

## 用于存储数据的方法和装置

### 相关申请的交叉引用

本专利申请要求于 2021 年 02 月 18 日提交的、申请号为  
5 202110187129.8、发明名称为“用于存储数据的方法和装置”的中国专利申请  
的优先权，该申请的全文以引用的方式并入本公开中。

### 技术领域

本公开涉及计算机技术领域，具体涉及用于存储数据的方法和装置。  
10

### 背景技术

随着数据时代的到来，数据成为重要的资产之一。现有的存储数据的方法  
是基于数据写入周期对各个写入周期内的数据进行逐一存储。

然而，现有的数据存储的方法存在存储效率低、以及浪费存储资源的问题。  
15

### 发明内容

本公开提供了一种用于存储数据的方法、装置、电子设备、计算机可读  
存储介质以及计算机程序产品。

20 根据本公开的第一方面，提供了一种用于存储数据的方法，包括：获取  
待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，  
存储模型中包括目标存储参数；响应于确定源存储参数与目标存储参数一  
致，获取待存储数据集的存储键值；获取首次写入待存储数据集中的与源  
存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变  
25 化数据；以及根据存储键值，将原始数据以及变化数据存储于存储模型中。

在一些实施例中，用于存储数据的方法包括：响应于确定源存储参数  
与目标存储参数不一致，采用源存储参数以及目标存储参数确定缺失参  
数，其中，缺失参数包括：待存储数据集中存在、且存储模型中不存在的  
存储参数，或者待存储数据集中不存在、且存储模型中存在的存储参数。

30 在一些实施例中，用于存储数据的方法包括：响应于确定存储模型中

不存在缺失参数，将缺失参数添加至存储模型；或者，响应于确定待存储数据集中不存在缺失参数，利用待存储数据集的原始获取方式，获取与缺失参数对应的待存储数据，并将缺失参数以及与缺失参数对应的待存储数据添加至待存储数据集中。

5        在一些实施例中，获取待存储数据集的存储键值，包括：针对至少一个预设存储键值中的每一个预设存储键值，获取根据该预设存储键值对待存储数据集中的待存储数据进行聚类后的、待存储数据的分布结果；以及响应于确定分布结果满足预设条件，确定该预设存储键值为目标存储键值。

10       在一些实施例中，用于存储数据的方法包括：将待存储数据集中的、除原始数据以及变化数据之外的数据标记为第一数据；以及根据存储键值，将原始数据以及变化数据存储于存储模型中，包括：根据存储键值，将待存储数据集中、除被标记为第一数据之外的数据存储于存储模型中。

15       在一些实施例中，根据存储键值，将原始数据以及变化数据存储于存储模型中，包括：根据存储键值，将原始数据、原始数据的存续时间，以及变化数据、变化数据的存续时间存储于存储模型中，并获得存储数据集，其中，存续时间用于表征数据存续的时间周期；以及方法包括：根据原始数据的存续时间从存储数据集中提取原始数据；根据变化数据的存续时间从存储数据集中提取变化数据；以及响应于确定提取的原始数据以及提取  
20       的变化数据构成的数据集，与待存储数据集一致，确定存储数据集为目标存储数据集。

25       在一些实施例中，用于存储数据的方法包括：根据待存储集中的数据的数据特征，确定数据库参数；将原始数据以及变化数据存储于存储模型中，包括：利用数据库参数，更新存储模型；以及将原始数据以及变化数据  
25       存储于更新后的存储模型中。

      根据本公开的第二方面，提供了一种用于存储数据的方法，包括：响应于确定待存储数据集满足预设条件，执行如第一方面提供的用于存储数据的方法；其中，预设条件包括以下至少一项：待存储数据集中的数据变化速度小于预设阈值，或者待存储数据集中的数据量大于预设数据量。

30       根据本公开的第三方面，提供了一种用于存储数据的方法，包括：响

应于确定当前时刻处于预设数据处理时间段，执行如第一方面提供的用于存储数据的方法。

根据本公开的第四方面，提供了一种用于存储数据的装置，包括：第一获取单元，被配置为获取待存储数据集以及存储模型，其中，待存储数据集包括源存储参数，存储模型中包括目标存储参数；第二获取单元，被配置为响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值；第三获取单元，被配置为获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据；以及存储单元，被配置为根据存储键值，将原始数据以及变化数据存储于存储模型中。

在一些实施例中，用于存储数据的装置包括：数据校验单元，被配置为响应于确定源存储参数与目标存储参数不一致，采用源存储参数以及目标存储参数确定缺失参数，其中，缺失参数包括：待存储数据集中存在、且存储模型中不存在的存储参数，或者待存储数据集中不存在、且存储模型中存在的存储参数。

在一些实施例中，用于存储数据的装置包括：第一数据校验模块，被配置为响应于确定存储模型中不存在缺失参数，将缺失参数添加至存储模型；或者，第二数据校验模块，被配置为响应于确定待存储数据集中不存在缺失参数，利用待存储数据集的原始获取方式，获取与缺失参数对应的待存储数据，并将缺失参数以及与缺失参数对应的待存储数据添加至待存储数据集中。

在一些实施例中，第二获取单元，包括：获取模块，被配置为针对至少一个预设存储键值中的每一个预设存储键值，获取根据该预设存储键值对待存储数据集中的待存储数据进行聚类后的、待存储数据的分布结果；以及确定模块，被配置为响应于确定分布结果满足预设条件，确定该预设存储键值为目标存储键值。

在一些实施例中，装置包括：标记模块，被配置为将待存储数据集中的、除原始数据以及变化数据之外的数据标记为第一数据；以及存储单元，包括：第一存储模块，被配置为根据存储键值，将待存储数据集中、除被标记为第一数据之外的数据存储于存储模型中。

在一些实施例中，存储单元，包括：第二存储模块，被配置为根据存储键值，将原始数据、与原始数据对应的至少一条存续时间，以及变化数据、与变化数据对应的至少一条存续时间存储于存储模型中，并获得存储数据集，其中，存续时间用于表征数据存续的时间周期；以及装置包括：

