

数据访问方法及系统、硬件卸载设备、电子设备及介质

本申请要求于 2022 年 03 月 25 日提交中国专利局、申请号为 202210307679.3、申请名称为“数据访问方法及系统、硬件卸载设备、电子设备及介质”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

本申请涉及计算机技术领域，尤其涉及一种数据访问方法及系统、硬件卸载设备、电子设备及介质。

背景技术

对象存储系统（Object-Based Storage System）是一种 Key-Value（键值对）形式的存储系统，能够提供高持久、高可用、高性能的对象存储服务。对象存储系统采用对象（Object）文件作为存储数据的基本单元，用户在使用对象存储系统的过程中，除了需要支付数据持久化的存储费用，还需要支付对象文件访问的请求费用。该请求费用一般是按照对象文件访问次数进行计费，对象文件访问次数越多，产生的请求费用越多。

在一些应用场景中，会产生海量的数据量较小的各种小文件，各种小文件例如包括但不限于：文本文件、图片文件、音频文件和视频文件等。目前，将每个小文件作为一个对象文件写入至对象存储系统中，这样，在将海量小文件写入对象存储系统时，会产生较多的对象文件访问次数，产生的请求费用较多，数据访问成本较高。

发明内容

本申请的多个方面提供一种数据访问方法及系统、硬件卸载设备、电子设备及介质，用以减少随着在对象存储系统中访问对象文件所产生的请求费用，降低数据访问成本。

本申请实施例提供一种数据访问方法，应用于硬件卸载设备上运行的客户端，硬件卸载设备通过总线与电子设备通信连接，该方法包括：获取电子设备上应用程序发送的待写入文件，并将待写入文件写入至硬件卸载设备上的缓存中；若缓存中已缓存的待写入文件满足文件合并条件，则将缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；将第一对象文件写入至客户端具有访问权限的对象存储系统中。

本申请实施例还提供一种硬件卸载设备，硬件卸载设备通过总线与电子设备通信连接，硬件卸载设备包括主处理器和缓存，主处理器运行客户端程序，以用于：获取电子设备上应用程序发送的待写入文件，并将待写入文件写入至硬件卸载设备上的缓存中；若缓存中已缓存的待写入文件满足文件合并条件，则将缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；将第一对象文件写入至客户端具有访问权限的对象存储系统中。

本申请实施例还提供一种电子设备，包括：处理器和上述硬件卸载设备，处理器上运行有至少一个应用程序，处理器通过总线与硬件卸载设备通信连接。

本申请实施例还提供一种数据访问系统，包括：电子设备、硬件卸载设备和对象存储系统；电子设备与硬件卸载设备通过总线通信连接，硬件卸载设备与对象存储系统通信连接；电子设备上运行有至少一个应用程序，用于通过应用程序向硬件卸载设备发送待写入文件；硬件卸载设备，用于获取待写入文件，并将待写入文件写入至硬件卸载设备上的缓存中；若缓存中已缓存的待写入文件满足文件合并条件，则将缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；以及向对象存储系统发送包括第一对象文件的写请求；对象存储系统，用于响应写请求存储第一对象文件。

本申请实施例还提供一种存储有计算机程序的计算机存储介质，当计算机程序被处理器执行时，致使处理器能够实现数据访问方法中的步骤。

在本申请实施例中，一方面，将针对对象存储系统的数据访问任务从电子设备转移至硬件卸载设备上执行，可以减轻电子设备的处理压力，提高电子设备的处理性能，且能够增强数据访问性能。另一方面，硬件卸载设备在每接收到来源于应用程序上的待写入文件后，并不直接将待写入文件写入至对象存储系统中，而是首先在本地进行缓存，在缓存到多个待写入文件后，再将多个待写入文件合并成一个新的对象文件写入至客户端具有访问权限的对象存储系统。这样，能够极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本，特别是针对海量小文件场景，可以极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本。另外，极大地减少了因大量的数据访问请求触发对象存储系统出现流量限速的情形发生，增强了对对象存储系统的存储性能和访问性能。

30

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1 为本申请实施例提供的一种数据访问系统的结构示意图；

图 2 为本申请实施例提供的一种数据访问方法的流程图；

5 图 3 为本申请实施例提供的另一种数据访问系统的结构示意图；

图 4 为本申请实施例提供的另一种数据访问方法的流程图；

图 5 为本申请实施例提供的一种硬件卸载设备的结构示意图；

图 6 为本申请实施例提供的一种电子设备的结构示意图。

10 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，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

15 另外，为了便于清楚描述本申请实施例的技术方案，在本申请的实施例中，采用了“第一”、“第二”等对功能和作用基本相同的相同项或相似项进行区分。本领域技术人员可以理解“第一”、“第二”等并不对数量和执行次序进行限定，并且“第一”、“第二”等也并不限定一定不同。

