

分布式存储系统的数据处理方法、装置及电子设备

本公开要求于 2019 年 10 月 31 日提交中国国家知识产权局、申请号为 201911051737.5、发明名称为“分布式存储系统的数据处理方法、装置及电子设备”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。

5

技术领域

本公开涉及分布式存储技术领域，更具体地，涉及一种分布式存储系统的数据处理方法、一种分布式存储系统的数据处理装置、一种电子设备及一种分布式存储系统。

10

背景技术

分布式存储是将数据分散存储在多台独立的设备上的存储方案。分布式网络存储系统采用可扩展的系统结构，利用多台存储服务器分担存储负荷，它不但提高了系统的可靠性、可用性和存取效率，还易于扩展。

15

在分布式存储系统中，通常由专门的控制服务器来协调数据在多台数据服务器上的存储。控制服务器中存储有用于描述数据属性的元数据，可以实现存储位置记录、历史数据记录、资源查找等功能。为了提高系统的可靠性，控制服务器的数目通常为多个，由其中的主控制服务器提供控制服务，从控制服务器作为备份。

20

主控制服务器和从控制服务器可以通过数据同步的方式保持元数据的一致性。但是，如果在数据同步过程中出现主控制服务器死机等情况，会导致部分元数据更新信息丢失，影响分布式存储系统的运行。

发明内容

25

本公开实施例的一个目的是提供一种分布式存储系统数据处理的新技术方案。

根据本公开实施例的第一方面，提供了一种分布式存储系统的数据处理方法，由所述分布式存储系统的从控制服务器实施，包括：

分别向每一存储服务器发送获取数据副本信息的第一请求；

根据每一所述存储服务器响应于所述第一请求返回的所述数据副本信息，更新用于反映所述数据副本信息和存储服务器对应关系的映射数据。

在一实施方式中，所述方法还包括：检测所述主控制服务器是否处于

5 不可用状态；

所述分别向每一存储服务器发送获取数据副本信息的第一请求，包括：

在所述主控制服务器处于不可用状态的情况下，分别向每一存储服务器发送获取数据副本信息的请求。

在一实施方式中，所述方法还包括：

10 基于更新后的所述映射数据，取代所述主控制服务器为所述分布式存储系统提供控制服务。

在一实施方式中，所述分别向每一存储服务器发送获取数据副本信息的第一请求，还包括：

按照设定的时间，定期向每一存储服务器发送获取数据副本信息的第一请求。

15 在一实施方式中，所述根据每一所述存储服务器响应于所述第一请求返回的所述数据副本信息，更新用于反映所述数据副本信息和存储服务器对应关系的映射数据，包括：

根据每一所述存储服务器响应于所述第一请求返回的所述数据副本信息，对现有的所述映射数据进行修改，获得更新后的所述映射数据。

20 根据本公开实施例的第二方面，提供了一种分布式存储系统的数据处理方法，由所述分布式存储系统中的存储服务器实施，包括：

接收从控制服务器发送的获取数据副本信息的请求；

25 响应于所述获取数据副本信息的请求，向所述从控制服务器发送自身对应的数据副本信息。

根据本公开实施例的第三方面，提供了一种分布式存储系统的数据处理装置，应用于所述分布式存储系统的从控制服务器，包括：

请求发送模块，设置为分别向每一存储服务器发送获取数据副本信息的第一请求；

数据更新模块，设置为根据每一所述存储服务器响应于所述第一请求返回的所述数据副本信息，更新用于反映所述数据副本信息和存储服务器对应关系的映射数据。

在一实施方式中，所述装置还可以包括：

5 检测模块：设置为检测主控制服务器是否处于不可用状态；

所述请求发送模块设置为：

在所述主控制服务器处于不可用状态的情况下，分别向每一存储服务器发送获取数据副本信息的第一请求。

10 在一实施方式中，所述装置还包括服务提供模块，所述服务提供模块设置为：

基于更新后的所述映射数据，取代所述主控制服务器为所述分布式存储系统提供控制服务。

在一实施方式中，所述请求发送模块还设置为：

15 按照设定的时间，定期向每一存储服务器发送获取数据副本信息的第一请求。

在一实施方式中，所述数据更新模块设置为：

根据每一所述存储服务器响应于所述第一请求返回的所述数据副本信息，对现有的所述映射数据进行修改，获得更新后的所述映射数据。

20 根据本公开实施例的第四方面，提供了一种分布式存储系统的数据处理装置，设置于所述分布式存储系统中的存储服务器，包括：

请求接收模块，用于接收从控制服务器发送的获取数据副本信息的第一请求；

数据发送模块，用于响应于所述第一请求，向所述从控制服务器发送自身存储的数据副本信息。

25 根据本公开实施例的第五方面，提供了一种电子设备，包括处理器和存储器，所述存储器存储有能够被所述处理器执行的机器可执行指令，所述处理器执行所述机器可执行指令以实现本公开实施例第一方面或者第二方面所述的分布式存储系统的数据处理方法。

