括第二客户端，内存中还存储写锁请求队列，写锁请求队列用于记录集群中针对共享资源处于写锁等待状态和获得写锁的客户端的标识；当写锁分配标识表明已经分配了写锁，写锁请求队列队首的内存地址中存储第二客户端的标识。锁服务器内存中写锁请求队列记录获得写锁的客户端，因为已经分配了写锁，读锁请求队列中的客户端处于读锁等待状态，因此，锁服务器中记录了获得写锁的客户端的信息，同时也记录了处于读锁等待状态的客户端的信息。

[0030] 结合第三方面，在第四种可能的实现方式中，多个客户端中还包括第三客户端，读锁请求队列队首的内存地址存储第三客户端的标识；

[0031] 则第一读锁请求携带读锁请求队列的第一内存地址为从读锁请求队列队首到队尾的顺序中，可写入第一客户端的标识的内存地址。锁服务器内存中的读锁请求队列可记录多个客户端，因此，锁服务器可以确认多个客户端的读锁状态，即是处于读锁请求状态或获得读锁。可选的，第一内存地址为从读锁请求队列队首到队尾的顺序中，第一个可写入第一客户端的标识的内存地址。

[0032] 结合第三方面，在第五种可能的实现方式中，第一读锁请求还携带锁租约；所锁租约用于表明读锁的有效期，在锁租约到期，锁服务器可及时释放读锁。

[0033] 结合第三方面第二种可能的实现方式，在第六种可能的实现方式中，接口卡接收第一客户端发送的读锁释放请求；读锁请释放求携带读锁请求队列中存储第一客户端的标识的第一内存地址；

[0034] 接口卡根据为第一客户端分配的读锁数量减少内存中的读锁分配计数；

[0035] 接口卡将读锁请求队列中第一内存地址中的第一客户端的标识删除。第一客户端释放读锁，从而锁服务器可以分配写锁。可选的，第一客户端根据其他申请写锁的客户端的通知发送读锁释放请求。

[0036] 结合第三方面的第二种可能的实现方式中，在第七种可能的实现方式中，当第一客户端获得读锁后发生故障，

[0037] 接口卡根据为第一客户端分配的读锁数量减少内存中的读锁分配计数；接口卡将读锁请求队列中第一内存地址中的第一客户端的标识删除。当第一客户端获得读锁后发生故障，锁服务器接口卡将第一客户端的读锁释放，防止故障客户端长期持有读锁。可选的，接口卡可以根据其他客户端的通知确定第一客户端故障后，释放第一客户端持有的读锁。

[0038] 第四方面，本发明实施例提供了另一种集群中锁管理方法，集群包含多个客户端和锁服务器，锁服务器包含接口卡和内存，内存中存储有读锁请求队列和写锁分配标识；其中，读锁请求队列用于记录集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁；锁管理方法包括：多个客户端中的第一客户端向接口卡发送第一读锁请求；第一读锁请求携带第一客户端的标识和读锁请求队列的第一内存地址；第一客户端接收接口卡发送的第一消息，第一消息携带写锁分配标识。

[0039] 结合第四方面，在第一种可能的实现方式中，当写锁分配标识表明已经分配了写锁，则第一消息表明第一客户端处于读锁等待状态。

[0040] 结合第四方面，在第二种可能的实现方式中，当所述写锁分配标识表明没有分配写锁，则所述第一消息表明所述第一客户端获得读锁。

[0041] 结合第四方面，在第三种可能的实现方式中，集群还包括第二客户端，读锁请求队列队首的内存地址存储第二客户端的标识；

[0042] 则第一读锁请求携带读锁请求队列的第一内存地址为从读锁请求队列队首到队尾的顺序中，可写入第一客户端的标识的内存地址。可选的，第一内存地址为从读锁请求队列队首到队尾的顺序中，第一个可写入第一客户端的标识的内存地址。

[0043] 结合第四方面，在第四种可能的实现方式中，第一读锁请求还携带锁租约；锁租约用于表明读锁的有效期。

[0044] 结合第四方面第二种可能的实现方式，在第五种可能的实现方式中，当第一客户端获得读锁，第一客户端向接口卡发送读锁释放请求；读锁释放请求携带读锁请求队列中存储第一客户端的标识的内存地址。可选的，第一客户端接收其他申请写锁的客户端发送的通知进行读锁释放操作。

[0045] 第五方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，写锁请求队列还存储第二客户端的标识，写锁分配标识表明已经分配写锁，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端的写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；接口卡根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识；接口卡接收第一客户端发送的写锁请求查询消息，写锁请求查询消息携带写锁请求队列存储第二客户端的标识的内存地址；接口卡向第一客户端发送写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；接口卡接收第一客户端发送的写锁占用请求，写锁占用请求携带第二客户端的标识和写锁请求队列队首的内存地址；接口卡将第二客户端的标识写入写锁请求队列队首的内存地址。本发明实施例中，第一客户端在释放写锁的同时，直接为写锁请求队列中记录的处于写锁等待状态的一个客户端分配写锁。

[0046] 结合第五方面，在第一种可能的实现方式中，写锁请求队列存储第二客户端的标识的内存地址为从写锁请求队列队首开始的下一个内存地址。因此，从写锁请求队列队首开始的下一个内存地址存储的客户端优先获得写锁。

[0047] 结合第五方面，在第二种可能的实现方式中，写锁请求队列还存储第二客户端的写锁优先级，其中，第二客户端的写锁优先级在写锁请求队列中记录的处于写锁等待状态的客户端中最高。因此，写锁请求队列记录的处于写锁等待状态的客户端中写锁优先级最高的客户端优先获得写锁。

[0048] 第六方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，读锁分配计数为0，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于读锁等待状态；锁管理方法包括：接口卡接收第一客户端发送的写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；接口卡将写锁分配标识置为未分配写锁状态；接口卡根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识；接口卡接收第一客户端发送的读锁请求查询消息；接口卡向第一客户端发送读锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；接口卡接收第一客户端发送的读锁分配计数修改请求；接口卡根据读锁分配计数修改请求增加内存中的读锁分配计数。在释放第一客户端写锁的同时，直接为读锁请求队列中记录的处于读锁等待状态的客户端分配读锁。

[0049] 第七方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，写锁请求队列还存储第二客户端的标识，写锁分配标识表明已经分配写锁，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：第一客户端向接口卡发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；当接口卡根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识，第一客户端向接口卡发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列存储第二客户端的标识的内存地址；第一客户端接收接口卡发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；第一客户端向接口卡发送写锁占用请求，写锁占用请求携带第二客户端的标识和写锁请求队列队首的内存地址；第一客户端接收接口卡发送的写锁占用请求响应，表明将第二客户端的标识写入写锁请求队列队首的内存地址；第一客户端向第二客户端发送通知，通知表明第二客户端获得写锁。第一客户端在释放写锁的同时，直接为写锁请求队列中记录的处于写锁等待状态的一个客户端分配写锁。

[0050] 结合第七方面，在第一种可能的实现方式中，写锁请求队列存储第二客户端的标识的内存地址为从写锁请求队列队首开始的下一个内存地址。因此，从写锁请求队列队首开始的下一个内存地址存储的客户端优先获得写锁。

[0051] 结合第七方面，在第二种可能的实现方式中，写锁请求队列还存储第二客户端的写锁优先级，其中，第二客户端的写锁优先级在写锁请求队列中记录的处于写锁等待状态的客户端中最高。因此，写锁请求队列记录的处于写锁等待状态的客户端中写锁优先级最

高的客户端优先获得写锁。

[0052] 第八方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，读锁分配计数为0，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于读锁等待状态；锁管理方法包括：第一客户端向接口卡发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；当接口卡将写锁分配标识置为未分配写锁状态并根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识时，第一客户端向接口卡发送读锁请求查询消息；第一客户端接收接口卡发送的读锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；第一客户端向接口卡发送读锁分配计数修改请求。第一客户端在释放写锁的同时，直接为读锁请求队列中记录的处于读锁等待状态的客户端分配读锁。

[0053] 第九方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，写锁请求队列还存储第二客户端的标识，写锁分配标识表明已经分配写锁，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：第一客户端向接口卡发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；当接口卡将写锁分配标识置为未分配写锁状态并根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识时，第一客户端向接口卡发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列存储第二客户端的标识的内存地址；第一客户端接收接口卡发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；第一客户端根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端向接口卡发送写锁请求。第一客户端释放写锁的同时，通知处于写锁等待状态的客户端发送写锁请求。

[0054] 结合第九方面，在第一种可能的实现方式中，写锁请求队列存储第二客户端的标识的内存地址为从写锁请求队列队首开始的下一个内存地址。从写锁请求队列队首开始的下一个内存地址存储的客户端优先获得第一客户端发送的写锁请求通知。

[0055] 结合第九方面，在第二种可能的实现方式中，写锁请求队列还存储第二客户端的写锁优先级，其中，第二客户端的写锁优先级在写锁请求队列中记录的处于写锁等待状态的客户端中最高。因此，写锁请求队列记录的处于写锁等待状态的客户端中写锁优先级最高的客户端优先获得第一客户端发送的写锁请求通知。

[0056] 第十方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，读锁分配计数为0，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于读锁等待状态；锁管理方法包括：第一客户端向接口卡发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；当接口卡将写锁分配标识置为未分配写锁状态并根据写锁请求

