

(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
2020 年 1 月 2 日 (02.01.2020)



(10) 国际公布号
WO 2020/000725 A1

(51) 国际专利分类号:
G06F 16/21 (2019.01)

(21) 国际申请号: PCT/CN2018/107702

(22) 国际申请日: 2018 年 9 月 26 日 (26.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810694928.2 2018年6月29日 (29.06.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 阴旭(YIN, Xu); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, Guangdong 518000 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ,

(54) Title: ELECTRONIC APPARATUS, DATA MIGRATION METHOD, AND STORAGE MEDIUM

(54) 发明名称: 电子装置、数据迁移的方法及存储介质

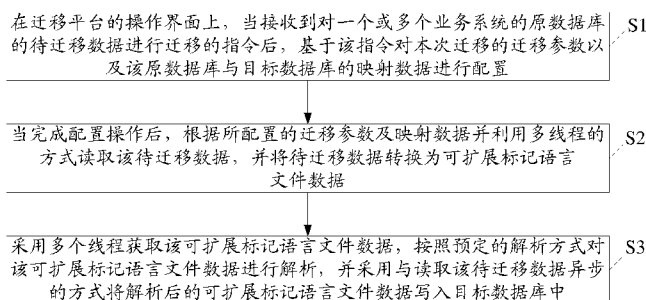


图 2

- S1 On an operation interface of a migration platform, after an instruction for migrating data to be migrated from an original database of one or more service systems is received, configure, based on the instruction, migration parameters of this instance of migration, and mapping data between the original database and a target database
- S2 After the configuration operation is completed, read the data to be migrated according to the configured migration parameters and mapping data and by means of multi-threading, and convert the data to be migrated into extensible markup language file data
- S3 Use multiple threads to acquire the extensible markup language file data, parse the extensible markup language file data according to a pre-set parsing method, and use a method which is asynchronous with reading of the data to be migrated to write the parsed extensible markup language file data in the target database

(57) Abstract: The present application relates to an electronic apparatus, a data migration method, and a storage medium. The method comprises: on an operation interface of a migration platform, after an instruction for migrating data to be migrated from an original database of one or more service systems is received, configuring, based on the instruction, migration parameters of this instance of migration, and mapping data between the original database and a target database; after the configuration operation is completed, reading the data to be migrated according to the configured migration parameters and mapping data and by means of multi-threading, and converting the data to be migrated into XML file data; and using multiple threads to acquire the XML file data, parsing the XML file data according to a pre-determined parsing method, and using a method which is asynchronous with reading of the data to be migrated to write the parsed XML file data in the target database. The present application can realize rapid and efficient migration of data at a higher amount level.

NA, RW, 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, 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))。

(57) 摘要：本申请涉及一种电子装置、数据迁移的方法及存储介质，该方法包括：在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为XML文件数据；采用多个线程获取该XML文件数据，按照预定的解析方式对该XML文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的XML文件数据写入目标数据库中。本申请能够实现快速、高效地迁移较大数量级别的数据。

电子装置、数据迁移的方法及存储介质

优先权申明

本申请基于巴黎公约申明享有 2018 年 06 月 29 日递交的申请号为 CN 201810694928.2、名称为“电子装置、数据迁移的方法及存储介质”中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

技术领域

本申请涉及通信技术领域，尤其涉及一种电子装置、数据迁移的方法及存储介质。

背景技术

目前，在数据量较大（例如，100GB 以上）的数据迁移技术中，例如，新旧数据库系统之间的数据迁移，现有的解决方案为由数据库工作人员主导，由开发人员协助，使用数据库层面的方式进行数据迁移。这种方式一般需要较长的时间才能完成数据的迁移，甚至出现无法实现迁移的情况。有鉴于此，在数据量较大时，需要一种快速而高效的数据迁移的解决方案。

发明内容

本申请的目的在于提供一种电子装置、数据迁移的方法及存储介质，旨在实现快速、高效地迁移较大数量级别的数据。

为实现上述目的，本申请提供一种电子装置，所述电子装置包括存储器及与所述存储器连接的处理器，所述存储器中存储有可在所述处理器上运行的处理系统，所述处理系统被所述处理器执行时实现如下步骤：

在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该

原数据库与目标数据库的映射数据进行配置；

当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

为实现上述目的，本申请还提供一种数据迁移的方法，所述数据迁移的方法包括：

S1，在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

S2，当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

S3，采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

本申请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上存储有处理系统，所述处理系统被处理器执行时实现步骤：

在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件

数据写入目标数据库中。

本申请的有益效果是：本申请在迁移平台的操作界面上，将迁移参数以及该原数据库与目标数据库的映射数据进行配置，在迁移数据时，根据配置的迁移参数及映射数据，并利用多线程的方式读取待迁移数据，采用 XML 文件作为中间存储，控制多线程读取 XML 文件，解析 XML 文件，将解析后得到的数据迁移至目标数据库中，本申请为平台化的数据迁移的方式，能够为业务系统提供 TB 级数据量的数据迁移服务，快速、高效地迁移不同数据库及不同数据表结构的数据。

