

存储系统及数据处理方法

5 本申请要求于 2022 年 08 月 24 日提交中国专利局、申请号为 202211019838.6、申请名称为“存储系统及数据处理方法”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

10 本说明书实施例涉及信息存储技术领域，特别涉及一种存储系统，数据处理方法。本说明书一个或者多个实施例同时涉及数据处理装置，一种计算设备，以及一种计算机可读存储介质。

背景技术

15 随着互联网技术的不断发展，为了实现可靠的海量数据存储，目前多采用分布式存储系统。分布式存储系统广泛采用三副本机制，例如键值对(key-value)分布式存储系统，从而在不同位置分别对数据进行存储以提高数据存储的可靠性。当前的分布式存储系统包括三层架构，底层的分布式文件系统、中层的 KV 存储引擎、上层的计算引擎层，在这种分布式存储系统的结构下，为了保证数据的可靠性，数据的删除、修改等都采用 Append-Only（仅追加）方式进行，于是在多层架构的分布式存储系统中，中间 KV 类存储引擎层和底部单机存储引擎层均采用上述方式处理数据。在后续进行内存释放时，需要每一层各自进行垃圾回收处理，导致带来更大的写放大，垃圾数据增加，从而影响整体系统性能。

20 因此，如何降低分布式存储系统进行垃圾回收处理过程中产生的写放大、减少垃圾数据、降低存储成本、提高系统性能是目前亟需解决的问题。

发明内容

25 有鉴于此，本说明书实施例提供了一种存储系统，该存储系统提供 Trim 能力，能够降低垃圾回收产生的写放大，提升整体系统的性能，减少垃圾数据，降低存储成本。本说明书一个或者多个实施例同时涉及数据处理方法，数据处理装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序，以解决现有技术中存在的技术缺陷。

30 根据本说明书实施例的第一方面，提供了一种存储系统，所述系统包括客户端和数据节点；所述客户端，用于接收针对数据文件的标记操作指令，解析所述标记操作指令，获得标记操作信息，根据所述标记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点；所述目标数据节点，用于接收所述数据操作请求，根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。

35 根据本说明书实施例的第二方面，提供了一种数据处理方法，应用于存储系统中的客户端，包括：

- 接收针对数据文件的标记操作指令；
- 解析所述标记操作指令，获得标记操作信息；
- 根据所述标记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点。
- 5 根据本说明书实施例的第三方面，提供了一种数据处理方法，应用于存储系统中的目标数据节点，包括：
- 接收针对数据文件数据操作请求；
- 根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。
- 10 根据本说明书实施例的第四方面，提供了一种数据处理装置，应用于存储系统中的客户端，包括：
- 接收模块，被配置为接收针对数据文件的标记操作指令；
- 解析模块，被配置为解析所述标记操作指令，获得标记操作信息；
- 发送模块，被配置为根据所述标记操作信息生成对应目标数据节点的数据操作请求，
- 15 并发送所述数据操作请求至所述目标数据节点。
- 根据本说明书实施例的第五方面，提供了一种数据处理装置，应用于存储系统中的目标数据节点，包括：
- 接收模块，被配置为接收针对数据文件数据操作请求；
- 标记模块，被配置为根据所述数据操作请求，对本地存储空间中关联所述数据文件的
- 20 待操作数据区间进行标记处理。
- 根据本说明书实施例的第六方面，提供了一种计算设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机指令，所述处理器执行所述计算机指令时实现所述数据处理方法的步骤。
- 根据本说明书实施例的第七方面，提供了一种计算机可读存储介质，其存储有计算机
- 25 指令，该计算机指令被处理器执行时实现所述数据处理方法的步骤。
- 根据本说明书实施例的第八方面，提供了一种计算机程序，其中，当所述计算机程序在计算机中执行时，令计算机执行上述数据处理方法的步骤。
- 本说明书提供的存储系统，所述系统包括客户端和数据节点；所述客户端，用于接收针对数据文件的标记操作指令，解析所述标记操作指令，获得标记操作信息，根据所述标
- 30 记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点；所述目标数据节点，用于接收所述数据操作请求，根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。
- 本说明书一实施例实现了通过客户端将针对数据文件的标记操作指令转化成针对数据节点的数据区间的数据操作请求，对数据节点中存储的不需要的数据进行标记，使得后
- 35 续在进行垃圾回收时，无需对无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据，也可以使得后续在新数据写入数据节点时，能够直接写入待操作数据区间，提高写入效率，优化存储数据节点，降低存储成本。

附图说明

图 1 是本说明书一个实施例提供的一种存储系统执行标记操作的场景示意图；

图 2a 是本说明书一个实施例提供的一种存储系统的架构示意图；

图 2b 是本说明书一个实施例提供的一种存储系统的结构示意图；

5 图 3 是本说明书一个实施例提供的一种应用于存储系统中的客户端的数据处理方法的流程图；

图 4 是本说明书一个实施例提供的一种应用于存储系统中的目标数据节点的数据处理方法的流程图；

10 图 5 是本说明书一个实施例提供的一种存储系统中客户端与目标数据节点的交互示意图；

图 6 是本说明书一实施例提供的一种应用于存储系统中的客户端的数据处理装置的结构示意图；

图 7 是本说明书一实施例提供的一种应用于存储系统中的目标数据节点的数据处理装置的结构示意图；

15 图 8 是本说明书一个实施例提供的一种计算设备的结构框图。

具体实施方式

在下面的描述中阐述了很多具体细节以便于充分理解本说明书。但是本说明书能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本说明书内涵的情况下做类似推广，因此本说明书不受下面公开的具体实施的限制。

20 在本说明书一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本说明书一个或多个实施例中使用的术语“和/或”是指包含一个或多个相关联的列出项目的任何或所有可能组合。

25 应当理解，尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本说明书一个或多个实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

30 首先，对本说明书一个或多个实施例涉及的名词术语进行解释。

Append-only：仅支持追加写，不支持随机修改的写入模式。

文件：由文件名或文件 Id 唯一标识的一段连续存储数据。每个文件由一个或多个 Chunk 组成。

35 Chunk：Chunk 是存储文件中的一段连续区间的数据，每个 Chunk 可以为多副本 Chunk，或者纠删码 Chunk。

Replica：Replica 是 Chunk 存储到单机存储引擎的一段连续数据的数据块。多副本 Chunk，每个 Replica 相同；纠删码 Chunk，每个 Replica 存储满足纠删码切片和排布规则的数据分

片，每个分片的数据不同。

KV 存储引擎：指在存储系统之上构建的以 Key-Value 方式存储数据，提供以 Key 方式访问数据的引擎。

5 ZNS-SSD: (Zoned Namespace, 分区命名空间) - (Solid State Disk, 固态硬盘) 是在 SSD 基础上发展而来的，实现了把 FTL(Flash Translation Layer, 闪存转换层)从 SSD 内部迁移到上层的 Host 端，把 SSD 内部开放给了 Host 端，实现用户可以根据自己需要比较灵活的拥有自己特定的 FTL。