- 5 第一提取模块，被配置为根据原始数据的存续时间从存储数据集中提取原始数据；第二提取模块，被配置为根据变化数据的存续时间从存储数据集中提取变化数据；以及验证模块，被配置为响应于确定提取的原始数据以及提取的变化数据构成的数据集，与待存储数据集中的待存储数据一致，确定存储数据集为目标存储数据集。

- 10 在一些实施例中，用于存储数据的装置包括：数据库参数获取单元，被配置为根据待存储集中的数据的数据特征，确定数据库参数；以及存储单元，包括：更新模块，被配置为利用数据库参数，更新存储模型；以及第三存储模块，被配置为将原始数据以及变化数据存储于更新后的存储模型中。

- 15 根据本公开的第五方面，提供了一种用于存储数据的装置，包括：预校验单元，被配置为响应于确定待存储数据集满足预设条件，执行如第一方面提供的用于存储数据的方法；其中，预设条件包括以下至少一项：待存储数据集中的数据变化速度小于预设阈值，或者待存储数据集中的数据量大于预设数据量。

- 20 根据本公开的第六方面，提供了一种用于存储数据的装置，包括：时间确认单元，被配置为响应于确定当前时刻处于预设数据处理时间段，执行如第一方面提供的用于存储数据的方法。

- 根据本公开的第七方面，本公开的实施例提供了一种电子设备，包括：至少一个处理器；以及与至少一个处理器通信连接的存储器；其中，存储器存储有可被至少一个处理器执行的指令，该指令被至少一个处理器执行，以使至少一个处理器实现如第一方面提供的用于存储数据的方法、或者实现如第二方面提供的用于存储数据的方法、或者实现如第三方面提供的用于存储数据的方法。
- 25

- 根据本公开的第八方面，本公开的实施例提供了一种计算机可读存储介质，其上存储有计算机指令，其中，计算机指令用于使计算机执行时实
- 30

现如第一方面提供的用于存储数据的方法、或者实现如第二方面提供的用于存储数据的方法、或者实现如第三方面提供的用于存储数据的方法。

根据本公开的第九方面，本公开的实施例提供了一种包括计算机程序的计算机程序产品，该计算机程序在被处理器执行时能够实现如第一方面提供的用于存储数据的方法、或者实现如第二方面提供的用于存储数据的方法、或者实现如第三方面提供的用于存储数据的方法。

根据本公开的技术解决了目前的存储数据的方法存在的存储效率低、以及浪费存储资源的问题。

应当理解，本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征，也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

#### 附图说明

附图用于更好地理解本方案，不构成对本公开的限定。其中：

图 1 是本公开的实施例可以应用于其中的示例性系统架构图；

图 2 是根据本公开的用于存储数据的方法的一个实施例的流程图；

图 3 是根据本公开的用于存储数据的方法的一个实施例的流程图；

图 4 是根据本公开的用于存储数据的方法的一个实施例的流程图；

图 5 是根据本公开的用于存储数据的方法的一个应用场景的流程图；

图 6 是根据本公开的用于存储数据的装置的一个实施例的结构示意图；

图 7 是根据本公开的用于存储数据的装置的一个实施例的结构示意图；

图 8 是根据本公开的用于存储数据的装置的一个实施例的结构示意图；

图 9 是用来实现本公开实施例的用于存储数据的方法的电子设备的框图。

#### 具体实施方式

以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

图1示出了可以应用本公开的用于存储数据的方法或用于存储数据的装置的实施例的示例性系统架构100。

如图1所示，系统架构100可以包括终端设备101、102、103，网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备101、102、103通过网络104与服务器105交互，以接收或发送消息等。终端设备101、102、103上可以安装有各种客户端应用，例如导航类应用、推荐类应用、聊天类应用、购物类应用、金融类应用、图像类应用、视频类应用等。

终端设备101、102、103可以是具有显示屏并且支持接收服务器消息的各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、电子播放器、便携式计算机和台式计算机等等。

终端设备101、102、103可以是硬件，也可以是软件。当终端设备101、102、103为硬件时，可以是各种电子设备，当终端设备101、102、103为软件时，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块（例如用来提供分布式服务的多个软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

服务器105可以通过终端设备获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型中包括目标存储参数；响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值，在待存储数据集中，获取与存储键值对应的原始数据、以及与该存储键值对应的变化数据，并根据存储键值，将原始数据以及变化数据存储于存储模型中。

需要说明的是，本公开的实施例所提供的用于存储数据的方法一般由

服务器 105 执行,相应地,用于存储数据的装置一般设置于服务器 105 中。

应该理解,图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器的。

本公开提供的用于存储数据的方法、装置,获取待存储数据集的存储  
5 键值,获取首次写入待存储数据集中的原始数据、以及待存储数据集中的  
变化数据,并根据存储键值,存储原始数据以及变化数据,可以对周期性  
写入的数据或者历史数据等数据进行轻量化的存储/归档,提高存储数据的  
效率、以及节约存储数据的存储资源。

继续参考图 2,示出了根据本公开的用于存储数据的方法的一个实施  
10 例的流程 200,包括以下步骤:

步骤 201,获取待存储数据集以及存储模型,其中,待存储数据集中  
包括源存储参数,存储模型中包括目标存储参数。

在本实施例中,用于存储数据的方法的执行主体(例如图 1 所示的服  
务器 105)可以通过有线或者无线的方式获取待存储数据集以及存储模型。  
15 其中,存储模型用于存储待存储数据,或者用于对待存储数据进行归档、  
整理或者持久化存储,存储模型可以是用于存储数据的数据库或者是用于  
存储数据的文件等等数据存储载体,存储模型中包括与待存储数据对应的  
目标存储参数,该目标存储参数可以是预先设置的、也可以是存储模型通  
过字段识别方法从待存储数据集中识别获取到的。

20 在本实施例中,待存储数据集中包括至少一个源存储参数以及与至少  
一个源存储参数中的每一个源存储参数分别对应的待存储数据/待归档数  
据,例如,源存储参数可以是物品编号、对应的待存储数据可以是编号  
“001”,又如,源存储参数可以是网络类型、对应的待存储数据可以是  
“无线网络/有线网络”等等。

