(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 105224677 B
(45)授权公告日 2018.10.30

(21)申请号 201510673209.9

(22)申请日 2015.10.16

(65)同一申请的已公布的文献号
申请公布号 CN 105224677 A

(43)申请公布日 2016.01.06

(73)专利权人 上海晶赞科技发展有限公司
地址 200072 上海市闸北区灵石路695号珠江创业园区3号楼1101室

(72)发明人 汤奇峰 粟超 李飞

(74)专利代理机构 北京集佳知识产权代理有限公司 11227
代理人 吴敏

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

(56)对比文件
CN 103593477 A,2014.02.19,
CN 1526107 A,2004.09.01,
CN 103577440 A,2014.02.12,
US 9053167 B1,2015.06.09,
US 8949180 B1,2015.02.03,
CN 101499022 A,2009.08.05,
郑燕玲.《空间数据库的分块多级索引机制的研究》.《微计算机信息》.2009,第25卷(第7-3期),

审查员 徐晓

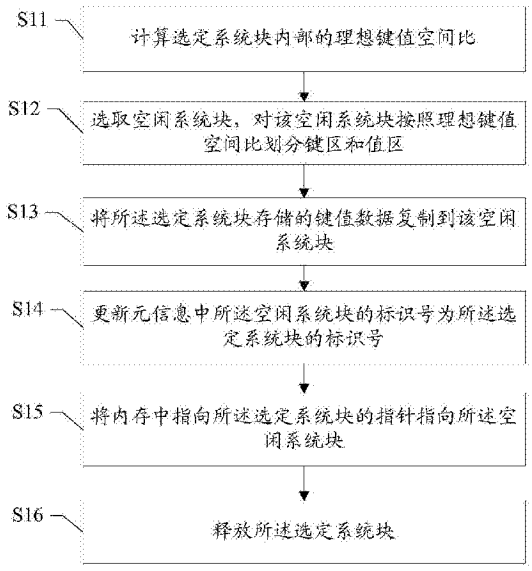
权利要求书2页 说明书7页 附图2页

(54)发明名称

一种数据库操作方法及装置

(57)摘要

一种数据库操作方法及装置,所述数据库为非关系型键值数据库;所述方法包括:计算选定系统块内部的理想键值空间比,所述理想键值空间比为键区与值区实际存储数据占用空间的比值;选取空闲系统块,对该空闲系统块按照理想键值空间比划分键区和值区;将所述选定系统块存储的键值数据复制到该空闲系统块;更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号;将内存中指向所述选定系统块的指针指向所述空闲系统块;释放所述选定系统块。所述方法及装置可以高效利用资源。



CN 105224677 B

1.一种数据库操作方法，其特征在于，所述数据库为非关系型键值数据库；

所述数据库操作方法包括：

计算选定系统块内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值；

选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；

将所述选定系统块存储的键值数据复制到该空闲系统块；

更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；

将内存中指向所述选定系统块的指针指向所述空闲系统块；

释放所述选定系统块。

2.根据权利要求1所述的数据库操作方法，其特征在于，在所述选取空闲系统块之前还包括：确定所述选定系统块内部的键区空间与值区空间的比值不同于所述理想键值空间比。

3.根据权利要求1所述的数据库操作方法，其特征在于，当所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

4.根据权利要求3所述的数据库操作方法，其特征在于，所述数据库存储于固态硬盘；所述将所述选定系统块存储的键值数据复制到该空闲系统块对应的写操作，和所述将所述待存入其他系统块的数据写入对应的键区和值区对应的写操作，均以如下方式完成：将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。

5.根据权利要求1所述的数据库操作方法，其特征在于，在所述计算选定系统块内部的理想键值空间比之前还包括：以固定的时间间隔计算选择系统块作为所述选定系统块。

6.一种数据库操作装置，其特征在于，所述数据库为非关系型键值数据库，所述数据库操作装置包括：理想键值空间比计算单元、键值区域划分单元、键值数据移动单元、标识号更新单元、指针更新单元以及释放单元；其中：所述理想键值空间比计算单元，适于计算选定系统块内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值；

所述键值区域划分单元，适于选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；

所述键值数据移动单元适于将所述选定系统块存储的键值数据复制到该空闲系统块；

所述标识号更新单元，适于更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；

所述指针更新单元，适于将内存中指向所述选定系统块的指针指向所述空闲系统块；

所述释放单元，适于释放所述选定系统块。

7.根据权利要求6所述的数据库操作装置，其特征在于，还包括：业务数据表存储单元，适于在所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