根据本公开实施例的第六方面，提供了一种分布式存储系统，包括主

控制服务器、实施本公开实施例第一方面所述方法的从控制服务器和实施本公开实施例第二方面所述方法的存储服务器，其中，所述存储服务器分别与所述主控制服务器和所述从控制服务器通信连接。

本公开一个实施例中的分布式存储系统的数据处理方法，无需主从控制服务器之间进行元数据同步，避免了因同步过程中主控制服务器故障导致元数据丢失而无法同步主、从控制服务器中的元数据，有利于提高分布式存储系统的可靠性。

通过以下参照附图对本公开的示例性实施例的详细描述，本公开的其他特征及其优点将会变得清楚。

10

附图说明

被结合在说明书中并构成说明书的一部分的附图示出了本公开的实施例，并且连同其说明一起用于解释本公开实施例的原理。

图 1 示出了可用于实现本公开的实施例的分布式存储系统的硬件配置示意图；

图 2 示出了可用于实现本公开实施例的服务器的结构示意图；

图 3 示出了根据本公开实施例的分布式存储系统的数据处理方法的流程图；

图 4 示出了根据本公开实施例的分布存储系统的数据处理方法实施的具体例子的流程图。

具体实施方式

现在将参照附图来详细描述本公开的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本公开实施例的范围。

以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本公开实施例及其应用或使用的任何限制。

对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，技术、方法和设备应当被视为说明书的一部分。

在这里示出和讨论的所有例子中，任何具体值应被解释为仅仅是示例性的，而不是作为限制。因此，示例性实施例的其它例子可以具有不同的值。

5 应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

<硬件配置>

图 1 示出了可用于实现本公开实施例的分布式存储系统的结构示意图。

10 如图 1 所示，分布式存储系统 100 包括用户代理服务器 1000、存储服务器 2001、存储服务器 2002、存储服务器 2003、以及控制服务器 3001、控制服务器 3002。其中，存储服务器和控制服务器的数目均可以为多个（两个以上）。在本公开实施例中，存储服务器、控制服务器、用户代理服务器的数量均为示例性的，本公开实施例对此不做限制。

15 本实施例中，分布式存储系统以数据块（Block）为基本单元存储目标数据。例如，将目标文件分解为多个数据块，并将每一数据块的多个副本分别存储在相互独立的多个存储服务器中。在一个例子中，每个数据块对应于三个数据块副本，其中一个数据块副本为主副本，其他两个数据块副本为从副本。主副本所在的存储服务器优先处理针对目标文件的读写请求，
20 从副本所在的存储服务器起到备份作用。

存储服务器 2001、存储服务器 2002 和存储服务器 2003 用于存储目标数据，例如存储目标数据对应的数据块副本。

控制服务器 3001 和控制服务器 3002 用于管理元数据，例如管理反映目标文件对应的数据块副本和存储服务器对应关系的映射数据。其中一个
25 控制服务器作为主控制服务器，为分布式存储系统提供控制服务。另外一个控制服务器作为从控制服务器，用于备份主控制服务器中的数据。本实施例中，控制服务器 3001 为主控制服务器，控制服务器 3002 为从控制服务器。主从控制服务器的角色可以相互转换，在主控制服务器出现故障时，从控制服务器可以取代主控制服务器为分布式存储系统提供控制服务。

用户代理服务器 1000 用于接收用户端发送的针对目标数据的数据读写请求，并将该数据读写请求转发至控制服务器。

主控制服务器 3001 接收到用户代理服务器 1000 发送的数据读写请求后，可以根据自身存储的元数据，查询目标数据对应的数据块副本所在的
5 存储服务器，并向用户代理服务器 1000 返回处理该数据读写请求的存储服务器的标识信息。用户代理服务器 1000 根据该标识信息与对应的存储服务器交互，完成对目标数据的读写操作。

用户代理服务器 1000、存储服务器 2001、存储服务器 2002、存储服务器 2003、控制服务器 3001、控制服务器 3002 之间可以通过有线网络或者无线网络进行通信。
10