25 步骤 202,响应于确定源存储参数与目标存储参数一致,获取待存储  
数据集的存储键值。

在本实施例中,若确定源存储参数与目标存储参数一致,则可以获取  
待存储数据集的存储键值。存储键值是指待存储数据的索引关键字、键值  
对数据库中的键值、或者将待存储数据划分为不同类别的数据标签,例如,  
30 若待存储数据集中的待存储数据为某电商平台的库存信息,则存储键值可

以是物品的类别或者物品的尺寸等；又如，若待存储数据集中的待存储数据为某学校的考试成绩，则存储键值可以是考试科目、或者考试成绩区间段。

步骤 203，获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据。

在本实施例中，获取首次写入待存储数据集中与源存储参数对应的原始数据，以及原始数据在被写入待存储数据集后、被执行了数据更新操作或者数据覆盖操作后的变化数据。

例如，若某仓库每天会向待存储数据集中写入当天的库存信息，第一天至第五天写入的与各个源存储参数对应的各条库存信息均相同，即，仓库的库存在第一天至第五天均未发生变化，第六天库存发生变化（如，存货增加或者存货减少等），第七天的库存信息相比于第六天的库存信息未发生变化。则第一天写入的库存信息为首次写入待存储数据集中的原始数据，第六天写入的库存信息为原始数据被执行了数据更新操作后的变化数据。所以获取的原始数据与变化数据，即为该仓库第一天写入的数据、以及第六天写入的数据。

又如，待存储数据集为某网络故障表单，待存储数据为该网络的故障信息，第一次进行网络检测时写入故障表单的故障信息与第二次、第三次进行网络检测时写入故障表单的故障信息均相同（即，该网络在第二次、第三次进行网络检测时检测出的网络故障，与第一次进行网络检测时检测出的网络故障均相同），第四次进行网络检测时检测出了前述检测过程中未出现的新故障。则第一次进行网络检测时写入故障表单的故障信息为原始数据，第四次进行网络检测时写入故障表单的故障信息为变化数据。所以获取的原始数据与变化数据，即为首次进行网络检测时写入待存储数据集的数据、以及第四次进行网络检测时写入待存储数据集中的数据。

步骤 204，根据存储键值，将原始数据以及变化数据存储于存储模型中。

在本实施例中，根据存储键值，将获取到的原始数据以及变化数据存储于存储模型中。由于存储模型中的目标存储参数与源存储参数相同，原始数据以及变化数据是与其各自的源存储参数对应的数据，在将原始数据

以及变化数据存储于存储模型时，可以将原始数据以及变化数据，与其各自对应的目标存储参数进行关联存储。

例如，源存储参数以及对应的原始数据为“库存剩余数目、100个”，源存储参数、以及对应的变化数据为“库存更新时间、星期一”，由于存储模型中的目标存储参数是与源存储参数一致的存储参数，则存储模型中包括目标存储参数“库存数目”以及“库存更新时间”，此时，在将原始数据以及变化数据存储于存储模型时，可以将原始数据“100个”与目标存储参数“库存数目”进行关联存储，将变化数据“星期一”与目标存储数据“库存更新时间”进行关联存储，以方便使用存储模型进行数据查询。也可以将原始数据以及变化数据直接存入存储模型中，而不进行与目标存储参数的关联存储，以简化数据存储步骤、提高数据存储效率。

在本实施例中，存储键值可以是多个目标存储参数中的一个存储参数，例如，可以将目标存储参数“物品名称”作为存储键值，之后，将包含相同物品名称的原始数据以及变化数据存储于同一存储模型中，或者存储于存储模型中的同一子单元中。存储键值也可以是与目标存储参数不相关的信息。例如，目标存储参数可以是“网络容量”、“网络编号”、“网络状态”，此时，可以将“网络类型”作为存储键值，之后，将与同一网络类型对应的原始数据以及变化数据存储于同一存储模型中，或者存储于存储模型中的同一子单元中。

本实施例提供的用于存储数据的方法，获取待存储数据集以及存储模型，若确定待存储数据集中的源存储参数与存储模型中的目标存储参数一致，获取待存储数据集的存储键值，获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中与源存储参数对应的变化数据，并根据存储键值，存储原始数据以及变化数据，可以对周期性写入的数据（如，每天写入数据集中的当天的网络检测数据）或者历史数据等数据进行轻量化的存储/归档，可以提高数据存储的效率、以及节约存储数据的软件/硬件资源。

可选地，用于存储数据的方法包括：响应于确定源存储参数与目标存储参数不一致，采用源存储参数与目标存储参数确定缺失参数，其中，缺失参数包括：待存储数据集中存在、且存储模型中不存在的存储参数，或

者待存储数据集中不存在、且存储模型中存在的存储参数。

在本实施例中，若确定源存储参数与目标存储参数不一致，则可以将源存储参数与目标存储参数进行逐一对比，若某一存储参数为源存储参数，但该存储参数不是目标存储参数，则将该存储参数确定为缺失参数；

- 5 或者，若某一存储参数为目标存储参数，但该存储参数不是源存储参数，则将该存储参数确定为缺失参数。即，将待存储数据集中存在、但存储模型中不存在的存储参数确定为缺失参数，或者将待存储数据集中不存在、但存储模型中存在的存储参数确定为缺失参数。

- 10 本实施例通过将源存储参数与目标存储参数进行一致性校验，可以确定出源存储参数中缺失的存储参数或者可以确定出目标存储参数中缺失的存储参数，可以在对数据进行归档存储的过程中获取缺失参数以及对应的数据，从而提高存储数据的准确性。

- 15 可选地，用于存储数据的方法包括：响应于确定存储模型中不存在缺失参数，将缺失参数添加至存储模型；或者，响应于确定待存储数据集中不存在缺失参数，利用待存储数据集的原始获取方式，获取与缺失参数对应的待存储数据，并将缺失参数以及与缺失参数对应的待存储数据添加至待存储数据集中。

- 20 在本实施例中，若确定存储模型中不存在缺失参数，即，存储模型中缺少包含于待存储数据集中的存储参数，则可以将该缺失参数添加至存储模型中，以确保存储模型中的目标存储参数、与待存储数据集中的源存储参数的一致性。

