(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
2011 年 10 月 6 日 (06.10.2011)

PCT

(10) 国际公布号
WO 2011/120335 A1

(51) 国际专利分类号:
G06F 12/02 (2006.01)

(21) 国际申请号: PCT/CN2011/000547

(22) 国际申请日: 2011 年 3 月 30 日 (30.03.2011)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201010137051.0 2010 年 3 月 30 日 (30.03.2010) CN

(71) 申请人 (对除美国外的所有指定国): **联想（北京）有限公司 (LENOVO (BEIJING) CO., LTD)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。 **北京联想软件有限公司 (BEIJING LENOVO SOFTWARE LTD.)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。

(72) 发明人; 及

(75) 发明人/申请人 (仅对美国): **汤良 (TANG, Liang)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。 **郭昌盛 (GUO, Changsheng)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。 **刘春梅 (LIU, Chunmei)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。 **郭轶尊 (GUO, Yizun)** [CN/CN]; 中国北京市海淀区上地信息产业基地创业路 6 号, Beijing 100085 (CN)。

(74) 代理人: **中科专利商标代理有限责任公司 (CHINA SCIENCE PATENT & TRADEMARK AGENT CO., LTD)**; 中国北京市海淀区王庄路 1 号清华同方科技大厦 B 座 25 层, Beijing 100083 (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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV,

[见续页]

(54) Title: METHOD, DEVICE AND COMPUTER FOR DATA OPERATION

(54) 发明名称 : 数据操作的方法、装置及计算机

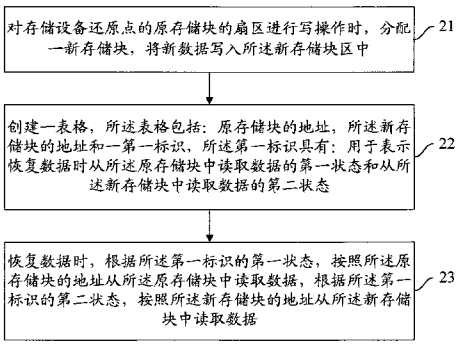


图 2 / Fig. 2

21 WHEN WRITING OPERATION TO SECTOR OF ORIGINAL STORAGE BLOCK OF STORAGE DEVICE REDUCTION POINT IS PERFORMED, NEW STROING BLOCK IS ALLOTTED AND NEW DATA IS WRITTEN INTO NEW STORAGE BLOCK AREA
22 TABLE IS CREATED, TABLE INCLUDES:ADDRESS OF ORIGINAL STORAGE BLOCK, ADDRESS OF NEW STORAGE BLOCK AND FIRST ID, FIRST ID HAS FIRST STATUS WHICH REPRESENTS READING DATA FROM ORIGINAL STORAGE BLOCK WHEN DATA IS RESUMED, AND SECOND STATUS OF READING DATA FROM NEW STORAGE BLOCK
23 WHEN DATA IS RESUMED, ACCORDING TO FIRST STATUS OF FIRST ID, DATA IS READ FROM ORIGINAL STORAGE BLOCK ACCORDING TO ADDRESS OF ORIGINAL STORAGE BLOCK, ACCORDING TO SECOND STATUS OF FIRST ID, DATA IS READ FROM NEW STORAGE BLOCK ACCORDING TO ADDRESS OF NEW STORAGE BLOCK

(57) **Abstract**: A method, a device and a computer for data operation are provided, wherein the method includes that: when a writing data operation to a sector of a storage device is performed, whether the sector belongs to an original storage block of a reduction point of the storage device is judged; if yes, then a new storage block is allotted and the data is written into the new storage block; a table is created, the table includes: an address of the original storage block, an address of the new storage block and a first ID, the first ID has the first ID elements that respectively correspond to the sectors in the original storage block and the new storage block, each first ID element has a first status which represents reading the data from the original storage block when the data of the corresponding sector is resumed, and a second status which represents reading the data from the new storage block when the data of the corresponding sector is resumed; and the address of the original storage block in an address index list of the reduction point is pointed to the table; if no, then a new storage block is allotted and the data is written into the new storage block; and an address of the original storage block in an address index list of the reduction point is modified to an address of the new storage block. The solution of the invention can save the time of operating the sector of the reduction point of the storage device, and can improve the capability of the computer.

[见续页]

WO 2011/120335 A1

SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

(57) **摘要**:

本发明提供一种数据操作的方法、装置及计算机，其中方法包括：在对存储设备的扇区进行写数据操作时，判断该扇区是否属于存储设备还原点的原存储块；如果是，分配新存储块，将数据写入所述新存储块中；创建表格，所述表格包括：原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态；以及将所述还原点的地址索引表中的所述原存储块的地址指向所述表格；如果否，分配新存储块，将数据写入所述新存储块中；以及将所述还原点的地址索引表中的所述原存储块的地址修改为新存储块的地址。本发明的方案节省了对存储设备的还原点的扇区的操作时间，提高计算机的性能。

数据操作的方法、装置及计算机

技术领域

本发明涉及计算机领域，特别是指一种数据操作的方法、装置及计算机。

背景技术

在计算机的存储设备的存储领域中，QCOW（QEMU Copy On Write）是一种支持能够回退到还原点的方法，在 QCOW 中，硬盘管理的单位为 Block（存储块），如图 1 所示，L1 对应于 QCOW 中一级硬盘地址页表，L2 对应于 QCOW 中二级硬盘地址页表，该 L1 和 L2 的作用是将一个虚拟扇区号转换成物理扇区号，其中，虚拟扇区号是 UOS（用户操作系统）访问虚拟硬盘的地址，通过虚拟扇区号读写虚拟硬盘内容，物理扇区号是物理硬盘扇区地址，Block offset 是虚拟扇区号在一个 Block 内部的偏移扇区号，还原点是用户可以回退的硬盘状态。

QCOW 采用 Copy On Write 存储方式进行扇区的写操作，在还原点建立后，如果是修改已有还原点里面的扇区，需要将该扇区所属的原 Block 里面的数据拷贝到新 Block 中，然后将需要写的数据写入新 Block 中，并记录该新 Block 的地址，这样每写入一次，就需要将要写入的原 Block 的内容拷贝到新 Block，再将需要写的内容写入新 Block，当需要频繁写扇区时，都要进行拷贝，这样会大大影响计算机的性能，造成性能损失。

一个还原点对应一个虚拟地址索引表，包括 L1 表和 L2 表，如图 1 所示。该 L1 表可称为该还原点的一级地址索引，L2 表可称为该还原点的二级地址索引，L1 表和 L2 表也可以合起来成为一张表，来表示该还原点的存储块的地址索引）。在恢复到还原点时，通过虚拟地址索引表，将虚拟硬盘地址访问转换成一个对物理硬盘地址的访问。当寻址转换的时候，先从一个最新的还原点的 L1，L2 表查找物理地址的相应的索引，如果最新的还原点的 L1，L2 表中没有找到相应的地址索引，通过递归查找前一个还原点的表，一直找到相应的索引，或者已经递归完所有还原点。

