

# (12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局



(43) 国际公布日  
2022 年 12 月 15 日 (15.12.2022)

WIPO | PCT

(10) 国际公布号  
WO 2022/257337 A1

- (51) 国际专利分类号:  
*G06F 11/14* (2006.01) *G06F 16/22* (2019.01)
- (21) 国际申请号: PCT/CN2021/127407
- (22) 国际申请日: 2021 年 10 月 29 日 (29.10.2021)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
202110646471.X 2021年6月10日 (10.06.2021) CN
- (71) 申请人: 苏州浪潮智能科技有限公司  
(INSPUR SUZHOU INTELLIGENT TECHNOLOGY CO., LTD.) [CN/CN]; 中国江苏省苏州市吴中区吴中经济开发区郭巷街道官浦路1号9幢, Jiangsu 215000 (CN)。

- (72) 发明人: 刘文国(LIU, Wenguo); 中国江苏省苏州市吴中区吴中经济开发区郭巷街道官浦路1号9幢, Jiangsu 215000 (CN)。王艳清(WANG, Yanqing); 中国江苏省苏州市吴中区吴中经济开发区郭巷街道官浦路1号9幢, Jiangsu 215000 (CN)。
- (74) 代理人: 北京三聚阳光知识产权代理有限公司 (SUNSHINE INTELLECTUAL PROPERTY INTERNATIONAL CO., LTD.); 中国北京市海淀区海淀南路甲21号中关村知识产权大厦A座5层503, Beijing 100080 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) Title: METHOD AND SYSTEM FOR RECOVERING MASTER METADATA IN REAL TIME, AND DEVICE AND MEDIUM

(54) 发明名称: 一种实时恢复主元数据的方法、系统、设备及介质

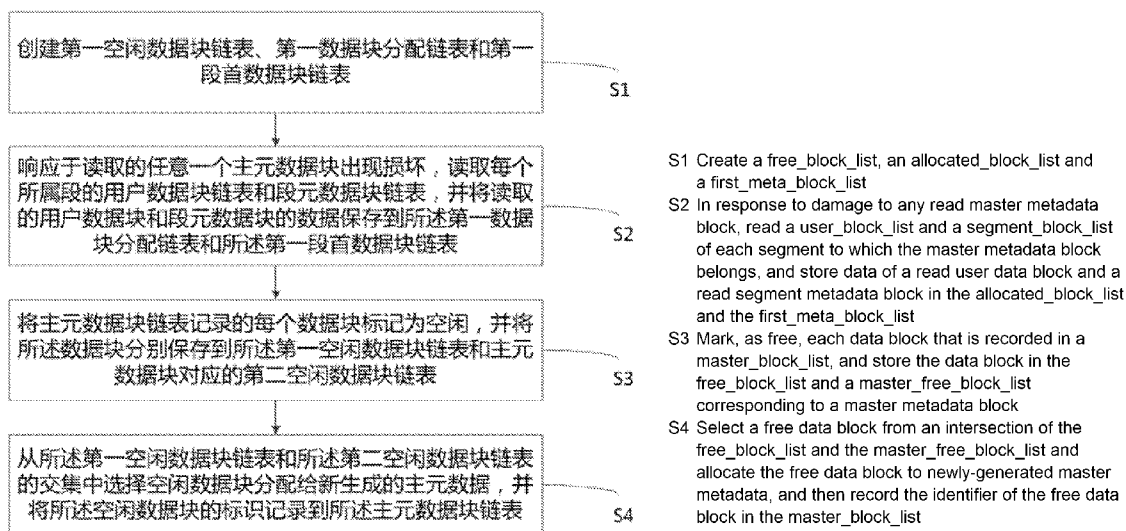


图 1

(57) Abstract: Disclosed in the present application are a method and system for recovering master metadata in real time, and a device and a storage medium. The method comprises: creating a free\_block\_list, an allocated\_block\_list and a first\_meta\_block\_list; in response to damage to any read master metadata block, reading a user\_block\_list and a segment\_block\_list of each segment to which the master metadata block belongs, and storing data of a read user data block and a read segment metadata block in the allocated\_block\_list and the first\_meta\_block\_list; marking, as free, each data block that is recorded in a master\_block\_list, and storing, in the free\_block\_list and a master\_free\_block\_list corresponding to a master metadata block, each data block that is recorded in the master\_block\_list; and selecting a free data block from an intersection of the free\_block\_list and the master\_free\_block\_list and allocating the free data block to newly-generated master metadata, and then recording the identifier of the free data block in the master\_block\_list.

LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,  
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,  
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,  
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,  
UZ, VC, VN, WS, ZA, ZM, ZW。

**(84)** 指定国 (除另有指明, 要求每一种可提供的地区  
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,  
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,  
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,  
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,  
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,  
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,  
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

根据细则4.17的声明:

— 发明人资格 (细则4.17(iv))

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

**(57) 摘要:** 本申请公开了一种实时恢复主元数据的方法、系统、设备和存储介质, 方法包括: 创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表; 响应于读取的任意一个主元数据块出现损坏, 读取每个所属段的用户数据块链表和段元数据块链表, 并将读取的用户数据块和段元数据块的数据保存到第一数据块分配链表和第一段首数据块链表; 将主元数据块链表记录的每个数据块标记为空闲, 并将主元数据块链表记录的每个数据块分别保存到第一空闲数据块链表和主元数据块对应的第二空闲数据块链表; 以及从第一空闲数据块链表和第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据, 并将空闲数据块的标识记录到主元数据块链表。

## 一种实时恢复主元数据的方法、系统、设备及介质

本申请要求在 2021 年 6 月 10 日提交中国专利局、申请号为 202110646471.X、发明名称为“一种实时恢复主元数据的方法、系统、设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

### 技术领域

本申请涉及数据中心领域，更具体地，特别是指一种实时恢复主元数据的方法、系统、计算机设备及可读介质。

### 背景技术

当今时代，随着数据的指数级增长，对存储容量的要求越来越高。通常有两种途径来应对存储容量的增长，一个是扩展存储容量，但这会带来成本的增加；另一个是对数据进行压缩，减少存储容量的消耗，节省成本。实时压缩存储系统在数据写入存储设备之前对数据进行压缩，不仅减少了存储空间的占用，更能进一步减少对存储设备（比如磁盘）的访问，从而提高存储系统的性能。

现有技术中，当实时压缩存储系统启动时，先从物理存储空间读取主元数据块，然后根据主元数据记录的段的首个段元数据块的地址，将段元数据块从物理存储空间读入内存，根据段元数据块记录的用户数据的存储位置，可以将所需的用户数据从物理存储空间读入内存。当实时压缩存储系统接收到用户写数据时，先将用户写数据在内存中压缩，当多个压缩数据的长度能够凑够一个数据块的长度时，由系统在物理存储空间中分配一个空闲数据块，然后将压缩的数据由内存写入该空闲数据块，同时更新主元数据，将刚分配的空闲数据块状态标记为“分配用户数据”，并更新段元数据，记录压缩的用户数据在物理存储空间的存储位置。

当前的实时压缩存储系统，当主元数据块损坏时，无法找到所有段的段元数据块，从而导致无法读取用户数据，且无法得知物理存储空间中每个数据块的状态，例如是空闲还是存储，是存储用户数据、段元数据还是主元数据，这样就需要遍历物理存储空间中每个数据块，根据每个数据块中记录的数据块状态、用户数据的存储位置、数据长度等信息重建段元数据和主元数据，严重影响了实时压缩存储系统的性能。

## 发明内容

有鉴于此，本申请实施例的目的在于提出一种实时恢复主元数据的方法、系统、计算机设备及计算机可读存储介质，本申请通过读取新增加的数据块链表，并遍历段元数据块而不是全部的数据块，从而能够迅速将主元数据恢复出来，提高实时压缩存储系统的性能。

基于上述目的，本申请实施例的一方面提供了一种实时恢复主元数据的方法，包括如下步骤：创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；以及从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

在一些实施方式中，所述从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据包括：读取所述第一空闲数据块链表，并将读取到的数据赋给所述第二空闲数据块链表。

在一些实施方式中，所述将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表包括：将读取的

用户数据块和段元数据块的标识及相应分配状态保存 to 所述第一数据块分配链表；以及将每个所属段的所述段元数据块链表存储的第一个段元数据块的标识存储到所述第一段首数据块链表。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识从所述第一空闲数据块链表和所述第二空闲数据块链表中删除。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识和分配状态加入所述主元数据块对应的第二数据块分配链表。

在一些实施方式中，方法还包括：响应于新增用户数据或段元数据，从所述第一空闲数据块链表中选择第二空闲数据块或第三空闲数据块分配给所述用户数据或所述段元数据。

在一些实施方式中，方法还包括：将所述第二空闲数据块和所述第三空闲数据块的标识分别记录到所属段的用户数据块链表和段元数据块链表中。

本申请实施例的另一方面，提供了一种实时恢复主元数据的系统，包括：创建模块，配置用于创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；第一链表模块，配置用于响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存 to 所述第一数据块分配链表和所述第一段首数据块链表；第二链表模块，配置用于将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存 to 所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；执行模块，配置用于从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