- 25 在本实施例中，若确定待存储数据集中不存在缺失参数，即，待存储模型中缺少包含于存储模型中的存储参数，则可以利用待存储数据集的原始获取方式，获取与缺失参数对应的待存储数据，并将该缺失参数以及对应的待存储数据添加至待存储数据集中。例如，若确定待存储数据集中不存在缺失参数“订单编号”，则可以利用待存储数据集的原始获取方式，获取与“订单编号”对应的编号“001”、编号“002”、编号“003”等待存储数据，并将缺失参数“订单编号”以及“001”、“002”“003”添加至待存储数据集中。本例中，待存储数据集的原始获取方式可以是如  
30 读取电商平台的订单流信息、可以是读取物流系统的客户订单信息等方

式。

本实施例通过将存储模型中的缺失参数进行补充，或者通过将待存储数据集中的缺失参数以及对应的数据补充至待存储/待归档的待存储数据集中，可以提高存储待存储数据的准确性以及确保存储数据的完整性。

- 5       可选地，获取待存储数据集的存储键值，包括：针对至少一个预设存储键值中的每一个预设存储键值，获取根据该预设存储键值对待存储数据集中的待存储数据进行聚类后的、待存储数据的分布结果；响应于确定分布结果满足预设条件，确定该预设存储键值为目标存储键值。

在本实施例中，可以预先设定多个预设存储键值，并针对多个预设存储键值中的每一个存储键值，获取根据该预设存储键值对待存储数据集中的待存储数据进行聚类后的、待存储数据的分布结果，若确定与该预设存储键值对应的分布结果满足预设条件，则可以将该预设存储键值确定为目标存储键值。

例如，可以预先设定第一存储键值、第二存储键值以及第三存储键值，  
15   根据这三个预设的存储键值分别对待存储数据集中的待存储数据进行聚类，并分别获得第一分布结果、第二分布结果以及第三分布结果。第一分布结果为待存储数据被划分为 10 类，第二分布结果为待存储数据被划分为 5 类，第三分布结果为待存储数据被划分为 2 类，若预设条件为聚类后的待存储数据的分布不超过 7 类，则第一存储键值与第二存储键值中的一项可以作为目标存储键值，若预设条件为根据该存储键值对待存储数据进行聚类后、待存储数据的分布在每一个分布结果中属于最集中的分布，则  
20   本例中，第一存储键值为目标存储键值。

本实施例中，根据待存储数据基于预设存储键值进行聚类、得到的聚类后的待存储数据的分布情况，确定目标存储键值，并基于目标存储键值  
25   进行数据存储，可以提高数据的压缩比例，进而节约数据存储资源。

可选地，用于存储数据的方法包括：将待存储数据集中的、除原始数据以及变化数据之外的数据标记为第一数据；根据存储键值，将原始数据以及变化数据存储于存储模型中，包括：根据存储键值，将待存储数据集中、除被标记为第一数据之外的数据存储于存储模型中。

30       在本实施例中，可以将待存储数据集中，除原始数据以及变化数据之

外的数据均标记为第一数据（例如“删除数据”或者“0”），并根据存储键值将待存储数据集中，除被标记为第一数据之外的数据存储于存储模型中。或者将待存储数据集中标记为第一数据的数据均执行删除操作，之后根据存储键值，存储执行删除操作后的待存储数据集中的数据。

- 5       本实施例中，对待存储数据集中的待存储数据进行标记，删除被标记的数据，并根据存储键值对待存储数据集中未被删除的数据进行排序以完成存储/归档，可以基于待存储数据已占用的写入空间对待存储数据进行数据整理/处理，可以节约系统资源。

- 10       在上述结合图2描述的实施例的一些可选的实现方式中，根据存储键值，将原始数据以及变化数据存储于存储模型中，包括：根据存储键值，将原始数据、与原始数据对应的至少一条存续时间，以及变化数据、与变化数据对应的至少一条存续时间存储于存储模型中，并获得存储数据集，其中，存续时间用于表征数据存续的时间周期；用于存储数据的方法包括：根据原始数据的存续时间从存储数据集中提取原始数据；根据变化数据的  
15       存续时间从存储数据集中提取变化数据；响应于确定提取的原始数据以及提取的变化数据构成的数据集，与待存储数据集一致，确定存储数据集为目标存储数据集。

在本实施例中，可以根据存储键值，存储原始数据和原始数据的存续时间，以及变化数据和变化数据的存续时间，以获得存储数据集。

- 20       在本实施例中，存续时间用于表征数据存续的时间周期。具体地，原始数据的存续时间是指原始数据存续的至少一个时间周期，变化数据的存续时间是指变化数据存续的至少一个时间周期。

- 25       例如，若某仓库每天会向待存储数据集中写入当天的库存信息，第一天至第五天写入的库存信息均相同，即，仓库的库存在第一天至第五天均未发生变化，第六天库存发生变化（如，存货增加或者存货减少等），第七天的库存信息相比于第六天的库存信息未发生变化。则，原始数据对应5个存续时间，分别为第一天、第二天、第三天、第四天以及第五天，变化数据对应2个存续时间，分别为第六天以及第七天。可以理解，当数据发生多次变化时，可以存在多组变化数据，每一组变化数据可以对应多个  
30       存续时间（如，第一组变化数据对应第六天以及第七天，第二组变化数据

对应第八天、第九天)。

在本实施例中，可以根据原始数据的各个存续时间从存储数据集中提取原始数据，以及根据变化数据的各个存续时间从存储数据集中提取变化数据。若确定提取出的原始数据以及变化数据构成的数据集、与待存储数据  
5 数据集一致，则可以确定存储数据集为目标存储数据集，即，可以确定存储数据集所存储的数据通过数据校验。

可以理解，待存储数据集中是周期性写入的待存储数据（如，每天写入数据集中的当天的网络检测数据，或者每天写入数据集中的当天的库存信息），存储数据集中存储的是原始数据和原始数据的存续时间，以及变  
10 化数据和变化数据的存续时间（如，初始库存和该初始库存对应的各个存续时间，以及变化后的库存以及变化后的库存对应的各个存续时间）。若根据原始数据的各个存续时间从存储数据集中提取出各条原始数据，以及根据变化数据的各个存续时间从存储数据集中提取出各条变化数据，拼接/构建成的数据集（如，将原始数据按照对应的各个存续时间，以及将变化  
15 数据按照对应的存续时间一条一条的拼接）、与待存储数据集一致，则可以确定存储数据集中存储了待存储数据集中的全部数据、存储数据集存储的信息准确、或者存储数据集中的数据通过校验。

