(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
2025 年 6 月 19 日 (19.06.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/123783 A1

(51) 国际专利分类号:
H04L 67/1097 (2022.01)

(21) 国际申请号: PCT/CN2024/114916

(22) 国际申请日: 2024 年 8 月 27 日 (27.08.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311713410.6 2023年12月12日 (12.12.2023) CN

(71) 申请人:阿里云计算有限公司(**ALIBABA CLOUD COMPUTING LTD.**) [CN/CN]; 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区1-2-A06室 310030 (CN)。

(72) 发明人:张烁(**ZHANG, Shuo**); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。吴东(**WU, Dong**); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。佟光勋(**TONG, Guangxun**); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区 310030 (CN)。张小路(**ZHANG, Xiaolu**); 中国北京市朝阳区广善路18号院-阿里巴巴北京朝阳科技园C区 100020 (CN)。吴均平(**WU, Junping**); 中国北京市朝阳区广善路18号院-阿里巴巴北京朝阳科技园C区 100020 (CN)。

(74) 代理人:北京同钧律师事务所(**BEIJING TONGJUN LAW FIRM**); 中国北京市海淀区西直门北大街32号院1号楼7层808室 100082 (CN)。

(54) **Title:** CLOUD DISK MANAGEMENT AND CONTROL METHOD, CLOUD DISK MANAGEMENT AND CONTROL SYSTEM, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 云盘管控方法、系统、电子设备及存储介质

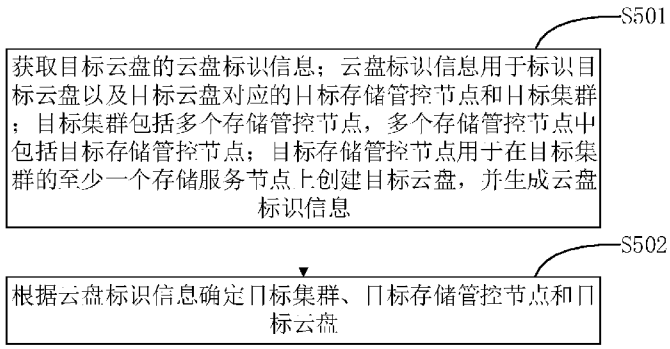


图 5

(57) **Abstract:** A cloud disk management and control method, a cloud disk management and control system, and an electronic device and a storage medium, which relate to the technical field of distributed storage. The method comprises: acquiring cloud disk identification information of a target cloud disk, wherein the cloud disk identification information is used for identifying the target cloud disk and a target storage master and a target cluster that correspond to the target cloud disk, the target cluster comprises a plurality of storage masters, the plurality of storage masters comprise the target storage master, and the target storage master is used for creating the target cloud disk on at least one storage server of the target cluster and generating the cloud disk identification information; and determining the target cluster, the target storage master and the target cloud disk on the basis of the cloud disk identification information. By means of the method, the resource utilization rate can be improved, and a large-scale cluster is supported.

(57) 摘要: 一种云盘管控方法、系统、电子设备及存储介质，涉及分布式存储技术领域。该方法包括：获取目标云盘的云盘标识信息；云盘标识信息用于标识目标云盘以及目标云盘对应的目标存储管控节点和目标集群；目标集群包括多个存储管控节点，多个存储管控节点中包括目标存储管控节点；目标存储管控节点用于在目标集群的至少一个存储服务节点上创建目标云盘，并生成云盘标识信息；根据云盘标识信息确定目标集群、目标存储管控节点和目标云盘。该方法可以提高资源利用率，支持大规模集群。

WO 2025/123783 A1

[见续页]

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

云盘管控方法、系统、电子设备及存储介质

本公开要求于2023年12月12日提交中国专利局、申请号为202311713410.6、申请名称为“云盘管控方法、系统、电子设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及分布式存储技术领域，尤其涉及一种云盘管控方法、系统、电子设备及存储介质。

背景技术

分布式存储是一种数据存储技术，它将数据分散存储在多个物理位置的网络中的设备上。块存储系统是一种低时延、持久性和高可靠的块级随机分布式存储系统。在块存储系统中，数据被分割成一块（block）一块的，每个块都作为独立的硬盘进行操作。这种方式提供了高性能和灵活性，特别适合处理大量数据，如数据库和大规模的虚拟机环境。在分布式块存储的应用场景下，云盘提供给用户数据块级别的块存储服务，它将物理硬盘的存储空间虚拟化，并通过网络提供给用户，用户可以像使用本地硬盘一样使用云盘，进行数据的读写操作。因此，在块存储系统中，如何实现对云盘的高效管控，减小块存储系统的故障域，显得十分重要。

发明内容

本公开实施例提供一种云盘管控方法、系统、电子设备及存储介质，以缓解或解决现有技术中存在的一项或多项技术问题。

第一方面，本公开实施例提供了一种云盘管控方法，包括：获取目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘以及所述目标云盘对应的目标存储管控节点和目标集群；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；所述目标存储管控节点用于在所述目标集群的至少一个存储服务节点上创建所述目标云盘，并生成所述云盘标识信息；根据所述云盘标识信息确定所述目标集群、所述目标存储管控节点和所述目标云盘。

第二方面，本公开实施例提供了一种云盘管控方法，包括：响应于云盘创建请求，在目标集群的至少一个存储服务节点上创建目标云盘；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；生成所述目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘、所述目标存储管控节点和所述目标集群。

第三方面，本公开实施例提供了一种云盘管控系统，包括：集群管控节点，用于执行本公开第一方面实施例提供的方法；至少一个集群，所述集群包括多个存储管控节点和多个存储服务节点，所述存储管控节点用于执行本公开第二方面实施例提供的方法。

第四方面，本公开实施例提供一种电子设备，包括存储器、处理器及存储在存储器上的计算机程序，处理器在执行计算机程序时实现本公开任一实施例提供的方法。

第五方面，本公开实施例提供一种计算机可读存储介质，计算机可读存储介质内存储有计算机程序，计算机程序被处理器执行时实现本公开任一实施例提供的方法。

第六方面，本公开实施例提供一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现本公开任一实施例提供的方法。

根据本公开实施例的云盘管控方法，通过对一个存储集群内的存储管控节点进行水平扩展，即一个存储集群包括多个存储管控节点，每个存储管控节点负责若干组云盘（例如为多个云盘集合）的管控。一方面，各存储管控节点同时工作，避免了一主多备架构造成的资源浪费，提高了资源利用率，并可以支持大规模集群；另一方面，在某一个存储管控节点故障，存储集群中受影响的云盘仅仅是一部分，因此，爆炸半径远小于集中式 BM 架构的爆炸半径。进一步地，可以针对分布式存储管控节点的云盘管控架构，提供行之有效的集群资源调度方案，使云盘调度不受架构更新的影响。

上述概述仅仅是为了说明书的目的，并不意图以任何方式进行限制。除上述描述的示意性的方面、实施方式和特征之外，通过参考附图和以下的详细描述，本公开进一步的方面、实施方式和特征将会是容易明白的。

附图说明

在附图中，除非另外规定，否则贯穿多个附图相同的附图标记表示相同或相似的部件或元素。这些附图不一定是按照比例绘制的。应该理解，这些附图仅描绘了根据本公开公开的一些实施方式，而不应将其视为是对本公开范围的限制。

图 1 示出相关技术中提供的云盘管控系统的架构图；

图 2 示出本公开实施例提供的云盘管控系统的架构图；

图 3 示出本公开实施例提供的云盘调度的一个应用示例图；

图 4 示出本公开实施例提供的云盘调度的另一个应用示例图；

图 5 示出本公开实施例提供的云盘管控方法的流程图；

图 6 示出本公开实施例提供的云盘管控方法的流程图；

图 7 示出本公开实施例提供的电子设备的框图。

具体实施方式

在下文中，仅简单地描述了某些示例性实施例。正如本领域技术人员可认识到的那样，在不脱离本公开的精神或范围的情况下，可通过各种不同方式修改所描述的实施例。因此，

附图和描述被认为本质上是示例性的而非限制性的。

为便于理解本公开实施例的技术方案，以下对本公开实施例的相关技术进行说明，以下相关技术作为可选方案与本公开实施例的技术方案可以进行任意结合，其均属于本公开实施例的保护范围。