本申请实施例的又一方面，还提供了一种计算机设备，包括：至少一个处理器；以及存储器，所述存储器存储有可在所述处理器上运行的计算机指

令，所述指令由所述处理器执行时实现如上方法的步骤。

本申请实施例的再一方面，还提供了一种计算机可读存储介质，计算机可读存储介质存储有被处理器执行时实现如上方法步骤的计算机程序。

本申请具有以下有益技术效果：增加了恢复主元数据所需的空闲数据块链表、数据块分配链表和段首数据块链表，在用户数据写入过程中，不仅更新主元数据、段元数据，而且更新增加的上述链表。当主元数据块发生损坏，需要恢复主元数据时，依次读取增加的空闲数据块链表、主元数据块链表、段元数据块链表和每个段相应的用户数据块链表等，从而重建主元数据。通过遍历段元数据块，从而能够迅速将主元数据恢复出来，提高实时压缩存储系统的性能。

## 附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的实施例。

图 1 为本申请提供的实时恢复主元数据的方法的实施例的示意图；

图 2 为本申请提供的实时恢复主元数据的计算机设备的实施例的硬件结构示意图；

图 3 为本申请提供的实时恢复主元数据的计算机存储介质的实施例的示意图。

## 具体实施方式

为使本申请的目的、技术方案和优点更加清楚明白，以下结合具体实施例，并参照附图，对本申请实施例进一步详细说明。

需要说明的是，本申请实施例中所有使用“第一”和“第二”的表述均是为了区分两个相同名称非相同的实体或者非相同的参量，可见“第一”“第

二”仅为了表述的方便，不应理解为对本申请实施例的限定，后续实施例对此不再一一说明。

本申请实施例的第一个方面，提出了一种实时恢复主元数据的方法的实施例。图 1 示出的是本申请提供的实时恢复主元数据的方法的实施例的示意图。如图 1 所示，本申请实施例包括如下步骤：

S1、创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；

S2、响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；

S3、将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；以及

S4、从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

可以将整个物理存储空间划分为若干个固定长度的数据块，当物理存储空间较大，数据块较多时，为便于管理，将数据块组成若干个段，每个段包含相同数量的数据块。根据用途，将数据块分为主元数据块、段元数据块和用户数据块。主元数据块记录主元数据，主元数据用于记录每个数据块的状态（空闲还是存储，存储用户数据、段元数据还是主元数据）和每个段的第一个段元数据块的首地址；段元数据块用于记录段元数据，段元数据用于记录压缩后的用户数据在物理存储空间的存储位置，用户数据块用于记录压缩后的用户数据。

创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表。第一空闲数据块链表用于记录空闲数据块的标识（block ID），第一数据块分配链表用于记录已分配数据块的标识及状态，第一段首数据块链表

用于记录每个段的首个段元数据块的标识。本申请实施例中，第一空闲数据块链表为 Free\_Block\_List，第一数据块分配链表为 Allocated\_Block\_List，第一段首数据块链表为 First\_Meta\_Block\_List。

另外，主元数据对应的链表包括：第二空闲数据块链表 Master\_Free\_Block\_List，第二数据块分配链表 Master\_Allocated\_Block\_List，第二段首数据块链表 Master\_First\_Meta\_Block\_List，以及主元数据块链表 Master\_Block\_List。段元数据对应的链表包括：组成该段的用户数据块链表 User\_Block\_List 和段元数据块链表 Segment\_Block\_List。

在一些实施方式中，方法还包括：响应于新增用户数据或段元数据，从所述第一空闲数据块链表中选择第二空闲数据块或第三空闲数据块分配给所述用户数据或所述段元数据。对于不同类型的数据可以分配不同的空闲数据块，并用分配的空闲数据块写入新增的数据。

在一些实施方式中，方法还包括：将所述第二空闲数据块和所述第三空闲数据块的标识分别记录到所属段的用户数据块链表和段元数据块链表中。当空闲数据块中写入对应的数据后，可以在所属段的对应链表中记录对应空闲数据块的标识，以明了数据块中记录数据的情况。同时，也可以将写入了数据的数据块的标识和状态记录到第一数据块分配链表和第二数据块分配链表。

当空闲数据块被分配后，将该数据块的标识从第一空闲数据块链表和第二空闲数据块链表中删除。若该数据块用于存储用户数据，将分配的数据块的 block ID 记录在所属段的 User\_Block\_List 上；若该数据块用于存储段元数据，将分配的数据块的 block ID 记录在所属段的 Segment\_Block\_List 上；若该数据块用于存储主元数据，将分配的数据块的 block ID 记录在 Master\_Block\_List 上。

响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表。

在一些实施方式中，所述将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表包括：将读取的用户数据块和段元数据块的标识及相应分配状态保存到所述第一数据块分配链表；以及将每个所属段的所述段元数据块链表存储的第一个段元数据块的标识存储到所述第一段首数据块链表。对于每个段，依次读取其 User\_Block\_List 和 Segment\_Block\_List，将其记录的用户数据块和段元数据块的 block ID 及相应的分配状态保存到第一数据块分配链表中；依次将每个段的 Segment\_Block\_List 存储的第一个段元数据块的 block ID 存储在第一段首数据块链表上。