用户代理服务器 1000、存储服务器 2001、存储服务器 2002、存储服务器 2003、控制服务器 3001、控制服务器 3002 都具有如图 2 所示的服务器 1100 的硬件配置。如图 2 所示，服务器 1100 可以包括处理器 1110、存储器 1120、接口装置 1130、通信装置 1140、显示装置 1150 和输入装置 1160。
15 处理器 1110 例如可以是中央处理器 CPU 等。存储器 1120 例如包括 ROM(Read-Only Memory, 只读存储器)、RAM(Random Access Memory, 随机存取存储器)、诸如硬盘的非易失性存储器等。接口装置 1130 例如包括 USB 接口、串行接口等。通信装置 1140 例如能够进行有线或无线通信。显示装置 1150 例如是液晶显示屏。输入装置 1160 例如可以包括触摸屏、
20 键盘等。

应用于本说明书的实施例中，服务器 1100 的存储器 1120 用于存储指令，该指令用于控制处理器 1110 进行操作以支持实现根据本说明书任意实施例的数据处理方法。技术人员可以根据本说明书所公开方案设计指令。指令如何控制处理器进行操作，这是本领域公知，故在此不再详细描述。

本领域技术人员应当理解，尽管在图 2 中示出了服务器 1100 的多个装置，但是，本说明书实施例的服务器 1100 可以仅涉及其中的部分装置，例如，只涉及处理器 1110、存储器 1120 和通信装置 1140。
25

图 1 所示的分布式存储系统 100 仅是解释性的，并且决不是为了要限制本说明书、其应用或用途。

<方法实施例>

本公开实施例提供了一种分布式存储系统的数据处理方法，该方法例如由图 1 中的从控制服务器 3002 实施。如图 3 所示，该方法包括以下步骤 S1100-S1200：

- 5 步骤 S1100，分别向每一存储服务器发送获取数据副本信息的第一请求。

本实施例中，控制服务器用于管理元数据，例如管理反映目标文件对应的数据块副本和存储服务器对应关系的映射数据。其中，主控制服务器为分布式存储系统提供控制服务，从控制服务器起到备份作用。

- 10 本实施例中，分布式存储系统以数据块为基本单元存储目标数据。例如，将目标文件分解为多个数据块，并将每一数据块的多个副本分别存储在相互独立的多个存储服务器中。在一个例子中，每个数据块对应于三个数据块副本，其中一个数据块副本为主副本，其他两个数据块副本为从副本。例如，三个副本分别为 A、B、C，分别存储到编号为 1、2、3 的存储服务器中。

15 本实施例中，从控制服务器向每一存储服务器发送获取数据副本信息的第一请求，以获取每一存储服务器中所存储的数据块副本的信息。在本步骤一实施方式中，主副本所在的存储服务器可以优先响应第一请求。

- 20 在一个实施例中，本公开实施例的方法还可以包括检测主控制服务器是否处于不可用状态的步骤。示例性地，从控制服务器可以向主控制服务器发送状态询问请求（例如，通过心跳交互的方式），如主控制服务器无响应，则可以确定主控制服务器当前为不可用状态。在此基础上，步骤 S1100 可以包括：在主控制服务器处于不可用状态的情况下，分别向每一存储服务器发送获取数据副本信息的第一请求。

- 25 本实施例中，主控制服务器处于不可用状态，例如包括主控制服务器出现死机故障、主控制服务器与外界通信中断等情况。容易理解，主控制服务器处于不可用状态，意味着主控制服务器不能正常为分布式存储系统提供控制服务。

在一个实施例中，从控制服务器检测主控制服务器是否处于不可用状

态的步骤可以包括：接收监听设备（例如，zookeeper）发送的关于主控制服务器是否处于不可用状态的通知；根据该通知，确定主控制服务器是否处于不可用状态。在一个例子中，监听设备通过接收主控制服务器定期发送的心跳包来获取主控制服务器的活动状态。示例性地，监听设备（例如，
5 zookeeper）可以定期与控制服务器（主控制服务器、从控制服务器）进行心跳信号连接，如果在预设时间段内没有收到主控制服务器的心跳信号，则可以确定主控制服务器处于不可用状态，进而向从控制服务器发送通知。从控制服务器可以根据监听设备的通知确定主控制服务器处于不可用状态。

10 本实施例中，监听设备用于对主控制服务器的活动状态进行监听。监听设备可以是额外的实体设备，也可以是从控制服务器自身，例如，从控制服务器中通过自身的专门进程来监听主控制服务器的活动状态。

