

(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
2023 年 10 月 5 日 (05.10.2023)



(10) 国际公布号
WO 2023/185802 A1

(51) 国际专利分类号:
G06F 16/27 (2019.01)

CN; 中国浙江省杭州市西湖区转塘科技经济区块12号, Zhejiang 310024 (CN)。

(21) 国际申请号: PCT/CN2023/084287

(72) 发明人: 李伟 (LI, Wei); 中国浙江省杭州市西湖区三墩镇灯彩街1008号云谷园区, Zhejiang 310030 (CN)。

(22) 国际申请日: 2023 年 3 月 28 日 (28.03.2023)

(25) 申请语言: 中文

(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, Beijing 100082 (CN)。

(26) 公布语言: 中文

(30) 优先权:
202210326055.6 2022年3月29日 (29.03.2022) CN

(71) 申请人: 阿里巴巴 (中国) 有限公司 (ALIBABA (CHINA) CO., LTD.) [CN/CN]; 中国浙江省杭州市余杭区五常街道文一西路969号3幢5层554室, Zhejiang 311121 (CN)。 阿里云计算有限公司 (ALIBABA CLOUD COMPUTING LTD.) [CN/

(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,

(54) Title: DATA PROCESSING METHOD AND APPARATUS

(54) 发明名称: 数据处理方法及装置

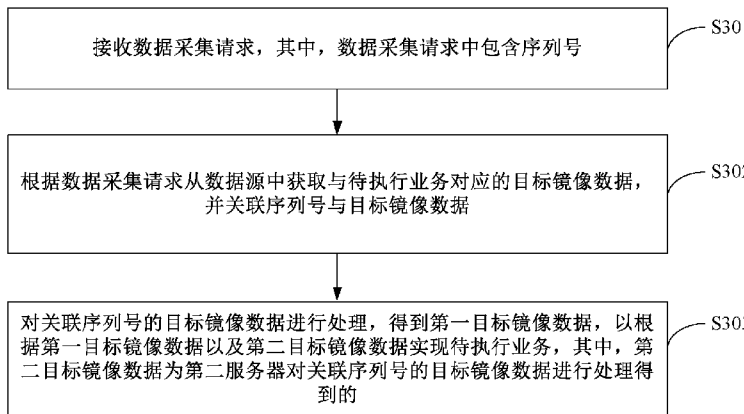


图 3

S301 Receive a data acquisition request, the data acquisition request comprising a serial number

S302 According to the data acquisition request, acquire from a data source target mirror data corresponding to a service to be executed, and associating the serial number with the target mirror data

S303 Process the target mirror data associated with the serial number to obtain first target mirror data, so as to implement said service according to the first target mirror data and second target mirror data, the second target mirror data being obtained by using a second server to process the target mirror data associated with the serial number

(57) Abstract: Provided in the present application are a data processing method and apparatus. The method is applied to a first server, and comprises: receiving a data acquisition request, the data acquisition request comprising a serial number (S301); according to the data acquisition request, acquiring from a data source target mirror data corresponding to a service to be executed, and associating the serial number with the target mirror data (S302); and processing the target mirror data associated with the serial number to obtain first target mirror data, so as to implement said service according to the first target mirror data and second target mirror data, the second

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))。

target mirror data being obtained by using a second server to process the target mirror data associated with the serial number (S303). The present application raises the accuracy of deduplication processing, thereby improving validity of data backup, and ensuring normal implementation of service.

(57) 摘要: 本申请提供了一种数据处理方法及装置, 方法应用于第一服务器, 包括: 接收数据采集请求, 其中, 数据采集请求中包含序列号 (S301), 根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据, 并关联序列号与目标镜像数据 (S302), 对关联序列号的目标镜像数据进行处理, 得到第一目标镜像数据, 以根据第一目标镜像数据以及第二目标镜像数据实现待执行业务, 其中, 第二目标镜像数据为第二服务器对关联序列号的目标镜像数据进行处理得到的 (S303)。本申请提高了去重处理的准确性, 进而提高了数据备份的有效性, 保证了业务的正常实现。

数据处理方法及装置

本申请要求于 2022 年 03 月 29 日提交中国专利局、申请号为 202210326055.6、申请名称为“数据处理方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请实施例涉及数据处理技术领域，尤其涉及一种数据处理方法及装置。

背景技术

随着网络技术的发展，在数字化背景下，开启了工厂数字化的进程。

现有技术中，工厂数字化进程中会涉及到众多业务的实现，为了保证业务的正常实现，可以对采集到的现场数据进行备份，即将数据源连接到至少两个数采系统对应的服务器中。每个数采系统对应的服务器均可以获取现场数据，并将获取到的现场数据均发送至业务方进行处理。业务方在对接收到的现场数据进行处理时，由于包含重复的现场数据，需要先对接收到的现场数据进行去重处理，再对去重处理之后的现场数据进行后续处理，进而实现相关业务。

然而，在对接收到的现场数据进行去重处理时，由于接收到的现场数据可能有多个，无法有效识别哪些数据是重复数据，降低了去重处理的准确性，进而降低了数据备份的有效性，影响了业务的正常实现。

发明内容

本申请实施例提供一种数据处理方法及装置，以提高去重处理的准确性，保证业务的正常实现。

第一方面，本申请实施例提供一种数据处理方法，应用于第一服务器，所述方法包括：接收数据采集请求，其中，所述数据采集请求中包含序列号；

根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据；

对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为所述第二服务器对关联所述序列号的目标镜像数据进行处理得到的。

可选的，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，包括：

对关联所述序列号的目标镜像数据进行存储处理，得到第一目标镜像数据；

5 将所述第一目标镜像数据分发至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，对所述第一目标镜像数据以及所述第二目标镜像数据进行去重处理，并根据去重处理之后的目标镜像数据实现所述待执行业务。

可选的，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，包括：

10 在获得资源锁权限之后，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据；

根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理；

15 释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据；

将所述第一目标镜像数据发送至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据和所述第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为所述第二服务器发送至所述第三方业务平台的。

20 可选的，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，包括：

对关联所述序列号的目标镜像数据进行存储处理和/或数据分发处理，得到第一目标镜像数据，其中，所述存储处理与所述分发处理对应不同的比特位图。

可选的，所述根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理，包括：

25 根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据确定处理失败的数据；

将所述比特位图中处理失败的数据对应的比特位设置为零，并将所述比特位图中处理成功的数据对应的比特位设置为一。

30 可选的，所述释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据，包括：

35 释放所述资源锁权限，以使所述第二服务器获取资源锁权限，并从关联所述序列号的目标镜像数据中获取目标子镜像数据，得到第二目标镜像数据，其中，所述关联所述序列号的目标镜像数据为所述第二服务器根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据得到的，所述目标子镜像

数据为从关联所述序列号的目标镜像数据中获取的与比特位图中比特位为零时对应的数据。

可选的，所述方法还包括：

确定所述目标子镜像数据的数量；

- 5 若所述目标子镜像数据的数量为零，则生成未发现异常提示，并释放所述资源锁权限；
若所述目标子镜像数据的数量不为零，则将所述目标子镜像数据设置为第二目标镜像数据发送至所述第三方业务平台。

可选的，所述释放所述资源锁权限，包括：

在置位处理完成之后，或处理时长超过预设时长阈值之后，释放所述资源锁权限。

- 10 可选的，所述接收数据采集请求，包括：

控制节拍模块每隔预设时长生成一帧同步信号，其中，所述帧同步信号中包含序列号；
接收所述帧同步信号，并根据所述帧同步信号生成包含所述序列号的数据采集请求。

第二方面，本申请实施例提供一种数据处理装置，包括：

接收模块，用于接收数据采集请求，其中，所述数据采集请求中包含序列号；

- 15 处理模块，用于根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据；

所述处理模块，还用于对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为第二服务器对关联所述序列号的目标镜像数据进行处理得到的。