在本实施例中，确定提取出的原始数据以及变化数据构成的数据集、与待存储数据集是否一致的方法可以是，确定二者之间每个存续时间对应的  
20 的数据的数据量是否一致，也可以是确定二者之间每个存续时间对应的数据所表征的信息是否一致。

本实施例通过对存储数据集进行数据校验，可以提高存储数据的准确性。

在上述结合图 2 描述的实施例的一些可选的实现方式中，用于存储数  
25 据的方法包括：根据待存储集中的数据的数据特征，确定数据库参数；将原始数据以及变化数据存储于存储模型中，包括：利用数据库参数，更新存储模型；将原始数据以及变化数据存储于更新后的存储模型中。

在本实施例中，可以根据待存储数据集中的数据的数据特征，确定用于存储具有该数据特征的数据的数据库参数，并利用该数据库参数更新用  
30 于存储具有该数据特征的数据的存储模型。

例如，若对待存储数据集中的数据的精度需求低，则可以将存储模型的存储压缩比这一数据库参数调高，以提高存储数据的效率以及节省存储空间。又如，若待存储数据集中的数据量大于预设阈值，则可以将存储模型的读写线程数目这一数据参数调高，以提高存储数据的效率。

- 5       本实施例中，根据待存储数据集中数据的数据特征确定存储模型的参数，以基于待存入存储模型的数据的数据特征优化存储模型的参数，可以提高存储数据的效率。

继续参考图 3，示出了根据本公开的用于存储数据的方法的一个实施例的流程 300，包括以下步骤：

- 10       步骤 301，响应于确定待存储数据集满足预设条件，获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型中包括目标存储参数；其中，预设条件包括以下至少一项：待存储数据集中的数据的变化速度小于预设阈值，或者待存储数据集中的数据量大于预设数据量。

- 15       在本实施例中，当用于存储数据的方法的执行主体（例如图 1 所示的服务器 105）在确定待存储数据集满足预设条件后，获取待存储数据集以及存储模型。

- 20       在本实施例中，预设条件可以是待存储数据集中的数据的变化速度小于预设阈值，例如，在待存储数据集中，每天的数据变化量占数据总量的比例小于预设比例（如 1%或者其他百分数值）。预设条件可以是待存储数据集中的数据量大于预设数据量，例如，待存储数据集中的数据量大于 40TB（或者其他数据量）。

步骤 302，响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值。

- 25       步骤 303，获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据。

步骤 304，根据存储键值，将原始数据以及变化数据存储于存储模型中。

- 30       本实施例中对步骤 302、步骤 303、步骤 304 的描述与步骤 202、步骤 203、步骤 204 的描述一致，此处不再赘述。

本实施例提供的用于存储数据的方法，在执行如图 2 实施例的用于存储数据的方法之前，需要先判断待存储数据集中的待存储数据是否满足预设条件，即待存储数据的变化速度是否小于预设阈值和/或待存储数据的数据量是否大于预设数据量。可以提高应用如图 2 实施例的用于存储数据的方法进行数据存储的存储资源利用率。

需要说明的是，在判断待存储数据集满足预设条件后，图 2 所述的各个实施例的方法适用于图 3 所述实施例的方法。

继续参考图 4，示出了根据本公开的用于存储数据的方法的一个实施例的流程 400，包括以下步骤：

10 步骤 401，响应于确定当前时刻处于预设数据处理时间段，获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型中包括目标存储参数。

在本实施例中，当用于存储数据的方法的执行主体（例如图 1 所示的服务器 105）确定当前时刻处于预设数据处理时间段内时，获取待存储数据集以及存储模型。其中，预设数据处理时间段为预设的、将待存储数据集中的数据存储于存储模型中的时间段，即，对待存储数据集中的数据执行数据归档操作的时间段。该预设数据处理时间段可以是与待存储数据集中的待存储数据的生成期间不同的时段。例如，若待存储数据为某电商平台的订单数据，则由于订单数据通常是在白天时段生成，此时可以将夜间时段作为预设数据处理时间段。又如，若待存储数据为某网络供应商升级设备时所生成的系统数据流，则由于网络供应商升级设备通常是夜间时段，此时可以将白天时段作为预设数据处理时段。

步骤 402，响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值。

25 步骤 403，获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据。

步骤 404，根据存储键值，将原始数据以及变化数据存储于存储模型中。

本实施例中对步骤 402、步骤 403、步骤 404 的描述与步骤 202、步骤 30 203、步骤 204 的描述一致，此处不再赘述。

本实施例提供的用于存储数据的方法，在执行如图 2 实施例的用于存储数据的方法之前，需要先判断当前时刻是否处于预设数据处理时段，以在预设数据处理时段内进行待存储数据集中数据的存储/归档，可以使数据存储/数据归档时段避开待存储数据集的数据写入时段，或者待存储数据的生成时段，避免数据存储操作对主业务（主业务运行时才生成待存储数据）运行所需资源（如服务器计算资源、或者通信资源等）的占用。

需要说明的是，在判断当前时刻处于预设数据处理时段后，图 2 所述的各个实施例的方法适用于图 4 实施例的方法。

在一些应用场景中，如图 5 所示，其示出了用于存储数据的方法的一个具体的应用场景 500，包括：

步骤 501，待存储数据集校验，即确定可执行对象，用于确定待存储数据集是否可使用该存储数据的方法进行数据存储。可以确定其数据存储量是否达到一定规模，例如超过 100TB，以及确定待存储数据集中的数据是否为缓慢变化的数据，例如每天变化的记录数占总记录数的比例小于 1%。

步骤 502，待存储数据校验，获取待存储数据集中的源存储参数，以及存储模型中的目标存储参数，并将二者之间不一致的存储参数确定为缺失参数，若待存储数据集中不存在缺失参数，则可以根据待存储数据集的原始获取方式，获取与缺失参数对应的待存储数据，并将该缺失参数以及对应的待存储数据补充至待存储数据集中，以在存储数据之前确保待存储数据集的完整性；若存储模型中不存在缺失参数，则可以将该缺失参数添加至存储模型中，以确定存储模型中的目标存储参数与源存储参数的一致性。