队列队首的内存地址删除写锁请求队列中第一客户端的标识时，第一客户端向接口卡发送读锁请求查询消息；第一客户端接收接口卡发送的读锁请求查询响应消息，读锁请求查询响应消息携带第二客户端的标识，第一客户端根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端申请读锁。第一客户端释放写锁的同时，通知处于读锁等待状态的第二客户端申请读锁，第二客户端向接口卡发送读锁分配计数修改请求，接口卡修改读锁分配计数，第二客户端获得读锁。

[0057] 第十一方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：第一客户端向接口卡发送读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；当接口卡将读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识时，第一客户端向接口卡发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；第一客户端接收接口卡发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；第一客户端根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端申请写锁。第一客户端释放读锁的同时，通知写锁请求队列队首记录的处于写锁等待状态的第二客户端申请写锁，第二客户端根据通知向接口卡发送写锁分配标识修改请求，接口卡将写锁分配标识置为已经分配写锁状态，第二客户端获得写锁。

[0058] 第十二方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：第一客户端向接口卡发送读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；当接口卡将内存中的读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识时，第一客户端向接口卡发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；第一客户端接收接口卡发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；第一客户端向接口卡发送写锁分配标识修改请求；当接口卡根据写锁分配标识修改请求将写锁分配标识置为已经分配写锁状态，第一客户端根据第二客户端的标识向第二客户端发送通知，通知表明第二客户端获得写锁。第一客户端在释放读锁的同时，直接为写锁请求队列队首记录的处于写锁等待状态的第二客户端分配写锁。

[0059] 第十三方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标

识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端发送的读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；接口卡将读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识；接口卡接收第一客户端发送的写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；接口卡向第一客户端发送写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识。在释放第一客户端的读锁的同时，接口卡查询写锁请求队列队首记录的处于写锁等待状态的第二客户端，以优先处理写锁。

[0060] 第十四方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端发送的读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；接口卡将读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识；接口卡接收第一客户端发送的写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；接口卡向第一客户端发送写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识，接口卡接收第一客户端发送的写锁分配标识修改请求；接口卡根据写锁分配标识修改请求将写锁分配标识置为已经分配写锁状态。在释放第一客户端的读锁的同时，为写锁请求队列队首记录的处于写锁等待状态的客户端分配写锁。

[0061] 第十五方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，写锁请求队列队首开始的下一个内存地址还存储第二客户端的标识，写锁分配标识表明已经分配写锁，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端的写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；接口卡根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识；接口卡查询写锁请求队列队首开始的下一个内存地址存储的第二客户端的标识，接口卡将第二客户端的标识写入写锁请求队列队首的内存地址，接口卡向第二客户端发送通知，通知第二客户端获得写锁。本发明实施例中，第一客户端在释放写锁的同时，锁服务器为写锁请求队列中记录的处于写锁等待状态的一个客户端分配写锁。

[0062] 第十六方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储第一客户端的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，读锁分配计数为0，第一客户端获得共享资源的写锁，第二客户端针对该共享资源处于读锁

等待状态；锁管理方法包括：接口卡接收第一客户端发送的写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址；接口卡将写锁分配标识置为未分配写锁状态；接口卡根据写锁请求队列队首的内存地址删除写锁请求队列中第一客户端的标识；接口卡查询读锁请求队列确定读锁请求队列存储的第二客户端的标识，接口卡增加内存中的读锁分配计数，接口卡向第二客户端发送通知表明第二客户端获得读锁。在释放第一客户端写锁的同时，锁服务器直接为读锁请求队列中记录的处于读锁等待状态的客户端分配读锁。

[0063] 第十七方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端发送的读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；接口卡将读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识；接口卡查询写锁请求队列队首的内存地址确定第二客户端的标识；接口卡根据第二客户端的标识向第二客户端发送通知，通知第二客户端向接口卡申请写锁。在释放第一客户端的读锁的同时，接口卡查询写锁请求队列队首记录的处于写锁等待状态的第二客户端，以优先处理写锁。

[0064] 第十八方面，本发明实施例提供了另外一种集群中锁管理的方法，集群包括锁服务器、第一客户端和第二客户端，其中，锁服务器包括接口卡和内存，内存中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储第一客户端的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为第一客户端获得的读锁数量，第一客户端获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；锁管理方法包括：接口卡接收第一客户端发送的读锁释放请求，读锁释放请求携带读锁请求队列存储第一客户端的标识的内存地址；接口卡将读锁分配计数减为0并根据读锁请求队列存储第一客户端的标识的内存地址删除读锁请求队列中第一客户端的标识；接口卡查询写锁请求队列队首的内存地址确定第二客户端的标识；接口卡将写锁分配标识置为已经分配写锁状态，接口卡根据第二客户端的标识向第二客户端发送通知，表明第二客户端获得写锁。在释放第一客户端的读锁的同时，为写锁请求队列队首记录的处于写锁等待状态的客户端分配写锁。

[0065] 根据上述各实施例，本发明实施例还提供了锁服务器，锁服务器包括内存和接口卡，用于实现上述各实施例中锁服务器的功能。

[0066] 根据上述各实施例，本发明实施例还提供了客户端，用于作为上述各实施例中描述的集群中的第一客户端，以实现本发明上述各实施例对应的方案。客户端包括实现本发明上述各实施例中集群中第一客户端的功能的结构单元，或实现本发明上述各实施例中集群中第一客户端的功能。相应地，本发明实施例还提供了非易失性计算机可读存储介质和计算机程序产品，当客户端的内存加载非易失性计算机可读存储介质和计算机程序产品中包含的计算机指令，客户端的中央处理单元(Center Processing Unit，CPU)执行该计算机指令时，使客户端实现本发明上述各实施例中集群中第一客户端的功能。

[0067] 结合上述各实施例，本发明实施例还提供了一种集群，集群中包括锁服务器和客户端，其中锁服务器和客户端可参考上述各实施例的描述。

[0068] 本发明各实施例提供的集群中一种集群中锁管理的方法、锁服务器及客户端，锁服务器与客户端之间处理读锁请求、写锁请求、写锁释放请求、读锁释放请求等操作，直接由锁服务器的接口卡与内存通信，不需要锁服务器的中央处理单元处理上述操作，因此，相对于使用(Transmission Control Protocol/Internet Protocol,TCP/IP)协议通信，减少了客户端与锁服务器之间的交互。

附图说明

[0069] 图1为现有技术集群结构图；

[0070] 图2为本发明实施例集群结构图；

[0071] 图3为本发明实施例锁服务器结构图；

[0072] 图4为本发明实施例锁服务器内存示意图；

[0073] 图5为本发明实施例请求写锁流程图；

[0074] 图6为本发明实施例锁服务器内存示意图；

[0075] 图7为本发明实施例锁服务器内存示意图；

[0076] 图8为本发明实施例请求读锁流程图；

[0077] 图9为本发明实施例锁服务器内存示意图；

[0078] 图10为为本发明实施例锁服务器内存示意图；

[0079] 图11为本发明实施例客户端结构示意图。

具体实施例

[0080] 如图2所示，本发明实施例中的集群包括锁服务器、客户端A和客户端B，客户端A和客户端B分别与锁服务器使用RDMA协议通信，其中锁服务器内存中存储锁请求队列、读锁分配计数和写锁分配标识，具体请见下面描述。

[0081] 如图3所示，锁服务器包括内存301和接口卡302，其中，接口卡302可以为支持以太协议(Ethenet)网络接口卡(Network Interface Card,NIC)，或者基于InfiniBand协议的接口卡或者其他可实现RDMA等可实现远程直接访问内存的协议的接口卡。本发明实施例中的内存301可以是动态随机存取存储器(Dynamic Random Access Memory,DRAM)、双列直接式存储模块(Dual-Inline-Memory-Modules,DIMM)、半导体存储器(Semiconductor Memory,SCM)或者非易失性存储器(Non-volatile memory,NVM)等，也可以是上述的组合，本发明实施例对此不作限定。客户端可以为服务器、个人计算机以及其他终端，客户端通常包括CPU和内存，客户端通常还包括接口卡，用于与锁服务器通信，CPU、内存和接口卡之间通过总线通信。根据图2所示的集群中，如图4所示，本发明实施例中，锁服务器内存301中的锁请求队列为写锁请求队列；另一种实现，锁服务器的内存301中的锁请求队列为读锁请求队列；另一种实现，锁服务器的内存301中的锁请求队列为写锁请求队列和读锁请求队列。

[0082] 本发明实施例中，写锁请求队列和读锁请求队列是内存301中的数据结构，其中，写锁请求队列用于存储处于写锁等待状态或获得写锁的客户端的标识；读锁请求队列用于存储处于读锁等待状态或获得读锁的客户端的标识。本发明实施例中，描述的写锁请求队

列和读锁请求队列存储请求同一共享资源的客户端的标识，请求同一共享资源是指针对同一共享资源申请写锁或读锁。即写锁请求队列用于记录集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，读锁请求队列用于记录集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，写锁分配标识用于标识是否分配针对同一共享资源的写锁，读锁分配计数用于标识针对同一共享资源分配读锁的数量相应的。内存301的读锁分配计数具体实现，可以分配一个内存地址，记录锁服务器分配的读锁的数量或者用多个内存地址，每一个内存地址记录一个读锁请求队列中记录的一个客户端获得读锁的数量。本发明实施例具体实现中，写锁请求队列是内存301中的数据结构读锁请求队列是内存301中的数据结构，具体实现时，可以为数组或其他结构。

[0083] 本发明实施例中，写锁请求队列的内存地址存储客户端的标识与写锁请求队列记录客户端的标识具有相同的含义。同理，读锁请求队列的内存地址存储客户端的标识与读锁请求队列记录客户端的标识具有相同的含义。写锁分配标识用于标识是否分配针对同一共享资源的写锁与写锁分配标识用于表明是否分配写锁具有相同的含义。本发明实施例中集群包含多个客户端，多个客户端中的客户端与多个客户端中包含的客户端具有相同的含义，例如，多个客户端中的第二客户端与多个客户端包含第二客户端。