本实施例中，监听设备可以在检测到主控制服务器由可用状态变为不可用状态的情况下，向从控制服务器发送相应的通知。从控制服务器根据
15 该通知确定主控制服务器处于不可用状态。

在另一个实施例中，从控制服务器检测主控制服务器是否处于不可用状态的步骤包括：从控制服务器可以与主控制服务器进行心跳信号连接，如果在预设时间段内（例如，1 分钟）未收到主控制服务器的心跳信号，则可以确定主控制服务器是处于不可用状态的。

20 步骤 S1200，根据每一存储服务器响应于第一请求返回的数据副本信息，更新用于反映数据副本信息和存储服务器对应关系的映射数据。

本实施例中，从控制服务器中存储有反映数据副本与存储服务器对应关系的映射数据。

本实施例中，每一存储服务器响应于从控制服务器发送的第一请求，
25 向从控制服务器发送自身存储的数据副本信息。从控制服务器接收每一存储服务器发送自身存储的数据副本信息。并根据接收到的数据副本信息，建立数据副本信息与其所在的存储服务器之间的对应关系（映射关系），对映射数据（或者对应关系）表进行更新。其中，映射数据表或者对应关系表中可以包括一条以上的数据副本信息与其所在的存储服务器之间的对

应关系（或映射数据）。

通过执行步骤 S1200，从控制服务器可获得数据副本与存储服务器之间最新的对应关系。

5 本实施例中的分布式存储系统的数据处理方法，无需主从控制服务器之间进行元数据同步，避免了因同步过程中主控制服务器故障导致数据更新信息丢失，有利于提高分布式存储系统的可靠性。

10 在一个实施例中，分布式存储系统的数据处理方法还包括：在确定主控制服务器处于不可用状态时，从控制服务器可以基于更新后的映射数据，取代主控制服务器为分布式存储系统提供控制服务。主控制服务器可能因为掉电、宕机或者意外重启等而处于不可用状态。从控制服务器检测到主控制服务器不可用时，在更新其存储的映射数据后，及时取代主控制服务器为分布式存储系统提供控制服务。保证分布式存储系统的正常运行，为分布式存储系统的数据存储安全提供保证。

15 本实施例中，主控制服务器为分布式存储系统提供控制服务的过程例如是：主控制服务器接收到用户代理服务器发送的针对目标数据的数据读取请求，根据自身存储的映射数据查询目标数据对应的数据块副本所在的存储服务器，并向用户代理服务器返回处理该数据读写请求的存储服务器的标识信息，以供用户代理服务器根据该标识信息与对应的存储服务器交互，完成对目标数据的读写操作。

20 本实施例中，主从控制服务器的角色可以相互转换，在主控制服务器处于不可用状态的情况下，从控制服务器可以取代主控制服务器为分布式存储系统提供控制服务。

25 本实施例中，从控制服务器可自动取代主控制服务器为分布式存储系统提供控制服务（即自动切换），也可以响应于人工操作取代主控制服务器为分布式存储系统提供控制服务（即手动切换）。在本公开实施例中，以自动切换为优选实施例，从而保证数据存储的安全性和及时性，避免数据丢失。

本实施例中，从控制服务器根据更新后的映射数据为分布式存储系统提供控制服务，能够保证分布式存储系统的正常运行。

根据本公开实施例，前述步骤 S1100 可以实现为根据设定的时间，定期分别向每一存储服务器发送获取数据副本信息的第一请求（以下称为定期请求）。

在一实施例中，从控制服务器以设定的时间定期向每一存储服务器发送获取数据副本信息的第一请求。在另一些实施例中，可以为在主控制服务器不可用的状态下，从控制服务器按照设定的时间定期向每个存储服务器发送第一请求。示例性地，设定的时间 1 秒至 1 分钟中的任意时长，例如可以为 5 秒、10 秒、15 秒、20 秒、25 秒等，在本公开实施例中不对此进行限制，可以根据技术人员的经验值进行预设置

通过定期获取数据副本信息以及定期更新映射数据，可以减少每次更新需要处理的数据量，尤其有利于从控制服务器在主控制服务器处于不可用状态时快速更新映射数据，提高主从切换速度。