首先，对本申请实施例涉及的名词进行解释：

20 虚拟文件系统(Virtual Filesystem Switch, VFS)是一个内核软件层，为上层的应用程序提供 POSIX (Portable Operating System Interface of UNIX, 可移植操作系统接口)，以使上层的应用程序能够使用 POSIX 接口访问不同文件系统。

用户空间文件系统 (Filesystem in Userspace, FUSE)，是一个面向类 Unix 计算机操作系统的软件接口，它使无特权的用户能够无需编辑内核代码而创建自己的文件系统。用户空间文件系统提供内核模块和用户空间库 (libfuse) 模块，内核模块负责将文件操作命令封装成 FUSE 协议的文件操作请求，并通过传输通道发送给用户空间库模块；用户空间库模块接收并解析 FUSE 协议的文件操作请求，根据 FUSE 协议数据命令类型，调用对应的文件操作函数进行处理。有关用户空间文件系统中内核模块和用户空间库模块的更多介绍可以参见相关技术。

30 硬件卸载设备，是指具有硬件卸载 offload 功能的硬件设备。在本申请实施例中，硬件卸载设备能够承担原本由运行应用程序(Application, App)的电子设备访问对象存储系统的

数据访问任务，进而减轻电子设备的处理压力，提高电子设备的处理性能。

在一些应用场景中，会产生海量的数据量较小的各种小文件，各种小文件例如包括但不限于：文本文件、图片文件、音频文件和视频文件等。目前，将每个小文件作为一个对象文件写入至对象存储系统中，这样，在将海量小文件写入对象存储系统时，会产生较多的对象文件访问次数，产生的请求费用较多，数据访问成本较高。针对上述技术问题，本申请实施例提供一种数据访问方法及系统、硬件卸载设备、电子设备及介质。在本申请实施例中，一方面，将针对对象存储系统的数据访问任务从电子设备转移至硬件卸载设备上执行，可以减轻电子设备的处理压力，提高电子设备的处理性能，且能够增强数据访问性能。另一方面，硬件卸载设备在每接收到来源于应用程序上的待写入文件后，并不直接将待写入文件写入至对象存储系统中，而是首先在本地进行缓存，在缓存到多个待写入文件后，再将多个待写入文件合并成一个新的对象文件写入至客户端具有访问权限的对象存储系统。这样，能够极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本，特别是针对海量小文件场景，可以极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本。另外，极大地减少了因大量的数据访问请求触发对象存储系统出现流量限速的情形发生，增强了对对象存储系统的存储性能和访问性能。

图 1 为本申请实施例提供的一种数据访问系统的结构示意图。参见图 1，该系统可以包括：电子设备 10、硬件卸载设备 20 和对象存储系统 40。其中，电子设备 10 和硬件卸载设备 20 通过总线 30 通信连接，硬件卸载设备 20 可与对象存储系统 40 进行有线或无线通信连接。

其中，电子设备 10 上可以运行一个或多个 App11，任一 App11 在有数据访问需求时，可以调用虚拟文件系统 12 提供的 POSIX 接口向用户空间文件系统 13 发送诸如写请求、读请求等数据访问请求。用户空间文件系统 13 通过其具有的内核模块接收虚拟文件系统 12 发送的数据访问请求，以及将数据访问请求通过总线发送给硬件卸载设备 20。

硬件卸载设备 20 通过总线接口 21 接收电子设备 10 通过内核模块发送的数据访问请求，以及将数据访问请求发送至主处理器 22，主处理器 22 将数据访问请求提供至客户端，客户端响应数据访问请求与对象存储系统 40 进行数据交互，例如，客户端向对象存储系统 40 写入对象文件，或者，读取对象存储系统 40 中的对象文件。

进一步可选的，内核模块还可以具有 POSIX 文件协议的数据访问请求转换成具有与用户空间文件系统 13 适配的 FUSE 协议的数据访问请求，将 FUSE 协议的数据访问请求通过

总线发送给硬件卸载设备 20。相应地,硬件卸载设备 20 上的总线接口 21 对接收到的 FUSE 协议的数据访问请求进行转换,得到 POSIX 文件协议的数据访问请求,并将 POSIX 文件协议的数据访问请求发送至主处理器 22。

其中,总线接口 21 可以是 PCIE (peripheral component interconnect express, 高速串行计算机扩展总线标准) 接口、SPI (serial peripheral interface, 串行外围设备接口) 或者 AXI (Advanced eXtensible Interface, 高级可扩展接口)。