10 目前，在对象存储，文件存储，表格存储，块存储等各类分布式存储系统，以及分布式分析型数据库，分布式事务型数据库以及分析型事务型混合数据库等数据库系统中，广泛采用一种三层架构。底层采用 Append-Only 的分布式文件系统做高可靠，高可用的持久化存储，中间层采用各类 KV 存储作为行存引擎，或列存引擎，或行列混合引擎，上层是前端接入层或者计算引擎层。而在该类架构中，数据的删除、修改及更新，都采用以 Append-Only 的方式写 Log 来实现，以充分利用存储介质的访问特性，以及提升写入吞吐能力，然后在后台 GC (Garbage Collection, 垃圾回收) 流程，做到被覆盖或已删除数据的物理空间释放。因此，在该类架构中，KV 层的 GC 优化，是该技术架构的关键技术，也是相关技术人员持续优化和突破的方向。

20 而随着底层分布式文件系统越来越多的支持 ZNS SSD/SMR HDD 等仅支持 Append-only 的存储介质，分布式文件系统的单机存储层也采用类似 KV 的存储结构，数据的删除，修改和更新采用 Append-Only 写 Log 的方式来实现，依赖后台 GC 真正做到被覆盖或已删除数据的物理空间释放；因此，中间 KV 类存储引擎层及底层分布式文件系统层各自独立做 GC 的方式，缺乏协同，会带来更大的写放大，垃圾数据增加，从而影响整体系统性能、成本，降低系统的竞争力。

25 基于此在本说明书中，提供了一种存储系统，该存储系统提供 Trim 能力，能够降低垃圾回收产生的写放大，提升整体系统的性能，减少垃圾数据，降低存储成本。提供了两种数据处理方法，本说明书同时涉及两种数据处理装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序，在下面的实施例中逐一进行详细说明。

30 图 1 示出了根据本说明书一个实施例提供的一种存储系统执行标记操作的场景示意图；其中，存储系统中的客户端在接收到针对数据文件的标记操作指令之后，会将标记操作指令中的携带的标记操作信息转换成逻辑地址信息，逻辑地址信息即为数据文件在存储节点中数据块上的具体位置信息，后续数据节点可以根据逻辑地址信息进一步映射出数据文件对应的物理地址信息，即数据在磁盘上的具体位置。

35 参见图 1，在客户端接收到针对数据文件的标记操作指令，即针对数据文件进行 Trim 的指令之后，标记操作指令中携带有标记操作信息，标记操作信息包括有文件唯一标识 fileid，文件内偏移 Offset，Trim 的长度 Length 等关键信息，客户端将会根据标记操作指令转换成目标数据节点中的数据块级别的数据操作请求，并发送给目标数据节点，由目标数据节点根据数据操作请求中携带的操作描述信息进行实际的 Trim 操作，从而完成对数据文件的标记处理，在后续进行垃圾回收处理时，数据节点可以直接对已经标记过的数据进行清理，降低垃圾回收产生的写放大，提高处理效率。

图 2a 示出了根据本说明书一个实施例提供的一种存储系统的架构示意图, 存储系统的架构示意图中包括有元数据服务, 客户端, 单机存储引擎, 其中, 元数据是用来描述文件系统和文件特征的系统数据, 例如文件类型、文件大小、访问权限和数据索引信息等, 用户在访问文件数据之前, 需要访问文件的元数据以获得文件的基本属性信息和数据的索引信息。客户端负责发送读写请求至单机存储引擎, 如有三副本的数据文件需要写入, 则先发送至客户端, 由客户端转发至单机存储引擎; 单机存储引擎, 又称数据节点或数据服务器, 用于存放文件数据, 保证数据的可用性和完整性。存储系统中将数据分散存储到多个单机存储引擎上, 使得性能和容量能够同时拓展, 系统规模具有很强的伸缩性。

图 2b 示出了根据本说明书一个实施例提供的一种存储系统的结构示意图, 所述存储系统 200 包括客户端 210 和数据节点 220, 以下对存储系统执行 trim 的标记操作进行进一步说明。

所述客户端 210, 用于接收针对数据文件的标记操作指令, 解析所述标记操作指令, 获得标记操作信息, 根据所述标记操作信息生成对应目标数据节点的数据操作请求, 并发送所述数据操作请求至所述目标数据节点。

所述目标数据节点 220, 用于接收所述数据操作请求, 根据所述数据操作请求, 对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。

其中, 数据文件可以理解为项目不再需要的数据文件, 当中间 KV 类存储引擎层需要对项目中的无效数据进行删除时, 可以调用存储系统中提供的 trim 接口, 用于对存储系统中数据节点上所存储的无效数据进行标记, 在后续进行垃圾回收时, 能够根据标记信息直接对无效数据进行覆盖写入, 或者将已标记的无效数据进行删除, 便于新数据写入, 提高写入效率。标记操作指令可以理解为调用存储系统的 trim 接口, 执行 trim 操作的指令。具体地, 所述客户端 210 包括标记操作接口, 所述客户端 210 用于, 调用所述标记操作接口解析所述标记操作指令, 获得所述标记操作信息, 其中, 所述标记操作信息包括文件标识信息、文件偏移量信息、文件标记长度信息。

在实际应用中, 存储系统可以是分布式文件存储系统, 在分布式文件存储系统中的客户端接收到标记操作指令之后, 则调用标记操作接口用于解析标记操作指令, 标记操作接口即为提供 trim 能力的接口, 基于 trim 接口执行后续 trim 操作。调用 trim 接口解析标记操作指令中携带的标记操作信息, 标记操作信息包括有文件标识信息即文件唯一标识 fileId、文件偏移量信息即文件内偏移 Offset、文件标记长度信息 trim 的长度 Length 等关键信息, 标记操作信息可以理解为数据文件的逻辑地址信息, 为了使数据节点能够获取到数据文件在本次存储空间中的具体位置, 还需要根据标记操作信息转化成目标数据节点可处理的数据操作请求, 目标数据节点根据数据操作请求中携带的逻辑地址信息可以映射出数据文件所在磁盘的物理地址信息, 当数据文件关联的数据在该数据节点中时, 该数据节点即为本次 trim 所要进行操作的目标数据节点, 每个目标数据节点对应各自的数据操作请求, 每个目标数据节点对应的数据操作请求中包含有数据文件的数据在目标数据节点中的偏移量 Offset 和数据长度 Length, 根据数据操作请求即可以确定出需要进行 trim 操作的待操作数据区间, 待操作数据区间可以理解为需要进行数据删除操作的待删除数据区间, 则该数据区间中被标记过的数据可以被删除。