发明内容

本发明要解决的技术问题是提供一种数据操作的方法、装置及计算机，能够在对存储设备的扇区进行写操作时，不需要拷贝，节省读写操作时间，提高计算机的性能。

根据本发明实施例，提供一种数据操作的方法，包括:

在对存储设备的扇区进行写数据操作时，判断该扇区是否属于存储设备还原点的原存储块；

如果是，

分配新存储块，将数据写入所述新存储块中；

创建表格，所述表格包括：原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态；以及

将所述还原点的地址索引表中的所述原存储块的地址指向所述表格；

如果否，

分配新存储块，将数据写入所述新存储块中；以及

将所述还原点的地址索引表中的所述原存储块的地址修改为新存储块的地址。

优选的，所述方法还包括：

在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

优选的，当扇区属于存储设备还原点的原存储块时，将新数据写入所述新存储块中之后，所述方法还包括：

针对所述还原点的地址索引表中的所述原存储块的地址，添加第二标识，所述第二标识具有第一状态和第二状态，所述第一状态用于表示在恢复数据时直接根据原存储块的地址读取数据，所述第二状态用于表示在恢复数据时根据所述表格中所述第一标识的第一标识元素的状态读取数据。

优选的，所述方法还包括：

恢复数据时，查找存储设备的还原点的地址索引表，得到要从中读取数据的存储块的地址；

若针对所述地址的第二标识处于第一状态，则直接根据该存储块的地址读取数据；

若针对所述地址的第二标识处于第二状态，根据所述表格中第一标识的相应第一标识元素的状态读取数据。

优选的，当扇区属于存储设备还原点的原存储块时，将新数据写入所述新存储块

中之后，所述方法还包括:

在对所述新存储块中的扇区又进行写数据操作时，直接将此次写操作的数据写入所述新存储块中，同时将所述表格中的对应于该扇区的第一标识元素置于第二状态。

优选的，若所述第一标识的第一标识元素全为第二状态，恢复对应扇区的数据时直接根据所述新存储块的地址从所述新存储块中读取数据。

优选的，若所述第一标识的第一标识元素全为第二状态，释放所述表格。

优选的，按照所述原存储块的地址从所述原存储块中读取数据的步骤包括:

将所述原存储块的地址加上所述原存储块中扇区的偏移地址，得到所述原存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据;

按照所述新存储块的地址从所述新存储块中读取数据的步骤包括:

将所述新存储块的地址加上所述新存储块中扇区的偏移地址，得到所述新存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据。

优选的，所述表格由位图表格或结构体数据表示，所述第一标识由位图表示。

根据本发明实施例，提供一种数据写操作的方法，包括:

对存储设备还原点的原存储块的扇区进行写操作时，分配新存储块，将数据写入所述新存储块中;

创建表格，所述表格包括: 原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态; 以及

将所述还原点的地址索引表中的所述原存储块的地址指向所述表格。

根据本发明实施例，提供一种数据操作的装置，包括:

写操作模块，用于对存储设备还原点的原存储块的扇区进行写操作时，分配新存储块，将数据写入所述新存储块中;

创建模块，用于创建表格，所述表格包括: 原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态; 以及

读操作模块，用于在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第

一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

根据本发明实施例，提供一种计算机，包括:

存储设备;

虚拟机管理器;

第一操作系统模块，运行在所述虚拟机管理器上;

第二操作系统模块，通过所述第一操作系统模块和虚拟机管理器实现对所述存储设备的读/写操作;

所述第一操作系统模块包括:

写操作模块，用于对所述存储设备还原点的原存储块的扇区进行写操作时，分配新存储块，将数据写入所述新存储块中;

创建模块，用于创建表格，所述表格包括: 原存储块的地址，所述新存储块的地址和第一标识; 所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态;

读操作模块，用于在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

优选地，第一操作系统模块还包括:

释放模块，用于在所述表格中的第一标识的第一标识元素全为第二状态时，直接根据所述新存储块的地址从所述新存储块中读取数据，并释放所述表格。

根据本发明实施例，在对一个已有还原点的原存储块的扇区进行写操作时，创建一个表格（如 bitmap 表格），表述原存储块（原 Block）的地址，新分配的存储块（新 Block）的地址，恢复数据时，新写的扇区需要从新 Block 中读取，原存储块的扇区从原存储块中读取,这样在写入一个新扇区时,不用拷贝原Block的扇区的内容到新Block中，可以在新 Block 里面直接写入并做相应的记录第一标识的状态即可，若该 Block 都更新为新写入扇区的时候，就可以直接将索引改为新的 Block 地址，节省写入时间，提高计算机的性能。

附图说明

图 1 为现有的还原点的地址索引表;

图 2 为本发明的实施例数据操作的方法的流程图;

图 3 为本发明的还原点的地址索引表;

图 4 为图 2 所示方法中，存储设备的还原点的数据写操作方法流程图;

图 5 为图 2 所示方法中，存储设备的还原点的数据读操作方法流程图;

图 6 为本发明的实施例计算机的具体实现结构图。

具体实施方式

为使本发明的实施例要解决的技术问题、技术方案和优点更加清楚，下面将结合附图及具体实施例进行详细描述。

本发明的实施例针对现有技术中按照 Copy on write 的方式对存储设备进行写入操作，当需要频繁写扇区时，都要进行拷贝，这样会大大影响计算机的性能，造成性能损失的问题，提供一种数据操作的方法、装置和计算机。

如图 2 所示，本发明的实施例数据操作的方法，包括:

步骤 21，对存储设备还原点的原存储块的扇区进行写操作时，分配一新存储块，将新数据写入所述新存储块中;

步骤 22，创建一表格，所述表格包括：原存储块的地址，所述新存储块的地址，和一第一标识，所述第一标识具有：用于表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态;

步骤 23，恢复数据时，根据所述第一标识的第一状态，按照所述原存储块的地址从所述原存储块中读取数据，根据所述第一标识的第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

其中，该实施例中，步骤 22 中创建的表格具体形式可以是一个 Bitmap 表格（即位图表格，该位图表格包括：原存储块的地址，所述新存储块的地址，和一位图，其中该位图相当于上述第一标识，该位图具有表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态），也可以是一个结构体的数据表示的该表格，其中该结构体数据的成员 1 表示的是原存储块的地址，成员 2 表示的是所述新存储块的地址，成员 3 表示的上述第一标识；当然，本发明的实施例也不排除用其它的方式表述该表格。

该实施例的方法，在对一个已有还原点的原存储块的扇区进行写操作时，创建一个表格，表述原存储块（原 Block）的地址，新分配的存储块（新 Block）的地址，表格中的第一标识描述哪些扇区是新写的扇区，哪些是原存储块的扇区，恢复数据时，新写的扇区需要从新 Block 中读取，原存储块的扇区从原存储块中读取，这样在写入一个新扇区时,不用拷贝原 Block 的扇区的内容到新 Block 中，可以在新 Block 里面直接写入并做相应的记录第一标识的状态即可，若该 Block 都更新为新写入扇区的时候，就可以直接将索引改为新的 Block 地址，节省写入时间，提高计算机的性能。

