

(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
2016年11月17日 (17.11.2016)



(10) 国际公布号
WO 2016/180168 A1

- (51) 国际专利分类号:
H04L 29/08 (2006.01)
- (21) 国际申请号: PCT/CN2016/079552
- (22) 国际申请日: 2016年4月18日 (18.04.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510236871.8 2015年5月11日 (11.05.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属开曼群岛大开曼乔治城资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 及
- (71) 申请人 (仅对美国): 董元元 (DONG, Yuanyuan) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。徐立 (XU, Li) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。闫卫斌 (YAN, Weibin) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969

号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。杨濮源 (YANG, Puyuan) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。李凯 (LI, Kai) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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。

[见续页]

(54) Title: DATA COPY METHOD AND DEVICE

(54) 发明名称: 一种数据复制方法和设备

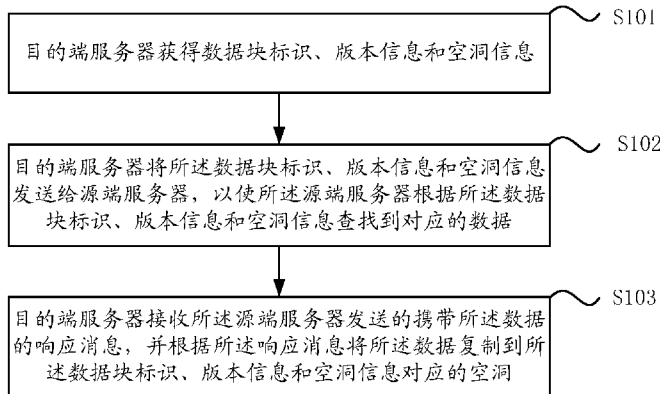


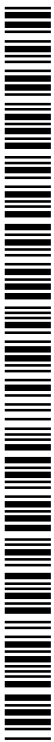
图 1

- S101 A destination server acquires a data block identifier, version information and hole information
- S102 The destination server transmits the data block identifier, the version information and the hole information to a source terminal server so as to enable the source terminal server to find corresponding data according to the data block identifier, the version information and the hole information
- S103 The destination server receives a response message carrying the data and transmitted by the source terminal server and according to the response message copies the data to the data block identifier, the version information and the hole corresponding to the hole information

(57) Abstract: Disclosed in embodiments of the present application are a data copy method and device. The method comprises: acquiring, by a destination server, a data block identifier, version information and hole information; transmitting, by the destination server, the data block identifier, the version information and the hole information to a source terminal server so as to enable the source terminal server to find corresponding data according to the data block identifier, the version information and the hole information; receiving, by the destination server, a response message carrying the data transmitted by the source terminal server and according to the response message copying the data to the data block identifier, the version information and the hole corresponding to the hole information, thereby copying data only to the hole in a random read/write manner, and eliminating the need to copy an entire amount of data, thereby reducing a data amount to be copied and copy time, improving data copy efficiency.

(57) 摘要: 本申请实施例公开了一种数据复制方法和设备, 包括: 目的端服务器获得数据块标识、版本信息和空洞信息; 所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器, 以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据;

所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息, 并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞, 从而实现了随机读写方式下只对空洞进行数据复制, 不需要进行数据的全量复制从而减少了待复制的数据量及复制时间, 提高了数据复制效率。



WO 2016/180168 A1



(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, 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))。

一种数据复制方法和设备

技术领域

本申请涉及通信技术领域，尤其涉及一种应用于分布式存储系统的数据复制方法和设备。

5

背景技术

在分布式存储系统中，会将数据块存储多个副本，每次数据块的更新需要在多个副本之间达成一致，当副本数存在不足，或者副本缺少一些版本的更新数据时，则需要从其它副本进行数据复制。

10 目前，现有方案实现数据复制的前提是：源端服务器或者目的端服务器需要能够根据本地维护的待复制数据的版本和对端维护的待复制数据的版本计算出数据增量或差量，进而能够进行从对端获得待复制数据的数据增量或差量进行复制。

15 在支持随机读写的分布式存储系统中，现有技术中采用从源端服务器向目的端服务器推送数据的模式，源端服务器需要将已有的数据块全量数据推送给目的端服务器，增加了复制的数据量，特别是在数据更新速度大于数据复制速度时，还会导致在预设值的数据复制时间内，数据复制过程无法完成的情况。