附图说明

图 1 为本申请电子装置一实施例的硬件架构的示意图；

图 2 为本申请数据迁移的方法第一实施例的流程示意图；

图 3 为本申请数据迁移的方法第二实施例的流程示意图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。基于本申请中的实施例，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。

需要说明的是，在本申请中涉及“第一”、“第二”等的描述仅用于描述目的，而不能理解为指示或暗示其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”的特征可以明示或者隐含地包括至少一个该特征。另外，各个实施例之间的技术方案可以相互结合，但是必须是以本领域普通技术人员能够实现为基础，当技术方案的结合出现相互矛

盾或无法实现时应当认为这种技术方案的结合不存在，也不在本申请要求的保护范围之内。

参阅图 1 所示，是本申请电子装置一实施例的硬件架构的示意图。电子装置 1 是一种能够按照事先设定或者存储的指令，自动进行数值计算和/或信息处理的设备。所述电子装置 1 可以是计算机、也可以是单个网络服务器、多个网络服务器组成的服务器组或者基于云计算的由大量主机或者网络服务器构成的云，其中云计算是分布式计算的一种，由一群松散耦合的计算机集组成的一个超级虚拟计算机。

在本实施例中，电子装置 1 可包括，但不仅限于，可通过系统总线相互通信连接的存储器 11、处理器 12、网络接口 13，存储器 11 存储有可在处理器 12 上运行的处理系统。需要指出的是，图 1 仅示出了具有组件 11-13 的电子装置 1，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。

其中，存储器 11 包括内存及至少一种类型的可读存储介质。内存为电子装置 1 的运行提供缓存；可读存储介质可为如闪存、硬盘、多媒体卡、卡型存储器（例如，SD 或 DX 存储器等）、随机访问存储器（RAM）、静态随机访问存储器（SRAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、可编程只读存储器（PROM）、磁性存储器、磁盘、光盘等的非易失性存储介质。在一些实施例中，可读存储介质可以是电子装置 1 的内部存储单元，例如该电子装置 1 的硬盘；在另一些实施例中，该非易失性存储介质也可以是电子装置 1 的外部存储设备，例如电子装置 1 上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。本实施例中，存储器 11 的可读存储介质通常用于存储安装于电子装置 1 的操作系统和各类应用软件，例如本申请一实施例中的处理系统的程序代码等。此外，存储器 11 还可以用于暂时地存储已

经输出或者将要输出的各类数据。

所述处理器 12 在一些实施例中可以是中央处理器 (Central Processing Unit, CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器 12 通常用于控制所述电子装置 1 的总体操作, 例如执行与其他设备进行数据交互或者通信相关的控制和处理等。本实施例中, 所述处理器 12 用于运行所述存储器 11 中存储的程序代码或者处理数据, 例如运行处理系统等。

所述网络接口 13 可包括无线网络接口或有线网络接口, 该网络接口 13 通常用于在所述电子装置 1 与其他电子设备之间建立通信连接。

所述处理系统存储在存储器 11 中, 包括至少一个存储在存储器 11 中的计算机可读指令, 该至少一个计算机可读指令可被处理器 12 执行, 以实现本申请各实施例的方法; 以及, 该至少一个计算机可读指令依据其各部分所实现的功能不同, 可被划为不同的逻辑模块。

在一实施例中, 上述处理系统被所述处理器 12 执行时实现如下步骤:

在迁移平台的操作界面上, 当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后, 基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置;

本实施例中, 通过访问浏览器登录迁移平台, 在迁移平台的操作界面上, 可以选择对一个或多个业务系统的原数据库的数据进行迁移。

其中, 待迁移数据可以是不同数据库的数据 (例如: MYSQL->ORACLE, SQLSERVER->ORACLE 之间的数据迁移)、不同表结构的数据或不同数据类型的数据, 本实施例可实现不同数据库的数据、不同表结构的数据或不同数据类型的数据的迁移。

其中, 迁移参数包括: 原数据库与目标数据库的数据库名、原数据库与目标数据库对应的服务器、原数据库与目标数据库对应的登录帐号及登录密码、数据表名、数据表属组、迁移开始时间点、迁移结束时间点、是否支持

回滚、线程数量（其中，最小 1 个线程，最大 100 个线程）等。原数据库与目标数据库的映射数据包括：原数据库的数据表及目标数据库的数据表在迁移时两表的字段、数据类型的对应关系等。

进一步地，为了实现大数据量迁移的灵活性，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，可以执行定时迁移，并在迁移过程中，可以执行迁移暂停、迁移停止及回滚的操作。

当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML（Extensible Markup Language，可扩展标记语言）文件数据；