将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表。将 Master\_Block\_List 记录的每个数据块标记为空闲，并将其分别保存到 Free\_Block\_List 和 Master\_Free\_Block\_List 上。只有将数据块标记为空闲才可以给该数据块分配新的数据，才可以记录主元数据，本申请实施例中为了恢复主元数据块，需要让主元数据块链表中原来记录数据的数据块都可以去记录新生成的主元数据，所以，将主元数据块链表记录的每个数据块标记为空闲。

从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。为新生成的主元数据分配新的空闲数据块并将数据写入，将这些数据块的 block ID 记录在 Master\_Block\_List 上，并更新 Free\_Block\_List、Master\_Free\_Block\_List 和 Master\_Allocated\_Block\_List。从第一空闲数据块链表和第二空闲数据块链表的交集中选择一个或多个空闲数据块作为主元数据块来接收新生成的主元数据以替换损坏的主元数据块。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识从所述第一空闲数据块链表和所述第二空闲数据块链表中删除。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识和分配状态加入所述主元数据块对应的第二数据块分配链表。

在一些实施方式中，所述从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据包括：读取所述第一空闲数据块链表，并将读取到的数据赋给所述第二空闲数据块链表。读取空闲数据块链表 `Free_Block_List`，将其赋值给 `Master_Free_Block_List`。

本实施例增加了恢复主元数据所需的空闲数据块链表、数据块分配链表和段首数据块链表，在用户数据写入过程中，不仅更新主元数据、段元数据，而且更新增加的上述链表。当主元数据块发生损坏，需要恢复主元数据时，依次读取增加的空闲数据块链表、主元数据块链表、段元数据块链表和每个段相应的用户数据块链表等，从而重建主元数据。通过遍历段元数据块，从而能够迅速将主元数据恢复出来，提高实时压缩存储系统的性能。

需要特别指出的是，上述实时恢复主元数据的方法的各个实施例中的各个步骤均可以相互交叉、替换、增加、删减，因此，这些合理的排列组合变换之于实时恢复主元数据的方法也应当属于本申请的保护范围，并且不应将本申请的保护范围局限在实施例之上。

基于上述目的，本申请实施例的第二个方面，提出了一种实时恢复主元数据的系统，包括：创建模块，配置用于创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；第一链表模块，配置用于响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；第二链表模块，配置用于将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；执行模块，配置用于从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，

并将所述空闲数据块的标识记录到所述主元数据块链表。

在一些实施方式中，所述执行模块配置用于：读取所述第一空闲数据块链表，并将读取到的数据赋给所述第二空闲数据块链表。

在一些实施方式中，所述第一链表模块配置用于：将读取的用户数据块和段元数据块的标识及相应分配状态保存到所述第一数据块分配链表；以及将每个所属段的所述段元数据块链表存储的第一个段元数据块的标识存储到所述第一段首数据块链表。

在一些实施方式中，所述执行模块配置用于：将所述空闲数据块的标识从所述第一空闲数据块链表和所述第二空闲数据块链表中删除。

在一些实施方式中，所述执行模块配置用于：将所述空闲数据块的标识和分配状态加入所述主元数据块对应的第二数据块分配链表。

在一些实施方式中，系统还包括分配模块，配置用于：响应于新增用户数据或段元数据，从所述第一空闲数据块链表中选择第二空闲数据块或第三空闲数据块分配给所述用户数据或所述段元数据。

在一些实施方式中，系统还包括记录模块，配置用于：将所述第二空闲数据块和所述第三空闲数据块的标识分别记录到所属段的用户数据块链表和段元数据块链表中。

基于上述目的，本申请实施例的第三个方面，提出了一种计算机设备，包括：至少一个处理器；以及存储器，存储器存储有可在处理器上运行的计算机指令，指令由处理器执行以实现如下步骤：S1、创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；S2、响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；S3、将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；以及 S4、从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲

数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

在一些实施方式中，所述从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据包括：读取所述第一空闲数据块链表，并将读取到的数据赋给所述第二空闲数据块链表。

在一些实施方式中，所述将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表包括：将读取的用户数据块和段元数据块的标识及相应分配状态保存到所述第一数据块分配链表；以及将每个所属段的所述段元数据块链表存储的第一个段元数据块的标识存储到所述第一段首数据块链表。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识从所述第一空闲数据块链表和所述第二空闲数据块链表中删除。

在一些实施方式中，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：将所述空闲数据块的标识和分配状态加入所述主元数据块对应的第二数据块分配链表。