其中，上述表格为位图表格时，该位图表格中的位图（即第一标识）的第一状态如可以是 0，第二状态可以是 1；或者第一状态是 1，第二状态可以是 0；分别表示存储块中哪些扇区是新写的扇区，哪些是原存储块的扇区；如位图是 1 时从新 Block 中读取数据；位图是 0 时从原 Block 中读取数据。

如图 3 所示，一个还原点对应一个虚拟地址索引表，该虚拟地址索引表包括：L1 表，L2 表及 Block offset，其中，L1 表中存放的是 L2 页表的地址，L2 表中存放的是存储块（Block）的地址，Block offset 是指虚拟扇区号在一个 Block 内部偏移扇区号；L1 和 L2 也可合并为一个表，表述还原点的虚拟地址索引；

其中，在对已有还有点的原 Block 中的扇区进行写操作时，除了要保留原 Block 的内容，还要记录写入的新数据，此时，需要分配一个新 Block，将需要写入的新数据写入新 Block 的扇区中，这样数据恢复时，可以从新 Block 中读取新写入的数据，从原 Block 中读取以前写入的数据，因此，在上述方法中，将新数据写入所述新存储块中之后还包括:

在所述还原点的地址索引表中的所述原存储块的地址中添加一第二标识（如图中 L2 表中的 Block 地址栏），所述第二标识也具有第一状态和第二状态，所述第二标识的第一状态用于表示在恢复数据时直接根据该原存储块的地址读取数据，所述第二标识的第二状态用于表示：在恢复数据时，根据所述表格中的第一标识的状态读取数据。

如果该第二标识用 Directly 这样一个二进制位表示，如 Directly 位为 1（如上述第二标识的第一状态，当然该第一状态也可以用除 1 之外的其它标识代替）时，表示可以直接按照该 Block 地址栏中的 Block 地址和 Block offset 找到对应的扇区的物理地址，并读出扇区中的内容；如果该 Directly 位为 0（如上述第二标识的第二状态，当然该第二状态也可以用除 0 之外的其它标识代替），表示需要根据下面的位图表格（Bitmap 表格）的位图的状态，从原 Block 中读取数据或者从新 Block 中读取数据；

在上述方法中，在将新数据写入所述新存储块中之后还可包括:

对所述新存储块中的扇区又进行写操作，直接将此次写操作的新数据写入所述新存储块中，同时将所述新存储块对应的表格中的第一标识置于第二状态；

如对新 Block 中的扇区又进行写操作，此时，就可以直接将要写入的数据写入该新 Block，从该 Block 中恢复数据时，直接根据该新 Block 地址和 Block offset 找到扇区的物理地址，并根据该扇区的物理地址读取扇区中的数据即可。

在上述方法中，若所述表格中的第一标识全为第二状态时，说明该还原点的存储块中的所有扇区均被执行重新写操作，恢复数据时直接根据所述新存储块的地址从所述新存储块的扇区中读取数据；

具体来讲，若所述位图表格中的位图位全为 1，说明该要恢复的 Block 中全部扇区是新写入数据的扇区，恢复数据时，直接根据该新 Block 的地址和 Block offset 找到扇区的物理地址，并根据该扇区的物理地址从扇区中读取数据。

另外，上述方法中，步骤 23 中，按照所述原存储块的地址从所述原存储块中读取数据的步骤具体为：

将所述原存储块的地址加上所述原存储块中扇区的偏移地址，得到所述原存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据；

按照所述新存储块的地址从所述新存储块中读取数据的步骤具体为：

将所述新存储块的地址加上所述新存储块中扇区的偏移地址，得到所述新存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据。

下面结合具体流程说明本发明的上述方法的具体实现，上述方法可以分为写操作流程和读操作流程，其中，

如图 4 所示，本发明的实施例还保护一种存储设备还原点的扇区的数据写操作的方法，该方法包括：

步骤 41，对存储设备还原点的原存储块的扇区进行写操作时，分配一新存储块，将新数据写入所述新存储块中；

步骤 42，创建一表格，所述表格包括：原存储块的地址，所述新存储块的地址，和一第一标识，所述第一标识具有表示恢复数据时从所述原存储块中读取数据的第一状态和恢复数据时从所述新存储块中读取数据的第二状态；

步骤 43，将所述还原点的地址索引表中的所述原存储块的地址指向所述表格。

其中，上述步骤 42 中创建的表格具体形式可以是一个 Bitmap 表格（即位图表格，该位图表格包括：原存储块的地址，所述新存储块的地址，和一位图，其中该位图相当于上述第一标识，该位图具有表示恢复数据时从所述原存储块中读取数据的第一状

态和从所述新存储块中读取数据的第二状态），也可以是一个结构体的数据表示的该表格，其中该结构体数据的成员 1 表示的是原存储块的地址，成员 2 表示的是所述新存储块的地址，成员 3 表示的上述第一标识；当然，本发明的实施例也不排除用其它的方式表述该表格。

下面举例说明上述写操作流程：

写扇区操作开始；

判断该扇区是否是已有还原点的原存储块（Block）的扇区，即该扇区是否写过数据；

如果是，而且不是新还原点第一次写的 BLOCK，则在内存中新建一位图表格（Bitmap 表格），将 L2 表中的 Block 地址栏指向该 Bitmap 表格；该 Bitmap 表格包括：原 Block 的地址，新 Block 的地址和位图；其中，新 Block 的地址指向新 Block，原 Block 的地址指向原 Block，数据恢复时，新写入的数据从新 Block 读取，原有的数据从原 Block 中读取；

分配新 Block，将新数据写入新 Block 中，并设置 Bitmap 表格中的位图，新 Block 对应设置为 1，原 Block 对应设置为 0；恢复数据时，根据该 Bitmap 表格中的位图的状态读取相应 Block 中的数据，位图的状态为 1，从新 Block 中读数据，位图的状态为 0，从原 Block 中读取数据；

写扇区完成后，判断位图的状态是否全部为 1，若是，将 L2 表中的 Block 地址修改为新 Block 地址，释放 Bitmap 表格；

如果新写入的扇区不是以前已有还原点的扇区，即从来没有写过的扇区，则分配新 Block，将新数据写入新 Block 中，并在 L2 表中将 Block 地址改为新 Block 地址。

如果写操作的扇区不是新还原点第一次写的 BLOCK，判断 Block 地址数据的第二标识位（Directly 位）；

若该 Directly 位为 1，说明可以直接写入该 Block 中；

若该 Directly 位为 0，根据 Block 地址栏中的 Block 地址指向，找到 Bitmap 表格在内存中的位置和新 Block 地址，将新数据写入新 Block 中，并设置 Bitmap 表格中的位图。

如图 5 所示，本发明的实施例还保护一种存储设备还原点的扇区的数据读操作的方法，包括：

步骤 51，查找存储设备的还原点的地址索引表，得到要恢复的存储块的地址；

步骤 52，若所述要恢复的存储块的地址中的第二标识表示该存储块的地址指向一

表格时，则根据所述表格中第一标识的第一状态，按照所述表格中原存储块的地址从所述原存储块中读取数据，根据所述第一标识的第二状态，按照所述新存储块的地址从所述新存储块中读取数据；进一步的，如果要恢复的存储块的第二标识表示该存储块的地址直接指向该存储块的物理扇区，则直接根据该存储块的地址读取该存储块对应的物理扇区中的数据。