本实施例利用 XML 文件作为中间存储，在完成配置操作后，利用多线程的方式并根据所配置的迁移参数及映射数据读取原数据库中的待迁移数据，将待迁移数据转换为 XML 文件数据。

其中，原数据库的数据表经 XML 文件后存储至目标数据库中的目标数据表中，原数据库的一个数据表可对应多个 XML 文件，XML 文件数据包括目标数据表的各项信息（例如，数据表名、字段名、数据类型、初始值等）。为了快速完成数据的迁移，原数据库的一个数据表可以使用多个线程来读取待迁移数据，例如，原数据库的数据表中有 1 亿的数据量，如果使用 1 个线程迁移耗时会很长，但使用 100 个线程进行迁移，则迁移速度提升 100 倍。

采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

其中，采用多个线程获取 XML 文件数据，以便快速进行数据迁移。预定的解析方式为 xml 简单应用程序接口的解析方式。具体地，xml 简单应用程序接口的解析方式适用于解析大型的 XML 文件数据，其对 XML 文件数据进行逐行解析，解析 XML 文件数据时，将目标数据表名称、字段名、数据类型、

初始值等需要写入目标数据表的数据记录在内存中，每当循环解析到一定数量时，例如 100 次时，会触发一次将记录在内存中的数据进行提交，将数据写入到目标数据表中。

其中，读取原数据库中的待迁移数据的操作和写入数据至目标数据表的操作作为异步处理，实现读/写分离，提高了迁移的效率。

与现有技术相比，本申请在迁移平台的操作界面上，将迁移参数以及该原数据库与目标数据库的映射数据进行配置，在迁移数据时，根据配置的迁移参数及映射数据，并利用多线程的方式读取待迁移数据，采用 XML 文件作为中间存储，控制多线程读取 XML 文件，解析 XML 文件，将解析后得到的数据迁移至目标数据库中，本申请为平台化的数据迁移的方式，能够为业务系统提供 TB 级数据量的数据迁移服务，快速、高效地迁移不同数据库、不同数据表结构的数据。

在一优选的实施例中，在上述实施例的基础上，所述处理系统被所述处理器执行时，还实现如下步骤：

按照预定的加密方式将原数据库的待迁移数据中每一单行数据进行加密后得到第一密文，按照该预定的加密方式将存储在该目标数据库中的每一单行数据进行加密后得到第二密文；

若该第一密文与对应的第二密文分别相同，则验证通过；

若有该第一密文与对应的第二密文不相同，将该第二密文对应的单行数据进行回滚清除，以重新迁移该单行数据。

其中，预定的加密方式为 MD5（Message-Digest Algorithm 5，信息-摘要算法）。本实施例将原数据库的待迁移数据按照逐行加密的方式进行加密，将目标数据库中的数据按照逐行加密的方式进行加密，将第一密文与对应的第二密文进行对比，如果两者相同，则验证通过，该第二密文对应的数据在目标数据库中是有效的；如果两者不相同，则验证失败，该第二密文对应的

数据需要从目标数据表中回滚清除，重新迁移该单行数据。通过这种方式，在快速、高效地迁移大数据量的数据的基础上，还能够进行单行数据严格的验证，以发现迁移过程出错的数据，确保迁移的数据的准确性。

在一优选的实施例中，在上述实施例的基础上，所述处理系统被所述处理器执行时，还实现如下步骤：

获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

其中，第一预设数量与第二预设数量可以相同或者不同，优选地，第一预设数量与第二预设数量相同，例如，第一预设数量与第二预设数量为 1000。主键是数据表中的一个或多个字段，主键标识用于唯一标识数据表中的某一条记录。本实施例以第一预设数量或第二预设数量的主键标识为一组（如果不够第一预设数量则以字符 0 补齐），以定长的格式存放第一索引文件或第二索引文件至磁盘文件中，每次按组取出主键数据进行处理。

其中，第一索引文件用于记录定位待迁移数据对应存放的 XML 文件。通常情况下，由于待迁移数据的数据量都非常大，其所生成的 XML 文件也非常多，所以迁移时需要先读取第一索引文件，通过第一索引文件的名称及主键数据进行定位，才能快速找到 XML 文件进行迁移。

第二索引文件，用于在将原数据库的待迁移数据写入目标数据库后，目标数据库中的每条数据在目标数据表中的主键标识。当迁移操作需要回滚时，可以通过该第二索引文件提取出目标数据表中对应的数据进行清除。