[0084] 具体实现中，锁服务器通过接口卡301与各客户端通信，将写锁请求队列队首内存地址及写锁请求队列的长度信息发送给各客户端，同理将读锁请求队列队首内存地址及读锁请求队列长度信息发送给各客户端。将写锁分配标识的内存地址信息以及读锁分配计数的内存地址信息发送给各客户端。

[0085] 客户端A和客户端B使用RDMA协议与锁服务器通信。客户端与锁服务器之间使用RADM协议通信，相对于使用TCP/IP协议通信，减少了客户端与锁服务器之间的交互。

[0086] 如图5所示，本发明实施例提供了客户端请求写锁流程。

[0087] 步骤501：接口卡302接收客户端A的写锁请求。

[0088] 客户端A向锁服务器发送写锁请求；写锁请求中携带客户端A的标识以及写锁请求队列队首的内存地址。

[0089] 内存301中的写锁请求队列，可以用于记录所有处于写锁等待状态和获得写锁的客户端的标识，则锁服务器根据写锁请求队列可确定集群中客户端申请写锁的状态。

[0090] 步骤502：接口卡302将客户端A的标识写入写锁请求队列队首的内存地址。

[0091] 锁服务器的接口卡302接收写锁请求，根据写锁请求携带的写锁请求队列队首的内存地址，将客户端A的标识写入写锁请求队列队首，如图6所示。

[0092] 如果成功将客户端A的标识写入写锁请求队列的队首，则表明锁服务器没有向集群中的客户端分配写锁，也没有处于写锁等待的客户端。可选的，接口卡302向客户端A返回消息，表明将客户端A的标识写入写锁请求队列队首的内存地址。

[0093] 可选的，接口卡302将客户端A的标识写入写锁请求队列队首失败，接口卡302向客户端A发送消息，表明将客户端A的标识写入写锁请求队列队首失败，这种情况下，客户端A可再次向接口卡302发送写锁请求，写锁请求携带客户端A的标识和写锁请求队列队首内存地址开始的下一个内存地址，接口卡302将客户端A的标识写入写锁请求队列的下一个内存地址。当客户端A的标识写入写锁请求队列的下一个内存地址失败，则客户端A依次选择写请求队列队首开始的第3个内存地址，重复上述操作，客户端A再次向接口卡302发送写锁请

求队列，写锁请求携带客户端A的标识和写锁请求队列的第3个内存地址，直到成功写入写锁请求队列的内存地址中为止，即写锁请求中携带了可写入客户端A的标识的写锁请求队列的内存地址。客户端A发送写锁请求可以间隔固定周期发送，也可以接收到消息表明写入写锁请求队列失败即发送。

[0094] 可选的，接口卡302将客户端A的标识写入写锁请求队列队首失败，接口卡302则从写锁请求队列队首开始，在写锁请求队列中依次查询可写入的内存地址，将客户端A的标识写入写锁请求队列，接口卡302向客户端A发送消息表明将客户端A的标识写入写锁请求队列的内存地址。可选的，消息中可携带写入写锁请求队列中存储客户端A的标识的内存地址。

[0095] 步骤503：查询是否已经分配读锁。

[0096] 本发明实施例以写锁分配标识为0表示锁服务器没有分配写锁，写锁分配标识为1表示锁服务器分配写锁为例。

[0097] 可选的，如果接口卡302成功将客户端A的标识写入写锁请求队列的队首，接口卡302查询读锁分配计数确定是否已经分配读锁。当锁服务器没有分配读锁时，根据写锁分配标识的内存地址将写锁分配标识置为已经分配写锁状态(本发明实施例中置为1)，接口卡302向客户端A发送消息，表明客户端A获得写锁。示例地，该消息中可携带写锁分配标识和读锁分配计数，其中，写锁分配标识为1，读锁分配计数为0。读锁分配计数为0表明没有分配读锁。

[0098] 可选的，如果接口卡302成功将客户端A的标识写入写锁请求队列的队首，接口卡302向客户端A发送消息表明将将客户端A的标识写入写锁请求队列的队首，客户端A向接口卡302发送写锁分配标识修改请求，写锁分配标识修改请求携带写锁分配标识的内存地址，接口卡302查询读锁分配计数，确定锁服务器是否已经分配读锁，当锁服务器没有分配读锁时，根据写锁分配标识的内存地址将写锁分配标识置为已经分配写锁状态(本发明实施例中置为1)，接口卡302向客户端A发送消息，表明客户端A获得写锁。具体地，该消息中可携带写锁分配标识和读锁分配计数，其中，写锁分配标识为1，读锁分配计数为0。

[0099] 接口卡302查询读锁分配计数确定锁服务器没有分配读锁，则接口卡302将写锁分配标识置为写锁分配状态，即将写锁分配标识置为1，表示客户端A成功获得写锁。即将写锁成功分配给客户端A。接口卡302查询读锁分配计数确定锁服务器已经分配读锁，则接口卡302不将写锁分配标识置为1，即写锁分配标识仍为未分配写锁状态(本发明实施例中，此时写锁分配标识为0)，客户端A处于写锁等待状态，接口卡302向客户端A发送消息表明客户端A处于写锁等待状态。示例地，该消息中可携带写锁分配标识和读锁分配计数，其中，写锁分配标识为0，读锁分配计数为不为0。读锁分配计数不为0表明已经分配读锁。

[0100] 本发明实施例中，接口卡302查询读锁分配计数确定锁服务器没有分配读锁(读锁分配计数为0)，接口卡302将写锁分配标识置为1。

[0101] 在集群中，当锁服务器已经分配了读锁，即客户端A要修改的共享资源此时正在被集群中另一个客户端读取，则不能够为客户端A分配写权限，即不能够分配写锁。因此，在客户端A的标识写入写锁请求队列队首时，需要确定锁服务器是否分配读锁。当接口卡302查询到锁服务器没有分配读锁，则处于写锁请求队列队首的客户端获得写锁，也即客户端A获得写锁。当接口卡302查询到锁服务器已经分配读锁，则客户端A处于写锁等待。

[0102] 当接口卡302根据客户端A发送的写锁请求携带的写锁请求队列队首的内存地址，将客户端A的标识写入写锁请求队列队首失败，则从写锁请求队列队首开始，在写锁请求队列中依次查询可写入的内存地址，将客户端A的标识写入写锁请求队列。因为写锁请求队列队首记录第一个申请写锁请求的客户端，并且一个写锁只能够分配给一个客户端。因此，当写锁请求队列队首的客户端已经获得写锁或处于写锁等待，则将客户端A的标识写入写锁请求队列队首失败。

[0103] 本发明实施例以将客户端A的标识成功写入写锁请求队列队首，并且内存301中的读锁分配计数记录没有分配读锁为例，即客户端A获得写锁。在此基础上，在另一个实施例中，锁服务器接口卡302接收客户端B发送的写锁请求，客户端B发送的写锁请求携带客户端B的标识以及写锁请求队列队首的内存地址。由于写锁请求队列的队首已经写入客户端A的标识，接口卡302将客户端B的标识写入写锁请求队列队首失败。可选的，接口卡302向客户端B发送消息表明将客户端B的标识写入写锁请求队列队首失败。接口卡302将客户端B的标识写入写锁请求队列从队首开始的下一个内存地址，即从写锁请求队列队首开始，客户端B的写锁请求处于写请求队列的第2位，如图7所示。具体将客户端B的标识写入写锁请求队列队首开始的下一个内存地址的方式请参见前面将客户端A的标识写入写锁请求队列中的描述，在此不再赘述。可选的，接口卡301查询写锁请求队列中从写锁请求队列队首到队尾的顺序中可写入客户端的标识的内存地址，向客户端B发送通知，通知客户端B该可写入客户端的标识的内存地址。可选的，该可写入客户端的标识的内存地址可以为从写锁请求队列队首到队尾的顺序中第一个可写入客户端的标识的内存地址。

[0104] 此时内存301中写锁分配标识置为分配写锁状态(本发明实施例中置为1)，即分配给写锁请求队列队首的客户端A，客户端B处于写锁等待状态，接口卡302不再修改内存301中写锁分配标识。可选的，接口卡302向客户端B发送消息表明客户端B处于写锁等待状态。具体地，该消息中可携带写锁请求队列中存储客户端B的标识的内存地址。客户端B根据该内存地址确定处于写锁等待状态。因此，写锁服务器确定此时集群中客户端写锁请求状态，即将写锁分配给客户端A，客户端B处于写锁等待状态。将写锁分配给客户端A也称为客户端A持有写锁，客户端A为写锁的持有者。可选的，客户端向锁服务器发送写锁请求还可携带写锁租约，即客户端持有写锁的有效时间，超出该有效时间，客户端未更新租约，则该写锁自动失效。

[0105] 如图8所示，本发明实施例还提供了客户端请求读锁流程。本发明实施例中，锁服务器还包含读锁请求队列，关于本发明实施列中写锁请求队列的描述请参见前面部分的描述。读锁请求队列用于存储处于读锁等待状态或获得读锁的客户端的标识，因此，锁服务器根据读锁请求队列可确定集群中客户端申请读锁的状态。根据读锁分配计数可确定已经分配的读锁的数量。