步骤 503，数据处理，主要包括步骤：输入参数确定、数据打标、数据重排和归档数据校验。

步骤 5031，输入参数确定，用于确定业务主键、数据处理时间范围、存储数据库的参数、归档天数和存储键值。

业务主键是根据与待存储数据集所关联的业务确定（例如，可以将订单编号确定为电商平台关联的待存储数据的业务主键），其是判断该记录是否发生变化的依据，其不同记录周期的信息是否一致，是否有过变化。

数据处理时间范围是用来判断用于存储数据的方法可以执行的时间范围。由于服务器集群的资源比较紧张，批量处理历史数据（即待存储数据）会占用服务器集群资源，所以要避开服务器集群生成业务数据的主要时段（例如 10:00-22:00），以免对生成业务数据造成影响。若发现存储任务在该时间段之外的时间启动，则会强制停止执行该任务。

存储数据库（即存储模型）的参数是针对每个待存储数据集有针对性的优化存储数据库的参数，以保证存储数据的效率。

归档天数用于表征每次存储的数据量的时间段（如存储哪几天的数据/存储哪几个月的数据），该参数可以根据待存储数据集中的数据量或者存储数据库的存储空间大小来确定。

存储键值是针对常用字段进行抽样统计其分布，筛选出具有高压缩比的字段，其确定方式可以是根据字段对应的数据分布的集中情况，数据分布越集中，则对应的字段的压缩比越高。例如，在商品模型中，物品的一级品类的压缩比、比三级品类作为排序字段的压缩比更高。

步骤 5032，数据标记，保留待存储数据集中变化前最早的数据（即原始数据）和变化后的数据（即变化数据），在此中间的数据标记为删除。例如，若待存储数据集中的数据为库存量最近 7 天的数据情况为：第 1-5 天没有发生变化，第 6 天发生变化，第 7 天无变化，则处理的结果为将除第 1 天和第 6 天之外的、其他 5 天的数据标记为删除。

步骤 5033，数据重生成，根据数据标记的结果，将标记为删除的数据删除后，再根据存储键值对数据进行重新排序后，将数据写入对应的归档存储空间（对应的存储模型，或者存储模型中对应的存储单元）中，完成对待存储数据的存储。该步骤可以将数据进行 ORC 格式（Optimized Row Columnar，优化行列格式）的存储，以进一步降低存储模型的存储占比。

步骤 5034，归档数据校验，用于确保待存储数据集中的数据、与存储后的数据的一致性，该步骤可以根据通过判断数据存储前与数据存储后的数据量是否一致确定一致性（即，待存储数据集中的数据量与存储模型中的数据量是否一致），也可以通过判断存储前数据与存储后数据每一条记录是否一致确定一致性。

步骤 504，数据同步，判断待存储数据集是否在其他服务器集群中存

在备份数据，如存在，则将存储后的数据（或者将存储模型）也同样在该其他服务器集群中进行备份，以实现数据同步。

步骤 505，数据清理，经过上述步骤对待存储数据集中的数据完成存储/归档后，删除待存储数据集以释放服务器资源。

5 进一步参考图 6，作为对上述各图所示方法的实现，本公开提供了一种用于存储数据的装置的一个实施例，该装置实施例与图 2 所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图 6 所示，本实施例的用于存储数据的装置 600，包括：第一获取单元 601、第二获取单元 602、第三获取单元 603、存储单元 604。其中，  
10 第一获取单元，被配置为获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型中包括目标存储参数；第二获取单元，被配置为响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值；第三获取单元，被配置为获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的  
15 变化数据；存储单元，被配置为根据存储键值，将原始数据以及变化数据存储于存储模型中。

在一些实施例中，用于存储数据的装置包括：数据校验单元，被配置为响应于确定源存储参数与目标存储参数不一致，采用源存储参数以及目标存储参数确定缺失参数，其中，缺失参数包括：待存储数据集中存在、  
20 且存储模型中不存在的存储参数，或者待存储数据集中不存在、且存储模型中存在的存储参数。

在一些实施例中，用于存储数据的装置包括：第一数据校验模块，被配置为响应于确定存储模型中不存在缺失参数，将缺失参数添加至存储模型；或者，第二数据校验模块，被配置为响应于确定待存储数据集中不存在缺失参数，利用待存储数据集的原始获取方式，获取与缺失参数对应的  
25 待存储数据，并将缺失参数以及与缺失参数对应的待存储数据添加至待存储数据集中。

在一些实施例中，第二获取单元，包括：获取模块，被配置为针对至少一个预设存储键值中的每一个预设存储键值，获取根据该预设存储键值  
30 对待存储数据集中的待存储数据进行聚类后的、待存储数据的分布结果；

确定模块，被配置为响应于确定分布结果满足预设条件，确定该预设存储键值为目标存储键值。

5 在一些实施例中，装置包括：标记模块，被配置为将待存储数据集中的、除原始数据以及变化数据之外的数据标记为第一数据；存储单元，包括：第一存储模块，被配置为根据存储键值，将待存储数据集中、除被标记为第一数据之外的数据存储于存储模型中。

10 在一些实施例中，存储单元，包括：第二存储模块，被配置为根据存储键值，将原始数据、与原始数据对应的至少一条存续时间，以及变化数据、与变化数据对应的至少一条存续时间存储于存储模型中，并获得存储数据集，其中，存续时间用于表征数据存续的时间周期；装置包括：第一提取模块，被配置为根据原始数据的存续时间从存储数据集中提取原始数据；第二提取模块，被配置为根据变化数据的存续时间从存储数据集中提取变化数据；验证模块，被配置为响应于确定提取的原始数据以及提取的变化数据构成的数据集，与待存储数据集中的待存储数据一致，确定存储数据集为目标存储数据集。

15 在一些实施例中，用于存储数据的装置包括：数据库参数获取单元，被配置为根据待存储集中的数据的数据特征，确定数据库参数；存储单元，包括：更新模块，被配置为利用数据库参数，更新存储模型；第三存储模块，被配置为将原始数据以及变化数据存储于更新后的存储模型中。