下文将使用以下术语解释：

块存储（Block Storage）系统：一种低时延、持久性和高可靠的块级随机存储系统，将数据分为块（block），每个块都作为独立的硬盘驱动器进行操作。这意味着用户可以自由地在这些块上创建自己的文件系统，从而更好地控制数据环境，特别适合处理大量数据，如数据库和大规模的虚拟机环境。

云盘（Cloud Disk）：基于分布式存储架构，提供数据块级别的块存储服务。云盘的数据被持久化地存储在后端的分布式存储系统中，计算端任何节点的宕机不会对云盘产生影响。

爆炸半径：表征故障域的大小，即某一节点出故障时的影响面。爆炸半径越小，故障域越小；爆炸半径越大，故障域越大。

互联网数据中心（Internet Data Center，IDC）：是存放、管理和分发大量数据的设施。它提供了包括服务器托管、网络连接、数据存储和备份、安全管理等在内的一系列服务。

可用区（Availability Zone）：代表一个物理位置，这个物理位置拥有自己的电力供应、网络连接以及安全设施等。在同一地理区域（Region）内，可能存在一个或多个可用区。每个可用区都是相互独立的，与其他可用区隔离，以防止在一个可用区出现问题时影响其他可用区。比如，如果一个可用区因为电力故障、网络中断或者自然灾害等原因出现问题，其他可用区可以继续正常运行。在云计算环境中，一个可用区包含一个或多个互联网数据中心（Internet Data Center，IDC）。

集群（Cluster）：是一组相互独立的、通过高速计算机网络互联的计算机，它们构成了一个组，并以单一系统的模式加以管理。一个集群中的计算机或服务器通常通过高速网络连接在一起，共享存储、软件或其他资源，如计算集群、存储集群等。

集群管控节点：用于管控各个存储集群的服务节点，如资源管控节点（Resource Master，RM）。

存储管控节点：本公开实施例中，每个存储集群有若干个存储管控节点，如块存储后端控制节点（Block Master，BM），多个存储管控节点共同管控一个存储集群。

存储服务节点：每个存储集群有多个存储服务节点，每个存储服务节点用于根据存储管控节点的调度，进行数据写入或读出，如块存储后端服务节点（Block Server，BS）。

虚拟块设备（Device）：虚拟化的存储设备，它模拟了物理硬盘驱动器的功能，使得虚拟机可以像访问物理硬盘一样访问存储空间，从而读取或写入云盘中的数据。

创建云盘（创盘）：系统根据用户指定的云盘大小和类型以及当前集群的资源使用情

况，进行资源调度，包括：在选定的存储服务节点上创建一个新的云盘，并设置相关的元数据，如存储卷的大小、类型、状态等；将新创建的云盘注册到集群的元数据服务中，以便用户可以找到和使用它。

挂载云盘（开盘）：将云盘挂载到服务器实例，如弹性计算服务（Elastic Compute Service，ECS）实例，使该云盘作为该服务器实例的数据盘。

垃圾回收（Garbage Collection，GC）：一种自动内存管理技术。在计算机系统中，当对象不再被程序使用时，垃圾回收器会自动回收这些对象所占用的内存空间，以便重新利用。在集群环境中，GC 任务是指定期清理无用的数据块或对象，以释放存储空间并提高资源利用率。

相关技术中，存储集群采用集中式 BM 架构。图 1 示出相关技术中云盘管控系统的架构图。如图 1 所示，云盘管控系统通常包括客户端（Client）、集群管控节点 RM 和至少一个存储集群（图 1 以存储集群 1 和存储集群 2 为示例），存储集群包括存储管控节点 BM 和多个存储服务节点 BS（图 1 以 BS1、BS2、BS3……为示例）。

一方面，在集中式 BM 架构中，不管存储集群的规模大小，一个存储集群部署一组 BM，单组 BM 管理整个存储集群相关的客户端、BS 和云盘的元数据，负责调度云盘分片（即 Segment，图 1 以 SEG0、SEG1、SEG2……为示例）以及分配垃圾回收等后台任务，完成集群级别的所有管控工作。单组 BM 采用一主两备的服务方式，即在正常工作状态时，由一个主 BM 来完成存储集群的所有管控工作，当主 BM 宕机时，由备 BM 接管存储集群的管控工作。因此，对于单个集群来说，只有一台 BM 工作，其他 BM 处于休眠状态，造成了资源的浪费。而且，在某一个存储集群的一组 BM 异常的时候，另外一个存储集群即使有充足的机器，也没办法支持这个异常的存储集群。

另一方面，这种集中管理的方式架构相对简单，客户端通过 BM 获取 Segment 的位置，然后就可以去对应的 BS 上读写数据。集中式 BM 负责整个存储集群的管理、调度和运维相关工作，在单集群云盘数目较少、Segment 数目有限和集群规模较小的情况下可以比较高效的工作。但是，随着规模的增大，未来单集群预计需要支持更多的 Segment 和云盘，以未来要支持到 1000 万 Segment 来计算，那就需要单个 BM 提供 360G 内存，很明显现有的服务器是不能支持这一需求的，而且，集中式 BM 一旦发生宕机，整个集群粒度的管控操作都会失败，直到一主两备三个 BM 中的一个可以正常启动提供服务为止。因此，集中式 BM 架构的存储集群，其爆炸半径是整个存储集群的所有云盘。

本公开实施例旨在提供一种基于分布式 BM 的云盘管控方案。图 2 示出本公开实施例一种实施方式的云盘管控系统 200 的架构图。如图 2 所示，云盘管控系统 200 包括集群管控节点 201 和一个或多个集群 202。

其中，集群管控节点 201 为管控各个集群 202 的服务节点。示例性地，集群管控节点 201 可以为资源管控节点 RM。示例性地，当计算节点创建对应的存储资源时，由管理各个存储集群的集群管控节点 201 从各集群 202 中决策出其中一个作为目标集群。

集群 202 为存储集群。对于一个集群 202，其部署包括多个存储管控节点和多个存储服务节点。示例性地，存储管控节点可以为块存储后端控制节点（BM），如图 2 所示，多个存储管控节点例如为 BM1、BM2、BM3 和 BM4。存储服务节点可以为块存储后端服务节点（BS），如图 2 所示，多个存储服务节点例如为 BS1、BS2、BS3、BS4、BS5、BS6、BS7 和 BS8。需要说明的是，本公开实施例对集群 202 的个数，以及每个集群所部署的存储管控节点和存储服务节点的个数均不作具体限定。

每个存储管控节点用于管控多个云盘集合（partition）。例如，BM1 可以管控云盘集合 1 和云盘集合 8（对于云盘集合列表 1），BM2 可以管控云盘集合 2 和云盘集合 3（对于云盘集合列表 2），BM3 可以管控云盘集合 4 和云盘集合 6（对于云盘集合列表 3），BM4 可以管控云盘集合 5 和云盘集合 7（对于云盘集合列表 4）。对于一个集群 202 来说，多个存储管控节点与多个云盘集合列表（partition list）具有一一对应的映射关系，也即云盘集合与存储管控节点具有映射关系。

其中，云盘集合是一个逻辑概念，是一批云盘的集合，某一块云盘只能属于一个云盘集合。云盘集合被创建后，需要由对应的存储管控节点加载（load）后才能提供相应的服务。示例性地，为了逻辑简单，云盘集合（partition）支持动态创建，但是不能修改和删除。

存储管控节点对其相应云盘集合的管控，包括但不限于：创建云盘，并为创建的云盘生成相应的云盘标识信息（Volume ID）；基于云盘标识信息挂载云盘（开盘）、删除云盘、调度云盘等。

其中，云盘标识信息包括集群标识（Cluster ID）、云盘集合标识（Partition ID）和云盘标识（云盘在云盘集合上的位置）。云盘在云盘集合上的位置可以用云盘编号或云盘随机数表示，对于一个云盘集合来说，该云盘集合中的每个云盘具有唯一的云盘编号或唯一云盘随机数。示例性地，云盘标识信息还可以包括预留字段（或叫做预设字段），预留字段可以标识云盘的预留功能或预设功能或其他信息等。