在一个实施例中，前述步骤 S1100 可以实现为在初始启动阶段，分别向每一存储服务器发送获取数据副本信息的第一请求（以下称为初始请求）。

本实施例中，从控制服务器在初始启动阶段向每一存储服务器发送获取数据副本信息的第一请求（在之后的叙述中，可以称之为初始请求）。容易理解，初始请求区别于从控制服务器在主控制服务器处于不可用状态的情况下发送的请求，以及从控制服务器定期发送的请求。

进而执行步骤 S1200，即根据每一存储服务器响应于初始请求返回的数据副本信息，建立反映数据副本和存储服务器对应关系的映射数据。

本实施例中，每一存储服务器响应于初始请求，向从控制服务器发送自身存储的数据副本信息。从控制服务器接收每一存储服务器发送自身存储的数据副本信息，建立反映数据副本和存储服务器对应关系的映射数据。

从控制服务器在初始启动阶段新建反映数据副本和存储服务器对应关系的映射数据，可以为后续更新映射数据建立基础。

在一个实施例中，从控制服务器更新自身存储的映射数据的过程包括：对自身存储的当前的映射数据进行修改，获得更新后的映射数据。

本实施例中，从控制服务器基于自身存储的当前的映射数据进行数据更

新，可以降低每次更新需要处理的数据量，有利于提高更新速度。

在一个实施例中，反映数据副本和存储服务器对应关系的映射数据存储于从控制服务器的内存中。如此，有利于提高对映射数据的读写速度，进而提高数据更新速度。

- 5 前述方法实施例中，针对从控制服务器主动向存储服务器发送第一请求、在主控制服务器处于不可用状态下，向存储服务器发送第一请求、在从控制服务器初始启动阶段（即启动运行时）向存储服务器发送第一请求三种状态下的实施方法和对应的技术效果。在分布式存储系统的实际运行过程中，全周期可以包括从控制服务器初始启动阶段、从控制服务器正常工作阶段以及检测到主控制服务器处于不可用状态等。以及，全周期中不同的状态下，从控制服务器对数据的处理的作用可能有一定差异。
- 10

下面将基于分布式存储系统的实际运行过程可能出现的系统状态，对本公开提供的另一实施例进行描述：

- 阶段（1）：在从控制服务器的初始启动阶段，向每个存储服务器发送第三请求获取数据副本信息，并基于此阶段得到的数据副本信息来创建初始映射数据。
- 15

阶段（2）：在从控制服务器的正常工作阶段，可以定期发送第四请求获取数据副本信息，用来定期更新映射数据。

- 阶段（3）：在主控制服务器处于不可用状态的情况下，从控制服务器发送第四请求，获取当前的数据副本信息并更新映射数据，进而可以基于最新的映射数据替代主控制服务器提供服务。
- 20

- 其中，本公开实施例中的第一请求、第二请求、第三请求和第四请求，是为了区别从控制服务器在不同阶段（或者不同状态下）发出的读取请求，但第一请求、第二请求、第三请求和第四请求操作方法和实现原理都是一致的。
- 25

总体而言，从控制服务器从启动时起，其全生命周期中均可以通过与存储服务器的交互，不断更新其存储的映射数据。且采用前述实施例分阶段进行，与在例如处于不可用状态时再去更新映射数据的实施例相比较，其须更新的数据相对较少，因此数据更新速度更快。

图 4 示出了本公开实施例中分布式存储系统的数据处理方法实施的具体例子。如图 4 所示，从控制服务器通过自身的监听进程监听主控制服务器的心跳，从而检测主控服务器的活动状态，即执行步骤 S101。此外，从控制服务器还定期向存储服务器发送获取数据副本信息的第二请求，存储服务器响应于该第二请求向从控制服务器发送自身存储的数据副本的信息，即执行步骤 S102。从控制服务器根据存储服务器发送的数据副本信息，定期更新自身内存中存储的反映数据副本和存储服务器对应关系的映射数据，即执行步骤 S103。假设主控制服务器在运行过程中出现了死机故障，这时从控制服务器可能无法在预设时间内监听到主控制服务器的，可以确定主控制服务器处于不可用状态，即执行步骤 S104。之后，从控制服务器分别向每个存储服务器发送获取副本信息的第一请求，即执行步骤 S105。每个存储服务器响应于该第一请求向从控制服务器发送自身存储的数据副本的信息，即执行步骤 S106。从控制服务器根据存储服务器发送的数据副本信息，对最近一次定期更新所获得数据进行修正，获得更新后的映射数据，即执行步骤 S107。最后，从控制服务器自动取代主控制服务器为分布式存储系统提供控制服务，即执行步骤 S108。