[0106] 步骤801：接口卡302接收客户端C发送的读锁请求。

[0107] 本发明实施例中，内存301中还包括读锁请求队列，集群中还包括客户端C。锁服务器接口卡302接收客户端C发送的读锁请求，读锁请求携带客户端C的标识以及读锁请求队列队首的内存地址。

[0108] 步骤802：接口卡302将客户端C的标识写入读锁请求队列。

[0109] 具体的，接口卡302将客户端C的标识写入读锁请求队列队首的内存地址，当读锁

请求队列队首的内存地址没有写入集群中其他客户端的标识时，则将客户端C的标识成功写入读锁请求队列队首的内存地址，如图9所示。可选的，接口卡302向客户端A发送消息表明将客户端C的标识写入读锁请求队列队首的内存地址。

[0110] 当读锁请求队列队首的内存地址写入集群中其他客户端的标识时，则将客户端C的标识写入读锁请求队列队首的内存地址失败，可选的，接口卡302向客户端C发送消息，表明将客户端C的标识写入读锁请求队列队首失败，这种情况下，客户端C可再次向接口卡302发送读锁请求，读锁请求携带客户端C的标识和读锁请求队列队首开始的下一个内存地址，接口卡302将客户端C的标识写入读锁请求队列队首开始的下一个内存地址。当客户端C的标识写入读锁请求队列队首开始的下一个内存地址失败，重复上次操作，客户端C再次向接口卡302发送读锁请求队列，读锁请求携带客户端C的标识和读锁请求队列首开始的的第3个内存地址，直到成功写入读锁请求队列的内存地址中为止，即读锁请求中携带了可写入客户端C的标识的内存地址。客户端C发送读锁请求可以周期发送，也可以接收到消息表明写入读锁请求队列失败即发送。可选的，接口卡301查询读锁请求队列中从读锁请求队列队首到队尾的顺序中可写入客户端的标识的内存地址，向客户端B发送通知，通知客户端B该可写入客户端的标识的内存地址。可选的，该可写入客户端的标识的内存地址是从读锁请求队列队首到队尾的顺序中第一个可写入客户端的标识的内存地址。

[0111] 可选的，接口卡302将客户端C的标识写入读锁请求队列队首失败，接口卡302则从读锁请求队列队首开始，在读锁请求队列中依次查询可写入的内存地址，将客户端C的标识写入读锁请求队列，接口卡302向客户端C发送消息表明将客户端C的标识写入读锁请求队列的内存地址。可选的，消息中可携带写入读锁请求队列中存储客户端C的标识的内存地址。

[0112] 由于集群中共享资源可以被多个客户端同时读取，因此，读锁可以由多个客户端共享，即可由多个客户端同时持有读锁，则多个客户端同时为读锁的持有者。因此读锁分配计数可以使用计数来表示锁服务器是否分配读锁，如当读锁分配计数为0，则表示锁服务器没有分配读锁。锁服务器每分配一个读锁，则读锁分配计数加1。因此，在锁服务器没有分配写锁的情况下(本发明实施例中，写锁分配标识为0)，客户端即可获得读锁。

[0113] 步骤803：增加读锁分配计数。

[0114] 可选的，当接口卡302将客户端C的标识写入读锁请求队列，接口卡302将读锁分配计数加1，接口卡302向客户端C返回写锁分配标识和读锁分配计数。进一步地，当写锁分配标识为1时，客户端C向接口卡302发送读锁分配计数修改请求，接口卡302根据读锁分配计数修改请求将读锁分配计数减1。

[0115] 可选的，当接口卡302将客户端C的标识写入读锁请求队列，当写锁分配标识为1时，接口卡302向客户端C发送消息表明客户端C处于读锁等待状态。具体的，该消息中可以携带写锁分配标识，客户端C根据写锁分配标识确定处于读锁等待状态。

[0116] 通过上面描述可知，接口卡302向客户端C发送消息表明客户端C获得读锁或者处于读锁等待状态，都表明将客户端C的标识写入读锁请求队列的内存地址。

[0117] 可选的，当接口卡302将客户端C的标识写入读锁请求队列，并且写锁分配标识为0时，接口卡302将读锁分配计数加1。进一步地，接口卡302向客户端C发送消息表明客户端C获得读锁，例如消息中携带写锁分配标识，其中写锁分配标识为0。可选的，当锁服务器已经

分配写锁的情况下(写锁分配标识为1),则读锁请求队列中的客户端处于读锁等待状态,不增加读锁分配计数。进一步的,接口卡302向客户端C发送消息,表明客户端C处于读锁等待状态,例如消息中携带写锁分配标识,其中写锁分配标识为1。

[0118] 本发明实施例接口卡302将客户端C的标识成功写入读锁请求队列队首的内存地址为例,则集群中没有客户端获得读锁或处于读锁等待状态,写锁分配标识记录已经分配写锁(写锁分配标识为1),则客户端C处于读等待状态。可选的,客户端C向锁服务器发送读锁请求还可携带读锁租约,即客户端C持有读锁的有效时间,超出该有效时间,客户端C未更新租约,则该读锁自动失效。

[0119] 本发明实施例,集群还包括客户端D,锁服务器的接口卡302接收客户端D发送的读锁请求。客户端D发送的读锁请求携带客户端D的标识和读锁请求队列队首的内存地址。由于读锁请求队列队首的内存地址已经存储客户端C的标识,接口卡302将客户端D的标识写入读锁请求队列队首开始的下一个内存地址,如图10所示,具体方案请参考客户端C写入读锁请求队列的描述,在此不再赘述。由于客户端A持有写锁(写锁分配标识为1),客户端C处于读锁等待状态,客户端D也处于读锁等待状态。

[0120] 本发明实施例,客户端使用RDMA协议向锁服务器的接口卡302发送锁请求,锁服务器内存301记录集群中客户端的锁状态。由于锁服务器记录了所有锁的状态,同时锁服务器可以确定客户端状态,如正常状态或故障状态,因此,在持有锁的客户端发生故障,可及时释放故障客户端持有的锁。可选的,锁服务器确定客户端状态包括通过锁服务器与客户端的心跳检测确定客户端的状态。可选的,客户端之间通过心跳检测确定客户端状态,正常状态客户端向锁服务器报告故障客户端信息。锁服务器根据故障客户端信息,释放故障客户端持有的锁。可选的,当故障客户端持有写锁,则锁服务器的接口卡302将内存301中写锁分配标识置为0,从写锁请求队列队首删除故障客户端的标识;当故障客户端持有读锁,则锁服务器的接口卡302将内存中301读锁分配计数记数减1,从读锁请求队列中删除故障客户端的标识。可选的,其他正常客户端也可执行上述锁服务器操作,释放故障客户端持有的锁,例如其他正常客户端向接口卡302发送锁释放请求,锁释放请求中携带故障客户端的标识,接口卡302执行上述锁释放操作。可选的,锁服务器的接口卡302释放故障客户端持有的写锁后,接口卡302向其他客户端发送通知,通知表明已经释放了故障客户端持有的写锁。同理,锁服务器的接口卡302释放故障客户端持有的读锁后,接口卡302向其他客户端发送通知,通知表明已经释放了故障客户端持有的读锁。

[0121] 本发明实施例中客户端A持有写锁,还包括进行释放写锁操作。客户端A向接口卡302发送写锁释放请求,写锁释放请求携带写锁请求队列队首的内存地址。可选的,接口卡302从写锁请求队列队首的内存地址将客户端A的标识删除,接口卡302将写锁分配标识置为未分配写锁状态(即由1置为0)。可选的,接口卡302从写锁请求队列队首的内存地址将客户端A的标识删除,接口卡302通知客户端A将客户端A的标识删除,客户端A向接口卡302发送写锁分配标识修改请求,接口卡302将写锁分配标识置为未分配写锁状态(即由1置为0)。

[0122] 可选的,当客户端A释放写锁,还包括:客户端A向接口卡302发送写锁请求查询消息,写锁请求查询消息携带写锁请求队列队首开始的下一个内存地址,接口卡302查询写锁请求队列队首开始的下一个内存地址存储客户端B的标识,接口卡302向客户端A发送写锁请求查询响应消息,写锁请求查询响应消息中携带客户端B的标识,客户端A根据写锁请求

查询响应消息携带的客户端B的标识向客户端B发送通知，通知客户端B申请写锁，客户端B根据客户端A的通知向接口卡302发送写锁请求，具体请参见前面客户端发送写锁请求的描述，在此不再赘述。因此，可根据处于写锁等待状态的客户端在写锁请求队列中从队首开始的顺序申请写锁。可选的，客户端A释放写锁后，锁服务器的接口卡302查询写锁请求队列队首开始的下一个内存地址存储客户端B的标识，接口卡302向客户端B发送通知，通知客户端B申请写锁，客户端B根据接口卡302发送的通知向接口卡302发送写锁请求消息。

[0123] 可选的，上述实施例中，客户端向接口卡302发送写锁请求中还携带客户端的写锁优先级，写锁请求队列中存储客户端的标识和写锁优先级，当客户端A释放写锁，还包括：客户端A向接口卡302发送写锁请求查询消息，接口卡302向客户端A发送写锁请求查询响应消息，查询响应消息中携带写锁请求队列中记录的处于写锁等待状态的客户端的标识及写锁优先级，本发明实施例中以客户端B的写锁优先级最高为例，客户端A根据写锁请求查询响应消息确定写锁优先级最高的客户端，客户端A向客户端B发送通知，指示客户端B向接口卡302发送写锁请求，具体请参见前面客户端发送写锁请求的描述，在此不再赘述。因此，可根据处于写锁等待状态的客户端的写锁请求优先级申请写锁。可选的，上述实施例中，客户端向接口卡302发送写锁请求中还携带客户端的写锁优先级，写锁请求队列中存储客户端的标识和写锁优先级。客户端A释放写锁后，锁服务器的接口卡302查询写锁请求队列中记录的处于写锁等待状态的客户端的写锁优先级确定写锁优先级最高的客户端，例如客户端B，接口卡302向客户端B发送通知，通知客户端B申请写锁，客户端B根据接口卡302发送的通知向接口卡302发送写锁请求消息。