- 20 本申请实施例提供了一种数据处理方法及装置，采用上述方案后，第一服务器可以先接收包含序列号的数据采集请求，然后根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联序列号与目标镜像数据，第一服务器还可以对关联序列号的目标镜像数据进行处理，得到第一目标镜像数据，第二服务器可以对关联序列号的目标镜像数据进行处理，得到第二目标镜像数据，再根据第一目标镜像数据以及第二目标镜像数据
25 实现待执行业务，通过为同一数据关联相同的序列号的方式，使得在接收到多路服务器发送的重复数据之后，可以根据序列号有效的识别哪些数据为重复数据，提高了去重处理的准确性，进而提高了数据备份的有效性，保证了业务的正常实现。

附图说明

- 30 为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图1为本申请现有技术中数据处理过程的应用示意图；

图2为本申请实施例提供的数据处理方法的应用系统的架构示意图；

图 3 为本申请实施例提供的数据处理方法的流程示意图；

图 4 为本申请另一现有技术中数据处理过程的应用示意图；

图 5 为本申请实施例提供的数据处理方法的应用示意图；

图 6 为本申请实施例提供的数据处理装置的结构示意图；

5 图 7 为本申请实施例提供的电子设备的硬件结构示意图。

具体实施方式

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

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”“第四”等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的本申请的实施例还能够包
15 括除了图示或描述的那些实例以外的其他顺序实例。此外，术语“包括”和“具有”以及他们的任何变形，意图在于覆盖不排他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

现有技术中，工厂数字化进程中会涉及到众多业务的实现，在实现相关业务时，一般
20 是通过处理器采集现场数据，然后将采集的现场数据传输到数采系统对应的服务器中进行处理，进而实现相关的业务。工业场景下一般是通过 PLC(Programmable Logic Controller, 可编程逻辑控制器)采集现场数据，然后传输到数采系统中，数采系统负责数据的计算、存储，分发等工作。数采系统北向是各种业务平台，一般通过消息队列的 pull 或 push 模式接收数据，同时也可以进行数据的查询。然后，在保证数据从采集、计算、存储、分发
25 以及查询等一系列链路上的高可用性的过程中，可能出现以下问题：如果南向数据源部分出现异常，则会出现丢数据的问题。如果数据处理系统内部在传输数据时发生异常即会出现数据的丢失。北向的业务系统在获取数据时出现丢数据的情况。如果数据处理系统在存储数据时发生异常，也会造成北向业务系统使用数据的丢失。如果北向业务系统出现异常，在异常恢复后会希望从异常点开始继续消费数据，否则会出现异常。因此，为了解决前述
30 问题，且兼顾工业领域成本问题，双机备份的方式在成本和可用性的平衡上为较合适的。

传统的双机备份的高可用方案可以分为数据和程序两部分：程序高可用是将程序无状态化，状态数据存入数据库，程序通过集群方式实现高可用。双机情况下程序也是类似镜像方式存在，即一台机器的程序和另一台机器的程序是相同的，一台机器异常后，另一个台机器可以完全替代继续工作。数据高可用一般通过数据库双机互备，共享存储来解决。

一个数据库异常后，另一个数据库也有全量数据。如果将镜像数据源分别链接到双机边缘数采，同时运行，则在数据处理过程中，双机边缘数采系统会同时存储两份数据，分发两路数据。可能会造成存储资源浪费的问题，即相当于一个数据存储了两份。而且双机还要同步备份，这样对网络 I/O 和磁盘 I/O 都是压力。业务平台如果仅对接一路边缘数采系统，在出现异常时故障切换会有切换时间，不能实时热备。业务平台如果对接两路边缘数采系统，则会出现两路重复数据，需要业务方自己去重。对于两路数据，如果进行去重，由于接收到的现场数据可能有多个，对于识别哪些数据是同一个点位的镜像数据仍然存在问题。

示例性的，图 1 为本申请现有技术中数据处理过程的应用示意图，如图 1 所示，可以通过可编程逻辑控制器从镜像数据源中采集两路镜像数据，有两个边缘数采系统（也可以对应两个服务器），分别为边缘数采系统 A（即第一服务器）和边缘数采系统 B（即第二服务器），1、2、3 代表数据源产生的先后顺序，1'、2'、3' 为镜像的数据，实际数据中是没有此顺序标号来标识数据之间的顺序关系的，在边缘数采系统 A 和边缘数采系统 B 获取数据以及对数据进行处理和传输的过程中，各数据之间的顺序可能会发生变更，若业务平台接收到的数据为乱序的，无法有效识别哪些数据是重复数据，降低了去重处理的准确性，进而降低了数据备份的有效性，影响了业务的正常实现。

基于上述技术问题，本申请通过为同一数据关联相同的序列号的方式，使得在接收到多路服务器发送的重复数据之后，可以根据序列号有效的识别哪些数据为重复数据，并进行去重操作，达到了既提高了去重处理的准确性，又提高了数据备份的有效性，进而保证了业务的正常实现的技术效果。

图 2 为本申请实施例提供的数据处理方法的应用系统的架构示意图，如图 2 所示，在该应用系统中，可以包括第一服务器 101、第二服务器 102、可编程逻辑控制器 103、镜像数据源 104。第一服务器 101 以及第二服务器 102 可以为边缘数采系统对应的服务器，即可以实现边缘数采系统负责的数据计算、存储、分发等工作。可编程逻辑控制器 103 是工业现场对接工业设备的装置，其南向通过工业协议采集数据，北向对接边缘数采系统。

第一服务器 101 和第二服务器 102 可以分别接收包含序列号的数据采集请求，然后通过可编程逻辑控制器 103 从镜像数据源 104 中采集与待执行业务对应的目标镜像数据，并关联序列号与目标镜像数据。示例性的，若序列号为 8，则将第一服务器 101 和第二服务器 102 本次获取的目标镜像数据的标识均设置为 8。

此外，应用系统还可以包含第三方业务平台 105，在第一服务器 101 和第二服务器 102 均分别对目标镜像数据进行处理之后，可以得到第一目标镜像数据和第二目标镜像数据，然后将第一目标镜像数据和第二目标镜像数据发送至第三方业务平台 105，以使第三方业务平台 105 根据第一目标镜像数据和第二目标镜像数据实现待执行业务。其中，第三方业务平台 105 根据第一目标镜像数据和第二目标镜像数据实现待执行业务可以采用现有的方式实现，在此不再详细进行限定。

另外，应用系统还可以包含节拍器（或称为节拍模型）106，为定时同步模块，其通过周期发送帧同步信号标记镜像数据，节拍器的作用为对镜像数据源的数据建立映射关系，即在同一节拍下产生的数据其节拍序号相同，可以周期性的发送帧同步信号，帧同步信号携带一个递增序列号，两路数采系统在收到帧同步信号之后可以生成数据采集请求，并向数据源进行数据的采集，采集的数据使用一个序列号进行后续的传输，通过节拍器可以标识镜像数据源产生的两个相互关联的镜像数据，即具有相同序列号的数据是可以认为是镜像数据。

下面以具体地实施例对本申请的技术方案进行详细说明。下面这几个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不再赘述。

图3为本申请实施例提供的数据处理方法的流程示意图，本实施例的方法可以由第一服务器101执行。如图3所示，本实施例的方法，可以包括：

S301：接收数据采集请求，其中，数据采集请求中包含序列号。

在本实施例中，可以接收节拍模块周期性发送的数据采集请求，然后根据数据采集请求采集对应的目标镜像数据，并根据目标镜像数据实现相关业务。

进一步的，所述接收数据采集请求，具体可以包括：

控制节拍模块每隔预设时长生成一帧同步信号，其中，所述帧同步信号中包含序列号。

接收所述帧同步信号，并根据所述帧同步信号生成包含所述序列号的数据采集请求。

具体的，节拍模块可以每隔预设时长生成一帧同步信号，第一服务器的采集驱动在接收到帧同步信号之后，可以生成包含序列号的数据采集请求。其中，序列号可以为阿拉伯数字，还可以为大写字母，小写字母、特殊符号中的任意一种或多种，仅需保证每次生成的序列号不同即可。