20 上述装置 600 中的各单元与参考图 2 描述的方法中的步骤相对应。由此上文针对用于存储数据的方法描述的操作、特征及所能达到的技术效果同样适用于装置 600 及其中包含的单元，在此不再赘述。

25 进一步参考图 7，作为对上述各图所示方法的实现，本公开提供了一种用于存储数据的装置的一个实施例，该装置实施例与图 3 所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图 7 所示，本实施例的用于存储数据的装置 700，包括：预校验单元 701、第一获取单元 702、第二获取单元 703、存储单元 704。其中，预校验单元，被配置为响应于确定待存储数据集满足预设条件，获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型  
30 中包括目标存储参数；其中，预设条件包括以下至少一项：待存储数据集

中的数据的变化速度小于预设阈值，或者待存储数据集中的数据量大于预设数据量。第一获取单元，被配置为响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值；第二获取单元，被配置为获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据；存储单元，被配置为根据存储键值，将原始数据以及变化数据存储于存储模型中。

上述装置 700 中的各单元与参考图 3 描述的方法中的步骤相对应。由此上文针对用于存储数据的方法描述的操作、特征及所能达到的技术效果同样适用于装置 700 及其中包含的单元，在此不再赘述。

进一步参考图 8，作为对上述各图所示方法的实现，本公开提供了一种用于存储数据的装置的一个实施例，该装置实施例与图 4 所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图 8 所示，本实施例的用于存储数据的装置 800，包括：时间确认单元 801、第一获取单元 802、第二获取单元 803、存储单元 804。其中，时间确认单元，被配置为响应于确定当前时刻处于预设数据处理时间段，获取待存储数据集以及存储模型，其中，待存储数据集中包括源存储参数，存储模型中包括目标存储参数；第一获取单元，被配置为响应于确定源存储参数与目标存储参数一致，获取待存储数据集的存储键值；第二获取单元，被配置为获取首次写入待存储数据集中的与源存储参数对应的原始数据、以及待存储数据集中的与源存储参数对应的变化数据；以及存储单元，被配置为根据存储键值，将原始数据以及变化数据存储于存储模型中。

上述装置 800 中的各单元与参考图 4 描述的方法中的步骤相对应。由此上文针对用于存储数据的方法描述的操作、特征及所能达到的技术效果同样适用于装置 800 及其中包含的单元，在此不再赘述。

根据本公开的实施例，本公开还提供了一种电子设备和一种可读存储介质。

如图 9 所示，是根据本公开实施例的用于存储数据的方法的电子设备 900 的框图。电子设备旨在表示各种形式的数字计算机，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装

置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例，并且不意在限制本文中描述的和/或者要求的本公开的实现。

如图 9 所示，该电子设备包括：一个或多个处理器 901、存储器 902，  
5 以及用于连接各部件的接口，包括高速接口和低速接口。各个部件利用不同的总线互相连接，并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理，包括存储在存储器中或者存储器上以在外部输入/输出装置（诸如，耦合至接口的显示设备）上显示 GUI 的图形信息的指令。在其它实施方式中，若需要，可以将  
10 多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样，可以连接多个电子设备，各个设备提供部分必要的操作（例如，作为服务器阵列、一组刀片式服务器、或者多处理器系统）。图 9 中以一个处理器 901 为例。

存储器 902 即为本公开所提供的非瞬时计算机可读存储介质。其中，  
15 该存储器存储有可由至少一个处理器执行的指令，以使该至少一个处理器执行本公开所提供的用于存储数据的方法。本公开的非瞬时计算机可读存储介质存储计算机指令，该计算机指令用于使计算机执行本公开所提供的用于存储数据的方法。

存储器 902 作为一种非瞬时计算机可读存储介质，可用于存储非瞬时  
20 软件程序、非瞬时计算机可执行程序以及模块，如本公开实施例中的用于存储数据的方法对应的程序指令/模块（例如，附图 6 所示的第一获取单元 601、第二获取单元 602、第三获取单元 603、存储单元 604）。处理器 901 通过运行存储在存储器 902 中的非瞬时软件程序、指令以及模块，从而执行服务器的各种功能应用以及数据处理，即实现上述方法实施例中的用于  
25 存储数据的方法。

存储器 902 可以包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需要的应用程序；存储数据区可存储根据用于提取视频片段的电子设备的使用所创建的数据等。此外，存储器 902  
30 可以包括高速随机存取存储器，还可以包括非瞬时存储器，例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中，

存储器 902 可选包括相对于处理器 901 远程设置的存储器，这些远程存储器可以通过网络连接至用于提取视频片段的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

用于存储数据的方法的电子设备还可以包括：输入装置 903、输出装置 904 以及总线 905。处理器 901、存储器 902、输入装置 903 和输出装置 904 可以通过总线 905 或者其他方式连接，图 9 中以通过总线 905 连接为例。

输入装置 903 可接收输入的数字或字符信息，以及产生与用于提取视频片段的电子设备的用户设置以及功能控制有关的键信号输入，例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置 904 可以包括显示设备、辅助照明装置（例如，LED）和触觉反馈装置（例如，振动电机）等。该显示设备可以包括但不限于，液晶显示器（LCD）、发光二极管（LED）显示器和等离子体显示器。在一些实施方式中，显示设备可以是触摸屏。

此处描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、专用 ASIC（专用集成电路）、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括：实施在一个或者多个计算机程序中，该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释，该可编程处理器可以是专用或者通用可编程处理器，可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令，并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

这些计算程序（也称作程序、软件、软件应用、或者代码）包括可编程处理器的机器指令，并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的，术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置（例如，磁盘、光盘、存储器、可编程逻辑装置（PLD）），包括，接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

为了提供与用户的交互，可以在计算机上实施此处描述的系统和技术，该计算机具有：用于向用户显示信息的显示装置（例如，CRT（阴极射线管）或者 LCD（液晶显示器）监视器）；以及键盘和指向装置（例如，鼠标或者轨迹球），用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互；例如，提供给用户的反馈可以是任何形式的传感反馈（例如，视觉反馈、听觉反馈、或者触觉反馈）；并且可以用任何形式（包括声输入、语音输入或者、触觉输入）来接收来自用户的输入。