其中,主处理器 22 例如包括但不限于: DSP (Digital Signal Processing, 数字信号处理器)、NPU (Neural-network Processing Unit, 嵌入式神经网络处理器)、CPU (central processing unit, 中央处理器) 和 GPU (Graphic Processing Unit, 图形处理器)。

10 在本实施例中,在对象文件写入阶段,电子设备用于通过应用程序向硬件卸载设备发送待写入文件;硬件卸载设备,用于获取待写入文件,并将待写入文件写入至硬件卸载设备上的缓存中;若缓存中已缓存的待写入文件满足文件合并条件,则将缓存中已缓存的待写入文件进行文件合并,得到待写入的第一对象文件;以及向对象存储系统发送包括第一对象文件的写请求;对象存储系统,用于响应写请求存储第一对象文件。关于数据访问系
15 统在对象文件写入阶段的交互过程更多介绍可以参见后文。

在本实施例中,在对象文件读取阶段,电子设备,用于通过应用程序向硬件卸载设备发送第一读请求,第一读请求包括待读取的目标数据所属目标文件的文件名和其在目标文件中的第一位置信息;硬件卸载设备接收应用程序发送的第一读请求;向对象存储系统发送第二读请求,第二读请求包括目标文件的文件名;接收对象存储系统返回的第二对象文件,并从第二对象文件中获取目标文件;根据第一位置信息从目标文件中读取目标数据,并将目标数据发送给应用程序;对象存储系统,用于响应第二读请求从已存储的对象文件中获取包括目标文件的第二对象文件。关于数据访问系统在对象文件读取的交互过程更多
20 介绍可以参见后文。

以下结合附图,详细说明本申请各实施例提供的技术方案。

25 图 2 为本申请实施例提供的一种数据访问方法的流程图。该方法应用于硬件卸载设备 20 上运行的客户端,硬件卸载设备 20 通过总线 30 与电子设备 10 通信连接,如图 2 所示,该方法可以包括以下步骤:

201、获取电子设备 10 上应用程序发送的待写入文件,并将待写入文件写入至硬件卸载设备 20 上的缓存中。

30 202、若缓存中已缓存的待写入文件满足文件合并条件,则将缓存中已缓存的待写入文件进行文件合并,得到待写入的第一对象文件。

203、将第一对象文件写入至客户端具有访问权限的对象存储系统 40 中。

具体而言，电子设备 10 上任一应用程序可以在文件写入需求的触发下，向硬件卸载设备 20 发送待写入文件，硬件卸载设备 20 上的客户端获取来源于应用程序的待写入文件，并将待写入文件写入至硬件卸载设备 20 上的缓存 24 中。

5 在一可选实现方式中，客户端获取电子设备 10 上应用程序发送的待写入文件时，具体用于：通过总线接口 21 接收应用程序发送的写请求，写请求是应用程序调用用户空间文件系统 13 提供的内核模块发送的；调用用户空间文件系统 13 提供的用户空间库模块处理写请求，以得到待写入文件。

其中，客户端利用用户空间库模块处理写请求，除了可以获取到来源于电子设备 10

10 上的应用程序的待写入文件，还可以确定与对象存储系统 40 进行交互的 SDK（Software Development Kit，软件开发工具包），但并不限于此。

客户端在将待写入文件写入至缓存 24 后，检测缓存 24 中已缓存的待写入文件是否满足文件合并条件。若满足文件合并条件，客户端将已缓存的若干个待写入文件进行文件合并，得到一个新的对象文件，为了便于理解和区分，将合并所得的新的对象文件称作第一

15 对象文件。客户端得到第一对象文件后，将第一对象文件写入至客户端具有访问权限的对象存储系统 40 中。若不满足文件合并条件，客户端暂不执行将缓存 24 中已缓存的待写入文件进行文件合并操作。其中，客户端可以调用与对象存储系统 40 进行交互的 SDK 将第一对象文件写入至对象存储系统 40 中。

本实施例对文件合并条件不做限制，文件合并条件例如包括但不限于：缓存 24 中的

20 剩余缓存空间小于预设缓存空间，或者，缓存 24 中的待写入文件的数量大于或等于预设文件数量，或者，缓存时长到达缓存周期。其中，缓存周期例如为 1 个小时、一天或者一个月。在缓存 24 已经缓存 1 个小时、一天或者一个月期间的待写入文件时，确定满足文件合并条件。

进一步可选的，在文件合并时，还可以生成与第一对象文件相关的描述文件，描述文

25 件例如可以记录被合并的文件的文件名、文件大小、被合并的文件在第一对象文件中的位置信息。其中，基于被合并的文件在第一对象文件中的位置信息可以从第一对象文件中读取被合并的文件的文件数据。参见图 2，对象文件中包括所合并的 n 个文件和一个描述文件。其中，n 为正整数。