其中, 在所有数据节点中, 本次需要进行 trim 操作的为目标数据节点 220, 目标数据节点会接收到对应的数据操作请求, 数据节点即为存储系统中的单机存储引擎, 数据操作

请求中包括单机存储引擎管理的 Chunk 级的 ChunkId、ChunkOffset、Length，目标数据节点根据数据操作请求进一步映射出数据文件所在物理硬盘上的物理地址信息，并进行实际的 trim 操作，即对本地存储空间中管理数据文件的待操作数据区间进行标记处理。

在实际应用中，中间 KV 类引擎层及底层分布式文件系统各自独立做垃圾回收的方式，
5 缺乏协同处理，会带来更大的写放大，增加垃圾数据。本说明书实施例中，当中间 KV 类引擎层对于无效数据进行 trim 操作时，会调用底层分布式文件系统并行进行 trim 操作，使得分布式文件系统的数据节点进行 trim 操作，后续在垃圾回收时，可以直接对已进行 trim 操作的数据进行覆盖写入，提高数据释放效率。

在本说明书一实施例中，存储系统中包括有客户端和 10 个数据节点，客户端接收到
10 针对数据文件的标记操作指令，数据文件为非活跃用户 A 的用户数据文件，标记操作指令即为对非活跃用户的用户数据文件进行标记操作的指令，即 trim 指令。客户端解析该标记操作指令，获得标记操作信息，标记操作信息为用户数据文件的文件标识信息“A”，文件内偏移 Offset，trim 的长度 Length，客户端根据标记操作信息确定用户 A 的用户数据文件分别存储在数据节点 1 和数据节点 2 中，则生成对目标数据节点 1 的数据操作请求 1，和
15 目标数据节点 2 的数据操作请求 2，数据操作请求 1 中包括部分数据文件在目标数据节点 1 的数据块中的具体位置，数据操作请求 2 中包括部分数据文件在目标数据节点 2 的数据块中的具体位置，将每个目标数据节点对应的数据操作请求分别发送至对应的目标数据节点，即将数据操作请求 1 发送至目标数据节点 1，将数据操作请求 2 发送至目标数据节点 2。

20 通过客户端将数据文件的文件级别的标记操作信息转化成数据节点的 Chunk 级别的数据操作请求，使得后续目标数据节点可以根据数据操作请求确定出数据文件对应的待操作数据区间，从而进行标记处理。

在本说明书一实施例中，沿用上例，目标数据节点 1 接收到数据操作请求 1，数据操作请求中携带有逻辑地址信息，根据逻辑地址信息确定物理地址信息为偏移
25 Offset=1MB+1KB，长度 Length=1MB+1KB，则在目标数据节点 1 中确定待操作数据区间为<1MB, 1MB+1KB>，目标数据节点 1 对该待操作数据区间进行标记处理。目标数据节点 2 接收到数据操作请求 2，数据操作请求中携带有逻辑地址信息，根据逻辑地址信息确定物理地址信息为偏移 Offset=2MB+1KB，长度 Length=2MB+1KB，则在目标数据节点 2 中确定待操作数据区间为<2MB, 2MB+1KB>，目标数据节点 2 对该待操作数据区间进行
30 标记处理。

通过本说明书实施例提供的一种提供 trim 能力的存储系统，中间 KV 存储引擎可利用存储系统的 Trim 能力，消除以文件重写的方式进行 GC 流量，大幅降低 GC 产生的写放大，提升整体系统的性能，减少垃圾数据，降低数据存储成本，从而提升系统竞争力。

具体地，客户端根据标记操作信息生成对应目标数据节点的数据操作请求，可以先根据
35 标记操作指令确定出逻辑地址信息，再生成目标数据节点对应的数据操作请求。所述客户端，用于根据所述标记操作信息确定所述数据文件对应的逻辑地址信息，基于所述逻辑地址信息生成对应目标数据节点的数据操作请求。

在实际应用中，由于标记操作信息是相对于数据文件而言，而数据文件又被分别存储在不同的数据节点中的不同数据块中，因此当需要对数据文件进行 trim 操作时，需要先
40 确定出数据文件在数据块中的逻辑地址，然后基于逻辑地址找到数据块对应多个 Replica 的

数据节点，举例说明，标记操作指令中包括标记操作信息，客户端根据用户输入的标记操作信息：文件 Id、fileOffset、Length 确定出对应的逻辑地址信息为在目标数据节点 1 中的 ChunkId、ChunkOffset=1MB、Length=1MB，而在目标数据节点 1 中包括多个数据块，即确定出多个数据块所在的数据节点为目标数据节点，每个数据块的数据区间不同，为<0, 0>，<1MB, 1MB+1KB>，<2MB, 2MB+1KB>，从而客户端可以生成目标数据节点 1 对应的数据操作请求，后续数据节点 1 可以根据数据操作请求中携带的逻辑地址信息进一步确定出数据文件在硬盘中的物理地址信息，并进行实际的 trim 操作。

具体实施时，当目标数据节点 1 接收到数据操作请求之后，可以根据数据操作请求中携带的逻辑地址信息确定操作描述信息即物理地址信息，并确定出待操作数据区间，从而进行后续的 trim 操作。具体地，所述目标数据节点，用于根据所述数据操作请求中携带的逻辑地址信息确定操作描述信息，根据所述操作描述信息在本地存储空间中确定所述数据文件关联的待操作数据区间，修改并保存所述待操作数据区间的属性信息，作为对所述待操作数据区间的标记处理。

其中，操作描述信息可以理解为数据文件在目标数据节点中对应的物理地址信息，引用上例，逻辑地址信息为 ChunkId、ChunkOffset=1MB、Length=1MB，进一步映射出物理地址信息，Offset=1MB+1KB，Length=1MB+1KB，从而在目标数据节点 1 的多个数据块中确定数据块<1MB, 1MB+1KB>对应的区间为待操作数据区间，并对该数据区间进行标记处理。便于后续数据节点进行 GC 流程时，直接对已经标记处理过的数据区间进行存储空间释放处理。

在实际应用中，每个数据区间包括有对应的区间头数据，区间头数据中包括有对应的数据区间的属性信息，其中就有对应的标记信息，对数据区间进行标记处理即可以理解为修改数据区间的属性信息，将属性信息中的标记信息修改成已标记并保存，从而后续可以在识别到属性信息中的标记信息为已标记的情况下，判断该数据区间为已经进行过 trim 操作的数据区间，从而可以覆盖写入该数据区间中的数据。

在数据节点执行 trim 操作时，也存在有操作失败的情况，在该情况下客户端可以根据存储状态信息继续执行对应的操作处理，具体地，所述客户端，用于在确定所述数据操作请求对应的操作反馈信息为操作失败的情况下，查询所述标记操作指令的存储状态信息，并执行所述存储状态信息对应的操作处理。