在一些实施方式中，步骤还包括：响应于新增用户数据或段元数据，从所述第一空闲数据块链表中选择第二空闲数据块或第三空闲数据块分配给所述用户数据或所述段元数据。

在一些实施方式中，步骤还包括：将所述第二空闲数据块和所述第三空闲数据块的标识分别记录到所属段的用户数据块链表和段元数据块链表中。

如图 2 所示，为本申请提供的上述实时恢复主元数据的计算机设备的一个实施例的硬件结构示意图。

以如图 2 所示的装置为例，在该装置中包括一个处理器 201 以及一个存储器 202，并还可以包括：输入装置 203 和输出装置 204。

处理器 201、存储器 202、输入装置 203 和输出装置 204 可以通过总线

或者其他方式连接，图 2 中以通过总线连接为例。

存储器 202 作为一种非易失性计算机可读存储介质，可用于存储非易失性软件程序、非易失性计算机可执行程序以及模块，如本申请实施例中的实时恢复主元数据的方法对应的程序指令/模块。处理器 201 通过运行存储在存储器 202 中的非易失性软件程序、指令以及模块，从而执行服务器的各种功能应用以及数据处理，即实现上述方法实施例的实时恢复主元数据的方法。

存储器 202 可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储根据实时恢复主元数据的方法的使用所创建的数据等。此外，存储器 202 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在一些实施例中，存储器 202 可选包括相对于处理器 201 远程设置的存储器，这些远程存储器可以通过网络连接至本地模块。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

输入装置 203 可接收输入的用户名和密码等信息。输出装置 204 可包括显示屏等显示设备。

一个或者多个实时恢复主元数据的方法对应的程序指令/模块存储在存储器 202 中，当被处理器 201 执行时，执行上述任意方法实施例中的实时恢复主元数据的方法。

执行上述实时恢复主元数据的方法的计算机设备的任何一个实施例，可以达到与之对应的前述任意方法实施例相同或者相类似的效果。

本申请还提供了一种计算机可读存储介质，计算机可读存储介质存储有被处理器执行时执行如上方法的计算机程序。

如图 3 所示，为本申请提供的上述实时恢复主元数据的计算机存储介质的一个实施例的示意图。以如图 3 所示的计算机存储介质为例，计算机可读存储介质 3 存储有被处理器执行时执行如上方法的计算机程序 31。

最后需要说明的是，本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，可以通过计算机程序来指令相关硬件来完成，实时恢复主元数据的方法的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，程序的存储介质可为磁碟、光盘、只读存储记忆体（ROM）或随机存储记忆体（RAM）等。上述计算机程序的实施例，可以达到与之对应的前述任意方法实施例相同或者相类似的效果。

以上是本申请公开的示例性实施例，但是应当注意，在不背离权利要求限定的本申请实施例公开的范围的前提下，可以进行多种改变和修改。根据这里描述的公开实施例的方法权利要求的功能、步骤和/或动作不需以任何特定顺序执行。此外，尽管本申请实施例公开的元素可以以个体形式描述或要求，但除非明确限制为单数，也可以理解为多个。

应当理解的是，在本文中使用的，除非上下文清楚地支持例外情况，单数形式“一个”旨在也包括复数形式。还应当理解的是，在本文中使用的“和/或”是指包括一个或者一个以上相关联地列出的项目的任意和所有可能组合。

上述本申请实施例公开实施例序号仅仅为了描述，不代表实施例的优劣。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

所属领域的普通技术人员应当理解：以上任何实施例的讨论仅为示例性的，并非旨在暗示本申请实施例公开的范围（包括权利要求）被限于这些例子；在本申请实施例的思路下，以上实施例或者不同实施例中的技术特征之间也可以进行组合，并存在如上的本申请实施例的不同方面的许多其它变化，为了简明它们没有在细节中提供。因此，凡在本申请实施例的精神和原则之内，所做的任何省略、修改、等同替换、改进等，均应包含在本申请

实施例的保护范围之内。

## 权 利 要 求 书

1. 一种实时恢复主元数据的方法，其特征在于，包括以下步骤：

创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；

响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；

将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；以及

从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

2. 根据权利要求 1 所述的方法，其特征在于，所述从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据包括：

读取所述第一空闲数据块链表，并将读取到的数据赋给所述第二空闲数据块链表。

3. 根据权利要求 1 所述的方法，其特征在于，所述将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表包括：