为了便于理解，结合表 1 和海量小文件场景进行举例说明。

30

表 1

成本	现有方案	改进方案
存储费用	\$23	\$23
请求费用	\$1342	\$0.001

例如，对象存储系统 40 的存储费用的计费单位：\$0.023/GB/month，也即 1 GB 数据大小的数据每个月需要付费 0.023 美元。对象存储系统 40 的请求费用的计费单位：\$0.0005c/PUT，也即每个对象文件每次访问需要付费 0.0005 美分。由于每个对象文件按照访问次数计费，访问次数越多请求费用越高。

5 在现有方案中，每个小文件作为一个对象文件写入至对象存储系统 40 中。例如，1TB 的数据容量可以存储 268435456 个 4KB 数据大小的小文件，按照 100 次/秒的写入速度计算，产生的一个月费的存储费用大概为 $0.023 \times 1024 = \$23$ ；产生的请求费用为 $0.0005c \times 268435456 / 100 = \1342 。由此，可以发现请求费用远大于数据存储成本，大约占到总成本的 98%。

10 在改进方案中，多个小文件合并成一个对象文件写入至存储系统中。例如，1TB 的数据容量可以存储 205 个 5GB 小文件，产生的一个月费的存储费用大概为 $0.023 \times 1024 = \$23$ ；产生的请求费用为 $0.0005c \times 205 / 100 = \0.001 。

本申请实施例提供的数据访问方法，一方面，将针对对象存储系统 40 的数据访问任务从电子设备 10 转移至硬件卸载设备 20 上执行，可以减轻电子设备 10 的处理压力，提高电子设备 10 的处理性能，且能够增强数据访问性能。另一方面，硬件卸载设备 20 在每接收到来源于应用程序上的待写入文件后，并不直接将待写入文件写入至对象存储系统 40 中，而是首先在本地进行缓存，在缓存到多个待写入文件后，再将多个待写入文件合并成一个新的对象文件写入至客户端具有访问权限的对象存储系统 40。这样，能够极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本，特别是针对海量小文件场景，可以极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本。另外，极大地减少了因大量的数据访问请求触发对象存储系统出现流量限速的情形发生，增强了对对象存储系统的存储性能和访问性能。

25 在一些可选实施例中，客户端除了可以将新的对象文件写入至对象存储系统 40 进行存储，还可以将新的对象文件写入至硬件卸载设备 20 的存储器 23 中进行存储。在对象存储系统 40 和硬件卸载设备 20 的存储器 23 中均存储新的对象文件，除了可以增加对象文件的存储安全性，还可以优先在存储器 23 中访问对象文件，在存储器 23 当前未存储对象

文件的情况下，再到对象存储系统 40 中访问对象文件，能够进一步减少随着在对象存储系统 40 中访问对象文件所产生的请求费用。

于是，进一步可选的，参见图 3，可以在硬件卸载设备 20 上设置存储器 23，该存储器 23 例如包括但不限于：机械硬盘(Hard Disk Drive, HDD)和固态硬盘(Solid State Disk, SSD)等各种硬盘。基于上述，客户端除了将第一对象文件写入至对象存储系统 40 中，还可以将第一对象文件写入至存储器 23 中。

进一步可选的，客户端还可以将第一对象文件的索引信息写入至索引文件中，第一对象文件的索引信息包括第一对象文件的对象标识、存储状态和所合并的至少一个待写入文件的文件名，存储状态表征存储器 23 或者对象存储系统 40 中是否存储第一对象文件。进一步可选的，索引文件可以保存至硬件卸载设备 20 上的存储器 23 中。

值得注意的是，相比于可以存储海量数据的对象存储系统 40，硬件卸载设备 20 上设置的存储器 23 的数据容量较小。随着时间推移，在历史时间段内写入存储器 23 中的对象文件有的可能已经从存储器 23 中清除，有的可能仍然保留在存储器 23 中。

于是，客户端通过索引文件记录的第一对象文件的索引信息，可以准确获知存储器 23 或对象存储系统 40 是否保存了第一对象文件，并快速确定是对存储器 23 还是对象存储系统 40 进行数据访问，进而增加数据访问效率和减少对对象存储系统 40 的访问频次。

实际应用中，客户端可以直接将缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件。进一步可选的，为了减少数据传输所要消耗的资源，增强数据访问性能，还可以在硬件卸载设备 20 上设置具有数据压缩功能的压缩模块。于是，客户端将缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件时，具体用于：将已缓存的待写入文件发送至压缩模块，以使压缩模块对已缓存的待写入文件进行压缩处理得到第一对象文件；接收压缩模块返回的第一对象文件。