对于一个集群 202，其部署还可以包括中心管控节点（Central Master，CM），用于从集群 202 的全局角度负责旁路化的管控工作，包括：管理该集群 202 中 BM 的上下线；管理该集群 202 中 BS 的上下线；分配创建云盘集合，为其创建和配置相应的资源等。

示例性地，云盘管控系统 200 还可以包括客户端 203，可以部署在计算集群，通过与集群管控节点 201 交互，获取云盘标识信息，进而基于云盘标识信息访问相应的云盘，如挂载云盘。客户端 203 可以通过应用程序接口（Application Programming Interface，API）与存储管控节点进行交互，得到 Segment 所在的存储服务节点的信息，进而去该存储服务节点上读写数据。

下面结合图 2 描述基于本公开实施例的云盘管控系统 200 进行动态创盘流程的应用示例。创盘流程主要包括：

（1）集群管控节点 201 从集群 201 的多个存储管控节点中，选择一个作为目标存储

管控节点（例如为BM1），选择策略可以是随机选择，也可以是基于哈希（hash）规则；

（2）集群管控节点201向BM1发送云盘创建请求；

（3）BM1可以配置有代理或管控组件（proxy），BM1的proxy从多个云盘集合中选择一个作为目标云盘集合（例如为云盘集合8）；

（4）BM1完成目标云盘的创建，生成目标云盘的云盘标识信息（Volume ID），并通过原路径向集群管控节点201返回创盘成功的信息以及该目标云盘的云盘标识信息。

下面结合图2描述基于本公开实施例的云盘管控系统200进行动态开盘流程的应用示例。开盘流程主要包括：

（1）客户端203通过虚拟块设备的设备标识信息（DEVID），向集群管控节点201查询目标集群的信息、目标云盘的云盘标识信息和目标存储管控节点的信息；

（2）集群管控节点201基于客户端203给出的设备标识信息，解析出对应的云盘标识信息（Volume ID），通过解析Volume ID，得到集群标识（Cluster ID）、云盘集合标识（Partition ID）和云盘标识（云盘在云盘集合上的位置），然后基于云盘集合与存储管控节点的映射关系，确定与云盘集合标识（Partition ID）对应的存储管控节点（例如为BM1）作为目标存储管控节点；

（3）集群管控节点201向客户端203返回目标集群的集群标识（例如为集群名称）、BM1的互联网协议（Internet Protocol，IP）地址、Volume ID；

（4）客户端203基于BM1的IP地址，向BM1发送针对目标云盘的云盘挂载请求；

（5）BM1响应云盘挂载请求，挂载目标云盘。

下面结合图3描述基于本公开实施例的云盘管控系统进行云盘调度的一个应用示例。该应用示例中，云盘调度流程主要包括：

（1）集群中的每一个BM获取该集群各BS的流量信息。获取方式可以是集群中各BS向中心管控节点（CM）发送心跳，上报各自的流量信息；获取方式也可以是集群中各BM分别向集群中各BS发送慢速心跳，从而每个BM都能得到各BS的流量信息，慢速心跳例如为每10秒发送一次；

（2）各BM按照预设顺序轮流执行云盘调度，例如，BM1、BM2、BM3和BM4轮流执行云盘调度，包括流量均衡调度（动态或静态）、从而将云盘的云盘分片（Segment）调度到多个存储服务节点上。示例性地，CM用于从集群的全局角度按照全局协调机制，控制各BM按照预设顺序轮流执行云盘调度，并进行一定的干涉，如果CM宕机，则各BM一直轮流执行云盘调度。

下面结合图4描述基于本公开实施例的云盘管控系统进行云盘调度的另一个应用示例。该应用示例中，云盘调度流程主要包括：

（1）集群中各BS划分为多组，形成个存储服务节点组，各BM与各存储服务节点组一一对应，BM拥有与其对应的存储服务节点组的管辖权，包括对该存储服务节点组的BS发起宕机操作。例如：BS1和BS2形成组1，BM1与组1对应，拥有对BS1和BS2的管辖权；

BS3 和 BS4 形成组 2，BM2 与组 2 对应，拥有对 BS3 和 BS4 的管辖权；BS5 和 BS6 形成组 3，BM3 与组 3 对应，拥有对 BS5 和 BS6 的管辖权；BS7 和 BS8 形成组 4，BM4 与组 4 对应，拥有对 BS7 和 BS8 的管辖权；

（2）BM 向与其对应的 BS 发送快速心跳，如每 200 毫秒一次，从而得到相应 BS 的运行状态信息，用于对 BS 进行故障判断（宕机判断）；

（3）如果 BM 判断出组内 BS 中存在故障 BS，即存在宕机的 BS，则对该 BS 进行宕机操作，例如，BM1 判断出组 1 中的 BS2 宕机，则 BM1 会对 BS 进行宕机操作，实现宕机调度（也叫做故障调度或异常调度）。

下面结合图 2 描述本公开实施例的云盘管控系统的心跳部署的应用示例。如图 2 所示，针对一个集群 202 内，各 BM 向 RM 报心跳，上报各自的运行状态信息，即确定该 BM 未故障；集群粒度的信息设置由该集群中的一个 BM（例如默认为 BM1）上报。中心管控节点向各 BM 报心跳，获取各 BM 的运行状态信息，如是否故障，以及获取 BM 上各云盘集合的负载状况。

本公开实施例的云盘管控系统提供了一种基于分布式存储管控节点的云盘管控方案，对一个存储集群内的存储管控节点进行水平扩展，即一个存储集群包括多个存储管控节点，每个存储管控节点负责若干个云盘集合的管控。一方面，各存储管控节点同时工作，避免了一主多备架构造成的资源浪费，提高了资源利用率；另一方面，在某一个存储管控节点故障，由其他存储管控节点接管其负责的云盘集合，在此过程中，存储集群中受影响的云盘仅仅是一部分，因此，爆炸半径远小于集中式 BM 架构的爆炸半径；又一方面，可以支持大规模集群，经仿真本公开实施例的方案可以支持 200 台甚至 1000 台这样更大规模的集群；再一方面，本公开实施例还针对分布式存储管控节点的云盘管控架构，提供了行之有效的集群资源调度方案，使云盘调度不受架构更新的影响。

需要说明的是，本公开实施例中提供的上述应用场景是为了便于理解，本公开实施例对技术方案的应用场景不作具体限定。此外，本公开所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择或编辑授权或者拒绝。

下面以具体的实施例对本公开的技术方案以及本公开的技术方案如何解决前述技术问题进行详细说明。所列举的若干具体的实施例可以相互结合，对于相同或相似的概念或过程可能在某些实施例中不再赘述。

图 5 示出根据本公开实施例的云盘管控方法的流程图。该云盘管控方法可以应用于集群管控节点。如图 5 所示，该云盘管控方法包括：

步骤 S501：获取目标云盘的云盘标识信息；云盘标识信息用于标识目标云盘以及目标云盘对应的目标存储管控节点和目标集群；目标集群包括多个存储管控节点，多个存储管

控节点中包括目标存储管控节点；目标存储管控节点用于在目标集群的至少一个存储服务节点上创建目标云盘，并生成云盘标识信息。

本公开实施例中，目标集群为存储集群，例如为图 2 所示中的集群 202。目标集群包括多个存储管控节点和多个存储服务节点。例如，如图 2 所示，存储管控节点可以为块存储后端控制节点，如图 2 所示，多个存储管控节点例如为 BM1、BM2、BM3 和 BM4。存储服务节点可以为块存储后端服务节点，如图 2 所示，多个存储服务节点例如为 BS1、BS2、BS3、BS4、BS5、BS6、BS7 和 BS8。多个存储管控节点中包括目标存储管控节点，目标存储管控节点为创建目标云盘的存储管控节点，且生成目标云盘的云盘标识信息（Volume ID）。