该步骤 52 中的所述表格与上述图 4 的步骤 42 中所创建的表格相同。

下面举例说明上述读操作流程:

通过虚拟地址高位 L1，L2 表，得到 L2 表中 Block 地址栏;

判断 Block 地址数据的第二标识（Directly 位）;

若该 Directly 位为 1（即第二标识的第一状态），说明可以直接从该 Block 中读数据;

若该 Directly 位为 0（即第二标识的第二状态），根据 Block 地址栏中的 Block 地址指向，找到 Bitmap 表格在内存中的位置，读出 Bitmap 表格;

判断该 Bitmap 表格中的位图是否为 1，若为 1（即第一标识的第二状态），则根据 Bitmap 表格中的新 Block 地址读取该新 Block 中的数据；具体是：根据该新 Block 地址和 Block offset（Block 偏移）找到该新 Block 中扇区的物理地址，根据该扇区的物理地址读取扇区中的数据;

若位图为 0（即第一标识的第一状态），则根据 Bitmap 表格中的原 Block 地址读取该原 Block 中的数据；具体是：根据该原 Block 地址和 Block offset（Block 偏移）找到该原 Block 中扇区的物理地址，根据该扇区的物理地址读取扇区中的数据。

上述方案中，在写入一个新的扇区的时候,不用 copy 原有的 Block 内容到一个新 Block，可以在新 Block 里面直接写入并记录该 Block 地址和新写入的第一标识的状态就可以，节省了扇区写操作的时间，提高了计算机的性能；然后等待这个 Block 都更新为新的写入扇区的时候，就可以直接将索引 L2 表中的 Block 地址栏改为新的 Block 地址;同时可以在后台或者关机的时候进行更新的 Block 内容,然后将 L2 表中的 Block 地址索引改为新 Block 地址。

如图 6 所示，本发明的实施例还提供一种计算机，包括:

存储设备（如可以是硬盘存储设备）；虚拟机管理器（VMM）;

第一操作系统模块（如 Domain 0），运行在所述虚拟机管理器上;

第二操作系统模块（如用户操作系统 UOS），通过所述第一操作系统模块和虚拟机管理器实现对所述存储设备的读/写操作；其中，所述第一操作系统模块包括:

写操作模块，用于对所述存储设备还原点的原存储块的扇区进行写操作时，分配一新存储块，将新数据写入所述新存储块中；

创建模块，用于为创建一表格，所述表格包括：原存储块的地址，所述新存储块的地址和一第一标识；所述第一标识具有：用于表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态；

读操作模块，用于恢复数据时，根据所述第一标识的第一状态，按照所述原存储块的地址从所述原存储块中读取数据，根据所述第一标识的第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

其中，该创建模块创建的表格具体形式可以是一个 Bitmap 表格（即位图表格，该位图表格包括：原存储块的地址，所述新存储块的地址，和一位图，其中该位图相当于上述第一标识，该位图具有表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态），也可以是一个结构体的数据表示的该表格，其中该结构体数据的成员 1 表示的是原存储块的地址，成员 2 表示的是所述新存储块的地址，成员 3 表示的上述第一标识；当然，本发明的实施例也不排除用其它的方式表述该表格。

另外，所述第一操作系统模块还包括：

释放模块，用于在所述位图表格中的第一标识全为第二状态时，直接根据所述新存储块的地址从所述新存储块中读取数据，并释放所述表格。

具体来讲，上述写操作模块，创建模块，读操作模块和释放模块均可以是在 Domain 0 的数据恢复引擎模块中；具体实现流程如下：

写扇区流程包括：

UOS 进行扇区写访问；

QEMU 进行写命令模拟；

QEMU 进行虚拟扇区写操作；

QCOW（数据恢复引擎）在 L1，L2 表中查找虚拟扇区所属 Block 地址；

在 L2 表中找到该虚拟扇区所属的存储块的地址栏；

判断该虚拟扇区是否写过数据；

如果是，则在内存中新建一位图表格（Bitmap 表格），将 L2 表中的 Block 地址指向该 Bitmap 表格；

新分配 Block，将新数据写入新 Block 中，并设置 Bitmap 表格中的位图，新 Block 对应设置为 1，原 Block 对应设置为 0；

写扇区完成后，判断位图是否全部为 1，若是，将 L2 表中的 Block 地址修改为新 Block 地址，释放 Bitmap 表格；

如果新写入的扇区不是以前已有还原点的扇区，即从来没有写过的扇区，则新建 Block，将新数据写入新 Block 中，并在 L2 表中将 Block 地址改为新 Block 地址；

QCOW 写操作完成后，QEMU 返回相应的写操作结果给 UOS。

读扇区流程包括:

UOS 进行扇区读访问；

QEMU 进行读命令模拟；

QEMU 进行虚拟扇区读；

QCOW 对虚拟扇区地址进行索引，即在 L1，L2 表中查找该虚拟扇区所属 Block 的地址；

在 L2 表中找到该虚拟扇区所属的存储块的地址栏；

判断 Block 地址栏的第二标识（Directly 位）；

若该 Directly 位为 1，说明可以直接从该 Block 中的扇区读数据；

若该 Directly 位为 0，根据 Block 地址栏中的 Block 地址指向，找到 Bitmap 表格在内存中的位置，读出 Bitmap 表格；

判断该 Bitmap 表格中的位图是否为 1，若为 1，则根据 Bitmap 表格中的新 Block 地址读取该新 Block 中的数据，具体根据该新 Block 地址和 Block offset（Block 偏移）找到该新 Block 中扇区的物理地址，根据该扇区的物理地址读取扇区中的数据；

若位图为 0，则根据 Bitmap 表格中的原 Block 地址读取该原 Block 中的数据，具体根据该原 Block 地址和 Block offset（Block 中扇区的偏移量）找到该原 Block 中扇区的物理地址，根据该扇区的物理地址读取扇区中的数据；

QCOW 读操作完成后，QEMU 返回相应的结果给 UOS。

该实施例中，在写入一个新的扇区的时候，直接将新数据写入一个新 Block，并在 Bitmap 表格中，记录该新 Block 地址和新 Block 的位图的状态，恢复数据（即读扇区操作时），新数据从新 Block 中读出，原数据从原 Block 中读出，读出来的数据合并在一起可以使硬盘恢复到对该原 Block 的扇区写操作后的状态，在这个过程中，不用 copy 原 Block 内容到新 Block，节省了写操作时间，提高了计算机的性能；

然后等待这个 Block 都更新为新的写入扇区的时候，可以直接将索引 L2 表中的 Block 地址栏改为新的 Block 地址；同时可以在后台或者关机的时候进行更新的 Block 内容，然后将 L2 表中的 Block 地址索引改为新 Block 地址。

另外，本发明的实施列还提供一种与上述图 2 所示方法对应的存储设备的数据操作的装置，包括:

写操作模块，用于对存储设备还原点的原存储块的扇区进行写操作时，分配一新存储块，将新数据写入所述新存储块中;

创建模块，用于创建一表格，所述表格包括：原存储块的地址，所述新存储块的地址和一第一标识，所述第一标识具有：用于表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态;