本公开实施例还提供另一种分布式存储系统的数据处理方法，该方法例如由图 1 所示的任意的存储服务器实施。该方法包括：接收从控制服务器发送的获取数据副本信息的请求；响应于获取数据副本信息的请求，向从控制服务器发送自身对应的数据副本信息。

该方法的具体实施方式可参见前文对分布式存储系统的数据处理方法的描述，此处不再赘述。

<装置实施例>

本公开实施例提供一种分布式存储系统的数据处理装置，应用于分布式存储系统的从控制服务器，包括请求发送模块和数据更新模块。

请求发送模块，设置为分别向每一存储服务器发送获取数据副本信息的第一请求。

数据更新模块，设置为根据每一存储服务器响应于第一请求返回的数据副本信息，更新用于反映数据副本信息和存储服务器对应关系的映射数

据。

在一实施方式中，本公开实施例的分布式存储系统的数据处理装置还可以包括：检测模块：设置为检测主控制服务器是否处于不可用状态。在一实施方式中，请求发送模块设置为在主控制服务器处于不可用状态的情况下，分别向每一存储服务器发送获取数据副本信息的第一请求。

在一实施方式中，装置还包括服务提供模块，服务提供模块设置为：基于更新后的映射数据，取代主控制服务器为分布式存储系统提供控制服务。

在一实施方式中，请求发送模块还可以设置为：按照设定的时间，定期向每一存储服务器发送获取数据副本信息的第一请求。

在一实施方式中，数据更新模块可以设置为根据每一存储服务器响应于第一请求返回的数据副本信息，对现有的映射数据进行修改，获得更新后的映射数据。

本公开实施例还提供另外一种分布式存储系统的数据处理装置，应用于分布式存储系统中的存储服务器，包括请求接收模块和数据发送模块。

请求接收模块，设置为接收从控制服务器发送的获取数据副本信息的第一请求。

数据发送模块，设置为响应于第一请求，向从控制服务器发送自身存储的数据副本信息。

<电子设备实施例>

本公开实施例提供一种电子设备，包括处理器和存储器，存储器存储有能够被处理器执行的机器可执行指令，处理器执行机器可执行指令以实现本公开方法实施例描述的分布式存储系统的数据处理方法。

<系统实施例>

本公开实施例提供一种分布式存储系统，包括主控制服务器、实施本公开方式实施例描述的第一种方法的从控制服务器和实施本公开方法实施例描述的第二种方法的存储服务器，其中，存储服务器分别与主控制服务器和从控制服务器通信连接。

<计算机可读存储介质实施例>

本公开实施例提供一种机器可读存储介质，机器可读存储介质存储有机器可执行指令，机器可执行指令在被处理器调用和执行时，机器可执行指令促使处理器实现本公开方法实施例描述的分布式存储系统的服务协调方法。

- 5 本公开实施例可以是系统、方法和/或计算机程序产品。计算机程序产品可以包括计算机可读存储介质，其上载有用于使处理器实现本公开实施例的各个方面的计算机可读程序指令。

计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是——但不限于——电存储
10 设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、静态随机存取存储器（SRAM）、便携式压缩盘只读存储器（CD-ROM）、数字多功
15 能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

- 20 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读
25 程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的计算机可读存储介质中。

用于执行本公开实施例操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代

码，编程语言包括面向对象的编程语言—诸如 Smalltalk、C++等，以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列（FPGA）或可编程逻辑阵列（PLA），该电子电路可以执行计算机可读程序指令，从而实现本公开实施例的各个方面。

这里参照根据本公开实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本公开实施例的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制造品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它设备上，使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤，以产生计算机实现的过程，从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和/或框图中的一个或多个方框中规定的功能/动作。

附图中的流程图和框图显示了根据本公开的多个实施例的系统、方法

和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分，模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或动作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。对于本领域技术人员来说公知的是，通过硬件方式实现、通过软件方式实现以及通过软件和硬件结合的方式实现都是等价的。

以上已经描述了本公开的各实施例，上述说明是示例性的，并非穷尽性的，并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下，对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择，旨在最好地解释各实施例的原理、实际应用或对市场中的技术改进，或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。本公开实施例的范围由所附权利要求来限定。

工业实用性

基于本公开实施例提供的分布式存储系统的数据处理方法、装置及电子设备，无需主从控制服务器之间进行元数据同步，避免了因同步过程中主控制服务器故障导致元数据丢失而无法同步主、从控制服务器中的元数据，有利于提高分布式存储系统的可靠性。