将读取的用户数据块和段元数据块的标识及相应分配状态保存到所述第一数据块分配链表；以及

将每个所属段的所述段元数据块链表存储的第一个段元数据块的标识存储到所述第一段首数据块链表。

4. 根据权利要求 1 所述的方法，其特征在于，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：

将所述空闲数据块的标识从所述第一空闲数据块链表和所述第二空闲数据块链表中删除。

5. 根据权利要求 1 所述的方法，其特征在于，所述将所述空闲数据块的标识记录到所述主元数据块链表包括：

将所述空闲数据块的标识和分配状态加入所述主元数据块对应的第二数据块分配链表。

6. 根据权利要求 1 所述的方法，其特征在于，方法还包括：

响应于新增用户数据或段元数据，从所述第一空闲数据块链表中选择第二空闲数据块或第三空闲数据块分配给所述用户数据或所述段元数据。

7. 根据权利要求 6 所述的方法，其特征在于，方法还包括：

将所述第二空闲数据块和所述第三空闲数据块的标识分别记录到所属段的用户数据块链表和段元数据块链表中。

8. 一种实时恢复主元数据的系统，其特征在于，包括：

创建模块，配置用于创建第一空闲数据块链表、第一数据块分配链表和第一段首数据块链表；

第一链表模块，配置用于响应于读取的任意一个主元数据块出现损坏，读取每个所属段的用户数据块链表和段元数据块链表，并将读取的用户数据块和段元数据块的数据保存到所述第一数据块分配链表和所述第一段首数据块链表；

第二链表模块，配置用于将主元数据块链表记录的每个数据块标记为空闲，并将主元数据块链表记录的每个数据块分别保存到所述第一空闲数据块链表和主元数据块对应的第二空闲数据块链表；

执行模块，配置用于从所述第一空闲数据块链表和所述第二空闲数据块链表的交集中选择空闲数据块分配给新生成的主元数据，并将所述空闲数据块的标识记录到所述主元数据块链表。