S302：根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联序列号与目标镜像数据。

在本实施例中，第一服务器在接收到数据采集请求之后，可以根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，其中，数据源可以为镜像数据源。获取到目标镜像数据之后，可以将序列号与目标镜像数据进行关联。同理，第二服务器也可以接收到数据采集请求，然后根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并在获取到目标镜像数据之后，将序列号与目标镜像数据进行关联。通过序列号，可以将第一服务器获取到的目标镜像数据与第二服务器获取到的目标镜像数据建立映射关系，即在同一节拍下产生的数据其序列号相同。此外，第二服务器可以有一个或多个。在本实施例中，第二服务器为一个。

S303：对关联序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据第一目标镜像数据以及第二目标镜像数据实现待执行业务，其中，第二目标镜像数据为第二服务器对关联序列号的目标镜像数据进行处理得到的。

在本实施例中，在目标镜像数据关联序列号之后，第一服务器可以对关联序列号的目标镜像数据进行处理，得到第一目标镜像数据，第二服务器也可以对关联序列号的目标镜像数据进行处理，得到第二目标镜像数据。其中，处理过程可以为存储处理，分发处理、去重处理等。第一服务器与第二服务器进行的处理过程可以相同，也可以不同。

5 处理完成之后，第一服务器可以根据第一目标镜像数据以及第二目标镜像数据实现待执行业务。此外，第二服务器也可以根据第一目标镜像数据以及第二目标镜像数据实现待执行业务。或者，还可以将第一目标镜像数据以及第二目标镜像数据发送至第三方业务平台，使得第三方业务平台根据第一目标镜像数据以及第二目标镜像数据实现待执行业务。

10 采用上述方案后，第一服务器可以先接收包含序列号的数据采集请求，然后根据数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联序列号与目标镜像数据，第一服务器还可以对关联序列号的目标镜像数据进行处理，得到第一目标镜像数据，第二服务器可以对关联序列号的目标镜像数据进行处理，得到第二目标镜像数据，再根据第一目标镜像数据以及第二目标镜像数据实现待执行业务，通过为同一数据关联相同的序列号的方式，使得在接收到多路服务器发送的重复数据之后，可以根据序列号有效的识别
15 哪些数据为重复数据，并进行去重操作，提高了去重处理的准确性，进而提高了数据备份的有效性，保证了业务的正常实现。

基于图3的方法，本说明书实施例还提供了该方法的一些具体实施方案，下面进行说明。

20 在另一实施例中，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，具体可以包括：

对关联所述序列号的目标镜像数据进行存储处理，得到第一目标镜像数据。

25 将所述第一目标镜像数据分发至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，对所述第一目标镜像数据以及所述第二目标镜像数据进行去重处理，并根据去重处理之后的目标镜像数据实现所述待执行业务。

30 具体的，不同服务器获取的同一批次的目标镜像数据均已经与同一序列号进行关联，即可以确定哪些目标镜像数据为重复数据。在不同服务器将目标镜像数据分发至第三方业务平台之后，第三方业务平台可以根据第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，对第一目标镜像数据以及第二目标镜像数据进行去重处理，得到不包含重复镜像数据的去重处理之后的目标镜像数据，并根据去重处理之后的目标镜像数据实现待执行业务。

此外，还可以通过第一服务器或者第二服务器对第一目标镜像数据以及第二目标镜像数据进行去重处理，并根据去重处理之后的目标镜像数据实现待执行业务。

综上，通过将不同服务器获取的目标镜像数据与同一序列号进行关联，使得第三方业务平台可以根据第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，直接对第一目标镜像数据以及第二目标镜像数据进行去重处理，提高了去重处理的效率与准确性。

5 在另一实施例中，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，具体可以包括：

在获得资源锁权限之后，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据。

10 根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理。

释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据。

15 将所述第一目标镜像数据发送至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据和所述第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为所述第二服务器发送至所述第三方业务平台的。

20 在本实施例中，在将不同服务器获取的同一批次的目标镜像数据与同一序列号进行关联之后，若服务器都将关联之后的目标镜像数据发送至第三方业务平台，第三方业务平台接收到重复数据之后需要先进行去重操作，然后再实现相关的待执行业务。然而，每个序列号仅能关联一个数据，若有多个数据，需要分配不同的序列号，否则也无法实现重复数据的去重操作，数据处理过程繁琐，且增加了处理时长。

现有技术中，还可以将镜像数据源对接两路边缘数采系统（即两路服务器），但是一路边缘数采系统运行，另一路边缘数采系统保持待机，等到一路出现问题后再进行切换。图 4 为本申请另一现有技术中数据处理过程的应用示意图，如图 4 所示，有两个边缘数采系统（也可以对应两个服务器），分别为边缘数采系统 A（即第一服务器）和边缘数采系统 B（即第二服务器），边缘数采系统 A 运行，边缘数采系统 B 保持待机，等到边缘数采系统 A 出现问题后，再切换至边缘数采系统 B。然而，该方式需要依赖双机健康监测，监测异常时进行双机切换，如果有些问题监测不到则导致无法起到备份的作用。例如：一台机器丢包但是没有被检测到，此时如果不切换双机，丢包异常依然会发生，导致无法充分利用双机资源，一台机器运行，一台机器资源在基本处于静默状态，且业务平台在对接双机时需要感知双机状态，当异常发生时，双机切换，业务平台也要切换，此时故障切换需要时间，不能实时进行热备份。如果将镜像数据源的一个只对接一路边缘数采，当出现问题时在连接另一路镜像源，同样也会有故障切换时间长的问题。

35 本申请通过设置一交替加载器，针对同一个帧同步信号的数据，两个服务器交替进行处理，避免并发处理同一个目标镜像数据。交替处理时可以判断镜像数据前后是否有丢失

或其它异常，并针对异常进行修复（如，有数据丢失时，重新分发或存储丢失的数据。或者，没有数据丢失时，仅分发一份完整目标镜像数据至第三方业务平台，实现数据的去重操作），提高了异常数据补救的及时性。且同一批次的多个目标镜像数据可以分配相同的序列号，无需再局限于为每个数据分配不同的序列号，简化了业务实现的处理过程。

5 其中，交替加载器的作用为针对同一个帧同步信号的数据，多个服务器交替进行处理，避免并发处理同一个镜像数据。交替处理时可以先判断镜像数据前后是否有丢失或其它异常，针对异常进行修复。交替加载器的组成包括资源锁与比特位图（也可称为 bitmap）。资源锁为各服务器交替处理的令牌，谁得到此令牌即具有对一个帧同步信号的镜像数据的处理权限，没有得到的一方需要等待另一方释放此锁。目标镜像数据携带帧同步信号进行
10 传输，哪个服务器先得到此镜像数据后，获取此帧同步信号的互斥锁，得到资源锁的服务器开始发起数据的计算、存储以及分发等操作。另外的服务器后收到目标镜像数据，则进行等待，直到前一个服务器解锁，或者超时释放之后，后一台服务器才可以继续对目标镜像数据进行处理。

此外，第二服务器在获得资源锁权限之后，可以先判断是否已经存在序列号对应的比特位图，若不存在，则第二服务器会创建当前序列号对应的比特位图，并根据关联序列号的目标镜像数据对序列号对应的比特位图进行置位处理，后续若第一服务器恢复运行之后，可以根据置位处理之后的比特位图对关联序列号的目标镜像数据进行处理，得到第二目标
15 镜像数据。即两个服务器身份不是固定的，并不是一直都是第一服务器获取资源锁，第二服务器比对信息，第一服务器和第二服务器获取资源锁的机会是均等的，也可以由第二服务器先获取资源锁，由第一服务器进行信息的比对。
20

综上，通过交替加载器的设置，使得两个服务器可以交替对目标镜像数据进行处理，避免并发处理同一个镜像数据，且可以先判断镜像数据前后是否有丢失或其它异常，并针对异常进行修复，提高了数据处理的准确性，进而保证了业务的正常实现。

进一步的，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，具体可以包括：
25

对关联所述序列号的目标镜像数据进行存储处理和/或数据分发处理，得到第一目标镜像数据，其中，所述存储处理与所述分发处理对应不同的比特位图。

具体的，处理过程可以仅为存储处理，或者仅为数据分发处理，还可既包含存储处理又包含数据分发处理，且不同的处理过程均单独对应一个比特位图。对于第二服务器，原理类似，在此不再详细进行限定。
30