20 申请人在实现本发明的过程中发现，在支持随机读写的分布式存储系统中，现有技术存在以下问题，无法直接根据版本号的差值计算出数据增量或差量，只能进行数据的全量复制，增加了待复制的数据量及复制时间。

发明内容

本申请提供了一种数据复制方法及设备，用以实现随机读写的分布式存储系统的数据复制的问题。

25 为达到上述目的，本发明提供了一种数据复制的方法，至少包括以下步骤：
目的端服务器获得数据块标识、版本信息和空洞信息；

所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；

所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。

其中，所述目的端服务器获得数据块标识、版本信息和空洞信息，具体包括：

所述目的端服务器获得数据复制请求，所述数据复制请求中包括：数据块标识、版本信息和空洞信息。

其中，所述数据复制请求还包括：数据块复制起始版本号；

所述版本信息具体为：数据复制终止版本号；

所述目的端服务器获得数据块标识、版本信息和空洞信息之后还包括：

所述目的端服务器在判断所述数据块标识对应的数据块的数据复制终止版本号大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号一致时，执行将所述数据块标识、版本信息和空洞信息发送给源端服务器的步骤。

其中，所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号；

所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

其中，所述目的端服务器获得数据块标识、版本信息和空洞信息之后，所述方法还包括：

所述目的端服务器根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

其中，所述根据所述响应消息将所述数据复制到所述空洞信息对应的空洞之后，所述方法还包括：

所述目的端服务器将所述数据块标识、版本信息和空洞信息删除。

其中，所述空洞信息具体包括：数据偏移位置和数据偏移长度。

本申请实施例还提供了一种数据更新的方法，应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，包括：

源端服务器接收到目的端服务器发送的数据块标识、版本信息和空洞信息；

所述源端服务器在判断所述源端服务器的、所述数据块标识对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，获得所述数据块标识、版本信息和空洞信息对应的空洞所对应的数据；

所述源端服务器向所述目的端服务器发送携带所述数据的响应消息，以使所述目的端服务器根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。

其中，所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

其中，所述空洞信息具体包括：数据偏移位置和数据偏移长度。

本申请实施例还提供了一种目的端服务器，应用于包含目的端服务器和源端服务器的随机读写的分布式存储系统中，所述目的端服务器包括：

获得模块，用于获得数据块标识、版本信息和空洞信息；

发送模块，用于将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；

复制模块，用于接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到数据块标识对应的数据块。

其中，所述获得模块，具体用于获得数据复制请求，所述数据复制请求中包括：数据块标识、版本信息和空洞信息。

其中，所述数据复制请求还包括：数据块复制起始版本号；

所述版本信息具体为：数据复制终止版本号；

所述发送模块还用于：在判断所述数据块标识对应的数据块的数据复制终止版本号大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号一致时，将所述数据块标识、版本信息和空洞信息发送给源端服务器。

其中，所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号；

所述复制模块还用于，在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

其中，所述获得模块还用于：根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

其中，所述复制模块还用于，将所述数据块标识、版本信息和空洞信息删除。

其中，所述空洞信息具体包括：数据偏移位置和数据偏移长度。

本申请实施例还提供了一种源端服务器，应用于包含目的端服务器和源端服务器的随机读写的分布式存储系统中，包括：

接收模块，用于接收目的端服务器发送的数据块标识、版本信息和空洞信息；

5 获得模块，用于在判断所述源端服务器的、所述数据对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，获得所述数据块标识、版本信息和空洞信息对应空洞所对应的数据；

发送模块，用于向所述目的端服务器发送携带所述数据的响应消息，以使所述目的端服务器根据所述响应消息将所述数据复制到所述数据块标识、版本
10 信息和空洞信息对应的空洞。

其中，所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为
15 所述数据块的最新版本号。

其中，所述空洞信息具体包括：数据偏移位置和数据偏移长度。

与现有技术相比，本申请的上述实施例具有以下有益技术效果：

本申请实施例中，目的端服务器获得数据块标识、版本信息和空洞信息；所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，
20 以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞，从而实现了随机读写方式下只对空洞进行数据复制，不需要进行数据的全量复制从而减少了待复制的数据量及复制时间，提高了数据复制效率。