读操作模块，用于恢复数据时，根据所述第一标识的第一状态，按照所述原存储块的地址从所述原存储块中读取数据，根据所述第一标识的第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

其中，该创建模块创建的表格和上述计算机中的创建模块创建的表格相同，具体可以是一个 Bitmap 位图表格。

其中，该写操作模块具体可以执行图 4 所示的写操作流程；读操作模块具体可以执行图 5 所示的读操作流程，同样在写入一个新的扇区的时候，直接将新数据写入一个新 Block，并在 Bitmap 表格中，记录该新 Block 地址和新 Block 的位图的状态，恢复数据（即读扇区操作时），新数据从新 Block 中读出，原数据从原 Block 中读出，读出来的数据合并在一起可以使硬盘恢复到对该原 Block 的扇区写操作后的状态，在这个过程中，不用 copy 原 Block 内容到新 Block，节省了写操作时间，提高了计算机的性能。

以上所述是本发明的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明所述原理的前提下，还可以作出若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。

权 利 要 求

1. 一种数据操作的方法，包括:

在对存储设备的扇区进行写数据操作时，判断该扇区是否属于存储设备还原点的原存储块;

如果是，

分配新存储块，将数据写入所述新存储块中;

创建表格，所述表格包括: 原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态; 以及

将所述还原点的地址索引表中的所述原存储块的地址指向所述表格;

如果否，

分配新存储块，将数据写入所述新存储块中; 以及

将所述还原点的地址索引表中的所述原存储块的地址修改为新存储块的地址。

2. 根据权利要求 1 所述的方法，还包括:

在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

3. 根据权利要求 1 所述的方法，其中，当扇区属于存储设备还原点的原存储块时，将新数据写入所述新存储块中之后，所述方法还包括:

针对所述还原点的地址索引表中的所述原存储块的地址，添加第二标识，所述第二标识具有第一状态和第二状态，所述第一状态用于表示在恢复数据时直接根据原存储块的地址读取数据，所述第二状态用于表示在恢复数据时根据所述表格中所述第一标识的第一标识元素的状态读取数据。

4. 根据权利要求 3 所述的方法，还包括:

恢复数据时，查找存储设备的还原点的地址索引表，得到要从中读取数据的存储块的地址;

若针对所述地址的第二标识处于第一状态，则直接根据该存储块的地址读取数据;

若针对所述地址的第二标识处于第二状态，根据所述表格中第一标识的相应第一标识元素的状态读取数据。

5. 根据权利要求 1 所述的方法，其中，当扇区属于存储设备还原点的原存储块时，将新数据写入所述新存储块中之后，所述方法还包括:

在对所述新存储块中的扇区又进行写数据操作时，直接将此次写操作的数据写入所述新存储块中，同时将所述表格中的对应于该扇区的第一标识元素置于第二状态。

6. 根据权利要求 1 所述的方法，其中，若所述第一标识的第一标识元素全为第二状态，恢复对应扇区的数据时直接根据所述新存储块的地址从所述新存储块中读取数据。

7. 根据权利要求 1 所述的方法，其中，若所述第一标识的第一标识元素全为第二状态，释放所述表格。

8. 根据权利要求 2 所述的方法，其中，按照所述原存储块的地址从所述原存储块中读取数据的步骤包括:

将所述原存储块的地址加上所述原存储块中扇区的偏移地址，得到所述原存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据;

按照所述新存储块的地址从所述新存储块中读取数据的步骤包括:

将所述新存储块的地址加上所述新存储块中扇区的偏移地址，得到所述新存储块中扇区的物理地址，根据所述扇区的物理地址读出所述扇区中的数据。

9. 根据权利要求 1 所述的方法，其中，所述表格由位图表格或结构体数据表示，所述第一标识由位图表示。

10. 一种数据写操作的方法，包括:

对存储设备还原点的原存储块的扇区进行写操作时，分配新存储块，将数据写入所述新存储块中;

创建表格，所述表格包括: 原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态; 以及

将所述还原点的地址索引表中的所述原存储块的地址指向所述表格。

11. 一种数据操作的装置，包括:

写操作模块，用于对存储设备还原点的原存储块的扇区进行写操作时，分配新存

储块，将数据写入所述新存储块中；

创建模块，用于创建表格，所述表格包括：原存储块的地址，所述新存储块的地址和第一标识，所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态；以及

读操作模块，用于在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

12. 一种计算机，包括：

存储设备；

虚拟机管理器；

第一操作系统模块，运行在所述虚拟机管理器上；

第二操作系统模块，通过所述第一操作系统模块和虚拟机管理器实现对所述存储设备的读/写操作；

所述第一操作系统模块包括：

写操作模块，用于对所述存储设备还原点的原存储块的扇区进行写操作时，分配新存储块，将数据写入所述新存储块中；

创建模块，用于创建表格，所述表格包括：原存储块的地址，所述新存储块的地址和第一标识；所述第一标识具有与原存储块和新存储块中的扇区分别对应的第一标识元素，每一个第一标识元素具有用于表示恢复对应扇区的数据时从所述原存储块中读取数据的第一状态和用于表示恢复对应扇区的数据时从所述新存储块中读取数据的第二状态；

读操作模块，用于在恢复扇区的数据时，如果所述第一标识中对应于该扇区的第一标识元素处于第一状态，按照所述原存储块的地址从所述原存储块中读取数据，如果所述第一标识元素处于第二状态，按照所述新存储块的地址从所述新存储块中读取数据。