此外，数据分发处理可以为第一服务器或第二服务器通过消息队列方式向第三方业务平台分发数据。其中，用户还可以指定关注的的数据，按需进行分发。

进一步的，根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理，具体可以包括：

根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据确定处理失败的数据。

将所述比特位图中处理失败的数据对应的比特位设置为零，并将所述比特位图中处理成功的数据对应的比特位设置为一。

5 具体的，在根据第一目标镜像数据以及关联序列号的目标镜像数据对序列号对应的比特位图进行置位处理时，可以先确定处理失败的数据，示例性的，可以为存储失败，或者分发失败的数据，然后将该数据对应的比特位设置为零，其他数据对应的比特位设置为1。

10 其中，不同的序列号可以对应不同的比特位图。在序列号与目标镜像数据关联之后，可以生成该序列号对应的比特位图，且该比特位图中比特位的个数与目标镜像数据中包含的数据个数相同。

另外，释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据，具体可以包括：

15 释放所述资源锁权限，以使所述第二服务器获取资源锁权限，并从关联所述序列号的目标镜像数据中获取目标子镜像数据，得到第二目标镜像数据，其中，所述关联所述序列号的目标镜像数据为所述第二服务器根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据得到的，所述目标子镜像数据为从关联所述序列号的目标镜像数据中获取的与比特位图中比特位为零时对应的数据。

20 具体的，第一服务器在处理完成目标镜像数据之后，可以释放资源锁权限，第二服务器获得资源锁权限，然后判断目标镜像数据的比特位图是否已经存在，若存在，则对比其接收到的数据和比特位图中的置位数据，如果有差异，则针对差异的数据进行再处理，以此来保证在前一服务器出现异常时另一服务器可以进行补救。示例性的，以存储为例，当第一服务器调用数据库接口完成写入操作后，则将此数据身份ID（可以为序列号）对应的
25 比特位图中的各比特位进行置位。第二服务器获得资源锁权限之后，可以先判断是否存在该序列号对应的比特位图，若存在，则对比其接收到的数据和比特位图中的置位数据，如果有差异，则针对差异的数据重新进行写入操作。

30 综上，通过资源锁权限与比特位图的结合使用，可以实现镜像数据流的差异性对比，并针对差异数据进行实时补救，既实现了两个服务器实时热备份，一个服务器断开另一个服务器可以无缝切换的效果，且在异常切换时可以实现数据基本不丢失，还不依赖保活和健康检测，避免了现有技术中需要监测异常，并在监测到异常时进行切换，导致切换时间内无法提供服务或切换时间内会有数据丢失的情况。

此外，所述方法还可以包括：

确定所述目标子镜像数据的数量。

35 若所述目标子镜像数据的数量为零，则生成未发现异常提示，并释放所述资源锁权限。

若所述目标子镜像数据的数量不为零，则将所述目标子镜像数据设置为第二目标镜像数据发送至所述第三方业务平台。

具体的，在得到目标子镜像数据之后，可以先确定目标子镜像数据中包含的数据数量，若包含的数据数量为零，则表明第一服务器在处理数据时，没有出现异常数据，则可以生成未发现异常提示，并释放所述资源锁权限，进而提醒用户未发生数据处理异常，可以实时提醒管理人员，提高了处理数据的及时性。

若目标子镜像数据中包含的数据数量不为零，则表明第一服务器在处理数据时出现了异常数据，则可以将目标子镜像数据设置为第二目标镜像数据发送至第三方业务平台，即将出现异常的数据重新发送至第三方业务平台，保证了数据的完整性。

综上，本申请既提高了处理数据的及时性，又保证了数据的完整性。

另外，释放所述资源锁权限，可以包括：

在置位处理完成之后，或处理时长超过预设时长阈值之后，释放所述资源锁权限。

具体的，释放资源锁权限可以包含两种情况，一种为第一服务器完成置位处理之后，另一种为第一服务器的置位处理过程超过了预设时长阈值之后。其中，预设时长阈值可以根据实际应用场景自定义进行设置。

示例性的，图 5 为本申请实施例提供的数据处理方法的应用示意图，如图 5 所示，有两个边缘数采系统（也可以对应两个服务器），分别为边缘数采系统 A（即第一服务器）和边缘数采系统 B（即第二服务器），且两个边缘数采系统均保持运行状态。数据处理过程可以包含数据存储以及数据分发，且数据存储对应一比特位图，数据分发对应一比特位图。通过节拍器发送的帧同步信号，两套边缘数采系统从镜像数据源采集一对目标镜像数据，如下图帧同步信号序列为 8（即序列号为 8），目标镜像数据中包含 3 个数据分别为 A，B，C。边缘数采系统在做数据存储时，先获取存储交替加载器的资源锁，获取的一方发起此帧同步镜像数据的写入操作。例如左边的边缘数采系统 A 先获取资源锁权限，在写入帧同步序列号为 8 的目标镜像数据时发现数据 B 的写入异常，则在序列号为 8 的比特位图（也可称为 bitmap）中将数据 B 的比特位设置为 0，因为 A 和 C 设置成功则设置 A 和 C 对应的比特位为 1。右边的边缘数采系统 B 后获取资源锁权限，在确定序列号为 8 的 bitmap 已经存在之后，则发起比对，经过比较发现其接收到的镜像数据中的 B 在 bitmap 中异常，则其重新发起 B 数据的写入操作，操作完成后释放序列号为 8 的资源锁和 bitmap。另外，边缘数采系统在做数据分发时，此处以 push 类型的消息队列为例，先获取分发交替加载器的资源锁，获取的一方发起帧同步镜像数据的分发。例如，左边的边缘数采系统 A 先获取资源锁权限，在分发序列号为 8 的目标镜像数据时发现 A 数据分发失败，或者响应异常，则在序列号为 8 的 bitmap 中设置数据 A 对应的比特位为 0，因为 B 和 C 分发成功则将其对应的比特位设置为 1。右边的边缘数采系统 B 后获取资源锁权限，其确定序列号为 8 的 bitmap 已经存在，则发起比对，经过比较发现其接收到的镜像数据中 A 的 bitmap 位异常，

则重新发起 A 数据的分发，操作完成后释放序列号为 8 的资源锁和 bitmap。数据分发的 pull 模式与此使用类似，在此不再重复进行论述。

5 综上，通过双机实时热备份的方式，实现异常切换时间为 0，通过节拍器以及交替加载器等实现了帧同步交替加载方式，解决了传统领域中双机冷备或者温备方案中切换时间较长的问题，解决了异常切换期间丢数据问题，真正实现了双机下数据的完整性。

基于同样的思路，本说明书实施例还提供了上述方法对应的装置，图 6 为本申请实施例提供的数据处理装置的结构示意图，如图 6 所示，可以包括：

接收模块 601，用于接收数据采集请求，其中，所述数据采集请求中包含序列号。

在本实施例中，所述接收模块 601，还用于：

10 控制节拍模块每隔预设时长生成一帧同步信号，其中，所述帧同步信号中包含序列号。接收所述帧同步信号，并根据所述帧同步信号生成包含所述序列号的数据采集请求。

处理模块 602，用于根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据。

15 所述处理模块 602，还用于对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为第二服务器对关联所述序列号的目标镜像数据进行处理得到的。

在另一实施例中，所述处理模块 602，还用于：

对关联所述序列号的目标镜像数据进行存储处理，得到第一目标镜像数据。

20 将所述第一目标镜像数据分发至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，对所述第一目标镜像数据以及所述第二目标镜像数据进行去重处理，并根据去重处理之后的目标镜像数据实现所述待执行业务。

在另一实施例中，所述处理模块 602，还用于：

25 在获得资源锁权限之后，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据。

根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理。

30 释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据。

将所述第一目标镜像数据发送至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据和所述第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为所述第二服务器发送至所述第三方业务平台的。