9. 一种计算机设备，其特征在于，包括：

至少一个处理器；以及

存储器,所述存储器存储有可在所述处理器上运行的计算机指令,所述指令由所述处理器执行时实现权利要求 1-7 任意一项所述方法的步骤。

10. 一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求 1-7 任意一项所述方法的步骤。

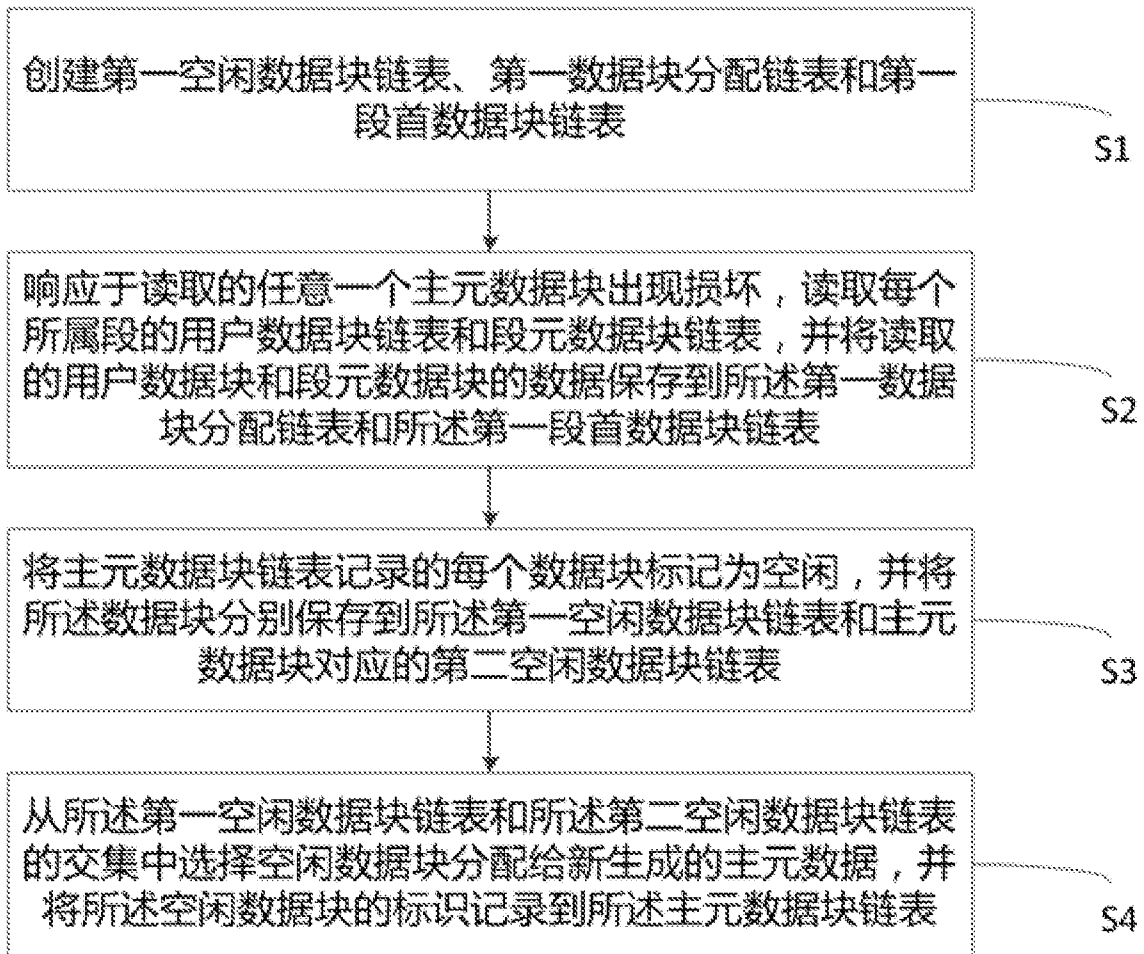


图 1

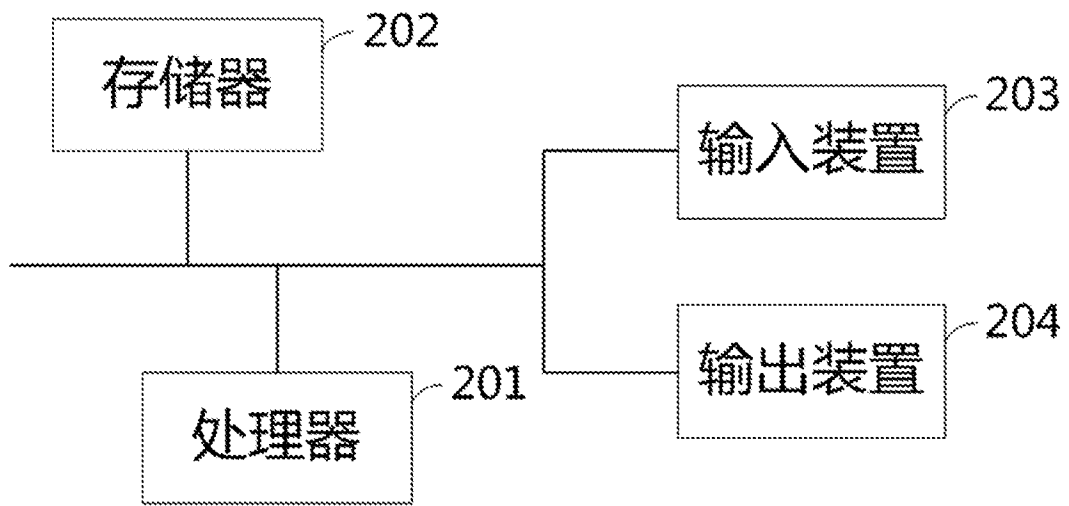


图 2

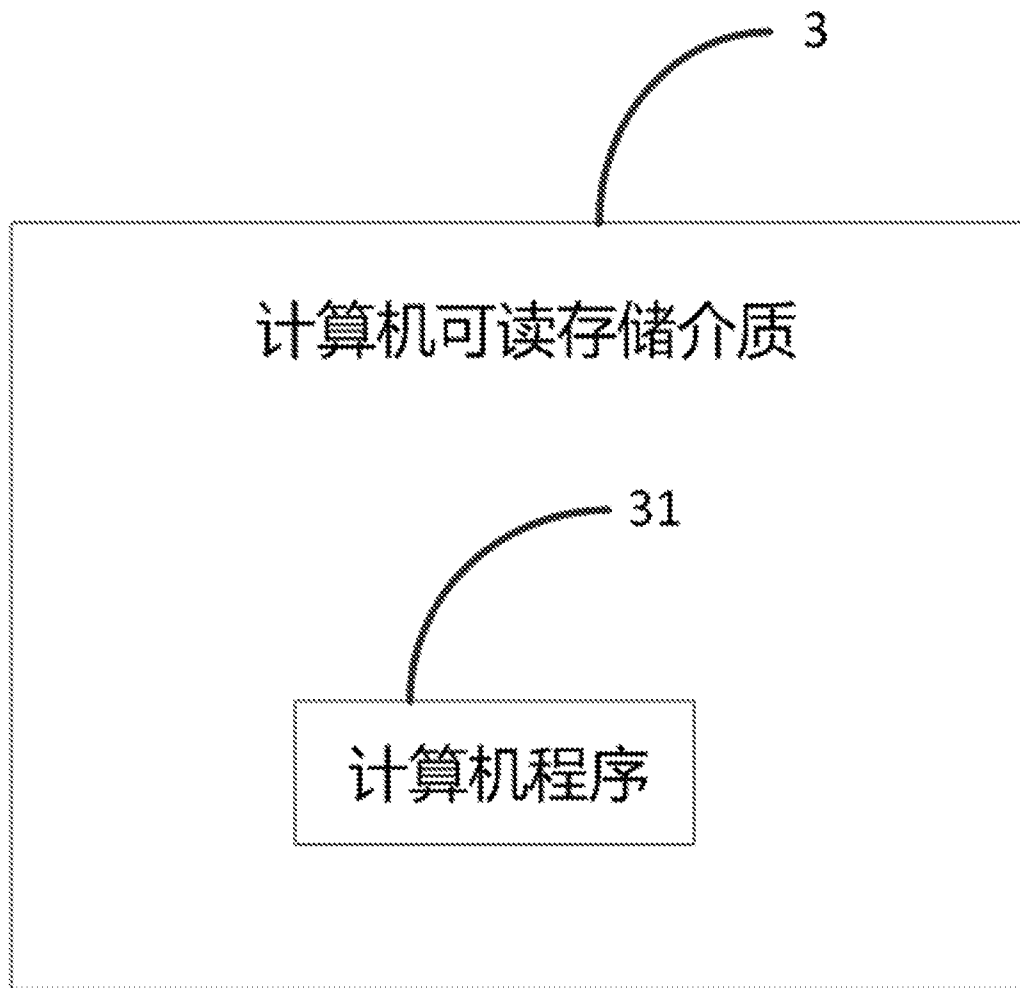


图 3

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/127407

**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 11/14(2006.01)i; G06F 16/22(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI, IEEE, GOOGLE: 元数据, 数据库, 链表, 段, 空闲, 分配, 记录, metadata, database, linked  
1w list, segment, idle, allocation, record**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                | Relevant to claim No. |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| PX        | CN 113254273 A (SUZHOU INSPUR INTELLIGENT TECHNOLOGY CO., LTD.) 13 August 2021 (2021-08-13)<br>claims 1-10                        | 1-10                  |
| A         | CN 109683825 A (GUANGDONG INSPUR BIG DATA RESEARCH CO., LTD.) 26 April 2019 (2019-04-26)<br>description, paragraphs [0052]-[0067] | 1-10                  |
| A         | CN 103488710 A (GUANGZHOU SEQUOIA SOFTWARE DEVELOPMENT CO., LTD.) 01 January 2014 (2014-01-01)<br>entire document                 | 1-10                  |
| A         | US 2002062422 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 23 May 2002 (2002-05-23)<br>entire document                              | 1-10                  |



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

22 February 2022

Date of mailing of the international search report

10 March 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/  
CN)  
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing  
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2021/127407**

| Patent document<br>cited in search report |            |    | Publication date<br>(day/month/year) | Patent family member(s) |         |    | Publication date<br>(day/month/year) |
|-------------------------------------------|------------|----|--------------------------------------|-------------------------|---------|----|--------------------------------------|
| CN                                        | 113254273  | A  | 13 August 2021                       | None                    |         |    |                                      |
| CN                                        | 109683825  | A  | 26 April 2019                        | None                    |         |    |                                      |