其中，操作反馈信息可以理解为执行 trim 操作之后，客户端接收到的针对该 trim 操作的反馈信息，根据该操作反馈信息可以获知到本次标记处理操作是否执行成功，当操作反馈信息操作失败时，客户端根据存储状态信息执行存储状态信息对应的操作处理。在实际应用中，只有关联数据文件的所有数据块都成功执行标记处理操作之后，才算是本次 trim 操作成功。

具体实施时，存储状态信息分为两种，分别是持久化存储状态和非持久化存储状态，在持久化存储状态下，客户端可以 trim 操作进行持久化，即将本次 trim 操作对应的数据操作请求缓存下来，重复执行发送数据操作请求至目标数据节点的步骤，直至本次 trim 操作成功执行。在存储状态信息为非持久化存储状态下，当本次 trim 操作失败时，客户端会反馈失败信息至中间 KV 类存储引擎层，由中间 KV 类存储引擎层对本次 trim 操作的失败进行记录，从而让中间 KV 类存储引擎层之后重新针对本次 trim 操作发送标记处理指令至客

户端，直至 trim 操作执行成功。具体地，所述客户端，用于在所述存储状态信息为持久化存储状态的情况下，基于持久化的数据操作请求，执行发送所述数据操作请求至所述目标数据节点的步骤；

5 在所述存储状态信息为非持久化存储状态的情况下，生成所述标记操作指令对应的处理结果信息，并发送至所述存储系统对应的索引系统，其中，所述索引系统用于重新发送所述标记操作指令至所述存储系统。

在实际应用中，在存储状态为持久化存储状态的情况下，存储系统的客户端可对数据操作请求进行持久化，即保存数据操作请求至本地，以方便进行 failover 处理。例如三副本情况下，部分数据节点执行 trim 操作成功，但客户端宕机；或部分数据节点宕机，导致无法立即执行 trim 操作。对数据操作请求持久化，在客户端重启时，可以重新发送数据操作请求至目标数据节点，由目标数据节点进行对应的 trim 操作；或，在本次宕机的数据节点重启时，客户端再发送对应的数据操作请求至该数据节点，使得该数据节点进行对应的 trim 操作，从而实现 trim 操作的正常执行。

15 在本说明书一实施例中，客户端在接收到标记操作指令之后，根据标记操作指令生成对应目标数据节点的数据操作请求，并将数据操作请求缓存下来，在客户端宕机重启之后，客户端从缓存数据中读取未执行的数据操作请求，并将该数据操作请求发送至对应的目标数据节点，使得数据节点能够正常执行对应的数据操作。

20 在本说明书另一实施例中，客户端在发送数据操作请求至目标数据节点时，由于目标数据节点宕机无法接收数据操作请求，从而无法执行对应的数据标记处理，客户端会在该目标数据节点重启时，再次发送对应的数据操作请求至该目标数据节点，使该目标数据节点完成数据操作请求对应的数据标记处理。

25 在实际应用中，在存储状态为非持久化存储状态的情况下，存储系统的客户端可不对 trim 操作持久化，即不保存数据操作请求至本地，在出现数据节点或者客户端出现 failover 时，对索引系统，即上层 KV 类引擎层返回 trim 失败信息，由 KV 引擎记录不成功的 Trim 操作，在之后仅重试处理。

在本说明书一实施例中，客户端在发送数据操作请求至目标数据节点时，由于目标数据节点宕机无法接收数据操作请求，从而无法执行对应的数据标记处理，客户端会将该数据操作请求反馈至上层 KV 类引擎层，由 KV 类引擎层记录本次不成功的数据操作请求，由 KV 类引擎层在合适的时机重新发送标记处理指令至客户端。

30 通过上述两种故障转移方式，在遇到未成功执行的 trim 操作时，能够记录执行失败 trim 操作，并在后续进行重试处理，直至完成 trim 操作，提供了可靠的 trim 执行方式，提高系统的可靠性。

35 在目标数据节点完成对无效数据的 trim 操作之后，当有新的数据节点需要写入到目标数据节点时，目标数据节点可以在进行 trim 标记过的数据区间中写入新数据，从而实现数据的快速写入，提高存储能力。具体地，所述目标数据节点，还用于接收针对目标数据的数据写入指令，响应于数据写入指令，在本地存储空间中确定待使用数据区间，删除所述待使用数据区间中的数据，并将所述目标数据写入所述待使用数据区间，其中，所述待使用数据区间属于所述待操作数据区间。

其中，数据写入指令可以理解为将目标数据写入至目标数据节点中的数据块的指令，

待使用数据区间可以理解为已经进行过 trim 操作的数据控件，待使用数据区间中的数据即被标记为无效数据，从而数据节点可以直接识别出无效数据，将目标数据覆盖写入至该数据区间，完成对目标数据的写入处理。

5 在实际应用中，由于 trim 操作只对数据节点中的待操作数据区间进行标记处理，未对待操作数据区间中的数据进行删除，则当有新数据需要写入时，可以先对待操作数据区间中的数据进行删除处理，再将目标数据写入至待操作数据区间。

10 在本说明书一实施例中，目标数据节点接收到针对目标数据的数据写入指令，根据数据写入指令在本地存储空间中确定待使用数据区间，待使用数据区间中的数据即为无效数据，删除待使用数据区间中的无效数据，并将目标数据写入待使用数据区间，以完成对目标数据的写入处理。

具体实施时，当目标数据的写入区间大于待操作数据区间时，需要确保待操作数据区间相邻的区间也为待操作数据区间或未存有数据的数据区间，使这两个数据区间组成连续的存储空间以供目标数据写入。

15 为确保系统运行的正确性，应用存储系统的项目系统，需要保证对已经执行 trim 操作的数据不再进行读取，但是当上层 KV 类存储引擎层对执行 trim 操作的数据进行读取的情况下，客户端会返回未定义的反馈结果给上层 KV 类存储引擎层。具体地，所述客户端，还用于接收针对所述数据文件的数据读取指令，根据所述数据读取指令向所述目标数据节点发送数据读取请求；所述目标数据节点，还用于接收所述数据读取请求，根据所述数据读取请求生成反馈结果，并发送至所述客户端。

20 其中，数据读取指令可以理解为读取数据节点中存储的数据的指令，客户端接收针对数据文件的数据读取指令之后，会向目标数据节点发送数据读取请求，与对数据文件执行标记处理指令同理，客户端也会将 file 级的数据读取指令转换成数据节点中 Chunk 级的数据读取请求，使得目标数据节点能够根据数据读取请求返回对应的数据读取结果。

25 在实际应用中，由于数据文件已经被执行过 trim 操作，从而目标数据节点返回给客户端的反馈结果均为未定义，但是分为三种不同的情况，分别是数据文件的所有数据均成功执行 trim 操作；数据文件的所有数据均未成功执行 trim 操作；数据文件的部分数据成功执行 trim 操作，部分数据未成功执行 trim 操作。