进一步的，所述处理模块 602，还用于：

对关联所述序列号的目标镜像数据进行存储处理和/或数据分发处理，得到第一目标镜像数据，其中，所述存储处理与所述分发处理对应不同的比特位图。

此外，所述处理模块 602，还用于：

5 根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据确定处理失败的数据。

将所述比特位图中处理失败的数据对应的比特位设置为零，并将所述比特位图中处理成功的数据对应的比特位设置为一。

此外，所述处理模块 602，还用于：

10 释放所述资源锁权限，以使所述第二服务器获取资源锁权限，并从关联所述序列号的目标镜像数据中获取目标子镜像数据，得到第二目标镜像数据，其中，所述关联所述序列号的目标镜像数据为所述第二服务器根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据得到的，所述目标子镜像数据为从关联所述序列号的目标镜像数据中获取的与比特位图中比特位为零时对应的数据。

15 此外，所述处理模块 602，还用于：

确定所述目标子镜像数据的数量。

若所述目标子镜像数据的数量为零，则生成未发现异常提示，并释放所述资源锁权限。

若所述目标子镜像数据的数量不为零，则将所述目标子镜像数据设置为第二目标镜像数据发送至所述第三方业务平台。

20 另外，所述处理模块 602，还用于：

在置位处理完成之后，或处理时长超过预设时长阈值之后，释放所述资源锁权限。

本申请实施例提供的装置，可以实现上述如图 3 所示的实施例的方法，其实现原理和技术效果类似，此处不再赘述。

25 图 7 为本申请实施例提供的电子设备的硬件结构示意图，如图 7 所示，本实施例提供的设备 700 包括：处理器 701，以及与所述处理器通信连接的存储器。其中，处理器 701、存储器 702 通过总线 703 连接。

在具体实现过程中，处理器 701 执行所述存储器 702 存储的计算机执行指令，使得处理器 701 执行上述方法实施例中的方法。

30 处理器 701 的具体实现过程可参见上述方法实施例，其实现原理和技术效果类似，本实施例此处不再赘述。

在上述的图 7 所示的实施例中，应理解，处理器可以是中央处理单元（英文：Central Processing Unit，简称：CPU），还可以是其他通用处理器、数字信号处理器（英文：Digital Signal Processor，简称：DSP）、专用集成电路（英文：Application Specific Integrated Circuit，简称：ASIC）等。通用处理器可以是微处理器或者该处理器也可以是任何常规

的处理器等。结合发明所公开的方法的步骤可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。

存储器可能包含高速 RAM 存储器，也可能还包括非易失性存储 NVM，例如至少一个磁盘存储器。

5 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。总线可以分为地址总线、数据总线、控制总线等。为便于表示，本申请附图中的总线并限定仅有一根总线或一种类型的总线。

10 本申请实施例还提供一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机执行指令，当处理器执行所述计算机执行指令时，实现上述方法实施例的数据处理方法。

本申请实施例还提供一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时，实现如上所述的数据处理方法。

15 上述的计算机可读存储介质，上述可读存储介质可以是由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器 (SRAM)，电可擦除可编程只读存储器 (EEPROM)，可擦除可编程只读存储器 (EPROM)，可编程只读存储器 (PROM)，只读存储器 (ROM)，磁存储器，快闪存储器，磁盘或光盘。可读存储介质可以是通用或专用计算机能够存取的任何可用介质。

20 一种示例性的可读存储介质耦合至处理器，从而使处理器能够从该可读存储介质读取信息，且可向该可读存储介质写入信息。当然，可读存储介质也可以是处理器的组成部分。处理器和可读存储介质可以位于专用集成电路 (Application Specific Integrated Circuits, 简称: ASIC) 中。当然，处理器和可读存储介质也可以作为分立组件存在于设备中。

25 本领域普通技术人员可以理解：实现上述各方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时，执行包括上述各方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

最后应说明的是：以上各实施例仅用以说明本申请的技术方案，而非对其限制；尽管
30 参照前述各实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的范围。

权利要求书

1、一种数据处理方法，其特征在于，应用于第一服务器，所述方法包括：

接收数据采集请求，其中，所述数据采集请求中包含序列号；

5 根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据；

对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为第二服务器对关联所述序列号的目标镜像数据进行处理得到的。

10 2、根据权利要求 1 所述的方法，其特征在于，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，包括：

对关联所述序列号的目标镜像数据进行存储处理，得到第一目标镜像数据；

15 将所述第一目标镜像数据分发至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据关联的序列号和第二目标镜像数据关联的序列号，对所述第一目标镜像数据以及所述第二目标镜像数据进行去重处理，并根据去重处理之后的目标镜像数据实现所述待执行业务。

3、根据权利要求 1 所述的方法，其特征在于，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，包括：

20 在获得资源锁权限之后，对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据；

根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理；

25 释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据；

将所述第一目标镜像数据发送至第三方业务平台，以使所述第三方业务平台根据所述第一目标镜像数据和所述第二目标镜像数据实现所述待执行业务，其中，所述第二目标镜像数据为所述第二服务器发送至所述第三方业务平台的。

30 4、根据权利要求 3 所述的方法，其特征在于，所述对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，包括：

对关联所述序列号的目标镜像数据进行存储处理和/或数据分发处理，得到第一目标镜像数据，其中，所述存储处理与所述分发处理对应不同的比特位图。

5、根据权利要求 3 所述的方法，其特征在于，所述根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据对所述序列号对应的比特位图进行置位处理，包括：

根据所述第一目标镜像数据以及关联所述序列号的目标镜像数据确定处理失败的数据；

将所述比特位图中处理失败的数据对应的比特位设置为零，并将所述比特位图中处理成功的数据对应的比特位设置为一。

- 5 6、根据权利要求 5 所述的方法，其特征在于，所述释放所述资源锁权限，以使所述第二服务器获得资源锁权限，并根据置位处理之后的比特位图对关联所述序列号的目标镜像数据进行处理，得到第二目标镜像数据，包括：

10 释放所述资源锁权限，以使所述第二服务器获取资源锁权限，并从关联所述序列号的目标镜像数据中获取目标子镜像数据，得到第二目标镜像数据，其中，所述关联所述序列号的目标镜像数据为所述第二服务器根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据得到的，所述目标子镜像数据为从关联所述序列号的目标镜像数据中获取的比特位图中比特位为零时对应的数据。

7、根据权利要求 6 所述的方法，其特征在于，所述方法还包括：

确定所述目标子镜像数据的数量；

- 15 若所述目标子镜像数据的数量为零，则生成未发现异常提示，并释放所述资源锁权限；
若所述目标子镜像数据的数量不为零，则将所述目标子镜像数据设置为第二目标镜像数据发送至所述第三方业务平台。

8、根据权利要求 3 所述的方法，其特征在于，所述释放所述资源锁权限，包括：

在置位处理完成之后，或处理时长超过预设时长阈值之后，释放所述资源锁权限。

- 20 9、根据权利要求 1-8 任一项所述的方法，其特征在于，所述接收数据采集请求，包括：

控制节拍模块每隔预设时长生成一帧同步信号，其中，所述帧同步信号中包含序列号；
接收所述帧同步信号，并根据所述帧同步信号生成包含所述序列号的数据采集请求。

10、一种数据处理装置，其特征在于，所述装置包括：

- 25 接收模块，用于接收数据采集请求，其中，所述数据采集请求中包含序列号；

处理模块，用于根据所述数据采集请求从数据源中获取与待执行业务对应的目标镜像数据，并关联所述序列号与所述目标镜像数据；

- 所述处理模块，还用于对关联所述序列号的目标镜像数据进行处理，得到第一目标镜像数据，以根据所述第一目标镜像数据以及第二目标镜像数据实现所述待执行业务，其中，
30 所述第二目标镜像数据为第二服务器对关联所述序列号的目标镜像数据进行处理得到的。