| CN                                        | 103488710  | A  | 01 January 2014                      | None                    |         |    |                                      |
| US                                        | 2002062422 | A1 | 23 May 2002                          | GB                      | 0028214 | D0 | 03 January 2001                      |
|                                           |            |    |                                      | GB                      | 2369206 | A  | 22 May 2002                          |

## 国际检索报告

国际申请号

PCT/CN2021/127407

## A. 主题的分类

G06F 11/14(2006.01)i; G06F 16/22(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

## B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT, WPI, EPDOC, CNKI, IEEE, GOOGLE: 元数据, 数据库, 链表, 段, 空闲, 分配, 记录, metadata, database, linked list, segment, idle, allocation, record

## C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落                                                                                   | 相关的权利要求 |
|------|-----------------------------------------------------------------------------------------------------|---------|
| PX   | CN 113254273 A (苏州浪潮智能科技有限公司) 2021年8月13日 (2021 - 08 - 13)<br>权利要求1-10                               | 1-10    |
| A    | CN 109683825 A (广东浪潮大数据研究有限公司) 2019年4月26日 (2019 - 04 - 26)<br>说明书第[0052]-[0067]段                    | 1-10    |
| A    | CN 103488710 A (广州巨杉软件开发有限公司) 2014年1月1日 (2014 - 01 - 01)<br>全文                                      | 1-10    |
| A    | US 2002062422 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2002年5月23日<br>(2002 - 05 - 23)<br>全文 | 1-10    |

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

\* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&amp;” 同族专利的文件

国际检索实际完成的日期

2022年2月22日

国际检索报告邮寄日期

2022年3月10日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

孙韬敏

电话号码 86-(10)-53961420

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2021/127407

| 检索报告引用的专利文件 |            |    | 公布日<br>(年/月/日) | 同族专利 |         |    | 公布日<br>(年/月/日) |
|-------------|------------|----|----------------|------|---------|----|----------------|
| CN          | 113254273  | A  | 2021年8月13日     | 无    |         |    |                |
| CN          | 109683825  | A  | 2019年4月26日     | 无    |         |    |                |
| CN          | 103488710  | A  | 2014年1月1日      | 无    |         |    |                |
| US          | 2002062422 | A1 | 2002年5月23日     | GB   | 0028214 | D0 | 2001年1月3日      |
|             |            |    |                | GB   | 2369206 | A  | 2002年5月22日     |