[0124] 可选的，当写锁请求队列中记录的处于写锁等待状态的客户端为多个时，当客户端A释放写锁，还包括：客户端A向接口卡302发送写锁请求查询消息，接口卡302查询写锁请求队列确定处于写锁等待状态的客户端，接口卡302向客户端A发送写锁请求查询响应消息，写锁请求查询响应消息中携带处于写锁等待状态的客户端的标识，客户端A根据写锁请求查询响应消息向处于写锁等待状态的客户端中的至少两个发送通知，指示客户端向接口卡302发送写锁请求。因此，写锁请求队列中记录的处于写锁等待状态的客户端都有相同的机会获得写锁。可选的，当客户端A释放写锁，还包括：接口卡302查询写锁请求队列确定处于写锁等待状态的客户端，接口卡302向处于写锁等待状态的客户端中的至少两个发送通知，指示至少两个客户端向接口卡302发送写锁请求。

[0125] 可选的，当读锁请求队列中存储客户端C和D的标识，即客户端C和D处于读锁等待状态，当客户端A释放写锁，还包括：客户端A向接口卡302发送读锁请求查询消息，接口卡302查询读锁请求队列确定处于读锁等待状态的客户端C和D，接口卡302向客户端A发送读锁请求查询响应消息，读锁请求查询响应消息中携带处于读锁等待状态的客户端C和D的标识。可选的，客户端A根据读锁请求查询响应消息向处于读锁等待状态的客户端C和D发送通知，指示客户端C和D申请读锁，则客户端C和D分别向接口卡302发送读锁分配计数修改请求，接口卡302根据客户端C和D的读锁分配计数修改请求增加读锁分配计数。可选的，户端A根据读锁请求查询响应消息向处于读锁等待状态的客户端C发送通知以申请读锁，则客户端C向接口卡302发送读锁分配计数修改请求，接口卡302根据客户端C的读锁分配计数修改请求增加读锁分配计数。因此，读锁请求队列中记录的处于读锁等待状态的客户端都可以获得读锁。可选的，当客户端A释放写锁，还包括：接口卡302查询读锁请求队列确定处于读

锁等待状态的客户端C和D，接口卡302向处于读锁等待状态的客户端C和D发送通知，指示客户端C和D申请读锁，则客户端C和D分别向接口卡302发送读锁分配计数修改请求，接口卡302根据客户端C和D的读锁分配计数修改请求增加读锁分配计数。

[0126] 可选的，当客户端A释放写锁，具体包括：客户端A向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，接口卡302从写锁请求队列队首的内存地址删除客户端A的标识。客户端A向接口卡302发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首开始的下一个内存地址，接口卡302查询从写锁请求队列队首开始的下一个内存地址存储客户端B的标识，接口卡302向客户端A发送写锁请求查询响应消息，写锁请求查询响应消息中携带客户端B的标识，客户端A根据写锁请求查询响应消息向接口卡302发送写锁占用请求，写锁占用请求携带写锁请求队列队首的内存地址和客户端B的标识，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端A发送写锁占用请求响应消息表明将客户端B的标识写入写锁请求队列队首的内存地址，客户端A根据写锁占用请求响应消息，向客户端B发送通知，表明客户端B获得写锁。因此，可保证从写锁请求队列队首开始的下一个内存地址记录的处于写锁等待的客户端能够获得写锁。可选的，当客户端A释放写锁，还包括：接口卡302查询从写锁请求队列队首开始的下一个内存地址存储客户端B的标识，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端B发送通知，表明客户端B获得写锁。

[0127] 可选的，客户端向接口卡302发送写锁请求中还携带客户端的写锁优先级，写锁请求队列中存储客户端的标识和写锁优先级，当客户端A释放写锁，具体包括：客户端A向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，接口卡302从写锁请求队列队首的内存地址删除客户端A的标识。客户端A向接口卡302发送写锁请求查询消息，接口卡302向客户端A发送写锁请求查询响应消息，写锁请求响应消息携带写锁请求队列中记录的处于写锁等待状态的客户端的标识及写锁优先级，本发明实施例中以客户端B的写锁优先级最高为例，客户端A根据写锁请求查询响应消息向接口卡302发送写锁占用请求，写锁占用请求携带写锁请求队列队首的内存地址和客户端B的标识，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端A发送写锁占用请求响应消息表明将客户端B的标识写入写锁请求队列队首的内存地址，客户端A根据写锁占用请求响应消息，向客户端B发送通知，表明客户端B获得写锁。可选的，客户端向接口卡302发送写锁请求中还携带客户端的写锁优先级，写锁请求队列中存储客户端的标识和写锁优先级，当客户端A释放写锁，接口卡302查询写锁请求队列中记录的处于写锁等待状态的客户端的标识及写锁优先级，本发明实施例中以客户端B的写锁优先级最高为例，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端B发送通知，表明客户端B获得写锁。

[0128] 可选的，写锁请求队列中记录多个处于写锁等待状态的客户端的标识，当客户端A释放写锁，具体包括：客户端A向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，接口卡302从写锁请求队列队首的内存地址删除客户端A的标识。客户端A向接口卡302发送写锁请求查询消息，接口卡302向客户端A发送写锁请求查询响应消息，写锁请求响应消息携带写锁请求队列中记录的处于写锁等待状态的客户端的标识，客户端A根据写锁请求查询响应消息随机确定客户端，例如客户端B，向接口卡302发送写锁占

用请求，写锁占用请求携带写锁请求队列队首的内存地址和客户端B的标识，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端A发送写锁占用请求响应消息表明将客户端B的标识写入写锁请求队列队首的内存地址，客户端A根据写锁占用请求响应消息，向客户端B发送通知，表明客户端B获得写锁。可选的，当客户端A释放写锁，还包括：接口卡302查询写锁请求队列中记录的处于写锁等待状态的客户端的标识，接口卡302从处于写锁等待状态的客户端中随机确定一个客户端，例如客户端B，接口卡302将客户端B的标识写入写锁请求队列队首的内存地址，接口卡302向客户端B发送通知，表明客户端B获得写锁。

[0129] 可选的，当读锁请求队列中存储客户端C和D的标识，当客户端A释放写锁，还包括：客户端A向接口卡302发送读锁请求查询消息，接口卡302查询读锁请求队列确定处于读锁等待状态的客户端C和D，接口卡302向客户端A发送读锁请求查询响应消息，读锁请求查询响应消息中携带处于读锁等待状态的客户端C和D的标识，客户端A根据读锁请求查询响应消息向接口卡302发送读锁分配计数修改请求，读锁分配计数修改请求携带分配读锁的数量，可选的，分配读锁的数量为2，接口卡302根据分配读锁的数量增加读锁分配计数，接口卡302向客户端A返回读锁分配计数修改请求响应，客户端A向客户C和D发送通知。可选的，分配读锁的数量为1，接口卡302根据分配读锁的数量增加读锁分配计数，接口卡302向客户端A返回读锁分配计数修改请求响应，客户端A通知客户端C或D获得读锁。可选的，当读锁请求队列中存储客户端C和D的标识，当客户端A释放写锁，还包括：接口卡302查询读锁请求队列确定处于读锁等待状态的客户端C和D，接口卡302根据客户端C和D分配读锁的数量增加读锁分配计数，接口卡302向客户端C或D发送通知，表明获得读锁。

[0130] 可选的，当获得读锁的客户端C和D释放读锁，客户端C向接口卡302发送读锁释放请求，读锁释放请求中携带客户端C的标识和读锁请求队列中存储客户端C的标识的内存地址，接口卡302删除读锁请求队列中客户端C的标识，减少读锁分配计数。同理，客户端D释放读锁，参考上述操作，读锁分配计数减为0。客户端D还向接口卡302发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址，接口卡302向客户端D返回写锁请求查询响应消息，写锁请求查询响应消息携带写锁请求队列队首的内存地址存储的客户端的标识，客户端D根据该客户端的标识通知该客户端向接口卡302申请写锁。客户端向接口卡302申请写锁包括：客户端向接口卡302发送写锁分配标识修改请求，接口卡302将写锁分配标识置为1。可选的，当获得读锁的客户端C和D释放读锁，接口卡302查询写锁请求队列队首的内存地址存储的客户端的标识，接口卡302根据该客户端的标识通知该客户端向接口卡302申请写锁。

[0131] 可选的，当获得读锁的客户端C和D释放读锁，客户端C向接口卡302发送读锁释放请求，读锁释放请求中携带客户端C的标识和读锁请求队列中存储客户端C的标识的内存地址，接口卡302删除读锁请求队列中客户端C的标识，减少读锁分配计数。同理，客户端D释放读锁，参考上述操作，读锁分配计数减为0。客户端D向接口卡302发送写锁分配标识修改请求，接口卡302将写锁分配标识修改为已经分配写锁(本发明实施例中置为1)，接口卡302向客户端D返回写锁分配标识修改请求响应，写锁分配标识修改请求响应携带写锁请求队列队首内存地址存储的客户端的标识，客户端D根据该客户端的标识通知该客户端获得写锁。可选的，当获得读锁的客户端C和D释放读锁，接口卡302将写锁分配标识修改为已经分配写