8.根据权利要求7所述的数据库操作装置，其特征在于，所述数据库存储于固态硬盘；所述数据库操作装置还包括：单次写入数据量控制单元，适于将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。

9.根据权利要求6所述的数据库操作装置，其特征在于，还包括：定时选取单元，适于以固定的时间间隔计算选择系统块作为所述选定系统块。

一种数据库操作方法及装置

技术领域

[0001] 本发明涉及数据库领域，尤其涉及一种数据库操作方法及装置。

背景技术

[0002] Nosql(not only sql，不仅仅是sql)是一种非关系型数据库，主要是用来解决半结构化数据和非结构化数据的存储问题。对数据库高并发读写的需求，关系型数据库能够应付每秒上万次的读请求，但是不能承受每秒上万次的写请求，或者是读写的混合请求。

[0003] 对海量数据的高效存储和访问的需求，Nosql数据库可以处理海量的数据，能够运行在大量便宜的普通服务器集群之上。对数据库的高可用和高可扩展的需求，关系型的数据库难以横向扩展，Nosql数据库能够通过增加硬件的数据和服务节点的数量来进行性能和负载能力的横向扩展。

[0004] 在web2.0时代，Nosql产品在互联网行业中的重要性随着互联网及其移动互联网的发展而日剧增大，大型互联网应用中，为了应对大规模、高并发访问，大多都引入了Nosql的产品。

[0005] 但是，现有的对非关系型数据库的操作方法的资源利用效率有待提高。

发明内容

[0006] 本发明解决的技术问题是提供一种高资源利用率的数据库操作方法。

[0007] 为解决上述技术问题，本发明实施例提供一种数据库操作方法，所述数据库为非关系型键值数据库；所述数据库操作方法包括：

[0008] 计算选定系统块内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值；

[0009] 选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；

[0010] 将所述选定系统块存储的键值数据复制到该空闲系统块；

[0011] 更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；

[0012] 将内存中指向所述选定系统块的指针指向所述空闲系统块；

[0013] 释放所述选定系统块。

[0014] 可选的，在所述选取空闲系统块之前还包括：确定所述选定系统块内部的键区空间与值区空间的比值不同于所述理想键值空间比。

[0015] 可选的，当所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

[0016] 可选的，所述数据库存储于固态硬盘；所述将所述选定系统块存储的键值数据复制到该空闲系统块对应的写操作，和所述将所述待存入其他系统块的数据写入对应的键区和值区对应的写操作，均以如下方式完成：将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固

态硬盘的闪存块。

[0017] 可选的，在所述计算选定系统块内部的理想键值空间比之前还包括：以固定的时间间隔计算选择系统块作为所述选定系统块。

[0018] 本发明实施例还提供一种数据库操作装置，所述数据库为非关系型键值数据库，所述数据库操作装置包括：理想键值空间比计算单元、键值区域划分单元、键值数据移动单元、标识号更新单元、指针更新单元以及释放单元；其中：

[0019] 所述理想键值空间比计算单元，适于计算选定系统块内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值；

[0020] 所述键值区域划分单元，适于选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；

[0021] 所述键值数据移动单元适于将所述选定系统块存储的键值数据复制到该空闲系统块；

[0022] 所述标识号更新单元，适于更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；

[0023] 所述指针更新单元，适于将内存中指向所述选定系统块的指针指向所述空闲系统块；

[0024] 所述释放单元，适于释放所述选定系统块。

[0025] 可选的，所述数据库操作装置还包括：业务数据表存储单元，适于在所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

[0026] 可选的，所述数据库存储于固态硬盘；所述数据库操作装置还包括：单次写入数据量控制单元，适于将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。

[0027] 可选的，所述数据库操作装置还包括：定时选取单元，适于以固定的时间间隔计算选择系统块作为所述选定系统块。

[0028] 与现有技术相比，本发明实施例的技术方案具有以下有益效果：

[0029] 通过计算选定系统块内部的理想键值空间比，选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；将所述选定系统块存储的键值数据复制到该空闲系统块；更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；将内存中指向所述选定系统块的指针指向所述空闲系统块；释放所述选定系统块，可以使得系统块中规划的键值空间比与实际键值空间比一致。因为若系统块内部划分的键区和值区的比例不当，也就是划分的键值空间比与实际存储的数据的键值空间比不一致，会出现规划的键区和值区中其中一个区已存储满，但另一个区还有空余的情况，会造成存储空间的浪费。故使得系统块中规划的键值空间比与实际键值空间比一致，可以节省数据库的资源，提高资源利用率。