30 在均成功执行 trim 操作的情况下，目标数据节点的反馈结果为全 0 数据；在均未成功执行 trim 操作的情况下，目标数据节点的反馈结果为 trim 前的原数据；在部分数据成功执行 trim 操作，部分数据未成功执行 trim 操作的情况下，目标数据节点的反馈结果为部分为 0 数据，部分为 trim 前的原数据。具体地，所述目标数据节点，用于在所述数据操作请求执行成功的情况下，生成表示数据为不可读取状态的反馈结果；在所述数据操作请求执行失败的情况下，将所述待操作数据区间中的原数据作为反馈结果；在所述数据操作请求部分执行失败的情况下，将所述待操作数据区间中对应所述数据操作请求执行失败的原数据，
35 作为第一反馈结果，以及根据所述待操作数据区间中对应所述数据操作请求执行成功的数据，生成第二反馈结果，基于所述第一反馈结果和所述第二反馈结果生成反馈结果。

具体实施时，不可读取状态的反馈结果可以理解为全为 0 的数据，在已经对数据进行 trim 操作之后，trim 后的数据不可再进行读取，则目标数据节点会反馈不可读取状态的结果给客户端；当 trim 执行失败之后，目标数据节点会反馈数据文件的原数据给客户端；当

部分数据成功 trim、部分数据 trim 失败时，目标数据节点会反馈部分不可读取的数据和原数据给客户端。

5 本说明书另一实施例中，所述客户端，用于接收针对数据文件的标记操作指令，并将所述标记操作指令关联的标记操作信息存储至操作日志，在所述操作日志满足日志信息处理条件的情况下，根据所述操作日志中记录的所述标记操作信息更新所述数据文件关联的数据操作表；基于更新后的数据操作表生成对应目标数据节点的数据操作请求，并将所述数据操作请求发送至所述目标数据节点。

10 具体的，数据文件是计算机系统中的一种文件类型，可以用于存储多种类型的数据。在存储系统中，数据文件具有文件标识，由一个或多个数据区间组成，每个数据区间又包括多个数据块。标记操作指令是指用于对数据文件中的数据进行标记操作的计算机指令，标记操作指令通常携带标记操作信息，标记操作信息包含需要进行标记操作的数据区间信息，以及操作日志中用于标记日志中存储的标记操作信息的序号，即，ID，其中，操作日志用于记录标记操作信息。

15 在上述将标记操作指令关联的标记操作信息存储至操作日志中之后，可以在操作日志满足日志信息处理条件的情况下，对操作日志中记录的标记操作信息进行处理，其中，日志信息处理条件是指预先设定的，用于确定是否可以处理操作日志中存储的标记操作信息的处理条件，在本实施例中，日志信息处理条件包括：在操作日志中存储的标记操作信息数量达到设定的数量阈值时，确定操作日志满足日志信息处理条件；预设日志处理周期，在达到预设日志处理周期时，确定操作日志满足日志信息处理条件。数据操作表用于存储与数据文件对应的待处理数据的相关信息，包括但不限于数据区间信息、偏移地址信息、长度信息，处理状态信息等。

20 基于此，在根据操作日志中记录的标记操作信息完成对数据文件关联的数据操作表的更新后，基于更新后的数据操作表生成对应目标数据节点的数据操作请求，并将数据操作请求发送至目标数据节点，由目标数据节点接收数据操作请求后，对数据节点中存储的数据进行标记操作。

25 本说明书提供的存储系统，所述系统包括客户端和数据节点：所述客户端，用于接收针对数据文件的标记操作指令，解析所述标记操作指令，获得标记操作信息，根据所述标记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点；所述目标数据节点，用于接收所述数据操作请求，根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。通过客户端将针对数据文件的标记操作指令转化成针对数据节点的数据区间的数据操作请求，对数据节点中存储的项目不需要的数据进行标记，使得后续在进行垃圾回收时，无需对无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据，也可以使得后续在新数据写入数据节点时，能够直接写入待操作数据区间，提高写入效率，优化存储数据节点，降低存储成本。

30 图3示出了根据本说明书一个实施例提供的一种应用于存储系统中的客户端的数据处理方法的流程图，包括步骤S302至步骤306。

- 步骤 S302: 接收针对数据文件的标记操作指令。
- 步骤 S304: 解析所述标记操作指令, 获得标记操作信息。
- 步骤 S306: 根据所述标记操作信息生成对应目标数据节点的数据操作请求, 并发送所述数据操作请求至所述目标数据节点。
- 5 可选的, 所述方法包括:
- 根据所述标记操作信息确定所述数据文件对应的逻辑地址信息, 基于所述逻辑地址信息生成对应目标数据节点的数据操作请求。
- 可选的, 所述方法还包括:
- 在确定所述数据操作请求对应的操作反馈信息为操作失败的情况下, 查询所述标记操作指令的存储状态信息, 并执行所述存储状态信息对应的操作处理。
- 10 可选的, 所述方法包括:
- 在所述存储状态信息为持久化存储状态的情况下, 基于持久化的数据操作请求, 执行发送所述数据操作请求至所述目标数据节点的步骤;
- 在所述存储状态信息为非持久化存储状态的情况下, 生成所述标记操作指令对应的处理结果信息, 并发送至所述存储系统对应的索引系统, 其中, 所述索引系统用于重新发送所述标记操作指令至所述存储系统。
- 15 可选的, 所述方法还包括:
- 接收针对所述数据文件的数据读取指令, 根据所述数据读取指令向所述目标数据节点发送数据读取请求。
- 20 可选的, 所述方法包括:
- 接收针对数据文件的标记操作指令, 并将所述标记操作指令关联的标记操作信息存储至操作日志, 在所述操作日志满足日志信息处理条件的情况下, 根据所述操作日志中记录的所述标记操作信息更新所述数据文件关联的数据操作表;
- 基于更新后的数据操作表生成对应目标数据节点的数据操作请求, 并将所述数据操作请求发送至所述目标数据节点。
- 25 可选地, 所述方法包括:
- 所述存储系统为分布式文件系统, 所述客户端包括标记操作接口;
- 所述客户端用于, 调用所述标记操作接口解析所述标记操作指令, 获得所述标记操作信息, 其中, 所述标记操作信息包括文件标识信息、文件偏移量信息、文件标记长度信息。
- 30 本实施例提供的数据处理方法与上述实施例提供的存储系统中的目标数据节点的执行过程相似, 相同或相应的描述内容均可参见上述实施例, 本实施例在此不作赘述。
- 本说明书提供的一种应用于存储系统中的客户端的数据处理方法, 包括接收针对数据文件的标记操作指令; 解析所述标记操作指令, 获得标记操作信息; 根据所述标记操作信息生成对应目标数据节点的数据操作请求, 并发送所述数据操作请求至所述目标数据节点。
- 35 通过客户端将针对数据文件的标记操作指令转化成针对数据节点的数据区间的数据操作请求, 对数据节点中存储的项目不需要的数据进行标记, 使得后续在进行垃圾回收时, 无