13. 根据权利要求 12 所述的计算机，所述第一操作系统模块还包括：

释放模块，用于在所述表格中的第一标识的第一标识元素全为第二状态时，直接根据所述新存储块的地址从所述新存储块中读取数据，并释放所述表格。

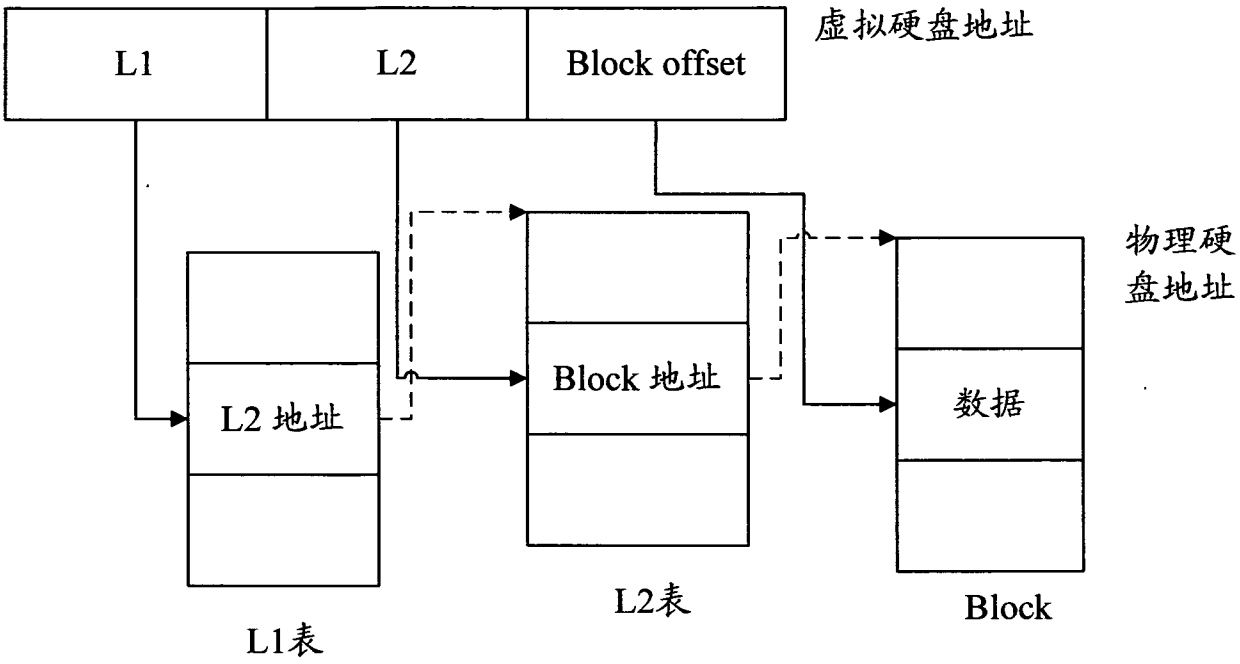

L1
L2
Block offset
虚拟硬盘地址
L2 地址
Block 地址
数据
物理硬盘地址
L1表
L2表
Block

图 1

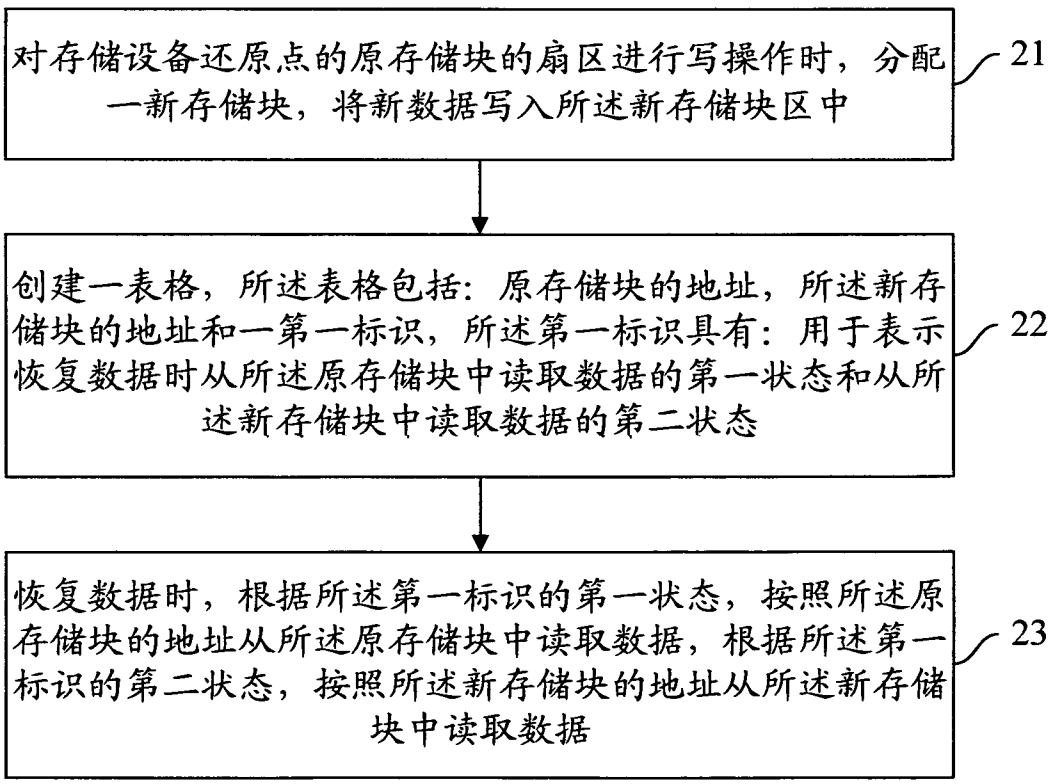

对存储设备还原点的原存储块的扇区进行写操作时，分配一新存储块，将新数据写入所述新存储块区中
21
创建一表格，所述表格包括：原存储块的地址，所述新存储块的地址和一第一标识，所述第一标识具有：用于表示恢复数据时从所述原存储块中读取数据的第一状态和从所述新存储块中读取数据的第二状态
22
恢复数据时，根据所述第一标识的第一状态，按照所述原存储块的地址从所述原存储块中读取数据，根据所述第一标识的第二状态，按照所述新存储块的地址从所述新存储块中读取数据
23