通过上面的描述可以看出，本实施例通过第一索引文件能够快速找到 XML 文件并进行数据的迁移，通过第二索引文件能够快速提取目标数据表中需要回滚的数据并进行清除。

如图 2 所示，图 2 为本申请数据迁移的方法一实施例的流程示意图，该数据迁移的方法包括以下步骤：

步骤 S1，在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

本实施例中，通过访问浏览器登录迁移平台，在迁移平台的操作界面上，可以选择对一个或多个业务系统的原数据库的数据进行迁移。

其中，待迁移数据可以是不同数据库的数据（例如：MYSQL->ORACLE，SQLSERVER->ORACLE 之间的数据迁移）、不同表结构的数据或不同数据类型的数据，本实施例可实现不同数据库的数据、不同表结构的数据或不同数据类型的数据的迁移。

其中，迁移参数包括：原数据库与目标数据库的数据库名、原数据库与目标数据库对应的服务器、原数据库与目标数据库对应的登录帐号及登录密码、数据表名、数据表属组、迁移开始时间点、迁移结束时间点、是否支持回滚、线程数量（其中，最小 1 个线程，最大 100 个线程）等。原数据库与目标数据库的映射数据包括：原数据库的数据表及目标数据库的数据表在迁移时两表的字段、数据类型的对应关系等。

进一步地，为了实现大数据量迁移的灵活性，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，可以执行定时迁移，并在迁移过程中，可以执行迁移暂停、迁移停止及回滚的操作。

步骤 S2，当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

本实施例利用 XML 文件作为中间存储，在完成配置操作后，利用多线程的方式并根据所配置的迁移参数及映射数据读取原数据库中的待迁移数据，

将待迁移数据转换为 XML 文件数据。

其中，原数据库的数据表经 XML 文件后存储至目标数据库中的目标数据表中，原数据库的一个数据表可对应多个 XML 文件，XML 文件数据包括目标数据表的各项信息（例如，数据表名、字段名、数据类型、初始值等）。为了快速完成数据的迁移，原数据库的一个数据表可以使用多个线程来读取待迁移数据，例如，原数据库的数据表中有 1 亿的数据量，如果使用 1 个线程迁移耗时会很长，但使用 100 个线程进行迁移，则迁移速度提升 100 倍。

步骤 S3，采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

其中，采用多个线程获取 XML 文件数据，以便快速进行数据迁移。预定的解析方式为 xml 简单应用程序接口的解析方式。具体地，xml 简单应用程序接口的解析方式适用于解析大型的 XML 文件数据，其对 XML 文件数据进行逐行解析，解析 XML 文件数据时，将目标数据表名称、字段名、数据类型、初始值等需要写入目标数据表的数据记录在内存中，每当循环解析到一定数量时，例如 100 次时，会触发一次将记录在内存中的数据进行提交，将数据写入到目标数据表中。

其中，读取原数据库中的待迁移数据的操作和写入数据至目标数据表的操作作为异步处理，实现读/写分离，提高了迁移的效率。

与现有技术相比，本申请在迁移平台的操作界面上，将迁移参数以及该原数据库与目标数据库的映射数据进行配置，在迁移数据时，根据配置的迁移参数及映射数据，并利用多线程的方式读取待迁移数据，采用 XML 文件作为中间存储，控制多线程读取 XML 文件，解析 XML 文件，将解析后得到的数据迁移至目标数据库中，本申请为平台化的数据迁移的方式，能够为业务系统提供 TB 级数据量的数据迁移服务，快速、高效地迁移不同数据库、

不同数据表结构的数据。

在一优选的实施例中，在上述实施例的基础上，如图 3 所示，所述步骤 S3 之后还包括：

步骤 S4，按照预定的加密方式将原数据库的待迁移数据中每一单行数据进行加密后得到第一密文，按照该预定的加密方式将存储在该目标数据库中的每一单行数据进行加密后得到第二密文；

步骤 S45，分析第一密文与对应的第二密文是否相同；

步骤 S5，若该第一密文与对应的第二密文分别相同，则验证通过，该第二密文对应的数据在目标数据库中为有效的数据；

步骤 S6，若有该第一密文与对应的第二密文不相同，将该第二密文对应的单行数据进行回滚清除，以重新迁移该单行数据。

其中，预定的加密方式为 MD5（Message-Digest Algorithm 5，信息-摘要算法）。本实施例将原数据库的待迁移数据按照逐行加密的方式进行加密，将目标数据库中的数据按照逐行加密的方式进行加密，将第一密文与对应的第二密文进行对比，如果两者相同，则验证通过，该第二密文对应的数据在目标数据中是有效的数据；如果两者不相同，则验证失败，该第二密文对应的数据需要从目标数据表中回滚清除，重新迁移该单行数据。通过这种方式，在快速、高效地迁移大数据量的数据的基础上，还能够进行单行数据严格的验证，以发现迁移过程出错的数据，确保迁移的数据的准确性。

在一优选的实施例中，在上述实施例的基础上，所述步骤 S2 之后还包括：获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