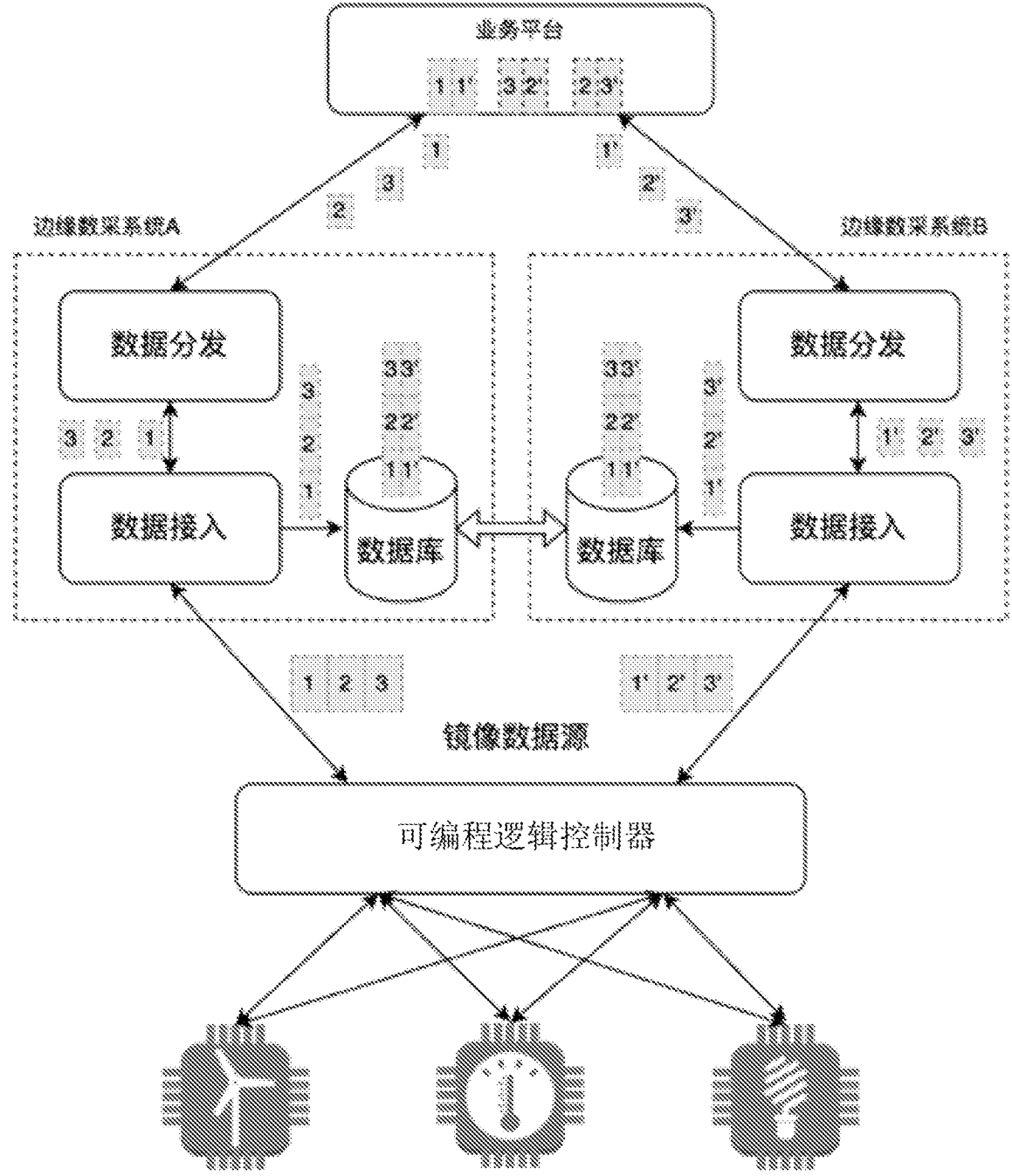


图 1

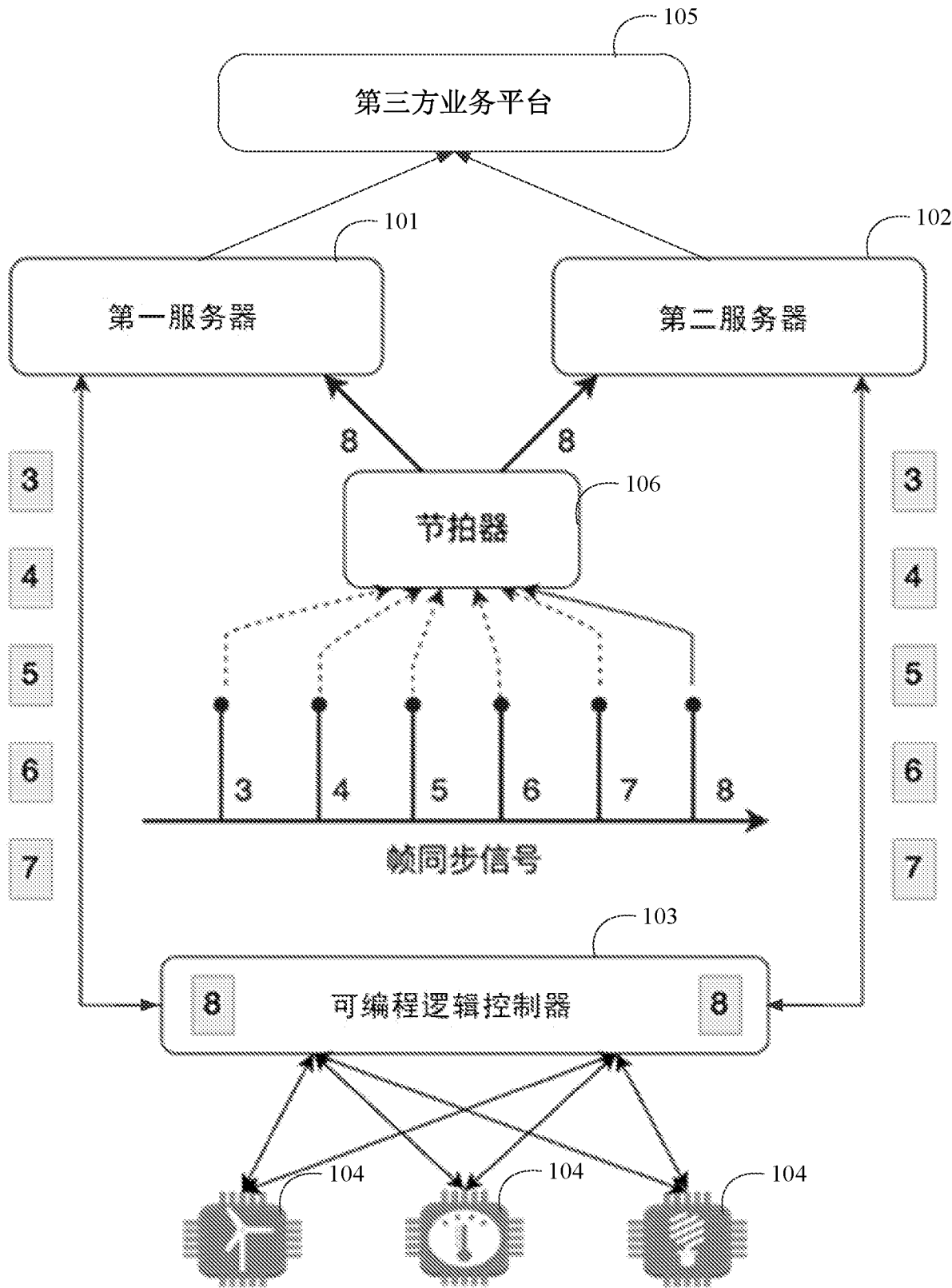


图 2

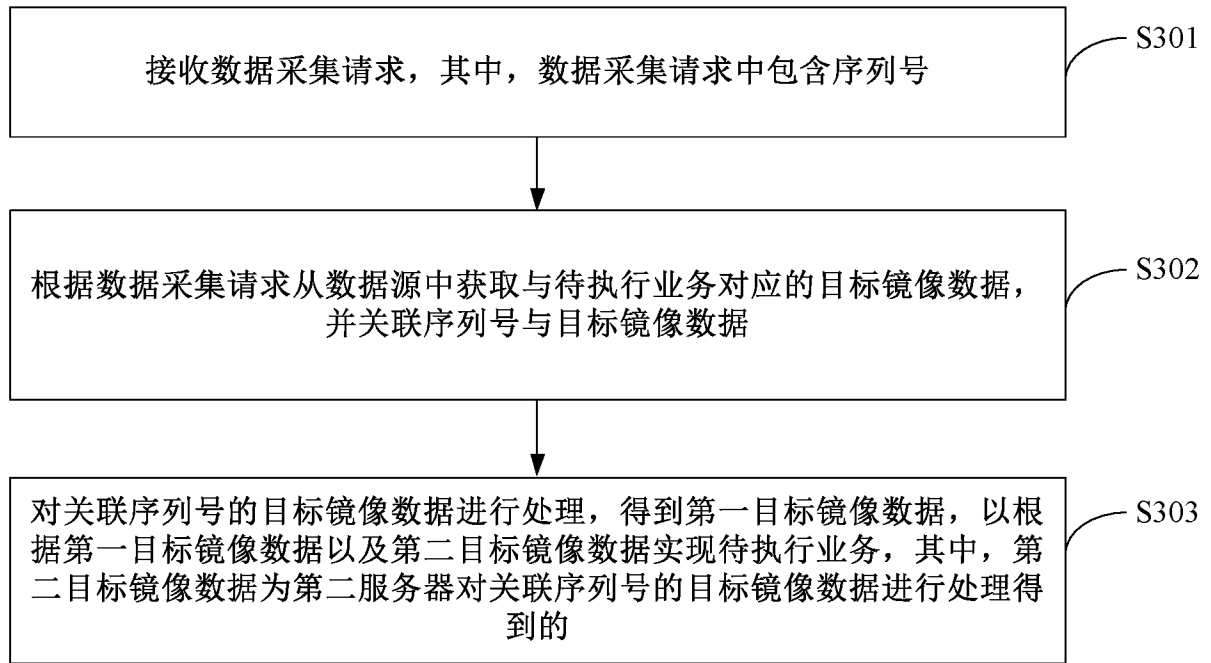


图 3

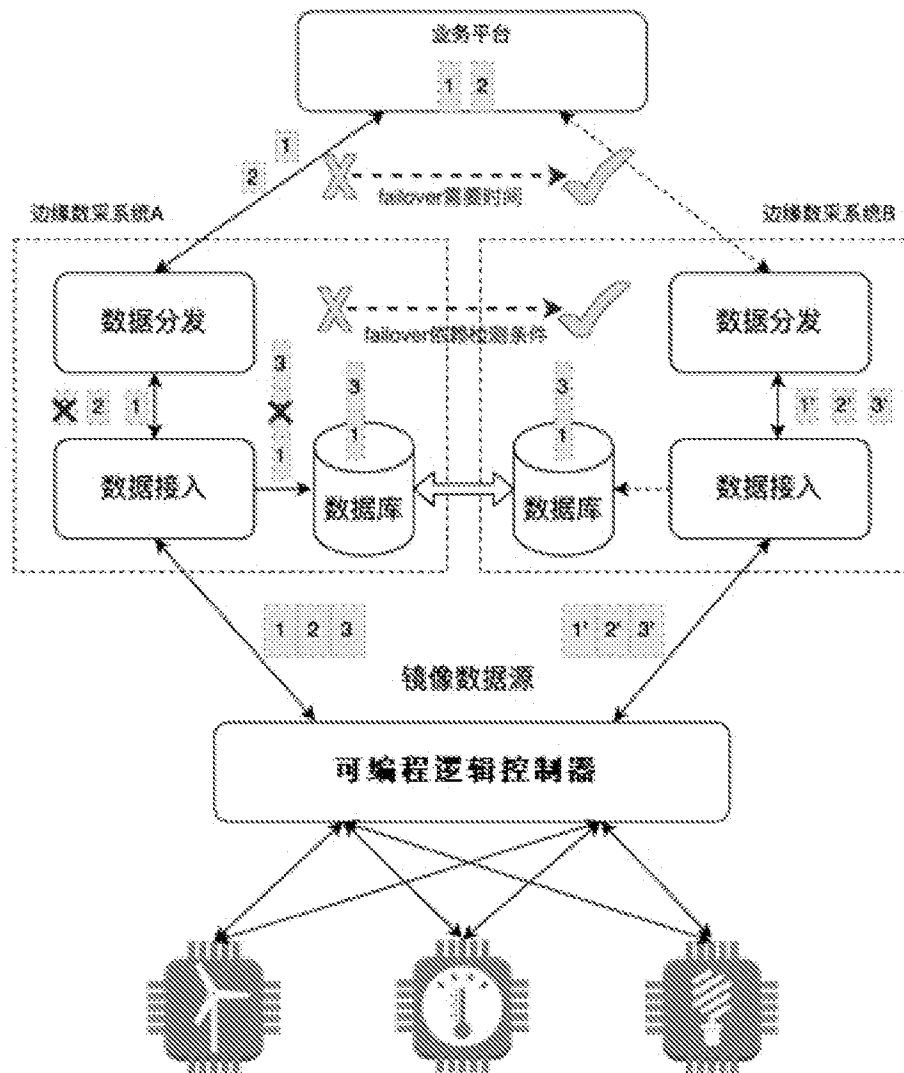


图 4

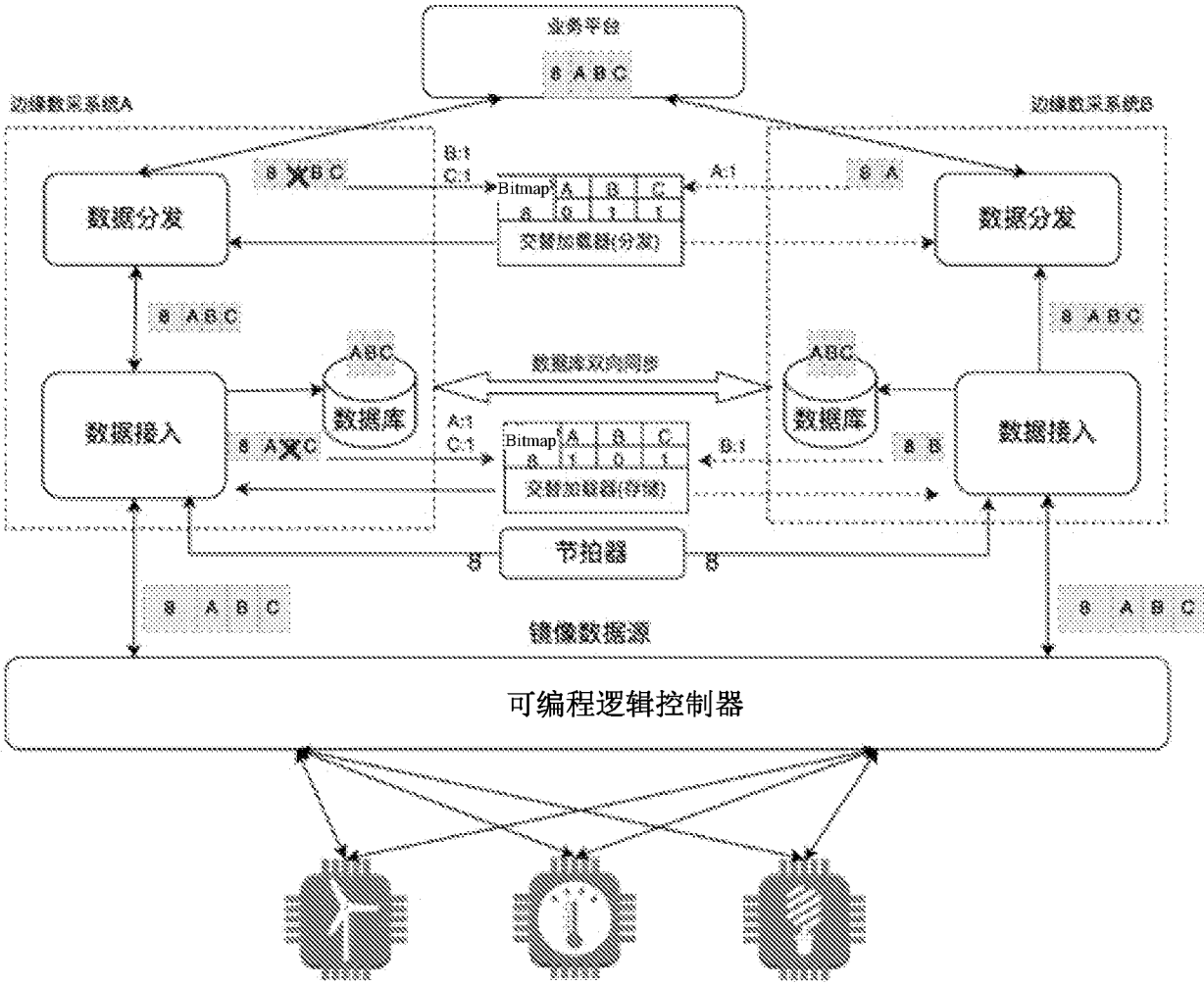


图 5

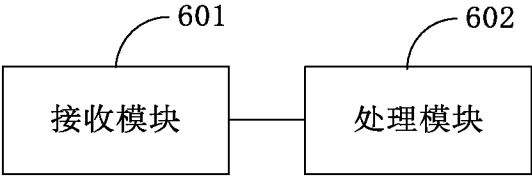


图 6

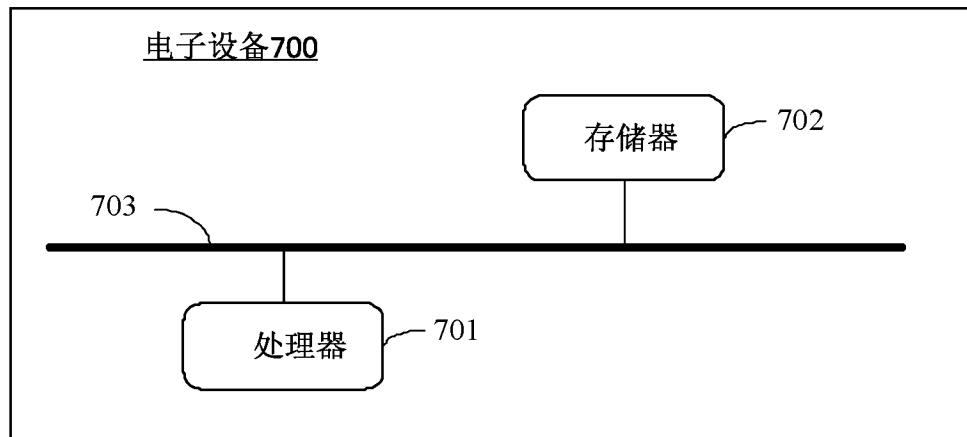


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/084287

A. CLASSIFICATION OF SUBJECT MATTER

G06F16/27(2019.01)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: G06F

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)

CNTXT; ENTXT; ENTXTC; VEN; CNKI: 序列号, 编号, 去重, 重复, 镜像, 备份, 关联, seq+, ID, ident+, repeat+, duplicat+, mitigat+, ease+, image, mirror, backup, associat+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 114554593 A (ALIBABA (CHINA) CO., LTD. et al.) 27 May 2022 (2022-05-27) entire document	1-10
X	CN 111274050 A (SHENZHEN ARCHFORCE FINANCIAL TECHNOLOGY CO., LTD.) 12 June 2020 (2020-06-12) description, paragraphs [0040]-[0095], and figures 2-4	1-10
A	US 2019081862 A1 (VERIZON DIGITAL MEDIA SERVICES INC.) 14 March 2019 (2019-03-14) entire document	1-10
A	US 2005256972 A1 (HEWLETT PACKARD DEVELOPMENT COMPANY, L.P.) 17 November 2005 (2005-11-17) entire document	1-10
A	US 2019207812 A1 (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 04 July 2019 (2019-07-04) entire document	1-10
A	WO 2020207267 A1 (ALIBABA GROUP HOLDING LTD.) 15 October 2020 (2020-10-15) entire document	1-10



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

27 May 2023

Date of mailing of the international search report

13 June 2023