锁(本发明实施例中置为1)，接口卡302查询写锁请求队列队首的内存地址存储的客户端的标识，通知该客户端获得写锁。

[0132] 可选的，本发明实施例还提供读锁召回策略。本发明实施例中，集群还包括客户端E。当客户端E向接口卡302发送写锁请求，具体操作参考前述客户端B发送写锁请求的操作。客户端E处于写锁等待状态，客户端C和D持有读锁，客户端E通知客户端C和D释放读锁。客户端C向接口卡302发送释放读锁请求，具体操作请参见前述客户端C释放读锁的操作。客户端D向接口卡302发送释放读锁请求，具体操作请参见前述客户端C释放读锁的操作。释放客户端C和D持有的读锁后，内存301中读锁分配计数为0。可选的，接口卡302向客户端E发送消息，通知客户端E申请写锁。可选的，接口卡302向客户端D发送消息，客户端D的读锁释放成功，客户端D向客户端E发送消息，通知客户端E申请写锁。

[0133] 本发明前面的实施例以一个客户端一次申请一个读锁为例进行描述，在客户端申请读锁时，当客户端申请的读锁数量为多个时，接口卡302根据申请的读锁的数量将增加内存301中的读锁分配计数，在读锁请求队列中记录多个客户端的标识，释放读锁的时候，根据申请的读锁的数量将减少内存301中的读锁分配计数，从读锁请求队列中删除多个客户端的标识。客户端发送写锁请求可以使用RDMA协议中的原子操作Compare-and-Swap，客户端发送读锁请求可以使用RDMA协议中的原子操作Fetch-and-Add。

[0134] 本发明实施例中，客户端向接口卡302发送查询消息，如写锁请求查询消息、读锁请求查询消息，可以在一次查询消息中携带多个待查询的内存地址，如写请求队列的内存地址或读锁请求队列的内存地址，也可以一次携带一个待查询的内存地址，通过多次查询消息完成查询。

[0135] 根据本发明实施例提供的集群中锁管理方案，相应提供了一种客户端1100，如图11所示，客户端1100应用于集群中作为集群中多个客户端中的第一客户端，所集群还包含本发明实施例描述的锁服务器，锁服务器包含接口卡302和内存301，内存301中存储有写锁请求队列、写锁分配标识和读锁分配计数；其中，写锁请求队列用于记录集群中针对同一共享资源处于写锁等待状态和获得写锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁，读锁分配计数用于表示针对共享资源分配读锁的数量；客户端1100包括发送单元1101和接收单元1102。其中，发送单元1101，用于向锁服务器的接口卡302发送第一写锁请求；第一写锁请求携带客户端1100的标识和写锁请求队列队首的内存地址；接收单元1102，用于接收接口卡302发送的第一消息，第一消息表明接口卡302将客户端1100的标识写入写锁请求队列队首的内存地址以及表明客户端1100获得写锁或处于写锁等待状态；发送单元1101，还用于当第一消息表明接口卡302将客户端1100的标识写入写锁请求队列队首的内存地址失败，向接口卡302发送第二写锁请求；第二写锁请求携带客户端1100的标识和写锁请求队列的第一内存地址；第一内存地址为从队首到队尾的顺序中可写入客户端1100的标识的内存地址。可选的，第一内存地址为从队首到队尾的顺序中第一个可写入客户端1100的标识的内存地址。

[0136] 本发明实施例提供的如图11所示的客户端1100，可选的，发送单元1101还用于当客户端1100获得写锁后，向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址。

[0137] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为多个客户端

中的第一客户端，集群还包含锁服务器，锁服务器包含接口卡302和内存301，内存301中存储有读锁请求队列和写锁分配标识；其中，读锁请求队列用于记录集群中针对同一共享资源处于读锁等待状态或获得读锁的客户端的标识，写锁分配标识用于表示是否分配针对共享资源的写锁；发送单元1101，用于向接口卡302发送第一读锁请求；第一读锁请求携带客户端1100的标识和读锁请求队列的第一内存地址；接收单元1102，用于接收接口卡302发送的第一消息，所述第一消息携带写锁分配标识。可选的，当写锁分配标识表明已经分配了写锁，则第一消息表明客户端1100处于读锁等待状态。可选的，当写锁分配标识表明没有分配写锁，则第一消息表明客户端1100获得读锁。可选的，多个客户端中还包括第二客户端，读锁请求队列队首的内存地址存储第二客户端的标识；则第一读锁请求携带所述读锁请求队列的第一内存地址为为从读锁请求队列队首到队尾的顺序中，可写入客户端1100的标识的内存地址。可选的，第一内存地址为为从读锁请求队列队首到队尾的顺序中，第一个可写入客户端1100的标识的内存地址。可选的，发送单元1101向接口卡302发送第一读锁请求之前，还用于向接口卡302发送第二读锁请求；第二读锁请求携带1100客户端的标识和读锁请求队列队首的内存地址；当接口卡302向读锁请求队列队首的内存地址写入客户端1100的标识失败，接收单元1102还用于接收接口卡302发送的第二消息，第二消息表明读锁请求失败。可选的，当客户端1100获得读锁，发送单元1102，还用于向接口卡302发送读锁释放请求；读锁释放请求携带读锁请求队列中存储客户端1100的标识的内存地址。

[0138] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储客户端1100的标识，写锁请求队列还存储第二客户端的标识，写锁分配标识表明已经分配写锁，客户端1100获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；发送单元1101用于向所述接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，当接口卡302根据写锁请求队列队首的内存地址删除写锁请求队列中客户端1100的标识时，向所述接口卡302发送写锁请求查询消息，写锁请求查询消息携带所述写锁请求队列存储第二客户端的标识的内存地址；接收单元1102用于接收接口卡302发送的写锁请求查询响应消息，写锁请求查询响应消息携带所述第二客户端的标识；发送单元1101用于向接口卡302发送写锁占用请求，写锁占用请求携带第二客户端的标识和写锁请求队列队首的内存地址；接收单元1101还用于接收接口卡302发送的写锁占用请求响应，表明将第二客户端的标识写入写锁请求队列队首的内存地址；发送单元1101还用于向所述第二客户端发送通知，通知表明第二客户端获得写锁。可选的，写锁请求队列存储第二客户端的标识的内存地址为从写锁请求队列队首开始的下一个内存地址。可选的，写锁请求队列还存储第二客户端的写锁优先级，其中，第二客户端的写锁优先级在写锁请求队列中记录的处于写锁等待状态的客户端中最高。

[0139] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储客户端1100的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，读锁分配计数为0，客户端1100获得共享资源的写锁，第二客户端针对该共

享资源处于读锁等待状态；发送单元1101用于向接口卡302发送写锁释放请求写锁释放请求携带写锁请求队列队首的内存地址，当接口卡302将写锁分配标识置为未分配写锁状态并根据写锁请求队列队首的内存地址删除写锁请求队列中客户端1100的标识时，向接口卡302发送读锁请求查询消息；接收单元1102用于接收接口卡302发送的读锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；发送单元1101还用于向接口卡302发送读锁分配计数修改请求，从而使第二客户端获得读锁。

[0140] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列和写锁分配标识，写锁请求队列队首的内存地址存储客户端1100的标识，写锁请求队列还存储第二客户端的标识，写锁分配标识表明已经分配写锁，客户端1100获得共享资源的写锁，第二客户端针对该共享资源处于写锁等待状态；发送单元1101用于向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，当接口卡302将写锁分配标识置为未分配写锁状态并根据写锁请求队列队首的内存地址删除写锁请求队列中客户端1100的标识时，向接口卡302发送写锁请求查询消息，所写锁请求查询消息携带写锁请求队列存储第二客户端的标识的内存地址；接收单元1102用于接收接口卡302发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；发送单元1101根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端向接口卡302发送写锁请求。可选的，写锁请求队列存储第二客户端的标识的内存地址为从写锁请求队列队首开始的下一个内存地址。可选的，写锁请求队列还存储第二客户端的写锁优先级，其中，第二客户端的写锁优先级在写锁请求队列中记录的处于写锁等待状态的客户端中最高。

[0141] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，写锁请求队列队首的内存地址存储客户端1100的标识，读锁请求队列存储第二客户端的标识，写锁分配标识表明已经分配写锁，所述读锁分配计数为0，客户端1100获得共享资源的写锁，第二客户端针对该共享资源处于读锁等待状态；发送单元1101用于向接口卡302发送写锁释放请求，写锁释放请求携带写锁请求队列队首的内存地址，当接口卡302将写锁分配标识置为未分配写锁状态并根据写锁请求队列队首的内存地址删除写锁请求队列中客户端1100的标识时，向所述接口卡302发送读锁请求查询消息；接收单元1102用于接收接口卡302发送的读锁请求查询响应消息，读锁请求查询响应消息携带第二客户端的标识；发送单元1101还用于根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端申请读锁，第二客户端根据该通知向接口卡302发送读锁分配计数修改请求，接口卡302根据第二客户端分配的读锁数量增量读锁分配计数。

[0142] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储客户端1100的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为客户端1100获得的读锁数量，客户端1100获得共享资源的读

锁，第二客户端针对该共享资源处于写锁等待状态；发送单元1101用于向接口卡302发送读锁释放请求，读锁释放请求携带读锁请求队列存储客户端1100的标识的内存地址，当接口卡302将读锁分配计数减为0并根据读锁请求队列存储客户端1100的标识的内存地址删除读锁请求队列中客户端1100的标识时，向接口卡302发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；接收单元1102用于接收接口卡302发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；发送单元1101还用于根据第二客户端的标识向第二客户端发送通知，通知用于指示第二客户端申请写锁，第二客户端根据该通知向接口卡302发送写锁分配标识修改请求，接口卡302将写锁分配标识置为已经分配写锁状态。