[0030] 进一步，在所述将所述选定系统块存储的键值数据复制到该空闲系统块时，对所述选定系统块读锁定，其他线程可以在复制过程中正常对所述选定系统块进行读操作，而不必等待复制过程的结束，从而可以减少等待时间，进而可以提升数据库整体的操作效率。

附图说明

[0031] 图1是本发明实施例中一种数据库操作方法的流程图；

[0032] 图2是本发明实施例中一种数据库操作装置的结构示意图；

[0033] 图3是本发明实施例中另一种数据库操作装置的结构示意图。

具体实施方式

[0034] 如前所述，在web2.0时代，Nosql产品在互联网行业中的重要性随着互联网及其移动互联网的发展而日剧增大，大型互联网应用中，为了应对大规模、高并发访问，大多都引入了Nosql的产品。但是，现有的对非关系型数据库的操作方法的资源利用效率有待提高。

[0035] 经发明人研究发现，非关系型键值数据库的存储过程中，若系统块内部划分的键区和值区的比例不当，也就是划分的键值空间比与实际存储的数据的键值空间比不一致，会出现规划的键区和值区中其中一个区已存储满，但另一个区还有空余的情况，会造成存储空间的浪费。

[0036] 本发明实施例通过计算选定系统块内部的理想键值空间比，选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区；将所述选定系统块存储的键值数据复制到该空闲系统块；更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；将内存中指向所述选定系统块的指针指向所述空闲系统块；释放所述选定系统块，可以使得系统块中规划的键值空间比与实际键值空间比一致，从而节省数据库的资源，提高资源利用率。

[0037] 为使本发明的上述目的、特征和有益效果能够更为明显易懂，下面结合附图对本发明的具体实施例做详细的说明。

[0038] 图1是本发明实施例中一种数据库操作方法的流程图。

[0039] S11，计算选定系统块内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值。

[0040] 数据库(DB)在物理空间上是由N个块(bucket)组成，即在系统初始化的时候会对整块disk做分割。系统块(HC bucket)既是逻辑也是物理方面的概念。在物理上它有自己的存储空间，更像一个块，比如32M或者64M的大小。每个系统块有head部分存储meta数据。除去meta的数据区由两个部分组成：键区(Key store area)程序加载的时候，可以直接从键区加载出bucket的index；值区(Value store area)只存储对应的值列表(value list)。

[0041] 可以看到是根据键(key)来做系统块内部的索引，这里直接使用映射地图(hash map)作为索引的查找结构，根据键的64位映射编码(hash code)来定位到索引中。默认系统块内部的索引都下载进入了内存中。

[0042] 可以看出，键值各自区域大小的设定非常重要，若将键值空间比设置成为固定值，将造成极大的资源浪费。

[0043] 在存入数据之前，会将系统块预先划分键区和值区，在存入数据后，实际占用的键区和值区的比值可能会不同于预先划分的键值空间比，所述理想键值空间比为实际存入键区的数据占用空间的比值。

[0044] 可以理解的是，所述选定系统块可以是一个或者多个系统块，可以将一个工作表

占用的系统块作为选定系统块，因为一个工作表往往对应相同的键值空间比。

[0045] 在具体实施中，所述数据库还可以包括系统块索引，所述系统块索引关联键值和系统块；所述选定系统块可以是与同一所述系统块索引关联的两个所述系统块。

[0046] 系统块索引可以包括在一个表(table)内，代表一种资源隔离。在一个table中包含多个表片(table slot)，根据key映射到各个table slot当中，table slot可以挂载系统块，从而通过系统块索引可以实现键值和系统块的关联。

[0047] 和大多关系型数据库类似，有表的元信息可以方便查询和修改。为了保持简洁的设计和性能，实现了模式自由(schema free)，应用层如果自己有需要，可以设计一些简易的格式协议来进行存储，而且性能不会有明显的损失。

[0048] 在table中对数据进行了分片处理，table slot代表每一个分片。每个table slot中会挂载一个或多个系统块，在表创立的时候，会根据初始化的table slot数目来选择加载系统块的数量，默认为每个slot加载一个系统块。Table slot是一个逻辑存储的单元，对于未来的数据迁移和并发操作都是一个必要的设计。