下面结合图 4，从数据读取角度介绍下本申请实施例提供的数据访问方法。图 4 为本申请实施例提供的一种数据访问方法的流程图。该方法应用于硬件卸载设备 20 上运行的客户端，硬件卸载设备 20 通过总线与电子设备 10 通信连接，如图 4 所示，该方法可以包括以下步骤：

401、接收应用程序发送的第一读请求，第一读请求包括待读取的目标数据所属目标文件的文件名和其在目标文件中的第一位置信息。

402、向对象存储系统 40 发送第二读请求，第二读请求包括目标文件的文件名，以供对象存储系统 40 从已存储的对象文件中获取包括目标文件的第二对象文件。

403、接收对象存储系统 40 返回的第二对象文件，并从第二对象文件中获取目标

文件。

404、根据第一位置信息从目标文件中读取目标数据，并将目标数据发送给应用程序。

在本实施例中，电子设备 10 上任一应用程序可以在文件读取需求的触发下，向硬件卸载设备 20 发送第一读请求，第一读请求包括待读取的目标数据所属目标文件的文件名和其在目标文件中的第一位置信息。其中，第一位置信息是目标数据在目标文件中的写入位置，根据第一位置信息可以从目标文件中读取目标数据。

客户端调用与对象存储系统 40 进行交互的 SDK 向对象存储系统 40 发送第二读请求。对象存储系统 40 响应第二读请求，根据第二读请求中目标文件的文件名，查询预先保存的元数据，确定包括目标文件的第二对象文件的对象名，并根据第二对象文件的对象名从已存储的对象文件中获取第二对象文件，并返回给客户端。

实际应用中，第二对象文件可能未经过数据压缩，针对这种情形，可以直接从第二对象文件对应的描述文件中获取目标文件在第二对象文件中的位置信息，为了便于理解和区分，将目标文件在第二对象文件中的位置信息称为第二位置信息。根据第二位置信息从第二对象文件中获取目标文件。第二对象文件可能经过数据压缩，针对这种情形，则需要对第二对象文件进行解压缩，并从解压缩后的第二对象文件中获取目标文件。在一可选实现方式中，可以在硬件卸载设备 20 上设置具有数据解压缩功能的解压缩模块，于是，客户端从第二对象文件中获取目标文件时，具体用于：将第二对象文件发送给解压缩模块，以供解压缩模块对第二对象文件进行解压缩处理得到解压缩后的第二对象文件；接收解压缩模块返回的解压缩后的第二对象文件，并根据目标文件的文件名查询解压缩后的第二对象文件对应的描述文件，得到目标文件在解压缩后的第二对象文件中的第二位置信息；根据第二位置信息，从解压缩后的第二对象文件中获取目标文件。

实际应用中，客户端在接收到应用程序发送的第一读请求后，可以直接将访问对象存储系统 40，以获取包括目标文件的第二对象文件。客户端还可以首先访问硬件卸载设备 20 上的存储器 23，在未从存储器 23 中获取到第二对象文件后，再访问对象存储系统 40。进一步可选的，为了增加数据访问效率和减少访问对象存储系统 40 的次数，客户端向对象存储系统 40 发送第二读请求之前，还可以根据目标文件的文件名查询索引文件，获取第二对象文件的存储状态；若第二对象文件的存储状态表征只有对象存储系统 40 中存储第二对象文件，则执行向对象存储系统 40 发送第二读请求的步骤。若第二对象文件的存储状态表征存储器 23 中存储第二对象文件，则从存储器 23

中获取第二对象文件。

值得注意的是，优先在存储器 23 中访问对象文件，在存储器 23 当前未存储对象文件的情况下，再到对象存储系统 40 中访问对象文件，能够进一步减少随着在对象存储系统 40 中访问对象文件所产生的请求费用。

5 本申请实施例提供的数据访问方法，一方面，将针对对象存储系统的数据访问任务从电子设备转移至硬件卸载设备上执行，可以减轻电子设备的处理压力，且能够增强数据访问性能。另一方面，在待读取数据是同一个对象文件中所包括的多个文件中的文件数据时，在需要读取同一个对象文件的多个文件的数据时，仅需对对象存储系统中同一个对象文件进行一次访问即可，无需多次访问对象存储系统中的多个对象文件。由此，能够极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本，特别是针对海量小文件场景，可以极大地减少针对对象存储系统的对象文件访问次数，减少针对对象存储系统的对象文件访问所产生的请求费用，降低数据访问成本。另外，极大地减少了因大量的数据访问请求触发对象存储系统出现流量限速的情形发生，增强了对对象存储系统的存储性能和访问性能。