步骤 S502：根据云盘标识信息确定目标集群、目标存储管控节点和目标云盘。

云盘标识信息包括：用于确定目标集群的信息，如集群标识（Cluster ID）；用于确定目标存储管控节点的信息；用于确定目标云盘的信息。基于目标云盘的云盘标识信息即可确定该目标云盘对应的目标集群、目标存储管控节点以及目标云盘，进而使目标存储管控节点对目标云盘进行相应的管控，包括但不限于挂载、删除、调度等管控。

在一种实施方式中，每个存储管控节点用于管控多个云盘集合（partition），云盘标识信息包括云盘标识、云盘集合标识和集群标识，在步骤 S501 中，根据云盘标识信息确定目标集群、目标存储管控节点和目标云盘，可以包括：利用目标云盘的集群标识、云盘集合标识和云盘标识，分别确定目标集群、目标云盘所属的目标云盘集合以及目标云盘在目标云盘集合上的位置；基于云盘集合与存储管控节点的映射关系，确定与目标云盘集合对应的存储管控节点作为目标存储管控节点。

其中，partition 是一个逻辑概念，是一批云盘的集合，某一块云盘只能属于一个 partition。partition 被创建后，需要由对应的存储管控节点加载（load）后才能提供相应的服务。示例性地，为了逻辑简单，partition 支持动态创建，但是不能修改和删除。

每个存储管控节点用于管控多个 partition，对于一个集群来说，partition 与存储管控节点具有映射关系。因此，利用目标云盘的 partition 标识即可确定目标云盘所属的 partition，进而基于 partition 与存储管控节点的映射关系，得到相应的目标存储管控节点的信息。示例性地，一个 partition 与 BM 的 IP 地址具有映射关系，基于目标云盘的 partition 标识，可以确定目标 BM 的 IP 地址。

在一种实施方式中，本公开实施例的方法还可以包括：从多个存储管控节点中，选择目标存储管控节点；向目标存储管控节点发送云盘创建请求，以使目标存储管控节点在至少一个存储服务节点上创建目标云盘，并生成云盘标识信息；接收目标存储管控节点返回的云盘标识信息。

示例性地，当计算节点创建对应的存储资源时，由集群管控节点从各集群中决策出其中一个作为目标集群；集群管控节点可以从目标集群的多个存储管控节点中，选择一个作为目标存储管控节点，选择策略可以是随机选择，也可以是基于哈希规则，本公开实施例对此不作具体限定；进一步地，集群管控节点向目标存储管控节点发送云盘创建请求，由

目标存储管控节点完成目标云盘的创建，生成目标云盘的云盘标识信息（Volume ID），并向集群管控节点返回该目标云盘的Volume ID。

在一种实施方式中，客户端可以通过虚拟块设备的设备标识信息（DEVID），向集群管控节点查询目标集群的信息、目标云盘的云盘标识信息和目标存储管控节点的信息。

示例性地，集群管控节点可以基于设备标识信息，解析出对应的 Volume ID，通过解析Volume ID，得到集群标识（Cluster ID）、云盘集合标识（Partition ID）和云盘标识（云盘在云盘集合上的位置），然后基于云盘集合与存储管控节点的映射关系，确定与云盘集合标识（Partition ID）对应的目标存储管控节点的 IP 地址，并向客户端返回目标集群的集群标识（例如为集群名称）、目标存储管控节点的 IP 地址以及 Volume ID；客户端基于这些信息，可以向目标存储管控节点发送针对目标云盘的管控请求，如云盘挂载请求、云盘删除请求等。

图 6 示出根据本公开实施例的云盘管控方法的流程图。该云盘管控方法可以应用于存储管控节点，本公开实施例以目标存储管控节点为示例，即创建目标云盘的存储管控节点。如图 6 所示，该云盘管控方法包括：

步骤 S601：响应于云盘创建请求，在目标集群的至少一个存储服务节点上创建目标云盘；目标集群包括多个存储管控节点，多个存储管控节点中包括目标存储管控节点；

步骤 S602：生成目标云盘的云盘标识信息；云盘标识信息用于标识目标云盘、目标存储管控节点和目标集群。

其中，每个存储管控节点用于管控多个云盘集合（partition），云盘标识信息包括集群标识、partition 标识和云盘标识；其中，目标云盘的集群标识用于确定目标集群；目标云盘的 partition 标识用于标识目标云盘所属的目标 partition，以基于 partition 与存储管控节点的映射关系，确定与目标 partition 对应的存储管控节点作为目标存储管控节点；目标云盘的云盘标识用于确定目标云盘在目标 partition 上的位置。

在步骤 S601 中，在目标集群的至少一个存储服务节点上创建目标云盘，包括：基于预设的负载策略，从目标存储管控节点对应的多个 partition 中确定目标 partition；从目标 partition 中选择目标云盘，并在至少一个存储服务节点上创建目标云盘。

示例性地，存储管控节点响应于云盘创建请求，从多个 partition 选择一个作为目标 partition。选择策略可以是负载均衡策略，即根据各 partition 已经创建云盘的数量，选择负载少的 partition 作为目标 partition。选择策略也可以是随机选择策略，选择策略也可以是基于距离优先的策略，即选择与计算节点网络拓扑距离较近的 partition 作为目标 partition，本公开实施例对此不作限定。

其中，存储管控节点可以配置有代理或管控组件，用于从多个 partition 选择一个作为目标 partition，并将相应的操作请求发给对应的存储管控节点。具体可参见上述创盘流程的应用示例，在此不再赘述。

在一种实施方式中，本公开实施例的方法还可以包括：响应于针对目标云盘的云盘挂

载请求，挂载目标云盘。具体可参见上述动态开盘流程的应用示例，在此不再赘述。

根据本公开实施例的云盘管控方法，通过对一个存储集群内的存储管控节点进行水平扩展，即一个存储集群包括多个存储管控节点，每个存储管控节点负责若干组云盘（例如为多个云盘集合）的管控。一方面，各存储管控节点同时工作，避免了一主多备架构造成的资源浪费，提高了资源利用率，并可以支持大规模集群；另一方面，在某一个存储管控节点故障的情况下，存储集群中受影响的云盘仅仅是一部分，因此，爆炸半径远小于集中式BM（一主多备）架构的爆炸半径。

在一种实施方式中，多个存储管控节点按照预设顺序轮流执行云盘调度，本公开实施例由目标存储管控节点执行的方法还可以包括：获取目标集群中各存储服务节点的流量信息；响应于轮到目标存储管控节点执行云盘调度，基于各存储服务节点的流量信息，对目标存储管控节点对应的至少一个云盘进行调度。

示例性地，目标集群中的每一个存储管控节点获取目标集群各存储服务节点的流量信息。获取方式可以是目标集群中各存储服务节点向中心管控节点（CM）发送心跳，上报各自的流量信息；获取方式也可以是目标集群中各存储管控节点分别向目标集群中各存储服务节点发送慢速心跳，从而目标存储管控节点可以得到目标集群中每个存储服务节点的流量信息。进一步地，各存储管控节点按照预设顺序轮流执行云盘调度，当轮到目标存储管控节点进行云盘调度时，目标存储管控节点可以对其管控的各partition上的云盘进行云盘调度，包括动态流量均衡调度或静态流量均衡调度。

在一种实施方式中，目标集群的各存储服务节点划分为多个存储服务节点组，多个存储管控节点与多个存储服务节点组一一对应，本公开实施例由目标存储管控节点执行的方法还可以包括：基于目标存储管控节点对应的存储服务节点组中各存储服务节点的运行状态信息，判断目标存储管控节点对应的存储服务节点组中是否存在故障存储服务节点；响应于存在故障存储服务节点，对故障存储服务节点执行宕机操作。具体可参见结合图4的云盘调度应用示例，在此不再赘述。

在一种实施方式中，本公开实施例由目标存储管控节点执行的方法还可以包括：响应于目标集群中存在故障存储管控节点，管控故障存储管控节点对应的各云盘。也就是说，在某一个存储管控节点故障，由其他存储管控节点接管其负责的云盘。