所述步骤 S3 之后还包括：获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

其中，第一预设数量与第二预设数量可以相同或者不同，优选地，第一预设数量与第二预设数量相同，例如，第一预设数量与第二预设数量为 1000。主键是数据表中的一个或多个字段，主键标识用于唯一标识数据表中的某一条记录。本实施例以第一预设数量或第二预设数量的主键标识为一组（如果不够第一预设数量则以字符 0 补齐），以定长的格式存放第一索引文件或第二索引文件至磁盘文件中，每次按组取出主键数据进行处理。

其中，第一索引文件用于记录定位待迁移数据对应存放的 XML 文件。通常情况下，由于待迁移数据的数据量都非常大，其所生成的 XML 文件也非常多，所以迁移时需要先读取第一索引文件，通过第一索引文件的名称及主键数据进行定位，才能快速找到 XML 文件进行迁移。

第二索引文件，用于在将原数据库的待迁移数据写入目标数据库后，目标数据库中的每条数据在目标数据表中的主键标识。当迁移操作需要回滚时，可以通过该第二索引文件提取出目标数据表中对应的数据进行清除。

通过上面的描述可以看出，本实施例通过第一索引文件能够快速找到 XML 文件并进行数据的迁移，通过第二索引文件能够快速提取目标数据表中需要回滚的数据并进行清除。

本申请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上存储有处理系统，所述处理系统被处理器执行时实现上述的数据迁移的方法的步骤。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如 ROM/RAM、磁碟、

光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例,并非因此限制本申请的专利范围,凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本申请的专利保护范围内。

权利要求书

1.一种电子装置，其特征在于，所述电子装置包括存储器及与所述存储器连接的处理器，所述存储器中存储有可在所述处理器上运行的处理系统，所述处理系统被所述处理器执行时实现如下步骤：

在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

2.根据权利要求 1 所述的电子装置，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

按照预定的加密方式将原数据库的待迁移数据中每一单行数据进行加密后得到第一密文，按照该预定的加密方式将存储在该目标数据库中的每一单行数据进行加密后得到第二密文；

若该第一密文与对应的第二密文分别相同，则验证通过；

若有该第一密文与对应的第二密文不相同，将该第二密文对应的单行数据进行回滚清除，以重新迁移该单行数据。

3.根据权利要求 1 所述的电子装置，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标

识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

4.根据权利要求 2 所述的电子装置，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

5.根据权利要求 1 所述的电子装置，其特征在于，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，所述处理系统被所述处理器执行时，还实现如下步骤：

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

6.根据权利要求 2 所述的电子装置，其特征在于，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，所述处理系统被所述处理器执行时，还实现如下步骤：

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

7.根据权利要求 1 或 2 所述的电子装置，其特征在于，所述预定的解析方式为 xml 简单应用程序接口的解析方式。

8.一种数据迁移的方法，其特征在于，所述数据迁移的方法包括：

S1，在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

S2，当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

S3, 采用多个线程获取该 XML 文件数据, 按照预定的解析方式对该 XML 文件数据进行解析, 并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

9.根据权利要求 8 所述的数据迁移的方法, 其特征在于, 所述步骤 S3 之后还包括:

按照预定的加密方式将原数据库的待迁移数据中每一单行数据进行加密后得到第一密文, 按照该预定的加密方式将存储在该目标数据库中的每一单行数据进行加密后得到第二密文;

若该第一密文与对应的第二密文分别相同, 则验证通过;

若有该第一密文与对应的第二密文不相同, 将该第二密文对应的单行数据进行回滚清除, 以重新迁移该单行数据。

10.根据权利要求 8 所述的数据迁移的方法, 其特征在于, 所述步骤 S2 之后还包括: 获取该待迁移数据的主键标识, 以第一预设数量的主键标识为一组, 基于各组主键标识生成该待迁移数据的第一索引文件;

所述步骤 S3 之后还包括: 获取存储在目标数据库中的数据的主键标识, 以第二预设数量的主键标识为一组, 基于各组主键标识生成目标数据库中的数据的第二索引文件。

11.根据权利要求 9 所述的数据迁移的方法, 其特征在于, 所述步骤 S2 之后还包括: 获取该待迁移数据的主键标识, 以第一预设数量的主键标识为一组, 基于各组主键标识生成该待迁移数据的第一索引文件;

所述步骤 S3 之后还包括: 获取存储在目标数据库中的数据的主键标识, 以第二预设数量的主键标识为一组, 基于各组主键标识生成目标数据库中的数据的第二索引文件。

12.根据权利要求 8 所述的数据迁移的方法, 其特征在于, 在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置, 该方法还包括:

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

13.根据权利要求 9 所述的数据迁移的方法，其特征在于，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，该方法还包括：

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

14.根据权利要求 8 或 9 所述的数据迁移的方法，其特征在于，所述预定的解析方式为 xml 简单应用程序接口的解析方式。

15.一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有处理系统，所述处理系统被处理器执行时实现步骤：