可以将此处描述的系统和技术实施在包括后台部件的计算系统（例如，作为数据服务器）、或者包括中间件部件的计算系统（例如，应用服务器）、或者包括前端部件的计算系统（例如，具有图形用户界面或者网络浏览器的用户计算机，用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互）、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信（例如，通信网络）来将系统的部件相互连接。通信网络的示例包括：局域网（LAN）、广域网（WAN）和互联网。

计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。

应该理解，可以使用上面所示的各种形式的流程，重新排序、增加或删除步骤。例如，本公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行，只要能够实现本公开公开的技术方案所期望的结果，本文在此不进行限制。

上述具体实施方式，并不构成对本公开保护范围的限制。本领域技术人员应该明白的是，根据设计要求和因素，可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等，均应包含在本公开保护范围之内。

## 权 利 要 求 书

1、一种用于存储数据的方法，包括：

5 获取待存储数据集以及存储模型，其中，所述待存储数据集中包括源存储参数，所述存储模型中包括目标存储参数；

响应于确定所述源存储参数与所述目标存储参数一致，获取所述待存储数据集的存储键值；

10 获取首次写入所述待存储数据集中的与所述源存储参数对应的原始数据、以及所述待存储数据集中的与所述源存储参数对应的变化数据；以及

根据所述存储键值，将所述原始数据以及所述变化数据存储于所述存储模型中。

2、根据权利要求 1 所述的方法，其中，所述方法包括：

15 响应于确定所述源存储参数与所述目标存储参数不一致，采用所述源存储参数以及所述目标存储参数确定缺失参数，其中，所述缺失参数包括：所述待存储数据集中存在、且所述存储模型中不存在的存储参数，或者所述待存储数据集中不存在、且所述存储模型中存在的存储参数。

20 3、根据权利要求 2 所述的方法，其中，所述方法包括：

响应于确定所述存储模型中不存在所述缺失参数，将所述缺失参数添加至所述存储模型；或者，

25 响应于确定所述待存储数据集中不存在所述缺失参数，利用所述待存储数据集的原始获取方式，获取与所述缺失参数对应的待存储数据，并将所述缺失参数以及与所述缺失参数对应的待存储数据添加至所述待存储数据集中。

4、根据权利要求 1 所述的方法，其中，所述获取待存储数据集的存储键值，包括：

30 针对至少一个预设存储键值中的每一个预设存储键值，获取根据该预

设存储键值对所述待存储数据集中的待存储数据进行聚类后的、所述待存储数据的分布结果；以及

响应于确定所述分布结果满足预设条件，确定该预设存储键值为目标存储键值。

5

5、根据权利要求1所述的方法，其中，所述方法包括：

将所述待存储数据集中的、除所述原始数据以及所述变化数据之外的数据标记为第一数据；以及

所述根据所述存储键值，将所述原始数据以及所述变化数据存储于所述存储模型中，包括：

根据所述存储键值，将所述待存储数据集中、除被标记为所述第一数据之外的数据存储于所述存储模型中。

6、根据权利要求1所述的方法，其中，所述根据所述存储键值，将所述原始数据以及所述变化数据存储于所述存储模型中，包括：

根据所述存储键值，将所述原始数据、与所述原始数据对应的至少一条存续时间，以及所述变化数据、与所述变化数据对应的至少一条存续时间存储于所述存储模型中，并获得存储数据集，其中，所述存续时间用于表征数据存续的时间周期；以及

20 所述方法包括：

根据所述原始数据的存续时间从所述存储数据集中提取原始数据；

根据所述变化数据的存续时间从所述存储数据集中提取变化数据；以及

25 响应于确定提取的原始数据以及提取的变化数据构成的数据集，与所述待存储数据集中的待存储数据一致，确定所述存储数据集为目标存储数据集。

7、根据权利要求1所述的方法，其中，所述方法包括：

根据所述待存储集中的数据的数据特征，确定数据库参数；以及

30 所述将所述原始数据以及所述变化数据存储于所述存储模型中，包

括:

利用所述数据库参数,更新所述存储模型;以及  
将所述原始数据以及所述变化数据存储于更新后的存储模型中。

5       8、一种用于存储数据的方法,包括:

响应于确定待存储数据集满足预设条件,执行权利要求 1-7 中任一项所述的方法;

其中,所述预设条件包括以下至少一项:所述待存储数据集中的数据的变化速度小于预设阈值,或者所述待存储数据集中的数据量大于预设数据量。  
10

9、一种用于存储数据的方法,包括:

响应于确定当前时刻处于预设数据处理时间段,执行权利要求 1-7 中任一项所述的方法。

15

10、一种用于存储数据的装置,包括:

第一获取单元,被配置为获取待存储数据集以及存储模型,其中,所述待存储数据集中包括源存储参数,所述存储模型中包括目标存储参数;

第二获取单元,被配置为响应于确定所述源存储参数与所述目标存储参数一致,获取所述待存储数据集的存储键值;  
20

第三获取单元,被配置为获取首次写入所述待存储数据集中的与所述源存储参数对应的原始数据、以及所述待存储数据集中的与所述源存储参数对应的变化数据;以及

存储单元,被配置为根据所述存储键值,将所述原始数据以及所述变化数据存储于所述存储模型中。  
25

11、一种用于存储数据的装置,包括:

预校验单元,被配置为响应于确定待存储数据集满足预设条件,执行权利要求 1-7 中任一项所述的方法;

其中,所述预设条件包括以下至少一项:所述待存储数据集中的数据  
30

的变化速度小于预设阈值，或者所述待存储数据集中的数据量大于预设数据量。

12、一种用于存储数据的装置，包括：

- 5       时间确认单元，被配置为响应于确定当前时刻处于预设数据处理时间段，执行权利要求 1-7 中任一项所述的方法。

13、一种电子设备，包括：

至少一个处理器；以及

- 10       与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求 1-9 中任一项所述的方法。

- 15       14、一种存储有计算机指令的非瞬时计算机可读存储介质，其中，所述计算机指令用于使所述计算机执行权利要求 1-9 中任一项所述的方法。

15. 一种计算机程序产品，包括计算机程序，所述计算机程序在被处理器执行时实现根据权利要求 1-9 中任一项所述的方法。

20