根据本公开实施例的云盘管控方法，可以针对分布式存储管控节点的云盘管控架构，提供行之有效的集群资源调度方案，使云盘调度不受架构更新的影响。

与本公开实施例提供的方法相对应地，本公开实施例提供一种云盘管控装置，该云盘管控装置应用于集群管控节点，可以包括：云盘标识信息获取模块，用于获取目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘以及所述目标云盘对应的目标存储管控节点和目标集群；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；所述目标存储管控节点用于在所述目标集群的至少一个存储服务节点上创建所述目标云盘，并生成所述云盘标识信息；目标存储管控节点确定模块，

用于根据所述云盘标识信息确定所述目标集群、所述目标存储管控节点和所述目标云盘。

在一种实施方式中，每个存储管控节点用于管控多个云盘集合，所述云盘标识信息包括云盘标识、云盘集合标识和集群标识，目标存储管控节点确定模块具体用于：利用所述目标云盘的集群标识、云盘集合标识和云盘标识，分别确定所述目标集群、所述目标云盘所属的目标云盘集合以及所述目标云盘在所述目标云盘集合上的位置；基于云盘集合与存储管控节点的映射关系，确定与所述目标云盘集合对应的存储管控节点作为所述目标存储管控节点。

在一种实施方式中，该装置还可以包括：目标存储管控节点选择模块，用于从所述多个存储管控节点中，选择所述目标存储管控节点；云盘创建请求发送模块，用于向所述目标存储管控节点发送云盘创建请求，以使所述目标存储管控节点在所述至少一个存储服务节点上创建所述目标云盘，并生成所述云盘标识信息；云盘标识信息接收模块，用于接收所述目标存储管控节点返回的所述云盘标识信息。

与本公开实施例提供的方法相对应地，本公开实施例提供一种云盘管控装置，该云盘管控装置应用于目标存储管控节点，可以包括：目标云盘创建模块，用于响应于云盘创建请求，在目标集群的至少一个存储服务节点上创建目标云盘；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；云盘标识信息生成模块，用于生成所述目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘、所述目标存储管控节点和所述目标集群。

在一种实施方式中，每个存储管控节点用于管控多个云盘集合，所述云盘标识信息包括集群标识、云盘集合标识和云盘标识；其中，所述目标云盘的集群标识用于确定所述目标集群；所述目标云盘的云盘集合标识用于标识所述目标云盘所属的目标云盘集合，以基于云盘集合与存储管控节点的映射关系，确定与所述目标云盘集合对应的存储管控节点作为所述目标存储管控节点；所述目标云盘的云盘标识用于确定所述目标云盘在所述目标云盘集合上的位置。

在一种实施方式中，目标云盘创建模块具体用于：基于预设的负载策略，从所述目标存储管控节点对应的多个云盘集合中确定所述目标云盘集合；从所述目标云盘集合中选择所述目标云盘，并在所述至少一个存储服务节点上创建所述目标云盘。

在一种实施方式中，该装置还可以包括挂载模块，用于响应于针对所述目标云盘的云盘挂载请求，挂载所述目标云盘。

在一种实施方式中，所述多个存储管控节点按照预设顺序轮流执行云盘调度，该装置还可以包括：流量信息获取模块，用于获取所述目标集群中各存储服务节点的流量信息；调度模块，用于响应于轮到所述目标存储管控节点执行云盘调度，基于各存储服务节点的流量信息，对所述目标存储管控节点对应的至少一个云盘进行调度。

在一种实施方式中，所述目标集群的各存储服务节点划分为多个存储服务节点组，所述多个存储管控节点与所述多个存储服务节点组一一对应，该装置还可以包括：判断模块，

用于基于所述目标存储管控节点对应的存储服务节点组中各存储服务节点的运行状态信息，判断所述目标存储管控节点对应的存储服务节点组中是否存在故障存储服务节点；宕机操作模块，用于响应于存在所述故障存储服务节点，对所述故障存储服务节点执行宕机操作。

在一种实施方式中，该装置还包括故障管控模块，用于响应于所述目标集群中存在故障存储管控节点，管控所述故障存储管控节点对应的各云盘。

本公开实施例各装置中的各模块的功能可以参见上述方法中的对应描述，并具备相应的有益效果，在此不再赘述。

图 7 为用来实现本公开实施例的电子设备的框图。如图 7 所示，该电子设备包括：存储器 701 和处理器 702，存储器 701 内存储有可在处理器 702 上运行的计算机程序。处理器 702 执行该计算机程序时实现上述实施例中的方法。存储器 701 和处理器 702 的数量可以为一个或多个。

该电子设备还包括：通信接口 703，用于与外界设备进行通信，进行数据交互传输。

如果存储器 701、处理器 702 和通信接口 703 独立实现，则存储器 701、处理器 702 和通信接口 703 可以通过总线相互连接并完成相互间的通信。该总线可以是工业标准体系结构（Industry Standard Architecture，ISA）总线、外部设备互连（Peripheral Component Interconnect，PCI）总线或扩展工业标准体系结构（Extended Industry Standard Architecture，EISA）总线等。该总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 7 中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

可选地，在具体实现上，如果存储器 701、处理器 702 及通信接口 703 集成在一块芯片上，则存储器 701、处理器 702 及通信接口 703 可以通过内部接口完成相互间的通信。

本公开实施例提供了一种计算机可读存储介质，其存储有计算机程序，该程序被处理器执行时实现本公开任一实施例中提供的方法。

本公开实施例提供了一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现本公开任一实施例中提供的方法。

本公开实施例还提供了一种芯片，该芯片包括，包括处理器，用于从存储器中调用并运行存储器中存储的指令，使得安装有芯片的通信设备执行本公开任一实施例提供的方法。

本公开实施例还提供了一种芯片，包括：输入接口、输出接口、处理器和存储器，输入接口、输出接口、处理器以及存储器之间通过内部连接通路相连，处理器用于执行存储器中的代码，当代码被执行时，处理器用于执行申请任一实施例提供的方法。

应理解的是，上述处理器可以是中央处理器（Central Processing Unit，CPU），还可以是其他通用处理器、数字信号处理器（Digital Signal Processing，DSP）、专用集成电路（Application Specific Integrated Circuit，ASIC）、现场可编程门阵列（Field Programmable gate Array，FPGA）或者其他可编程逻辑器件、分立门或者晶体 管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者是任何常规的处理器等。值得说明

的是，处理器可以是支持进阶精简指令集机器（Advanced RISC Machines，ARM）架构的处理器。

可选地，上述存储器可以包括只读存储器和随机存取存储器，还可以包括非易失性随机存取存储器。该存储器可以是易失性存储器或非易失性存储器，或可包括易失性和非易失性存储器两者。其中，非易失性存储器可以包括只读存储器（Read-Only Memory，ROM）、可编程只读存储器（Programmable ROM，PROM）、可擦除可编程只读存储器（Erasable PROM，EPROM）、电可擦除可编程只读存储器（Electrically EPROM，EEPROM）或闪存。易失性存储器可以包括随机存取存储器（Random Access Memory，RAM），其用作外部高速缓存。通过示例性但不是限制性说明，许多形式的RAM可用。例如：静态随机存取存储器（Static RAM，SRAM）、动态随机存取存储器（Dynamic Random Access Memory，DRAM）、同步动态随机存取存储器（Synchronous DRAM，SDRAM）、双倍数据速率同步动态随机存取存储器（Double Data Rate SDRAM，DDR SDRAM）、增强型同步动态随机存取存储器（Enhanced SDRAM，ESDRAM）、同步连接动态随机存取存储器（Sync Link DRAM，SLDRAM）和直接内存总线随机存取存储器（Direct Rambus RAM，DRRAM）。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行计算机程序指令时，全部或部分地产生按照本公开的流程或功能。计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输。

需要说明的是，本公开所涉及的用户信息（包括但不限于用户设备信息、用户个人信息、用户操作信息等）和数据（包括但不限于用于处理的数据、分析的数据、存储的数据、展示的数据等）等，均为经用户授权或者经过各方充分授权的信息和数据，并且相关信息和数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包括于本公开的至少一个实施例或示例中。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或隐含地包括至少一个该特征。在本公开的描述中，“多个”的含义是两个或两个以上，除非另有明确具体的限定。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分。并且本公开的优选实施方式的范围包括另外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按基本同时的方式或按相反的顺序，来执行功能。