在迁移平台的操作界面上，当接收到对一个或多个业务系统的原数据库的待迁移数据进行迁移的指令后，基于该指令对本次迁移的迁移参数以及该原数据库与目标数据库的映射数据进行配置；

当完成配置操作后，根据所配置的迁移参数及映射数据并利用多线程的方式读取该待迁移数据，并将待迁移数据转换为 XML 文件数据；

采用多个线程获取该 XML 文件数据，按照预定的解析方式对该 XML 文件数据进行解析，并采用与读取该待迁移数据异步的方式将解析后的 XML 文件数据写入目标数据库中。

16.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

按照预定的加密方式将原数据库的待迁移数据中每一单行数据进行加密后得到第一密文，按照该预定的加密方式将存储在该目标数据库中的每一单行数据进行加密后得到第二密文；

若该第一密文与对应的第二密文分别相同，则验证通过；

若有该第一密文与对应的第二密文不相同，将该第二密文对应的单行数

据进行回滚清除，以重新迁移该单行数据。

17.根据权利要求 15 所述的计算机可读存储介质，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

18.根据权利要求 16 所述的计算机可读存储介质，其特征在于，所述处理系统被所述处理器执行时，还实现如下步骤：

获取该待迁移数据的主键标识，以第一预设数量的主键标识为一组，基于各组主键标识生成该待迁移数据的第一索引文件；

获取存储在目标数据库中的数据的主键标识，以第二预设数量的主键标识为一组，基于各组主键标识生成目标数据库中的数据的第二索引文件。

19.根据权利要求 15 所述的计算机可读存储介质，其特征在于，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，所述处理系统被所述处理器执行时，还实现如下步骤：

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

20.根据权利要求 16 所述的计算机可读存储介质，其特征在于，在该操作界面上提供定时迁移、迁移暂停、迁移停止及回滚的配置，所述处理系统被所述处理器执行时，还实现如下步骤：