15 值得注意的是，硬件卸载设备 20 可以在一个处理器上运行客户端、压缩模块和解压模块，也可以在多个处理器上运行客户端、压缩模块和解压模块。参见图 3 所示，硬件卸载设备 20 设置了一个主处理器 22 和协处理器 25。其中，主处理器 22 和协处理器 25 例如包括但不限于：DSP (Digital Signal Processing, 数字信号处理器)、NPU (Neural-network Processing Unit, 嵌入式神经网络处理器)、CPU (central processing unit, 中央处理器) 和 GPU (Graphic Processing Unit, 图形处理器)。主处理器 22 上运行客户端，协处理器 25 上提供压缩模块和解压模块，在主处理器 22 和协处理器 25 相互配合工作，能够提高整个数据访问性能。

需要说明的是，上述实施例所提供方法的各步骤的执行主体均可以是同一设备，或者，该方法也由不同设备作为执行主体。比如，步骤 201 至步骤 203 的执行主体可以为设备 A；又比如，步骤 201 和 202 的执行主体可以为设备 A，步骤 203 的执行主体可以为设备 B；等等。

另外，在上述实施例及附图中的描述的一些流程中，包含了按照特定顺序出现的多个操作，但是应该清楚了解，这些操作可以不按照其在本文中出现的顺序来执行或并行执行，操作的序号如 201、202 等，仅仅是用于区分开各个不同的操作，序号本身不代表任何的执行顺序。另外，这些流程可以包括更多或更少的操作，并且这些操作可以按顺序执行或并行执行。需要说明的是，本文中的“第一”、“第二”等描述，是用于区分不同的消息、设

备、模块等，不代表先后顺序，也不限定“第一”和“第二”是不同的类型。

图 5 为本申请实施例提供的一种硬件卸载设备 20 的结构示意图。该硬件卸载设备 20 通过总线与电子设备 10 通信连接，如图 5 所示，该硬件卸载设备 20 包括主处理器 22 和缓存，主处理器 22 运行客户端程序，以用于：获取电子设备 10 上应用程序发送的待写入文件，并将待写入文件写入至硬件卸载设备 20 上的缓存 24 中；若缓存 24 中已缓存的待写入文件满足文件合并条件，则将缓存 24 中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；将第一对象文件写入至客户端具有访问权限的对象存储系统 40 中。

进一步可选的，硬件卸载设备 20 还包括：存储器 23；主处理器 22 还用于：将第一对象文件写入至存储器 23 中，以及将第一对象文件的索引信息写入至索引文件中，第一对象文件的索引信息包括第一对象文件的对象标识、存储状态和所合并的至少一个待写入文件的文件名，存储状态表征存储器 23 或者对象存储系统 40 中是否存储第一对象文件。

进一步可选的，硬件卸载设备 20 还包括：协处理器 25；主处理器 22 进行文件合并时，具体用于：将已缓存的待写入文件发送至协处理器 25，并接收协处理器 25 发送的第一对象文件；协处理器 25，用于对已缓存的待写入文件进行压缩处理得到第一对象文件，并将第一对象文件发送给主处理器 22。

进一步可选的，主处理器 22 还用于：接收应用程序发送的第一读请求，第一读请求包括待读取的目标数据所属目标文件的文件名和其在目标文件中的第一位置信息；向对象存储系统 40 发送第二读请求，第二读请求包括目标文件的文件名，以供对象存储系统 40 从已存储的对象文件中获取包括目标文件的第二对象文件；接收对象存储系统 40 返回的第二对象文件，并从第二对象文件中获取目标文件；根据第一位置信息从目标文件中读取目标数据，并将目标数据发送给应用程序。

进一步可选的，主处理器 22 获取目标文件时，具体用于：将第二对象文件发送给协处理器 25，并接收协处理器 25 返回的目标文件，并根据目标文件的文件名查询解压缩后的第二对象文件对应的描述文件，得到目标文件在解压缩后的第二对象文件中的第二位置信息；根据第二位置信息，从解压缩后的第二对象文件中获取目标文件；协处理器 25，还用于对第二对象文件进行解压缩处理得到解压缩后的第二对象文件，并将解压缩后的第二对象文件返回至主处理器 22。

图 5 所示的硬件卸载设备 20 其中各个模块、单元执行操作的具体方式已经在有关数据访问方法的实施例中进行了详细描述，此处将不做详细阐述说明。

图 6 为本申请实施例提供的一种电子设备的结构示意图。参见图 6，该电子设备

包括：处理器 62 和前述实施例任一项所述的硬件卸载设备 61，处理器 62 上运行有至少一个应用程序，处理器 62 通过总线与硬件卸载设备通信连接。值得注意的是，处理器 62 上的应用程序可与硬件卸载设备进行交互，以实现访问对象存储系统。硬件卸载设备 61 接收来源于处理器 62 上的应用程序发送的数据访问请求，并执行本申请实施例提供的