[0049] 通过查找存储于内存中的系统块索引，可以快速的判断数据是否存在及数据存在的位置，从而节省检索时间，提升系统效率。

[0050] 当所述选定系统块是与同一所述系统块索引关联的两个所述系统块时，可以对系统块进行整理，节省系统空间。

[0051] S12，选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区。

[0052] 空闲系统块指当前没有存储数据的系统块，对该空闲系统块按照理想键值空间预先划分键区和值区，便于最大化利用该系统块的资源。

[0053] 在具体实施中，在步骤S12之前，还可以对所述选定系统块内部的键区空间与值区空间的比值也就是键值空间比进行判断，当选定系统块的键值空间比不同于所述理想键值空间比时，执行步骤S12。

[0054] S13，将所述选定系统块存储的键值数据复制到该空闲系统块。

[0055] 由于理想键值比是根据所述选定系统块中存储的键值数据占用空间的比例确定的，该空闲系统块预先划分键区和值区时是按照理想键值比划分的，从而将所述选定系统块存储的键值数据复制到该空闲系统块，使得预先划分的键值比例与实际存储键值数据所占空间的比例匹配，避免存储空间的资源浪费。

[0056] 在具体实施中，在所述将所述选定系统块存储的键值数据复制到该空闲系统块时，对所述选定系统块读锁定，其他线程可以在复制过程中正常对所述选定系统块进行读操作，而不必等待复制过程的结束，从而可以减少等待时间，进而可以提升数据库整体的操作效率。

[0057] S14，更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号。

[0058] 元信息即是head部分存储meta数据，将所述空闲系统块的标识号更新为所述选定系统块的标识号，可以利用所述空闲系统块代替所述选定系统块。

[0059] S15，将内存中指向所述选定系统块的指针指向所述空闲系统块。

[0060] 将内存中指向所述选定系统块的指针指向所述空闲系统块后，所述空闲系统块对存储系统来说，已经替代所述选定系统块。

[0061] 在具体实施中，在所述将内存中指向所述选定系统块的指针指向所述空闲系统块

时，对所述选定系统块读写锁定。仅在指针swap的瞬间进行读写锁定，大大减小了对并发读操作的影响。

[0062] S16，释放所述选定系统块。

[0063] 释放所述选定系统块后，对数据库系统来说，获取了更多的存储资源。

[0064] 在具体实施中，当所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

[0065] 也就是说，硬盘存储再分配中，系统会根据key和value使用的情况来调整之后的空间分配。就是根据历史来预测未来。因为在本法发明实施例数据库系统中key和value是存储在一起的，会存在分配大小和比例的问题，在这个机制下，会自适应的调整物理存储的分配，达到最优化利用硬盘空间，减少硬盘碎片的目的。

[0066] 在具体实施中，所述数据库可以存储于固态硬盘；所述将所述选定系统块存储的键值数据复制到该空闲系统块对应的写操作，和所述将所述待存入其他系统块的数据写入对应的键区和值区对应的写操作，均以如下方式完成：将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。

[0067] 目前主流的固态硬盘(SSD)，使用的存储颗粒是NAND FLASH，主要由如下几个关键的组件组成：

[0068] 协议转换层(可选)，主要是把磁盘读写协议转换为针对NAND FLASH的读写访问请求；

[0069] FLASH芯片控制层，是整个SSD最重要的部分，直接决定了SSD的一切。有一些实现中，需要SDRAM来做缓存支持，并有自己独立的CPU；

[0070] NAND FLASH芯片，存数据的地方。

[0071] 由于固态硬盘不能进行原地更新，因此需要通过擦除+写入的方式来更新数据。而SSD写入性能比读性能至少降一个数量级，因此在极端情况下，每次数据的插入都是擦除+写入的时间总和，延迟比单写增加近一个数量级。写入放大倍数＝闪存中实际写入的数据量/用户请求写入的数据量。

[0072] 将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。根据固态硬盘的物理特点，因为其最小的读写单位是块，按照ssd的块大小，进行地址对齐后完整写入。不会因为小量数据或者一些跨越不同硬盘块的数据而去引起额外的写放大。

[0073] 可以看出，步骤S11至S16可以是系统数据库整理的过程，通过对系统数据库的整理，可以提升数据库系统的资源利用效率。

[0074] 在具体实施中，步骤S11之前还可以判断数据库状态，以确认所述数据库处于非繁忙状态。在判断所述数据库处于非繁忙状态时对数据库进行整理，可以均衡数据库的系统资源，减少整理过程对数据库读写的效率的影响。