[0143] 本发明另一实施例中，如图11所示的客户端1100，应用于集群中作为第一客户端，集群还包括锁服务器和第二客户端。其中，锁服务器包括接口卡302和内存301，内存301中存储写锁请求队列、读锁请求队列、读锁分配计数和写锁分配标识，读锁请求队列存储客户端1100的标识，写锁请求队列队首的内存地址存储第二客户端的标识，写锁分配标识表明未分配写锁，读锁分配计数为客户端1100获得的读锁数量，客户端1100获得共享资源的读锁，第二客户端针对该共享资源处于写锁等待状态；发送单元1101用于向接口卡302发送读锁释放请求，读锁释放请求携带读锁请求队列存储客户端1100的标识的内存地址，当接口卡302将读锁分配计数减为0并根据读锁请求队列存储客户端1100的标识的内存地址删除读锁请求队列中客户端1100的标识时，向接口卡302发送写锁请求查询消息，写锁请求查询消息携带写锁请求队列队首的内存地址；接收单元1102接收接口卡302发送的写锁请求查询响应消息，写锁请求查询响应消息携带第二客户端的标识；发送单元1102还用于向接口卡302发送写锁分配标识修改请求，当接口卡302根据写锁分配标识修改请求将写锁分配标识置为已经分配写锁状态时，根据第二客户端的标识向第二客户端发送通知，通知表明第二客户端获得写锁。

[0144] 本发明实施例提供的客户端1100实现的功能可参考前面实施例中客户端A、客户端B、客户端C和客户端D对应的描述。客户端1100的实现的效果请参见前面实施列的描述。

[0145] 如图11所示的客户端1100，一种实现方式为客户端1100安装上有述单元，上述单元可被加载到客户端1100的内存中，由客户端1100的CPU执行内存中的指令，实现本发明对应的实施例中的功能；另一种实现，客户端1100中包含的单元可以由硬件来实现，或者由软件与硬件组合实现。上述单元也称为结构单元。

[0146] 本发明实施例，还提供了非易失性计算机可读存储介质和计算机程序产品，当图11所示的客户端1100的内存加载非易失性计算机可读存储介质和计算机程序产品中包含的计算机指令，CPU执行内存中加载的该计算机指令，以实现本发明各实施中对应的功能。

[0147] 在本发明所提供的几个实施例中，应该理解到，所公开的装置、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。

[0148] 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。

[0149] 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存在，也可以两个或两个以上单元集成在一个单元中。

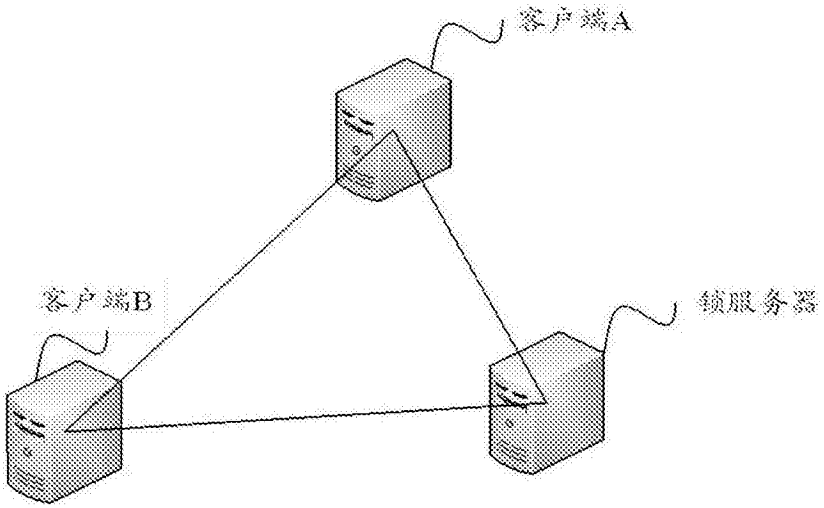

客户端A
客户端B
锁服务器

图1(现有技术)