5 例提供的数据访问方法响应数据访问请求。

其中，处理器 62 例如包括但不限于：DSP（Digital Signal Processing，数字信号处理器）、NPU（Neural-network Processing Unit，嵌入式神经网络处理器）、CPU（central processing unit，中央处理器）和 GPU（Graphic Processing Unit，图形处理器）。

进一步，如图 6 所示，该电子设备还包括：通信组件 63、显示器 64、电源组件 10 65、音频组件 66 等其它组件。图 6 中仅示意性给出部分组件，并不意味着电子设备只包括图 6 所示组件。另外，图 6 中虚线框内的组件为可选组件，而非必选组件，具体可视电子设备的

产品形态而定。本实施例的电子设备可以实现为台式电脑、笔记本电脑、智能手机或 IOT 设备等终端设备，也可以是常规服务器、云服务器或服务器阵列等服务端设备。若本实施例的电子设备实现为台式电脑、笔记本电脑、智能手机等终端设备，可以包含图 6 中虚线框内的组件；若本实施例的电子设备实现为常规服务器、云服务器或服务器阵列等服务端设备，则可以不包含图 6 中虚线框内的组件。

15

相应地，本申请实施例还提供一种存储有计算机程序的计算机可读存储介质，计算机程序被执行时能够实现上述数据访问方法中的各步骤。

相应地，本申请实施例还提供一种计算机程序产品，包括计算机程序/指令，当计算机程序/指令被处理器执行时，致使处理器能够实现上述数据访问方法中的各步骤。

20

上述通信组件被配置为便于通信组件所在设备和其他设备之间有线或无线方式的通信。通信组件所在设备可以接入基于通信标准的无线网络，如 WiFi，2G、3G、4G/LTE、5G 等移动通信网络，或它们的组合。在一个示例性实施例中，通信组件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，通信组件还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

25

上述显示器包括屏幕，其屏幕可以包括液晶显示器（LCD）和触摸面板（TP）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与触摸或滑动操作相关的持续时间和压力。

30

上述电源组件，为电源组件所在设备的各种组件提供电力。电源组件可以包括电源管理系统，一个或多个电源，及其他与为电源组件所在设备生成、管理和分配电力相关联的组件。

上述音频组件，可被配置为输出和/或输入音频信号。例如，音频组件包括一个麦克风
5 (MIC)，当音频组件所在设备处于操作模式，如呼叫模式、记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器 23 或经由通信组件发送。在一些实施例中，音频组件还包括一个扬声器，用于输出音频信号。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实
10 施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流
15 程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理器或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

20 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机
25 或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，电子设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

30 内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

- 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被电子设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。
- 10 还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括要素的过程、方法、商品或者设备中还存在另外的相同要素。
- 15 以上仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的权利要求范围之内。

权 利 要 求 书

1、一种数据访问方法，其特征在于，应用于硬件卸载设备上运行的客户端，所述硬件卸载设备通过总线与电子设备通信连接，所述方法包括：

5 获取所述电子设备上应用程序发送的待写入文件，并将所述待写入文件写入至所述硬件卸载设备上的缓存中；

若所述缓存中已缓存的待写入文件满足文件合并条件，则将所述缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；

将所述第一对象文件写入至所述客户端具有访问权限的对象存储系统中。

2、根据权利要求1所述的方法，其特征在于，所述硬件卸载设备还包括：存储器，
10 所述方法还包括：

将所述第一对象文件写入至所述存储器中，以及将所述第一对象文件的索引信息写入至索引文件中，所述第一对象文件的索引信息包括所述第一对象文件的对象标识、存储状态和所合并的至少一个待写入文件的文件名，所述存储状态表征所述存储器或者所述对象存储系统中是否存储所述第一对象文件。

3、根据权利要求1所述的方法，其特征在于，所述硬件卸载设备还包括：压缩模块，所述将所述缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件，包括：