[0075] 在具体实施中，还以固定的时间间隔计算选择系统块作为所述选定系统块，也就是可以预设固定的时间点对数据库进行整理。

[0076] 本发明实施例通过计算选定系统块内部的理想键值空间比，选取空闲系统块，对

该空闲系统块按照理想键值空间比划分键区和值区；将所述选定系统块存储的键值数据复制到该空闲系统块；更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号；将内存中指向所述选定系统块的指针指向所述空闲系统块；释放所述选定系统块，可以使得系统块中规划的键值空间比与实际键值空间比一致。因为若系统块内部划分的键区和值区的比例不当，也就是划分的键值空间比与实际存储的数据的键值空间比不一致，会出现规划的键区和值区中其中一个区已存储满，但另一个区还有空余的情况，会造成存储空间的浪费。故使得系统块中规划的键值空间比与实际键值空间比一致，可以节省数据库的资源，提高资源利用率。

[0077] 本发明实施例还提供一种数据库操作装置，结构示意图如图2所示。

[0078] 本发明实施例中的数据库为非关系型键值数据库，所述数据库操作装置包括：理想键值空间比计算单元21、键值区域划分单元22、键值数据移动单元23、标识号更新单元24、指针更新单元25以及释放单元26，其中：

[0079] 所述理想键值空间比计算单元21，适于计算选定系统块B1内部的理想键值空间比，所述理想键值空间比为键区与值区实际存储数据占用空间的比值；

[0080] 所述键值区域划分单元22，适于选取空闲系统块B2，对该空闲系统块B2按照理想键值空间比划分键区和值区；

[0081] 所述键值数据移动单元23适于将所述选定系统块B1存储的键值数据复制到该空闲系统块B2；

[0082] 所述标识号更新单元24，适于更新元信息中所述空闲系统块B2的标识号为所述选定系统块B1的标识号；

[0083] 所述指针更新单元25，适于将内存中指向所述选定系统块B1的指针指向所述空闲系统块B2；

[0084] 所述释放单元26，适于释放所述选定系统块B1。

[0085] 在具体实施中，数据库操作装置还可以包括：比较单元27(参见图3)，适于确定所述选定系统块内部的键区空间与值区空间的比值不同于所述理想键值空间比。

[0086] 在具体实施中，数据库操作装置还可以包括：读锁定单元28(参见图3)，适于在所述将所述选定系统块存储的键值数据复制到该空闲系统块时，对所述选定系统块读锁定。

[0087] 在具体实施中，数据库操作装置还可以包括：读写锁定单元29(参见图3)，适于在所述将内存中指向所述选定系统块的指针指向所述空闲系统块时，对所述选定系统块读写锁定。

[0088] 在具体实施中，数据库操作装置还可以包括：业务数据表存储单元，适于在所述选定系统块存储的数据属于某项业务数据表，且该业务数据表尚有待存入其他系统块的数据时，以所述理想键值空间比划分所述其他系统块的键区和值区，将所述待存入其他系统块的数据写入对应的键区和值区。

[0089] 在具体实施中，所述数据库可以存储于固态硬盘；所述数据库操作装置还包括：单次写入数据量控制单元，适于将待写入的数据暂存于内存，当所述暂存于内存的数据大小等于所述固态硬盘闪存块的大小时，将所述暂存于内存的数据一次写入所述固态硬盘的闪存块。

[0090] 在具体实施中，数据库操作装置还可以包括：系统块索引单元，适于关联键值和系

统块;所述选定系统块为与同一所述系统块索引关联的两个所述系统块。

[0091] 在具体实施中,数据库操作装置还可以包括:数据库状态判断单元,适于判断数据库状态,确认所述数据库处于非繁忙状态。

[0092] 在具体实施中,数据库操作装置还可以包括:定时选取单元,适于以固定的时间间隔计算选择系统块作为所述选定系统块。

[0093] 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过程序来指令相关的硬件来完成,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。

[0094] 虽然本发明披露如上,但本发明并非限定于此。任何本领域技术人员,在不脱离本发明的精神和范围内,均可作各种更动与修改,因此本发明的保护范围应当以权利要求所限定的范围为准。