需对无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据。

图4示出了根据本说明书一个实施例提供的一种应用于存储系统中的目标数据节点的数据处理方法的流程图，包括步骤S402至步骤S404。

步骤S402：接收针对数据文件数据操作请求。

5 步骤S404：根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。

可选的，所述方法包括：

根据所述数据操作请求中携带的逻辑地址信息确定操作描述信息，根据所述操作描述信息在本地存储空间中确定所述数据文件关联的待操作数据区间，修改并保存所述待操作数据区间的属性信息，作为对所述待操作数据区间的标记处理。

10

可选的，所述方法还包括：

接收针对目标数据的数据写入指令，响应于数据写入指令，在本地存储空间中确定待使用数据区间，删除所述待使用数据区间中的数据，并将所述目标数据写入所述待使用数据区间，其中，所述待使用数据区间属于所述待操作数据区间。

15

可选的，所述方法还包括：

接收针对所述数据文件的数据读取指令，根据所述数据读取指令向所述目标数据节点发送数据读取请求；

所述目标数据节点，还用于接收所述数据读取请求，根据所述数据读取请求生成反馈结果，并发送至所述客户端。

20

可选的，所述方法还包括：

在所述数据操作请求执行成功的情况下，生成表示数据为不可读取状态的反馈结果；

在所述数据操作请求执行失败的情况下，将所述待操作数据区间中的原数据作为反馈结果；

在所述数据操作请求部分执行失败的情况下，将所述待操作数据区间中对应所述数据操作请求执行失败的原数据，作为第一反馈结果，以及根据所述待操作数据区间中对应所述数据操作请求执行成功的数据，生成第二反馈结果，基于所述第一反馈结果和所述第二反馈结果生成反馈结果。

25

本实施例提供的数据处理方法与上述实施例提供的存储系统中的目标数据节点的执行过程相似，相同或相应的描述内容均可参见上述实施例，本实施例在此不作赘述。

本说明书提供的一种应用于存储系统中的目标数据节点的数据处理方法，包括：接收针对数据文件数据操作请求；根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。通过数据节点根据数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理，使得后续在进行垃圾回收时，无需对无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据，也可以使得后续在新数据写入数据节点时，能够直接写入待操作数据区间，提高写入效率，优化存储数据节点，降低存储成本。

30

35

图5示出了本说明书一个实施例提供的一种存储系统中客户端与目标数据节点的交互

示意图，具体步骤包括步骤 S502 至步骤 S526。

步骤 S502：客户端接收针对数据文件的标记操作指令。

客户端接收到的是对数据文件为用户的个人信息文件进行标记操作的指令，标记操作指令为针对用户个人信息文件进行 trim 操作的指令。

5 步骤 S504：客户端解析所述标记操作指令，获得标记操作信息。

客户端解析标记操作指令后，获得的标记操作信息为 fileId、fileOffset、Length。

步骤 S506：客户端根据所述标记操作信息确定所述数据文件对应的逻辑地址信息。

客户端从元数据管理节获得针对数据文件的标记操作指令之后，根据标记操作信息中携带的属于文件级的操作信息 fileId、fileOffset、Length，可以确定出数据文件的逻辑地址信息，逻辑地址信息为数据块级的操作信息 chunkId、chunkOffset、Length。

10 步骤 S508：客户端根据所述逻辑地址信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点。

客户端根据解析出的逻辑地址信息生成数据操作请求，数据操作请求是针对数据节点中的数据块上的数据进行操作的请求，并将数据操作请求发送至目标数据节点，使目标数据节点根据逻辑地址信息进一步映射出数据文件在硬盘上的物理地址信息，并在存储空间中确定出待操作数据空间，从而对待操作数据空间中的数据进行处理。

步骤 S510：目标数据节点接收针对数据文件数据操作请求。

20 步骤 S512：目标数据节点根据所述数据操作请求中携带的逻辑地址信息确定操作描述信息，并根据所述操作描述信息在本地存储空间中确定所述数据文件关联的待操作数据区间。

目标数据节点根据数据操作请求中携带的逻辑地址信息确定出操作描述信息，并根据操作描述信息确定出待操作数据区间，操作描述信息可以理解为物理地址信息，待操作数据区间为数据节点中数据块中的存储空间。

25 步骤 S514：目标数据节点修改并保存所述待操作数据区间的属性信息，作为对所述待操作数据区间的标记处理。

目标数据节点修改待操作数据区间的属性信息，并保存修改记录。待操作数据区间的属性信息可以是该数据块的头数据，如数据块的容量信息、数据块的标记信息（是否被 trim 处理过）。

30 步骤 S516：客户端在确定所述数据操作请求对应的操作反馈信息为操作失败的情况下，查询所述标记操作指令的存储状态信息，并执行所述存储状态信息对应的操作处理。

客户端在确定本次 trim 操作失败的情况下，查询存储状态信息，存储状态信息分为持久化存储状态和非持久化存储状态，并执行与存储状态信息对应的操作处理。

步骤 S518：客户端接收针对所述数据文件的数据读取指令。

35 客户端在接收到读取数据文件的数据读取指令之后，会向目标数据节点发送对应的数据读取请求，数据读取指令可以理解为读取数据文件的指令，由于数据文件在之前已经发起过 trim 操作，因此客户端读取的是 trim 过的数据，而在实际应用中，数据读取指令也可以是读取任意类型数据文件的指令，如读取未 trim 过的数据。

步骤 S520: 客户端根据所述数据读取指令向所述目标数据节点发送数据读取请求。

步骤 S522: 目标数据节点接收所述数据读取请求。

步骤 S524: 目标数据节点根据所述数据读取请求生成反馈结果, 并发送至所述客户端。

- 5 目标数据节点接收到数据读取请求之后, 根据数据文件执行 trim 的状态生成反馈结果, 具体地, 在用户个人信息被成功执行 trim 时, 反馈结果中包括全 0 数据; 在用户个人信息执行 trim 失败时, 反馈结果中包括用户的个人信息; 在用户个人信息部分执行 trim 成功, 部分执行 trim 失败时, 反馈结果中包括部分全 0 数据和部分用户个人信息。

- 10 步骤 526: 目标数据节点接收针对目标数据的数据写入指令, 响应于数据写入指令, 在本地存储空间中确定待使用数据区间, 删除所述待使用数据区间中的数据, 并将所述目标数据写入所述待使用数据区间, 其中, 所述待使用数据区间属于所述待操作数据区间。

目标数据节点在接收到目标数据的数据写入指令时, 可以复用待操作数据区间, 目标数据可以是数据文件对应的更新数据, 则当这类数据写入时, 可以选择原有删除数据对应的数据区间进行写入; 目标数据也可以是其他数据文件, 当有新数据需要写入数据节点中时, 也可以选择写入至待操作数据区间中。