在流程图中表示或在此以其他方式描述的逻辑和/或步骤，例如：可以被认为是用于实现逻辑功能的可执行指令的定序列表，可以具体实现在任何计算机可读介质中，以供指令执行系统、装置或设备（如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统）使用，或结合这些指令执行系统、装置或设备而使用。

应理解的是，本公开的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。上述实施例方法的全部或部分步骤是可以通过程序来指令相关的硬件完成，该程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本公开各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。上述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读存储介质中。该存储介质可以是只读存储器，磁盘或光盘等。

以上所述，仅为本公开的具体实施方式，但本公开的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本公开揭露的技术范围内，可轻易想到其各种变化或替换，这些都应涵盖在本公开的保护范围之内。因此，本公开的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1. 一种云盘管控方法，应用于集群管控节点，所述方法包括：

获取目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘以及所述目标云盘对应的目标存储管控节点和目标集群；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；所述目标存储管控节点用于在所述目标集群的至少一个存储服务节点上创建所述目标云盘，并生成所述云盘标识信息；

根据所述云盘标识信息确定所述目标集群、所述目标存储管控节点和所述目标云盘。

2. 根据权利要求 1 所述的方法，其中，每个存储管控节点用于管控多个云盘集合，所述云盘标识信息包括云盘标识、云盘集合标识和集群标识，根据所述云盘标识信息确定所述目标集群、所述目标存储管控节点和所述目标云盘，包括：

利用所述目标云盘的集群标识、云盘集合标识和云盘标识，分别确定所述目标集群、所述目标云盘所属的目标云盘集合以及所述目标云盘在所述目标云盘集合上的位置；

基于云盘集合与存储管控节点的映射关系，确定与所述目标云盘集合对应的存储管控节点作为所述目标存储管控节点。

3. 根据权利要求 1 所述的方法，还包括：

从所述多个存储管控节点中，选择所述目标存储管控节点；

向所述目标存储管控节点发送云盘创建请求，以使所述目标存储管控节点在所述至少一个存储服务节点上创建所述目标云盘，并生成所述云盘标识信息；

接收所述目标存储管控节点返回的所述云盘标识信息。

4. 一种云盘管控方法，应用于目标存储管控节点，所述方法包括：

响应于云盘创建请求，在目标集群的至少一个存储服务节点上创建目标云盘；所述目标集群包括多个存储管控节点，所述多个存储管控节点中包括所述目标存储管控节点；

生成所述目标云盘的云盘标识信息；所述云盘标识信息用于标识所述目标云盘、所述目标存储管控节点和所述目标集群。

5. 根据权利要求 4 所述的方法，其中，每个存储管控节点用于管控多个云盘集合，所述云盘标识信息包括集群标识、云盘集合标识和云盘标识；其中，所述目标云盘的集群标识用于确定所述目标集群；所述目标云盘的云盘集合标识用于标识所述目标云盘所属的目标云盘集合，以基于云盘集合与存储管控节点的映射关系，确定与所述目标云盘集合对应的存储管控节点作为所述目标存储管控节点；所述目标云盘的云盘标识用于确定所述目标云盘在所述目标云盘集合上的位置。

6. 根据权利要求 5 所述的方法，其中，在目标集群的至少一个存储服务节点上

创建目标云盘，包括：

基于预设的负载策略，从所述目标存储管控节点对应的多个云盘集合中确定所述目标云盘集合；

从所述目标云盘集合中选择所述目标云盘，并在所述至少一个存储服务节点上创建所述目标云盘。

7. 根据权利要求4所述的方法，其中，还包括：

响应于针对所述目标云盘的云盘挂载请求，挂载所述目标云盘。

8. 根据权利要求4至7任一项所述的方法，其中，所述多个存储管控节点按照预设顺序轮流执行云盘调度，所述方法还包括：

获取所述目标集群中各存储服务节点的流量信息；

响应于轮到所述目标存储管控节点执行云盘调度，基于各存储服务节点的流量信息，对所述目标存储管控节点对应的至少一个云盘进行调度。

9. 根据权利要求4至7任一项所述的方法，其中，所述目标集群的各存储服务节点划分为多个存储服务节点组，所述多个存储管控节点与所述多个存储服务节点组一一对应，所述方法还包括：

基于所述目标存储管控节点对应的存储服务节点组中各存储服务节点的运行状态信息，判断所述目标存储管控节点对应的存储服务节点组中是否存在故障存储服务节点。

10. 根据权利要求9所述的方法，还包括：

响应于存在所述故障存储服务节点，对所述故障存储服务节点执行宕机操作。

11. 根据权利要求4至7任一项所述的方法，还包括：

响应于所述目标集群中存在故障存储管控节点，管控所述故障存储管控节点对应的各云盘。

12. 一种云盘管控系统，包括：

集群管控节点，用于执行权利要求1至3任一项所述的方法；

至少一个集群，所述集群包括多个存储管控节点和多个存储服务节点，所述存储管控节点用于执行权利要求4至11任一项所述的方法。

13. 根据权利要求12所述的云盘管控系统，其中，每个集群还包括中心管控节点，用于将集群中各存储服务节点的流量信息发送给各存储管控节点，并控制集群中各存储管控节点按照预设顺序轮流执行云盘调度；和/或，所述中心管控节点用于为每个存储管控节点分配创建多个云盘集合。

14. 一种电子设备，包括存储器、处理器及存储在存储器上的计算机程序，所述处理器在执行所述计算机程序时实现权利要求1至11中任一项所述的方法。

15. 一种计算机可读存储介质，计算机可读存储介质内存储有计算机程序，计算机程序被处理器执行时实现权利要求1至11中任一项所述的方法。

16. 一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现如权利要求 1 至 11 中任一项所述的方法。

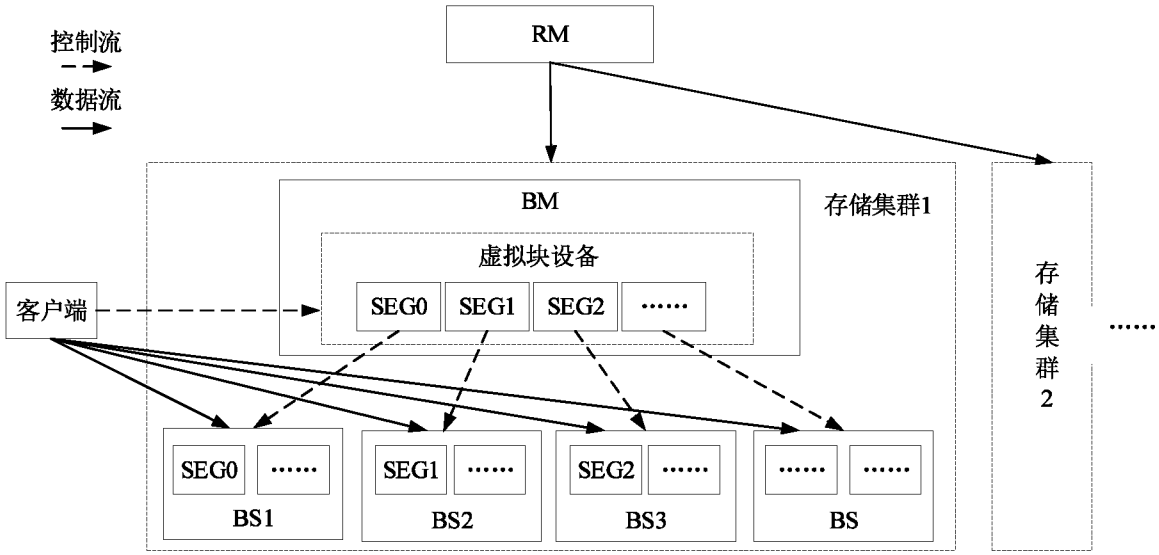

控制流
数据流
RM
BM
存储集群1
虚拟块设备
客户端
SEG0
SEG1
SEG2
……
SEG0
……
BS1
SEG1
……
BS2
SEG2
……
BS3
……
……
BS
存储集群2
……