将所述已缓存的待写入文件发送至所述压缩模块，以使所述压缩模块对所述已缓存的待写入文件进行压缩处理得到所述第一对象文件；

20 接收所述压缩模块返回的所述第一对象文件。

4、根据权利要求1所述的方法，其特征在于，获取所述电子设备上应用程序发送的待写入文件包括：

通过总线接口接收所述应用程序发送的写请求，所述写请求是所述应用程序调用用户空间文件系统提供的内核模块发送的；

25 调用所述用户空间文件系统提供的用户空间库模块处理所述写请求，以得到所述待写入文件。

5、根据权利要求2所述的方法，其特征在于，还包括：

接收所述应用程序发送的第一读请求，所述第一读请求包括待读取的目标数据所属目标文件的文件名和其在所述目标文件中的第一位置信息；

30 向所述对象存储系统发送第二读请求，所述第二读请求包括所述目标文件的文件

名，以供所述对象存储系统从已存储的对象文件中获取包括所述目标文件的第二对象文件；

接收所述对象存储系统返回的所述第二对象文件，并从所述第二对象文件中获取所述目标文件；

5 根据所述第一位置信息从所述目标文件中读取所述目标数据，并将所述目标数据发送给所述应用程序。

6、根据权利要求5所述的方法，其特征在于，向所述对象存储系统发送第二读请求之前，还包括：

10 根据所述目标文件的文件名查询所述索引文件，获取所述第二对象文件的存储状态；

若所述第二对象文件的存储状态表征只有所述对象存储系统中存储所述第二对象文件，则执行向所述对象存储系统发送第二读请求的步骤。

7、根据权利要求6所述的方法，其特征在于，还包括：

15 若所述第二对象文件的存储状态表征所述存储器中存储所述第二对象文件，则从所述存储器中获取所述第二对象文件。

8、根据权利要求5所述的方法，其特征在于，所述硬件卸载设备还包括：解压缩模块，所述从所述第二对象文件中获取所述目标文件，包括：

将所述第二对象文件发送给所述解压缩模块，以供所述解压缩模块对所述第二对象文件进行解压缩处理得到解压缩后的第二对象文件；

20 接收所述解压缩模块返回的所述解压缩后的第二对象文件，并根据所述目标文件的文件名查询所述解压缩后的第二对象文件对应的描述文件，得到所述目标文件在所述解压缩后的第二对象文件中的第二位置信息；

根据所述第二位置信息，从所述解压缩后的第二对象文件中获取所述目标文件。

25 9、一种硬件卸载设备，其特征在于，所述硬件卸载设备通过总线与电子设备通信连接，所述硬件卸载设备包括主处理器和缓存，所述主处理器运行客户端程序，以用于：

获取所述电子设备上应用程序发送的待写入文件，并将所述待写入文件写入至所述硬件卸载设备上的缓存中；

30 若所述缓存中已缓存的待写入文件满足文件合并条件，则将所述缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；

将所述第一对象文件写入至所述客户端具有访问权限的对象存储系统中。

10、根据权利要求9所述的硬件卸载设备，其特征在于，还包括：存储器；

所述主处理器还用于：将所述第一对象文件写入至所述存储器中，以及将所述第一对象文件的索引信息写入至索引文件中，所述第一对象文件的索引信息包括所述第一对象文件的对象标识、存储状态和所合并的至少一个待写入文件的文件名，所述存储状态表征所述存储器或者所述对象存储系统中是否存储所述第一对象文件。

11、根据权利要求9所述的硬件卸载设备，其特征在于，所述硬件卸载设备还包括：协处理器；

所述主处理器进行文件合并时，具体用于：将所述已缓存的待写入文件发送至所述协处理器，并接收所述协处理器发送的所述第一对象文件；

10 所述协处理器，用于对所述已缓存的待写入文件进行压缩处理得到所述第一对象文件，并将所述第一对象文件发送给所述主处理器。

12、一种电子设备，其特征在于，包括：处理器和如权利要求9至11任一项所述的硬件卸载设备，所述处理器上运行有至少一个应用程序，所述处理器通过总线与所述硬件卸载设备通信连接。

15 13、一种数据访问系统，其特征在于，包括：电子设备、硬件卸载设备和对象存储系统；所述电子设备与所述硬件卸载设备通过总线通信连接，所述硬件卸载设备与所述对象存储系统通信连接；

所述电子设备上运行有至少一个应用程序，用于通过所述应用程序向所述硬件卸载设备发送待写入文件；

20 所述硬件卸载设备，用于获取所述待写入文件，并将所述待写入文件写入至所述硬件卸载设备上的缓存中；若所述缓存中已缓存的待写入文件满足文件合并条件，则将所述缓存中已缓存的待写入文件进行文件合并，得到待写入的第一对象文件；以及向所述对象存储系统发送包括所述第一对象文件的写请求；

所述对象存储系统，用于响应所述写请求存储所述第一对象文件。

25 14、一种存储有计算机程序的计算机存储介质，其特征在于，当所述计算机程序被处理器执行时，致使所述处理器能够实现权利要求1-8任一项所述方法中的步骤。