- 15 综上所述, 本说明书实施例提供的存储系统, 通过客户端将针对数据文件的标记操作指令转化成针对数据节点的数据区间的数据操作请求, 对数据节点中存储的项目不需要的数据进行标记, 使得后续在进行垃圾回收时, 无需对无效数据再进行处理, 有效降低垃圾回收产生的写放大, 减少垃圾数据。也可以使得后续在新数据写入数据节点时, 能够直接写入待操作数据区间, 提高写入效率, 优化存储数据节点, 降低存储成本。

- 20 与上述方法实施例相对应, 本说明书还提供了数据处理装置实施例, 图 6 示出了本说明书一实施例提供的一种应用于存储系统中的客户端的数据处理装置的结构示意图。如图 6 所示, 该装置包括:

接收模块 602, 被配置为接收针对数据文件的标记操作指令;

解析模块 604, 被配置为解析所述标记操作指令, 获得标记操作信息;

- 25 发送模块 606, 被配置为根据所述标记操作信息生成对应目标数据节点的数据操作请求, 并发送所述数据操作请求至所述目标数据节点。

可选的, 所述解析模块 604 进一步被配置为:

根据所述标记操作信息确定所述数据文件对应的逻辑地址信息, 基于所述逻辑地址信息生成对应目标数据节点的数据操作请求。

- 30 可选的, 所述存储系统为分布式文件系统, 所述客户端包括标记操作接口, 所述解析模块 604 进一步被配置为:

调用所述标记操作接口解析所述标记操作指令, 获得所述标记操作信息, 其中, 所述标记操作信息包括文件标识信息、文件偏移量信息、文件标记长度信息。

可选的, 所述装置还包括查询模块, 被配置为:

- 35 在确定所述数据操作请求对应的操作反馈信息为操作失败的情况下, 查询所述标记操作指令的存储状态信息, 并执行所述存储状态信息对应的操作处理。

可选的, 所述查询模块, 进一步被配置为:

在所述存储状态信息为持久化存储状态的情况下，基于持久化的数据操作请求，执行发送所述数据操作请求至所述目标数据节点的步骤；

在所述存储状态信息为非持久化存储状态的情况下，生成所述标记操作指令对应的处理结果信息，并发送至所述存储系统对应的索引系统，其中，所述索引系统用于重新发送所述标记操作指令至所述存储系统。

5 可选的，所述装置还包括读取模块，被配置为：

接收针对所述数据文件的数据读取指令，根据所述数据读取指令向所述目标数据节点发送数据读取请求。

10 可选地，所述装置还包括存储模块，被配置为：

接收针对数据文件的标记操作指令，并将所述标记操作指令关联的标记操作信息存储至操作日志，在所述操作日志满足日志信息处理条件的情况下，根据所述操作日志中记录的所述标记操作信息更新所述数据文件关联的数据操作表；基于更新后的数据操作表生成对应目标数据节点的数据操作请求，并将所述数据操作请求发送至所述目标数据节点。

15 本说明书提供的数据处理装置，通过客户端将针对数据文件的标记操作指令转化成针对数据节点的数据区间的数据操作请求，对数据节点中存储的项目不需要的数据进行标记，使得后续在进行垃圾回收时，无需对无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据。

20 上述为本实施例的数据处理装置的示意性方案。需要说明的是，该数据处理装置的技术方案与上述的数据处理方法的技术方案属于同一构思，数据处理装置的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

与上述方法实施例相对应，本说明书还提供了数据处理装置实施例，图7示出了本说明书一实施例提供的一种应用于存储系统中的目标数据节点的数据处理装置的结构示意图。如图7所示，该装置包括：

接收模块702，被配置为接收针对数据文件数据操作请求；

25 标记模块704，被配置为根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。

可选的，所述标记模块704，进一步被配置为：

30 根据所述数据操作请求中携带的操作描述信息，在本地存储空间中确定所述数据文件关联的待操作数据区间，修改并保存所述待操作数据区间的属性信息，作为对所述待操作数据区间的标记处理。

可选的，所述装置还包括写入模块，被配置为：

接收针对目标数据的数据写入指令，响应于数据写入指令，在本地存储空间中确定待使用数据区间，删除所述待使用数据区间中的数据，并将所述目标数据写入所述待使用数据区间，其中，所述待使用数据区间属于所述待操作数据区间。

35 可选的，所述装置还包括反馈模块，被配置为：

接收所述数据读取请求，根据所述数据读取请求生成反馈结果，并发送至所述客户端。

可选的，所述装置还包括反馈模块，被配置为：

在所述数据操作请求执行成功的情况下，生成表示数据为不可读取状态的反馈结果；
在所述数据操作请求执行失败的情况下，将所述待操作数据区间中的原数据作为反馈结果；

在所述数据操作请求部分执行失败的情况下，将所述待操作数据区间中对应所述数据
5 操作请求执行失败的原数据，作为第一反馈结果，以及根据所述待操作数据区间中对应所
述数据操作请求执行成功的数据，生成第二反馈结果，基于所述第一反馈结果和所述第二
反馈结果生成反馈结果。

本说明书提供的数据处理装置，通过数据节点根据数据操作请求，对本地存储空间中
10 关联所述数据文件的待操作数据区间进行标记处理，使得后续在进行垃圾回收时，无需对
无效数据再进行处理，有效降低垃圾回收产生的写放大，减少垃圾数据，也可以使得后续
在新数据写入数据节点时，能够直接写入待操作数据区间，提高写入效率，优化存储数据
节点，降低存储成本。

上述为本实施例的数据处理装置的示意性方案。需要说明的是，该数据处理装置的技术
15 方案与上述的数据处理方法的技术方案属于同一构思，数据处理装置的技术方案未详细
描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

图 8 示出了根据本说明书一实施例提供的一种计算设备 800 的结构框图。该计算设备
800 的部件包括但不限于存储器 810 和处理器 820。处理器 820 与存储器 810 通过总线 830
相连接，数据库 850 用于保存数据。

计算设备 800 还包括接入设备 840，接入设备 840 使得计算设备 800 能够经由一个或
20 多个网络 860 通信。这些网络的示例包括公用交换电话网（PSTN）、局域网（LAN）、广
域网（WAN）、个域网（PAN）或诸如因特网的通信网络的组合。接入设备 840 可以包括
有线或无线的任何类型的网络接口（例如，网络接口卡（NIC））中的一个或多个，诸如
IEEE802.11 无线局域网（WLAN）无线接口、全球微波互联接入（Wi-MAX）接口、以太
25 网接口、通用串行总线（USB）接口、蜂窝网络接口、蓝牙接口、近场通信（NFC）接口，
等等。

在本说明书的一个实施例中，计算设备 800 的上述部件以及图 8 中未示出的其他部件
也可以彼此相连接，例如通过总线。应当理解，图 8 所示的计算设备结构框图仅仅是出于
示例的目的，而不是对本说明书范围的限制。本领域技术人员可以根据需要，增添或替换
其他部件。