接收定时迁移、迁移暂停、迁移停止或回滚的操作指令，根据该操作指令执行相应的定时迁移、迁移暂停、迁移停止或回滚的操作。

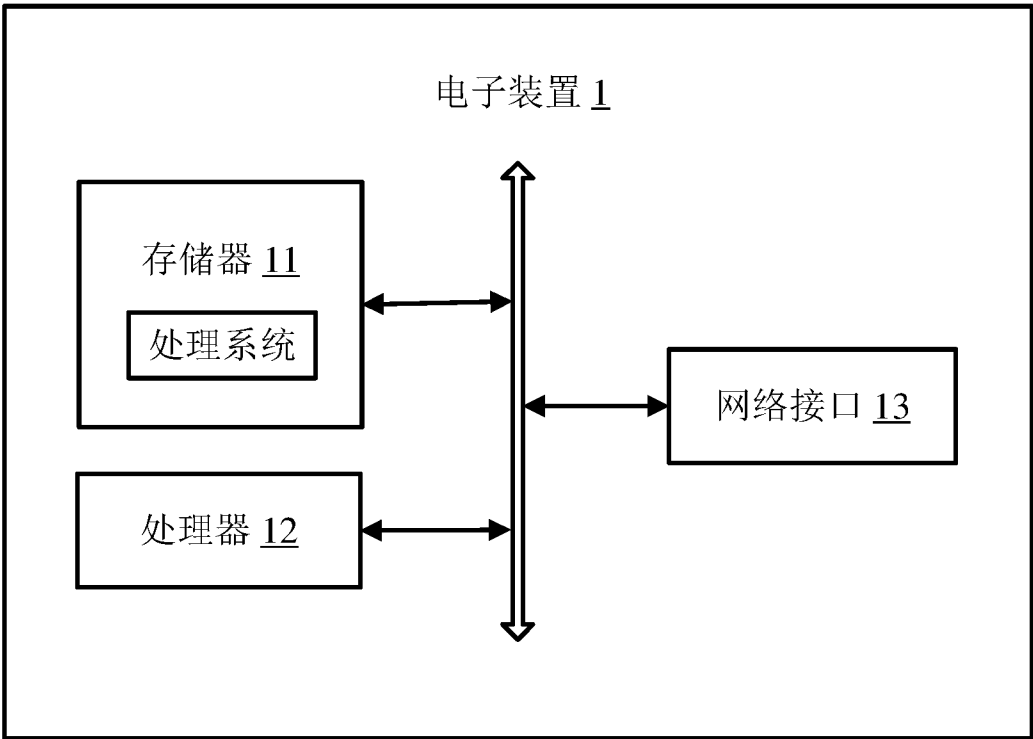


图 1

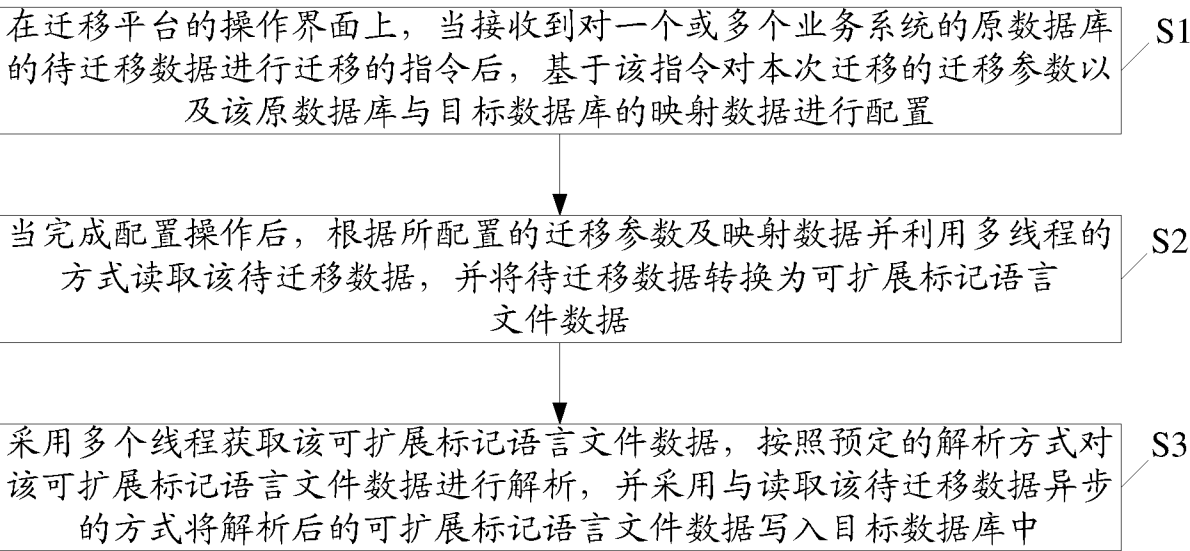


图 2

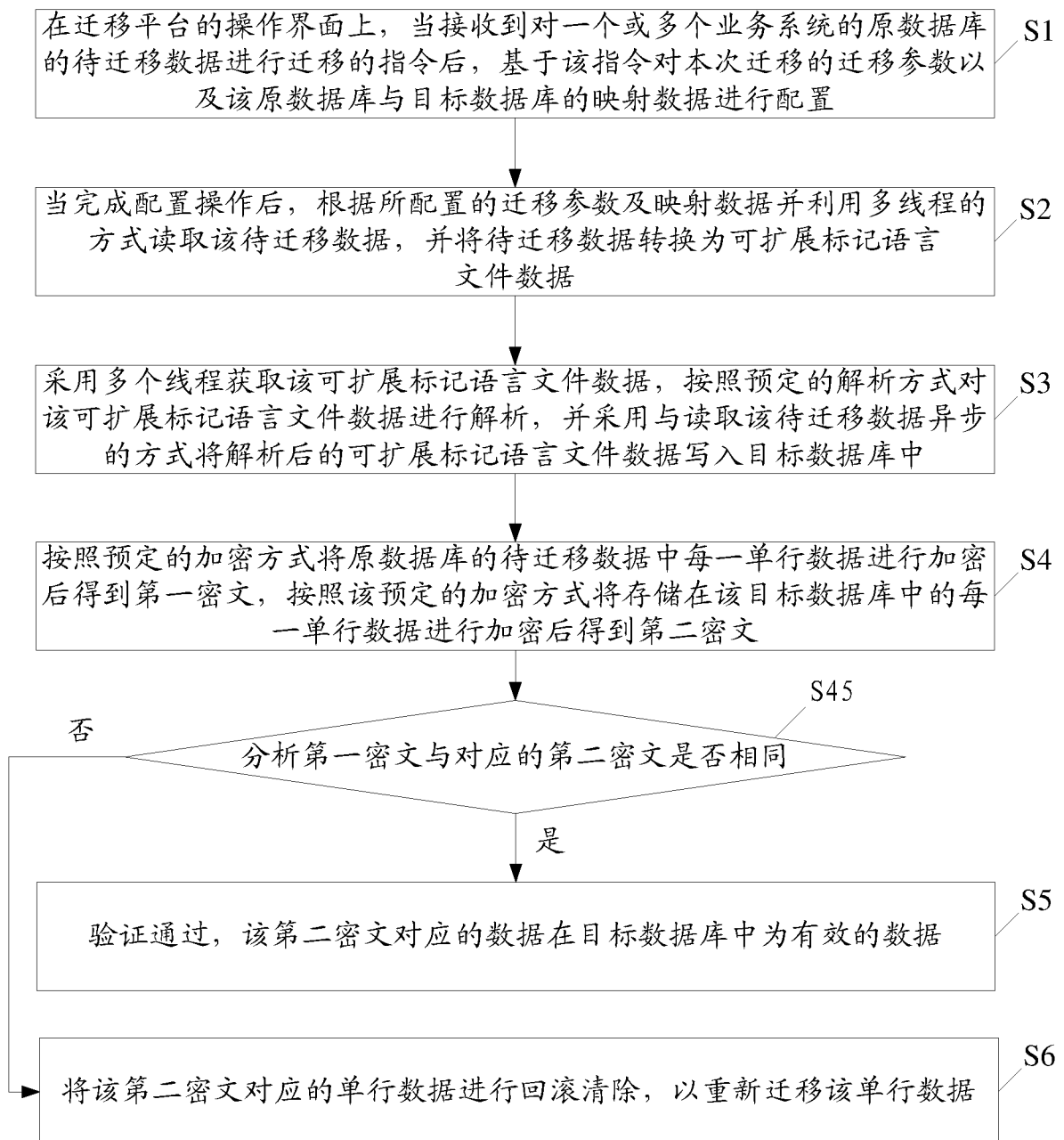


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/107702

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/21(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)

G06F 16/-; G06F 9/-

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)

DWPI; CNABS; CNTXT; CNKI: 数据库, 数据迁移, 中间文件, XML, 多线程; database?, data, migrat???, XML, multithread

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106598977 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD.) 26 April 2017 (2017-04-26) see description, paragraphs 24-32	1-20
Y	CN 104462119 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 25 March 2015 (2015-03-25) see description, paragraphs 26-32	1-20
Y	CN 106682534 A (ZHENGZHOU YUNHAI INFORMATION TECHNOLOGY CO., LTD.) 17 May 2017 (2017-05-17) see description, paragraphs 23 and 24	2, 4, 6, 7, 9, 11, 13, 14, 16, 18, 20
A	CN 102521318 A (HUAQIAO UNIVERSITY) 27 June 2012 (2012-06-27) entire document	1-20
A	CN 101739436 A (SUN, BIN ET AL.) 16 June 2010 (2010-06-16) entire document	1-20
A	CN 105760517 A (INSPUR (BEIJING) ELECTRONIC INFORMATION INDUSTRY CO., LTD.) 13 July 2016 (2016-07-13) entire document	1-20



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

“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

14 March 2019

Date of mailing of the international search report

09 April 2019

Name and mailing address of the ISA/CN

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

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/107702

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101789021 A (INSPUR COMMUNICATION INFORMATION SYSTEM LIMITED) 28 July 2010 (2010-07-28) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/107702