图 1

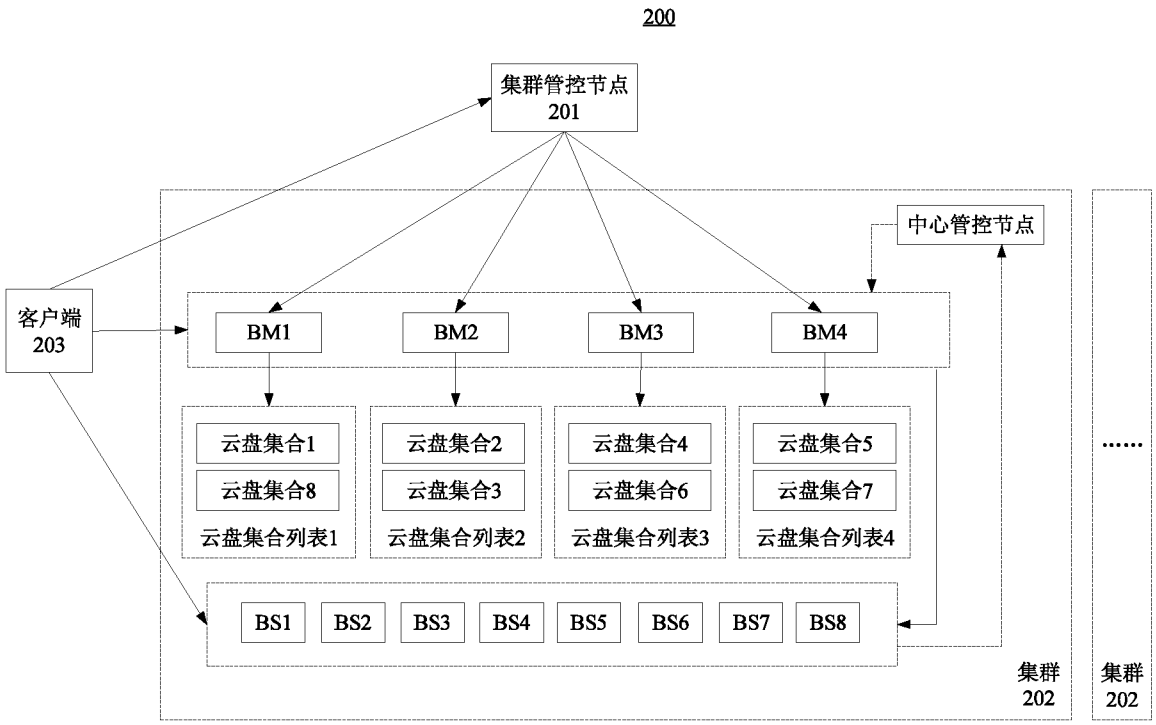

200
集群管控节点
201
中心管控节点
客户端
203
BM1
BM2
BM3
BM4
云盘集合1
云盘集合8
云盘集合列表1
云盘集合2
云盘集合3
云盘集合列表2
云盘集合4
云盘集合6
云盘集合列表3
云盘集合5
云盘集合7
云盘集合列表4
BS1
BS2
BS3
BS4
BS5
BS6
BS7
BS8
集群202
……
集群202

图 2

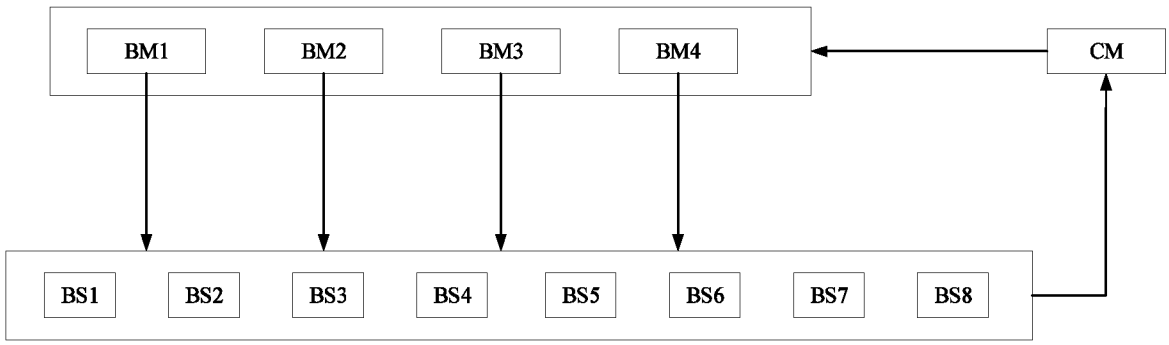

BM1
BM2
BM3
BM4
CM
BS1
BS2
BS3
BS4
BS5
BS6
BS7
BS8

图 3

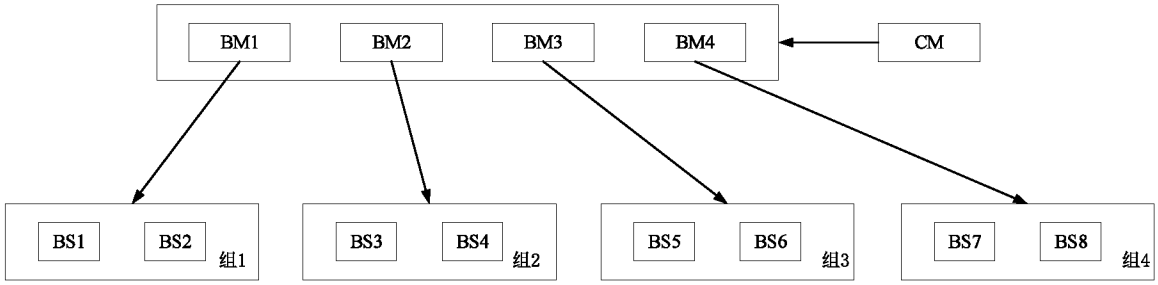

BM1
BM2
BM3
BM4
CM
BS1
BS2
组1
BS3
BS4
组2
BS5
BS6
组3
BS7
BS8
组4

图 4

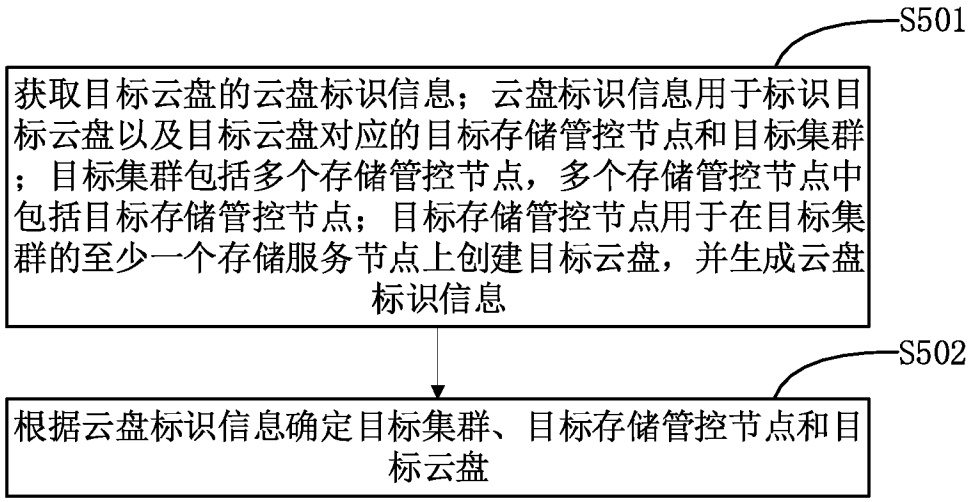

S501
获取目标云盘的云盘标识信息；云盘标识信息用于标识目标云盘以及目标云盘对应的目标存储管控节点和目标集群；目标集群包括多个存储管控节点，多个存储管控节点中包括目标存储管控节点；目标存储管控节点用于在目标集群的至少一个存储服务节点上创建目标云盘，并生成云盘标识信息
S502
根据云盘标识信息确定目标集群、目标存储管控节点和目标云盘