图 2

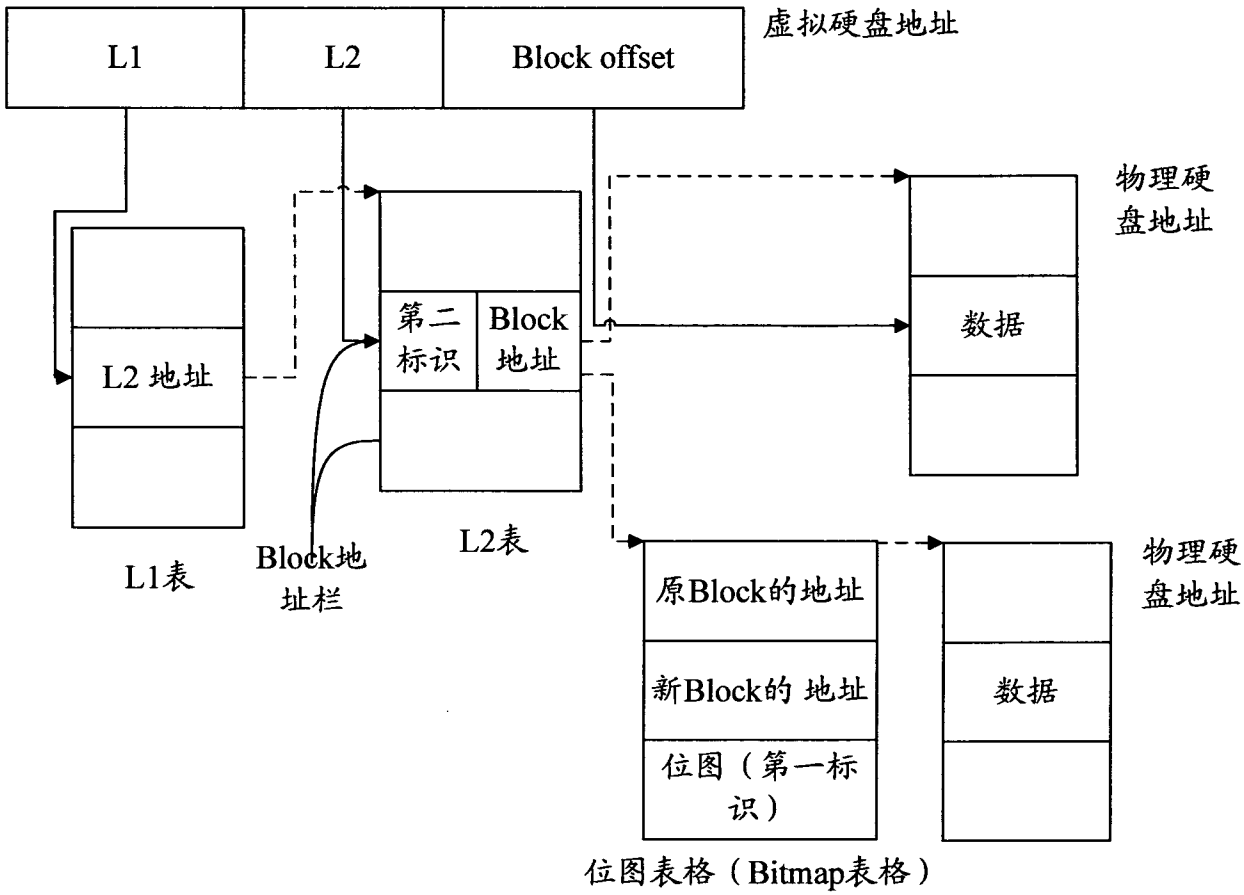

L1
L2
Block offset
虚拟硬盘地址
物理硬盘地址
L2 地址
第二标识
Block地址
数据
L1表
Block地址栏
L2表
原Block的地址
新Block的 地址
位图（第一标识）
数据
物理硬盘地址
位图表格（Bitmap表格）

图 3

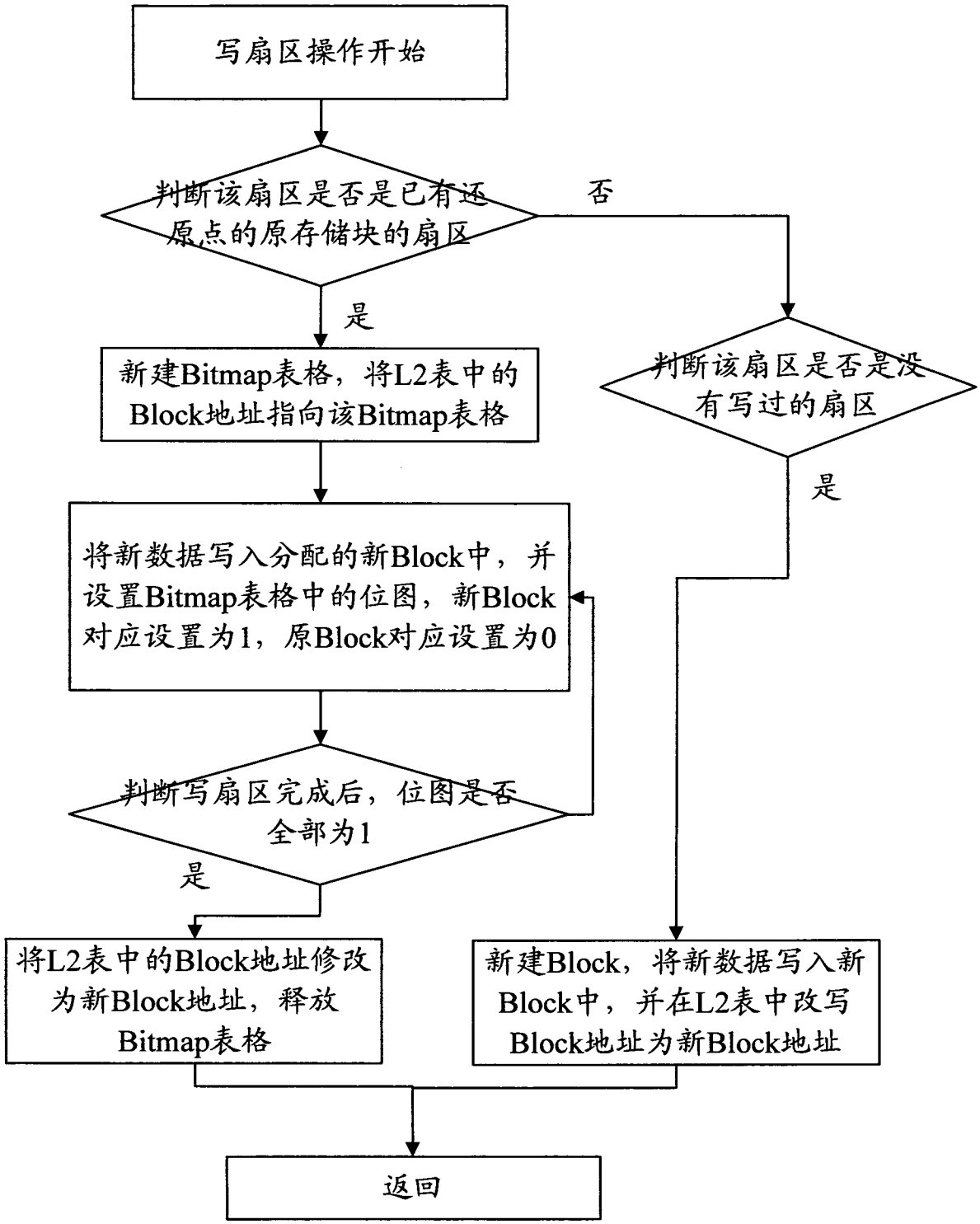

写扇区操作开始
判断该扇区是否是已有还原点的原存储块的扇区
否
是
新建Bitmap表格，将L2表中的Block地址指向该Bitmap表格
判断该扇区是否是没有写过的扇区
是
将新数据写入分配的新Block中，并设置Bitmap表格中的位图，新Block对应设置为1，原Block对应设置为0
判断写扇区完成后，位图是否全部为1
是
将L2表中的Block地址修改为新Block地址，释放Bitmap表格
新建Block，将新数据写入新Block中，并在L2表中改写Block地址为新Block地址
返回