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/084287

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	张灿阳等 (ZHANG, Canyang et al.). "基于改进Simhash的虚拟机镜像去重方法 (Virtual Machine Image Clustering Deduplication Algorithm Based on Improved Simhash)" 四川大学学报 (自然科学版) (<i>Journal of Sichuan University(Natural Science Edition)</i>), Vol. 57, No. 1, 20 March 2020 (2020-03-20), entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/084287

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114554593	A	27 May 2022	None			
CN	111274050	A	12 June 2020	None			
US	2019081862	A1	14 March 2019	US	10805161	B2	13 October 2020
US	2005256972	A1	17 November 2005	US	8185663	B2	22 May 2012
				JP	2005327283	A	24 November 2005
US	2019207812	A1	04 July 2019	WO	2018095416	A1	31 May 2018
				US	11088903	B2	10 August 2021
				CN	108111473	A	01 June 2018
				CN	108156002	A	12 June 2018
				CN	108156005	A	12 June 2018
				CN	108173892	A	15 June 2018
				CN	108111473	B	13 November 2020
				CN	108173892	B	13 November 2020
				CN	108156005	B	29 December 2020
				CN	108156002	B	06 April 2021
WO	2020207267	A1	15 October 2020	CN	111800442	A	20 October 2020
				CN	111800442	B	28 October 2022
				CN	115633050	A	20 January 2023

国际检索报告

国际申请号

PCT/CN2023/084287

A. 主题的分类

G06F16/27 (2019.01) 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: G06F

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))

CNTXT; ENTXT; ENTXTC; VEN; CNKI: 序列号, 编号, 去重, 重复, 镜像, 备份, 关联, seq+, ID, ident+, repeat+, duplicat+, mitigat+, ease+, image, mirror, backup, associat+