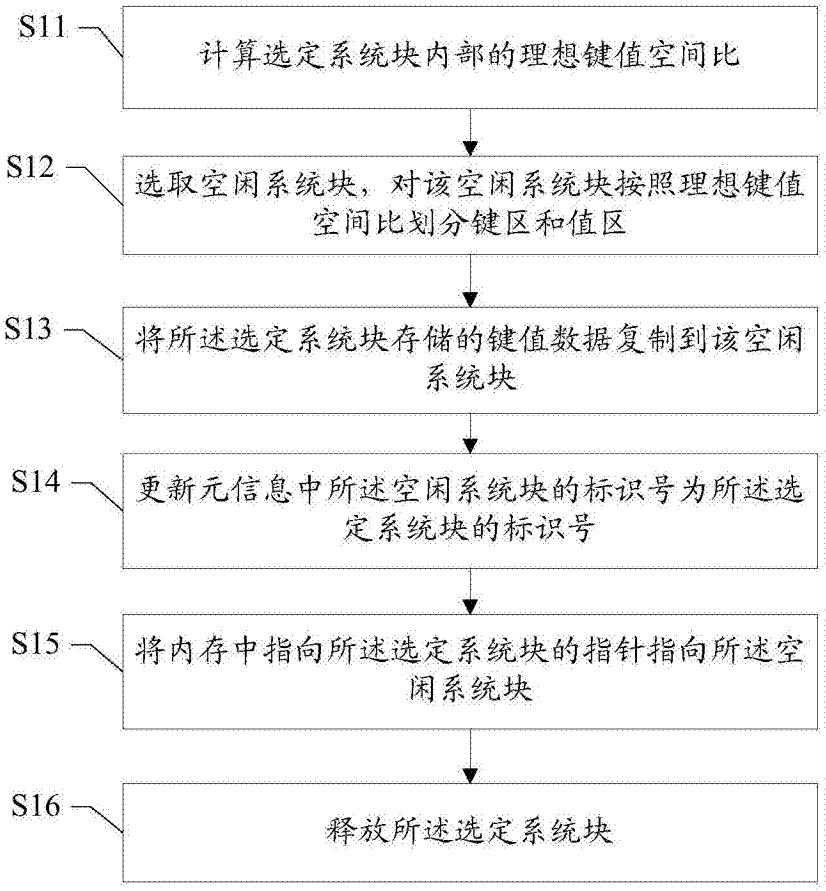

S11
计算选定系统块内部的理想键值空间比
S12
选取空闲系统块，对该空闲系统块按照理想键值空间比划分键区和值区
S13
将所述选定系统块存储的键值数据复制到该空闲系统块
S14
更新元信息中所述空闲系统块的标识号为所述选定系统块的标识号
S15
将内存中指向所述选定系统块的指针指向所述空闲系统块
S16
释放所述选定系统块