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106598977	A	26 April 2017	None			
CN	104462119	A	25 March 2015	WO	2015039564	A1	26 March 2015
CN	106682534	A	17 May 2017	None			
CN	102521318	A	27 June 2012	None			
CN	101739436	A	16 June 2010	CN	101739436	B	19 December 2012
CN	105760517	A	13 July 2016	None			
CN	101789021	A	28 July 2010	None			

国际检索报告

国际申请号

PCT/CN2018/107702

A. 主题的分类

G06F 16/21 (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)

G06F 16/-; G06F 9/-

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))

DWPI; CNABS; CNTXT; CNKI: 数据库, 数据迁移, 中间文件, XML, 多线程; database?, data, migrat???, XML, multithread

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
Y	CN 106598977 A (北京国双科技有限公司) 2017年 4月 26日 (2017 - 04 - 26) 参见说明书第24-32段	1-20
Y	CN 104462119 A (腾讯科技深圳有限公司) 2015年 3月 25日 (2015 - 03 - 25) 参见说明书第26-32段	1-20
Y	CN 106682534 A (郑州云海信息技术有限公司) 2017年 5月 17日 (2017 - 05 - 17) 参见说明书第23-24段	2, 4, 6-7, 9, 11, 13-14, 16, 18, 20
A	CN 102521318 A (华侨大学) 2012年 6月 27日 (2012 - 06 - 27) 全文	1-20
A	CN 101739436 A (孙彬等) 2010年 6月 16日 (2010 - 06 - 16) 全文	1-20
A	CN 105760517 A (浪潮北京电子信息产业有限公司) 2016年 7月 13日 (2016 - 07 - 13) 全文	1-20
A	CN 101789021 A (浪潮通信信息系统有限公司) 2010年 7月 28日 (2010 - 07 - 28) 全文	1-20

☐ 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

“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

2019年 3月 14日

国际检索报告邮寄日期

2019年 4月 9日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

郝晓丽

传真号 (86-10) 62019451

电话号码 62412331

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/107702

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	106598977	A	2017年 4月 26日	无			
CN	104462119	A	2015年 3月 25日	W0	2015039564	A1	2015年 3月 26日
CN	106682534	A	2017年 5月 17日	无			
CN	102521318	A	2012年 6月 27日	无			
CN	101739436	A	2010年 6月 16日	CN	101739436	B	2012年 12月 19日
CN	105760517	A	2016年 7月 13日	无			
CN	101789021	A	2010年 7月 28日	无			