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 114554593 A (阿里巴巴 (中国) 有限公司 等) 2022年5月27日 (2022 - 05 - 27) 全文	1-10
X	CN 111274050 A (深圳华锐金融技术股份有限公司) 2020年6月12日 (2020 - 06 - 12) 说明书第[0040]-[0095]段, 附图2-4	1-10
A	US 2019081862 A1 (VERIZON DIGITAL MEDIA SERVICES INC) 2019年3月14日 (2019 - 03 - 14) 全文	1-10
A	US 2005256972 A1 (HEWLETT PACKARD DEVELOPMENT CO) 2005年11月17日 (2005 - 11 - 17) 全文	1-10
A	US 2019207812 A1 (TENCENT TECH SHENZHEN CO LTD) 2019年7月4日 (2019 - 07 - 04) 全文	1-10
A	WO 2020207267 A1 (ALIBABA GROUP HOLDING LTD) 2020年10月15日 (2020 - 10 - 15) 全文	1-10

☒ 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

2023年5月27日

国际检索报告邮寄日期

2023年6月13日

ISA/CN的名称和邮寄地址

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

授权官员

陈园

电话号码 (+86) 027-59371341

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	张灿阳等. “基于改进Simhash的虚拟机镜像去重方法” 四川大学学报(自然科学版), 第57卷, 第1期, 2020年3月20日 (2020 - 03 - 20), 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/084287

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	114554593	A	2022年5月27日	无			
CN	111274050	A	2020年6月12日	无			
US	2019081862	A1	2019年3月14日	US	10805161	B2	2020年10月13日
US	2005256972	A1	2005年11月17日	US	8185663	B2	2012年5月22日
				JP	2005327283	A	2005年11月24日
US	2019207812	A1	2019年7月4日	WO	2018095416	A1	2018年5月31日
				US	11088903	B2	2021年8月10日
				CN	108111473	A	2018年6月1日
				CN	108156002	A	2018年6月12日
				CN	108156005	A	2018年6月12日
				CN	108173892	A	2018年6月15日
				CN	108111473	B	2020年11月13日
				CN	108173892	B	2020年11月13日
				CN	108156005	B	2020年12月29日
				CN	108156002	B	2021年4月6日
WO	2020207267	A1	2020年10月15日	CN	111800442	A	2020年10月20日
				CN	111800442	B	2022年10月28日
				CN	115633050	A	2023年1月20日