图1

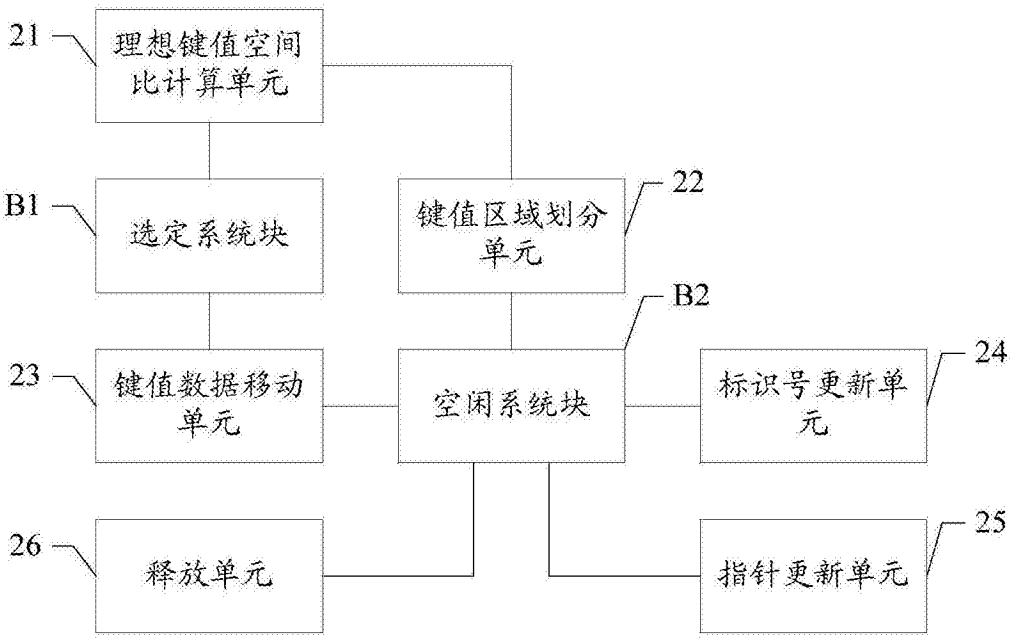

21
理想键值空间比计算单元
B1
选定系统块
22
键值区域划分单元
B2
23
键值数据移动单元
空闲系统块
24
标识号更新单元
26
释放单元
25
指针更新单元

图2

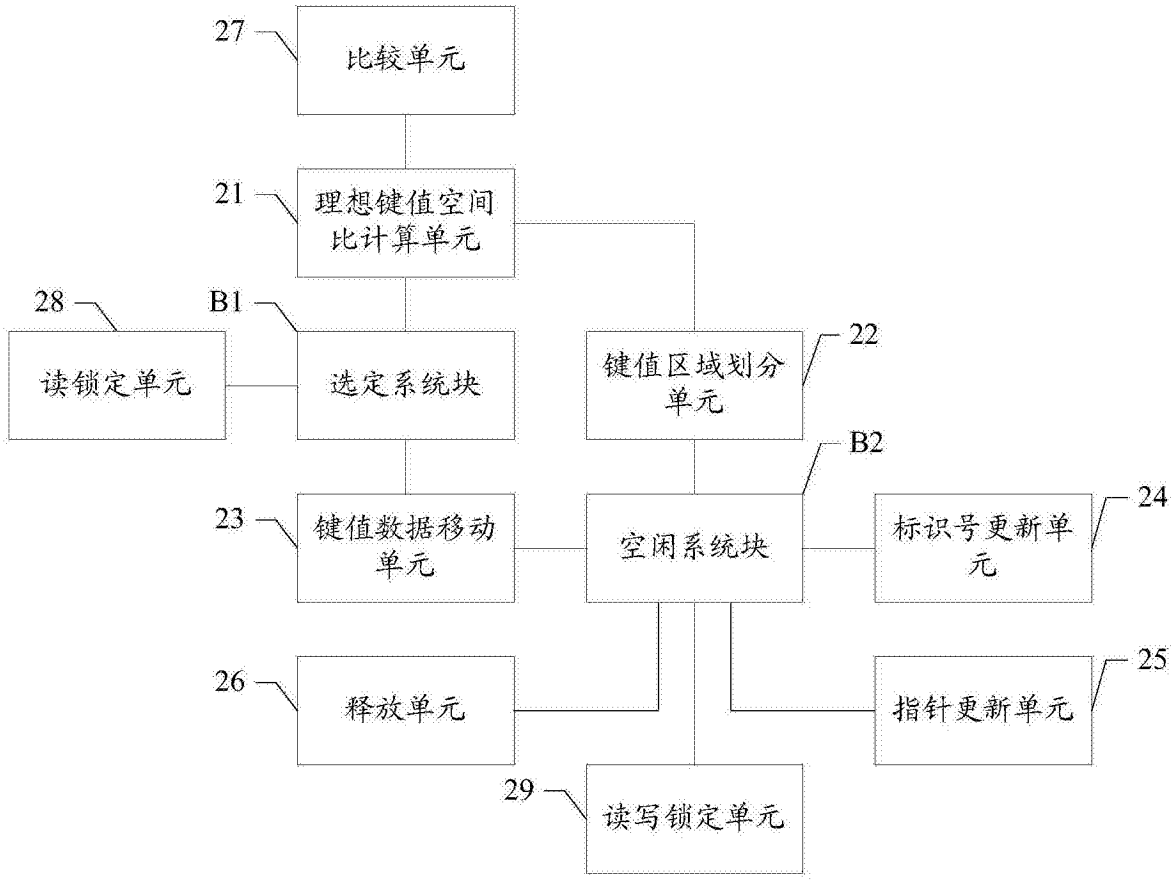

27
比较单元
21
理想键值空间比计算单元
28
B1
读锁定单元
选定系统块
键值区域划分单元
22
B2
23
键值数据移动单元
空闲系统块
标识号更新单元
24
26
释放单元
指针更新单元
25
29
读写锁定单元

图3