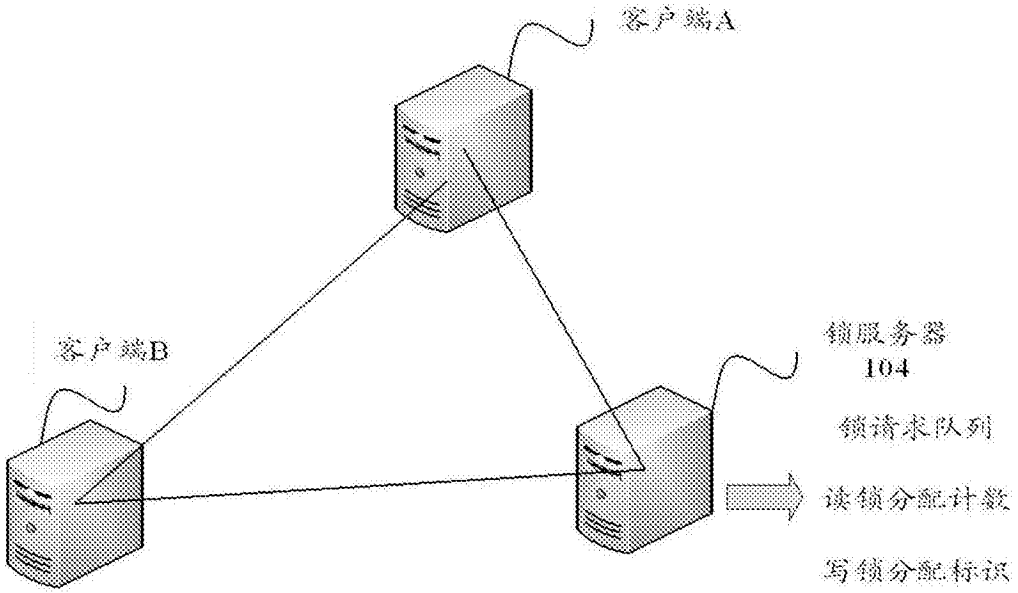

客户端A
客户端B
锁服务器
104
锁请求队列
读锁分配计数
写锁分配标识

图2

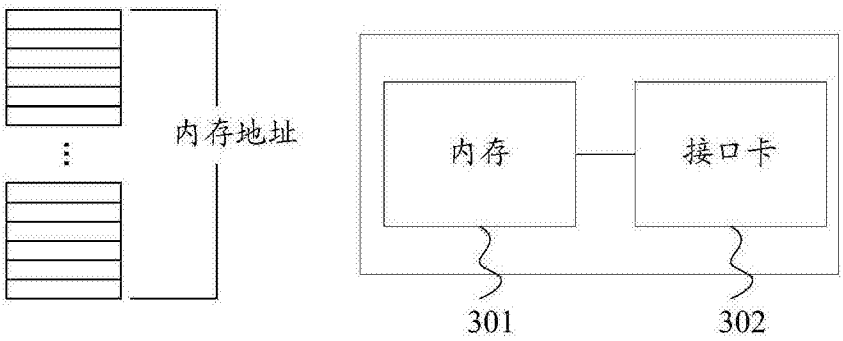

内存地址
内存
接口卡
301
302

图3

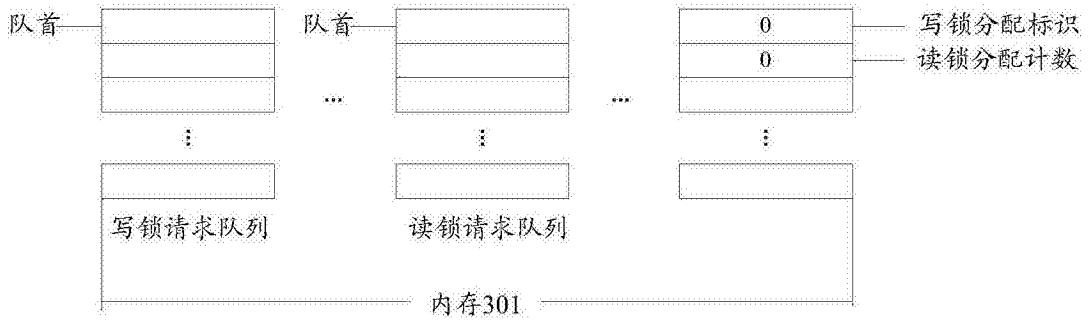

队首
队首
0
0
写锁分配标识
读锁分配计数
写锁请求队列
读锁请求队列
内存301

图4

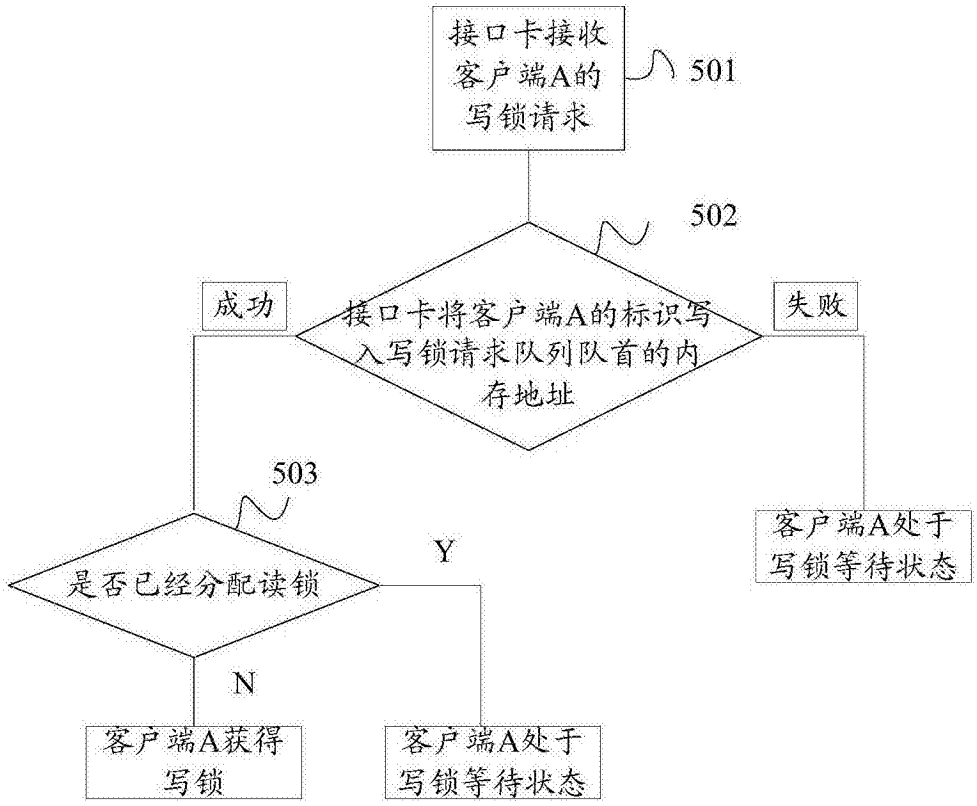

接口卡接收客户端A的写锁请求
501
502
接口卡将客户端A的标识写入写锁请求队列队首的内存地址
成功
失败
503
是否已经分配读锁
Y
N
客户端A处于写锁等待状态
客户端A获得写锁
客户端A处于写锁等待状态

图5

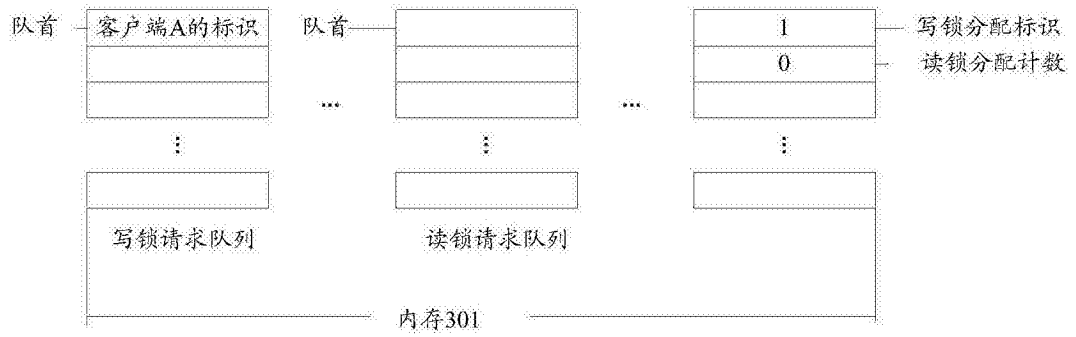

队首
客户端A的标识
队首
1
写锁分配标识
0
读锁分配计数
写锁请求队列
读锁请求队列
内存301

图6

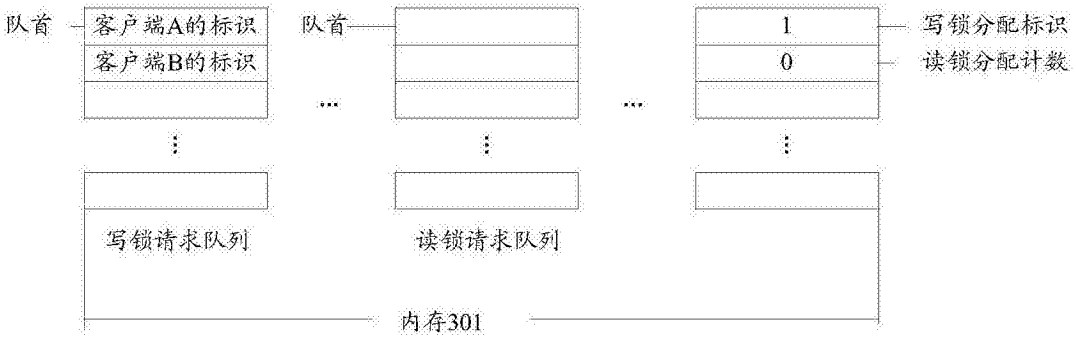

队首
客户端A的标识
客户端B的标识
队首
1
写锁分配标识
0
读锁分配计数
写锁请求队列
读锁请求队列
内存301

图7

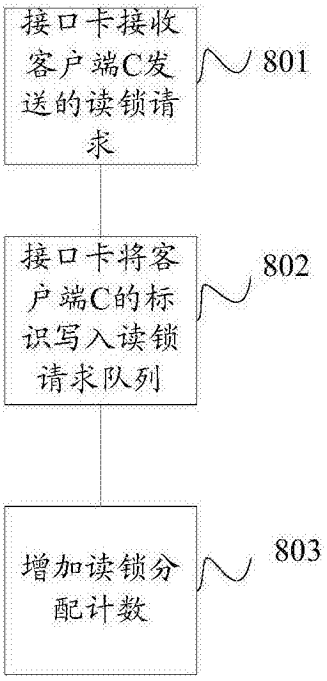

接口卡接收客户端C发送的读锁请求
801
接口卡将客户端C的标识写入读锁请求队列
802
增加读锁分配计数
803

图8

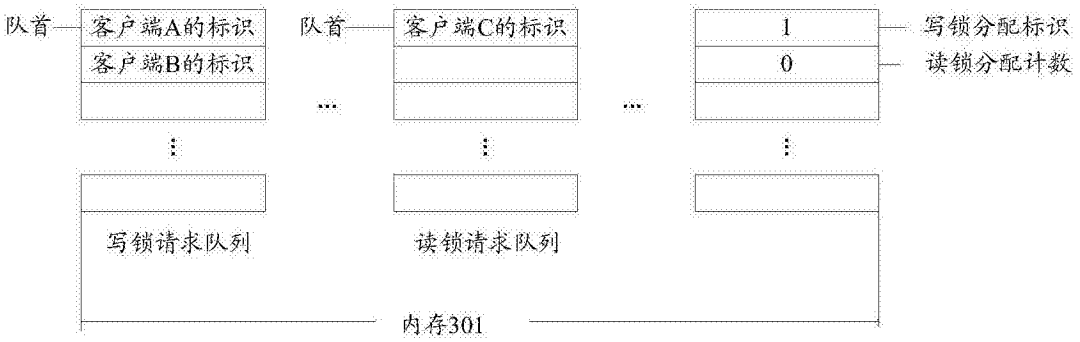

队首
客户端A的标识
客户端B的标识
队首
客户端C的标识
1
0
写锁分配标识
读锁分配计数
写锁请求队列
读锁请求队列
内存301

图9

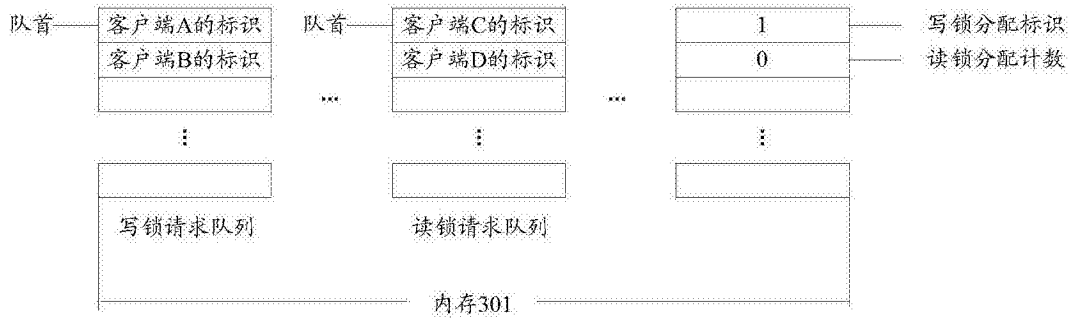

队首
客户端A的标识
客户端B的标识
队首
客户端C的标识
客户端D的标识
1
0
写锁分配标识
读锁分配计数
写锁请求队列
读锁请求队列
内存301

图10

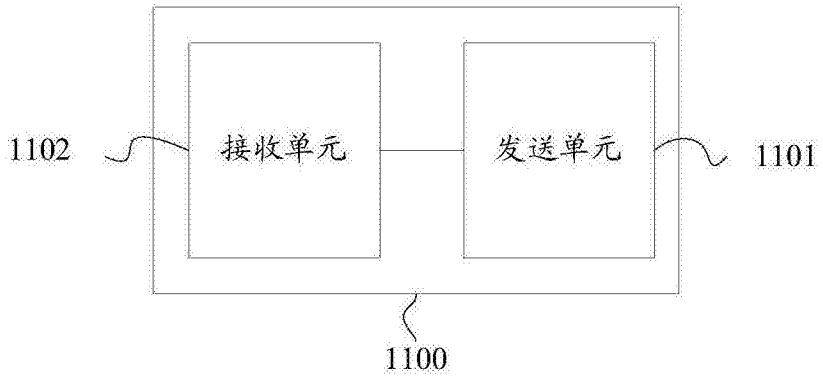

1102
接收单元
发送单元
1101
1100

图11