图 4

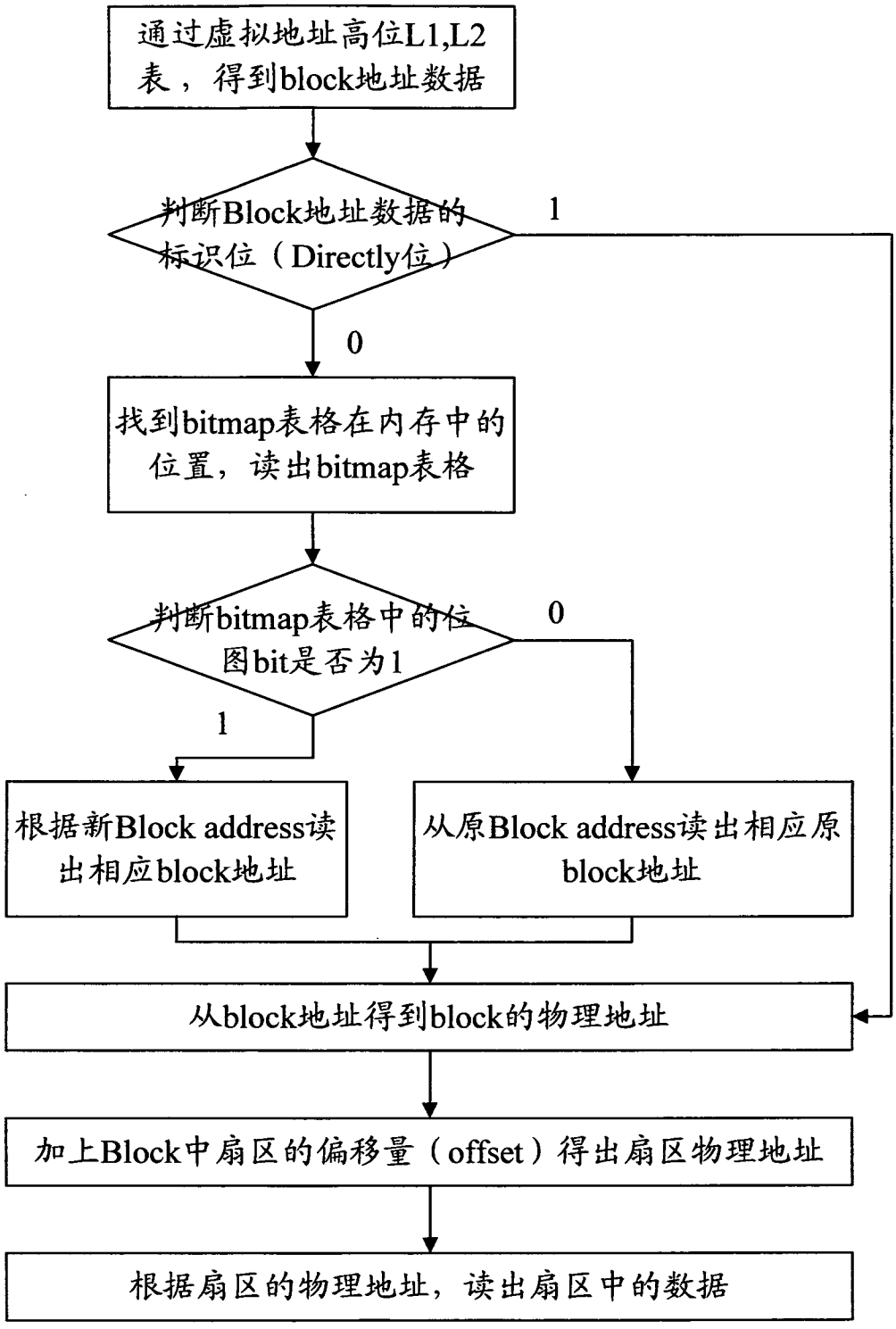

通过虚拟地址高位L1,L2表，得到block地址数据
判断Block地址数据的标识位（Directly位）
1
0
找到bitmap表格在内存中的位置，读出bitmap表格
判断bitmap表格中的位图bit是否为1
0
1
根据新Block address读出相应block地址
从原Block address读出相应原block地址
从block地址得到block的物理地址
加上Block中扇区的偏移量（offset）得出扇区物理地址
根据扇区的物理地址，读出扇区中的数据

图 5

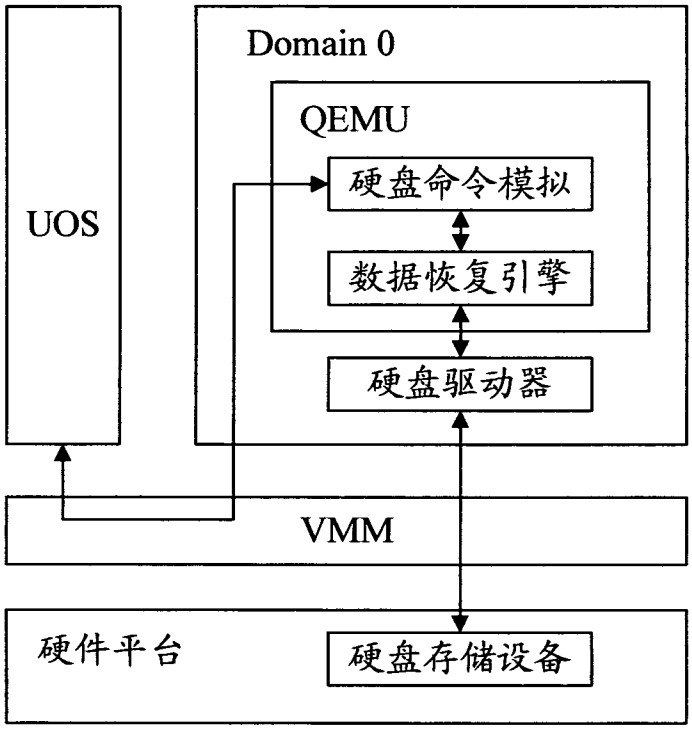


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2011/000547

A. CLASSIFICATION OF SUBJECT MATTER

G06F 12/02(2006.01)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)

IPC: G06F,G11C

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)
DWPI,SIPOABS,CPRSABS,CNTXT: Copy, on ,write, block, address, index, write operation, sector, Reduction, table

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN1700189A(RICHIP INC) 23 Nov. 2005 (23.11.2005) The whole document	1-13
A	CN101515276A(BEIJING WATCH DATA SYS CO LTD) 26 Aug. 2009 (26.08.2009) The whole document	1-13
A	US2009083503 A1(INVENTEC CORP) 26 Mar. 2009 (26.03.2009) The whole document	1-13

☐ 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	Date of mailing of the international search report
20 May 2011 (20.05.2011)	**07 Jul. 2011 (07.07.2011)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer TANG,Tiantian Telephone No. (86-10)62412065

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2011/000547

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN1700189A	23.11.2005	EP1736884A2	27.12.2006
		EP1736884A3	08.10.2008
		CN100353337C	05.12.2007
CN101515276A	26.08.2009	NONE	
US2009083503 A1	26.03.2009	NONE	

国际检索报告

国际申请号
PCT/CN2011/000547

A. 主题的分类

G06F 12/02(2006.01)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)

IPC: G06F,G11C

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）) DWPI,SIPOABS,CPRSABS,CNTXT: 存储，扇区，地址索引，写操作，还原，QCOW，存储块，表格,Copy, on ,write, block, address, index, write operation, sector, Reduction, table

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
A	CN1700189A（瑞程科技股份有限公司）　23.11 月 2005 (23.11.2005) 全文	1-13
A	CN101515276A(北京握奇数据系统有限公司)　26.8 月 2009 (26.08.2009)　全文	1-13
A	US2009083503 A1(英业达股份有限公司)　26.3 月 2009 (26.03.2009)　全文	1-13

☐ 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

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
20.5 月 2011（20.05.2011）	**07.7 月 2011 (07.07.2011)**
ISA/CN 的名称和邮寄地址： 中华人民共和国国家知识产权局 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 唐田田 电话号码：(86-10) **62412065**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2011/000547

检索报告中引用的专利文件	公布日期	同族专利	公布日期
CN1700189A	23.11.2005	EP1736884A2	27.12.2006
		EP1736884A3	08.10.2008
		CN100353337C	05.12.2007
CN101515276A	26.08.2009	无	
US2009083503 A1	26.03.2009	无	