30 计算设备 800 可以是任何类型的静止或移动计算设备，包括移动计算机或移动计算设
备（例如，平板计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等）、移
动电话（例如，智能手机）、可佩戴的计算设备（例如，智能手表、智能眼镜等）或其他
类型的移动设备，或者诸如台式计算机或 PC 的静止计算设备。计算设备 800 还可以是移
动式或静止式的服务器。

35 其中，处理器 820 执行所述计算机指令时实现所述的数据处理方法步骤。

上述为本实施例的一种计算设备的示意性方案。需要说明的是，该计算设备的技术方
案与上述的数据处理方法的技术方案属于同一构思，计算设备的技术方案未详细描述的细节
内容，均可以参见上述数据处理方法的技术方案的描述。

本说明书一实施例还提供一种计算机可读存储介质，其存储有计算机指令，该计算机

指令被处理器执行时实现如前所述数据处理方法的步骤。

上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是，该存储介质的技术方案与上述的数据处理方法的技术方案属于同一构思，存储介质的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

- 5 本说明书一实施例还提供一种计算机程序，其中，当所述计算机程序在计算机中执行时，令计算机执行上述数据处理方法的步骤。

上述为本实施例的一种计算机程序的示意性方案。需要说明的是，该计算机程序的技术方案与上述的数据处理方法的技术方案属于同一构思，计算机程序的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

- 10 上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

- 15 所述计算机指令包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读
20 介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括电载波信号和电信信号。

- 需要说明的是，对于前述的各方法实施例，为了简便描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本说明书实施例并不受所描述的动作顺序的限制，因为依据本说明书实施例，某些步骤可以采用其它顺序或者同时进行。其次，本领域
25 技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定都是本说明书实施例所必须的。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中未详述的部分，可以参见其它实施例的相关描述。

- 以上公开的本说明书优选实施例只是用于帮助阐述本说明书。可选实施例并没有详尽
30 叙述所有的细节，也不限制该发明仅为所述的具体实施方式。显然，根据本说明书实施例的内容，可作很多的修改和变化。本说明书选取并具体描述这些实施例，是为了更好地解释本说明书实施例的原理和实际应用，从而使所属技术领域技术人员能很好地理解和利用本说明书。本说明书仅受权利要求书及其全部范围和等效物的限制。

权 利 要 求 书

- 1.一种存储系统，所述系统包括客户端和数据节点；
- 所述客户端，用于接收针对数据文件的标记操作指令，解析所述标记操作指令，获得标记操作信息，根据所述标记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点；
- 5 所述目标数据节点，用于接收所述数据操作请求，根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。
- 2.如权利要求 1 所述的系统，所述客户端，用于根据所述标记操作信息确定所述数据文件对应的逻辑地址信息，基于所述逻辑地址信息生成对应目标数据节点的数据操作请求。
- 10 3.如权利要求 1 所述的系统，所述目标数据节点，用于根据所述数据操作请求中携带的逻辑地址信息确定操作描述信息，根据所述操作描述信息在本地存储空间中确定所述数据文件关联的待操作数据区间，修改并保存所述待操作数据区间的属性信息，作为对所述待操作数据区间的标记处理。
- 4.如权利要求 1 所述的系统，所述客户端，还用于在确定所述数据操作请求对应的操作反馈信息为操作失败的情况下，查询所述标记操作指令的存储状态信息，并执行所述存储状态信息对应的操作处理。
- 15 5.如权利要求 4 所述的系统，所述客户端，用于在所述存储状态信息为持久化存储状态的情况下，基于持久化的数据操作请求，执行发送所述数据操作请求至所述目标数据节点的步骤；
- 20 在所述存储状态信息为非持久化存储状态的情况下，生成所述标记操作指令对应的处理结果信息，并发送至所述存储系统对应的索引系统，其中，所述索引系统用于重新发送所述标记操作指令至所述存储系统。
- 6.如权利要求 1 所述的系统，所述目标数据节点，还用于接收针对目标数据的数据写入指令，响应于数据写入指令，在本地存储空间中确定待使用数据区间，删除所述待使用数据区间中的数据，并将所述目标数据写入所述待使用数据区间，其中，所述待使用数据区间属于所述待操作数据区间。
- 25 7.如权利要求 1 所述的系统，所述客户端，还用于接收针对所述数据文件的数据读取指令，根据所述数据读取指令向所述目标数据节点发送数据读取请求；
- 所述目标数据节点，还用于接收所述数据读取请求，根据所述数据读取请求生成反馈结果，并发送至所述客户端。
- 30 8.如权利要求 7 所述的系统，所述目标数据节点，用于在所述数据操作请求执行成功的情况下，生成表示数据为不可读取状态的反馈结果；
- 在所述数据操作请求执行失败的情况下，将所述待操作数据区间中的原数据作为反馈结果；
- 35 在所述数据操作请求部分执行失败的情况下，将所述待操作数据区间中对应所述数据操作请求执行失败的原数据，作为第一反馈结果，以及根据所述待操作数据区间中对应所述数据操作请求执行成功的数据，生成第二反馈结果，基于所述第一反馈结果和所述第二

反馈结果生成反馈结果。

9.如权利要求 1 所述的系统，所述客户端，用于接收针对数据文件的标记操作指令，并将所述标记操作指令关联的标记操作信息存储至操作日志，在所述操作日志满足日志信息处理条件的情况下，根据所述操作日志中记录的所述标记操作信息更新所述数据文件关

5 联的数据操作表；

基于更新后的数据操作表生成对应目标数据节点的数据操作请求，并将所述数据操作请求发送至所述目标数据节点。

10.如权利要求 1 至 9 任意一项所述的系统，所述存储系统为分布式文件系统，所述客户端包括标记操作接口；

10 所述客户端用于，调用所述标记操作接口解析所述标记操作指令，获得所述标记操作信息，其中，所述标记操作信息包括文件标识信息、文件偏移量信息、文件标记长度信息。

11.一种数据处理方法，应用于存储系统中的客户端，包括：

接收针对数据文件的标记操作指令；

解析所述标记操作指令，获得标记操作信息；

15 根据所述标记操作信息生成对应目标数据节点的数据操作请求，并发送所述数据操作请求至所述目标数据节点。

12.一种数据处理方法，应用于存储系统中的目标数据节点，包括：

接收针对数据文件数据操作请求；

20 根据所述数据操作请求，对本地存储空间中关联所述数据文件的待操作数据区间进行标记处理。

13.一种计算设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机指令，所述处理器执行所述计算机指令时实现权利要求 11 或 12 所述方法的步骤。

14.一种计算机可读存储介质，其存储有计算机可执行指令，该计算机指令被处理器执行时实现权利要求 11 或 12 所述方法的步骤。