图 5

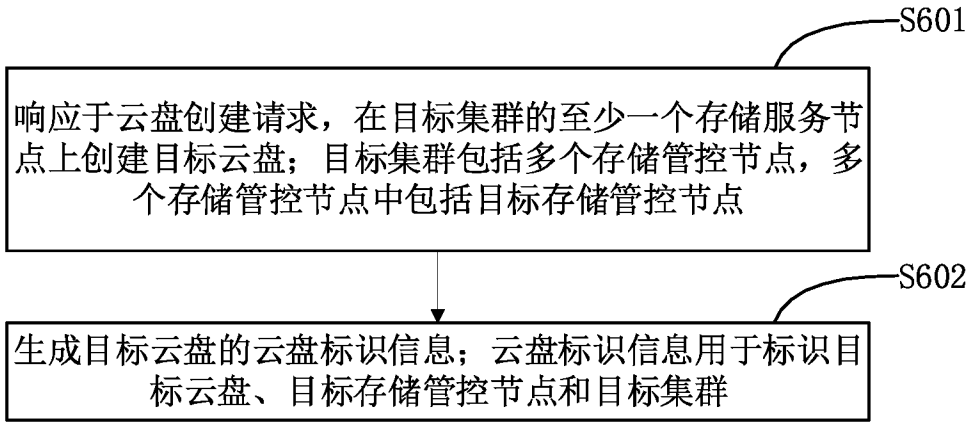

S601
响应于云盘创建请求，在目标集群的至少一个存储服务节点上创建目标云盘；目标集群包括多个存储管控节点，多个存储管控节点中包括目标存储管控节点
S602
生成目标云盘的云盘标识信息；云盘标识信息用于标识目标云盘、目标存储管控节点和目标集群

图 6

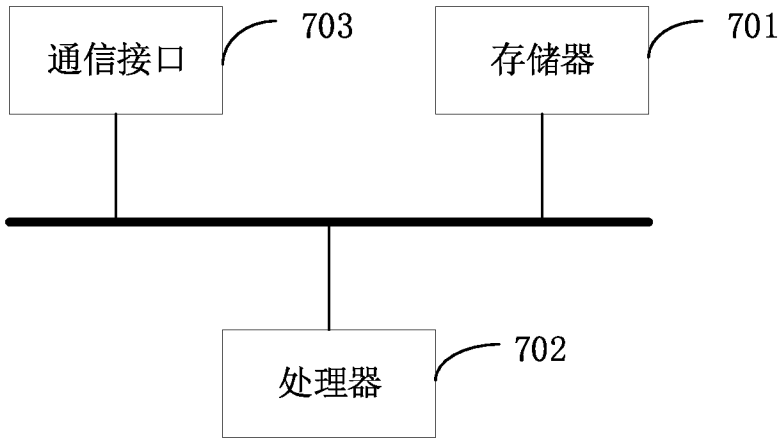

通信接口
703
存储器
701
处理器
702

图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/114916

A. CLASSIFICATION OF SUBJECT MATTER

H04L 67/1097(2022.01)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:H04L67/-

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)

CNABS, CNTXT, CNKI, 万方, WANFANG, 百度学术, BAIDU SCHOLAR: 云盘, 虚拟硬盘, 创建, 新建, 建立, 加载, 挂载, 标识, 标志, 管控, 管理, 控制, 集群, 分布式, BM; VEN, WOTXT, EPTXT, USTXT, IEEE, 3GPP: cloud drive, virtual hard disk, creat+, ID, label, indicator, colony, distributed, BM

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117714476 A (ALIBABA CLOUD COMPUTING LTD.) 15 March 2024 (2024-03-15) claims 1-15	1-16
Y	CN 115827148 A (HANGZHOU HIKVISION SYSTEM TECHNOLOGY CO., LTD.) 21 March 2023 (2023-03-21) description, paragraphs [0060]-[0078]	1-16
Y	CN 108241469 A (BEIJING KINGSOFT CLOUD NETWORK TECHNIQUE CO., LTD. et al.) 03 July 2018 (2018-07-03) description, paragraphs [0061]-[0112], and figure 3	1-16
A	CN 110708345 A (ALIBABA GROUP HOLDING LIMITED) 17 January 2020 (2020-01-17) entire document	1-16
A	CN 113282564 A (ALIBABA CLOUD COMPUTING LTD.) 20 August 2021 (2021-08-20) entire document	1-16
A	CN 114697331 A (INSPUR CLOUD INFORMATION TECHNOLOGY CO., LTD.) 01 July 2022 (2022-07-01) entire document	1-16

☑ 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
"D" document cited by the applicant in the international application
"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
05 November 2024	**05 November 2024**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 115168287 A (ALIBABA (CHINA) CO., LTD.) 11 October 2022 (2022-10-11) entire document	1-16
A	CN 115883662 A (ALIBABA (CHINA) CO., LTD.) 31 March 2023 (2023-03-31) entire document	1-16
A	CN 116436934 A (ALIBABA (CHINA) CO., LTD.) 14 July 2023 (2023-07-14) entire document	1-16
A	WO 2020088340 A1 (ALIBABA GROUP HOLDING LIMITED) 07 May 2020 (2020-05-07) entire document	1-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/114916

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 117714476 A	15 March 2024	None	
CN 115827148 A	21 March 2023	None	
CN 108241469 A	03 July 2018	None	
CN 110708345 A	17 January 2020	None	
CN 113282564 A	20 August 2021	None	
CN 114697331 A	01 July 2022	None	
CN 115168287 A	11 October 2022	None	
CN 115883662 A	31 March 2023	None	
CN 116436934 A	14 July 2023	None	
WO 2020088340 A1	07 May 2020	None	

国际检索报告

国际申请号
PCT/CN2024/114916

A. 主题的分类

H04L 67/1097(2022.01)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:H04L67/-

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））

CNABS;CNTXT;CNKI;万方;百度学术:云盘，虚拟硬盘，创建，新建，建立，加载，挂载，标识，标志，管控，管理，控制，集群，分布式，BM VEN;WOTXT;EPTXT;USTXT;IEEE;3GPP: cloud drive，virtual hard disk, creat+, ID, label，indicator，colony，distributed，BM.

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 117714476 A (阿里云计算有限公司) 2024年3月15日 (2024 - 03 - 15) 权利要求1-15	1-16
Y	CN 115827148 A (杭州海康威视系统技术有限公司) 2023年3月21日 (2023 - 03 - 21) 说明书第[0060]-[0078]段	1-16
Y	CN 108241469 A (北京金山云网络技术有限公司等) 2018年7月3日 (2018 - 07 - 03) 说明书第[0061]-[0112]段，说明书附图3	1-16
A	CN 110708345 A (阿里巴巴集团控股有限公司) 2020年1月17日 (2020 - 01 - 17) 全文	1-16
A	CN 113282564 A (阿里云计算有限公司) 2021年8月20日 (2021 - 08 - 20) 全文	1-16
A	CN 114697331 A (浪潮云信息技术股份公司) 2022年7月1日 (2022 - 07 - 01) 全文	1-16
A	CN 115168287 A (阿里巴巴（中国）有限公司) 2022年10月11日 (2022 - 10 - 11) 全文	1-16

☑ 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

“D” 申请人在国际申请中引证的文件

“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
2024年11月5日	2024年11月5日

ISA/CN的名称和邮寄地址	受权官员
中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	石琪琦 电话号码 (+86) 028-62969231

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	CN 115883662 A (阿里巴巴（中国）有限公司) 2023年3月31日 (2023 - 03 - 31) 全文	1-16
A	CN 116436934 A (阿里巴巴（中国）有限公司) 2023年7月14日 (2023 - 07 - 14) 全文	1-16
A	WO 2020088340 A1 (阿里巴巴集团控股有限公司) 2020年5月7日 (2020 - 05 - 07) 全文	1-16

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/114916

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	117714476	A	2024年3月15日	无	
CN	115827148	A	2023年3月21日	无	
CN	108241469	A	2018年7月3日	无	
CN	110708345	A	2020年1月17日	无	
CN	113282564	A	2021年8月20日	无	
CN	114697331	A	2022年7月1日	无	
CN	115168287	A	2022年10月11日	无	
CN	115883662	A	2023年3月31日	无	
CN	116436934	A	2023年7月14日	无	
WO	2020088340	A1	2020年5月7日	无	