25

附图说明

图 1 为本申请实施例提供一种数据复制的方法的流程示意图；

图 2 为本申请实施例提供的有重合空洞的整合示意图；

图 3 为本申请实施例提供的另一种数据复制的方法的流程示意图；

图 4 为本申请实施例提供的一种目的端服务器的结构示意图；

5 图 5 为本申请实施例提供的一种源端服务器的结构示意图。

具体实施方式

下面将结合本发明的实施例中的附图，对本发明的实施例中的技术方案进行清楚、完整的描述，显然，下面所描述的实施例仅仅是本发明一部分实施例，
10 而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有付出创造性劳动前提下所获得的所有其他实施例，都属于本发明的实施例保护的范围。

如图 1 所示，为实施例一所提出的一种数据复制的方法的流程示意图，该方法应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，下面将在目的端服务器侧对该方法进行详细说明。
15

该方法，具体包括以下步骤：

步骤 S101、目的端服务器获得数据块标识、版本信息和空洞信息。

在发起数据复制时，目的端服务器会获得数据复制发起端发送的数据复制请求，该数据复制请求中包含：数据块标识，版本信息，空洞信息以及数据块
20 复制起始版本号，其中，版本信息具体为数据复制终止版本号。

所述目的端服务器获得数据块标识、版本信息和空洞信息之后，目的端服务器判断数据块标识对应的数据块的数据复制终止版本号是否大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块
25 当前的数据块版本号是否一致。需要注意的是，如果目的端服务器不存在所述数据块标识对应的信息，则判断目的端服务器当前不存在该数据块标识对应的数据块，则需要对该数据块进行全量复制。

如果判断结果为是，则该数据块标识，版本信息，空洞信息为有效继续执行下面的步骤。

如果判断结果为否，则该数据块标识，版本信息，空洞信息为无效，不再对该数据块标识，版本信息，空洞信息进行处理，或者将其删除。

进一步的，空洞信息具体包括：数据偏移位置和数据偏移长度，所述目的端服务器根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

目的端服务器获得数据复制请求之后，所述目的端服务器根据所述数据复制请求中的数据块标识、版本信息和空洞信息确定待复制数据的空洞，查询是否存在与所述空洞有重合的数据复制请求，就是查询是否存在所述数据复制请求中的数据块标识相同、版本信息小于所述数据复制请求中的版本信息，且空洞信息中的数据偏移位置和数据偏移长度有重合区间的数据库复制请求，如果存在，则将有重合的空洞整合成一个空洞，就是将有重合的空洞所对应的数据块标识相同、版本信息和空洞信息合并成一个空洞所对应的数据块标识相同、版本信息和空洞信息。其中，需要注意的是，合并成的空洞所对应的版本信息为最后一个数据库复制请求的版本信息对应的版本号。

具体地，如图 2 所示，将有重合的空洞整合成一个空洞，具体包括：部分覆盖、中间覆盖和全部覆盖。根据数据偏移位置和数据偏移长度可以确定待复制数据的空洞，当目的端服务器获得的数据块标识为 1，版本信息为 16，空洞信息确定的空洞为 V2，如果查询到存在数据块标识为 1，版本信息为 13，空洞信息确定的空洞为 V1，且 V2 和 V1 有重合时，则将 V2 和 V1 整合成一个空洞，并将该空洞对应的空洞信息更新为原来空洞为 V2 的空洞信息，即将整合成的空洞更新为新收到的数据库复制请求中的空洞。

具体地，如图 2 所示，由于版本 16 所对应的空洞比版本 13 所对应的空洞数据更新，将 V2 和 V1 整合成一个空洞，就是“V2 交 V1”，具体如图 2 中所述的三种合并方式。

步骤 S102、目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据。

目的端将所述数据块标识、版本信息和空洞信息发送给源端服务器后，所述源端服务器根据数据块标识、版本信息和空洞信息，可以查找到数据块标识对应的数据块在版本信息对应的版本下的空洞信息所对应的数据。

5 步骤 S103、目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。

所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，所述响应消息中包括：所述数据块标识、版本信息和空洞信息和所述数据，目的端服务器根据所述数据块标识、版本信息和空洞信息将所述数据复制到所述
10 空洞信息对应的空洞。

当根据所述响应消息将所述数据复制到所述空洞信息对应的空洞之后，所述目的端服务器将所述数据块标识、版本信息和空洞信息删除。

其中，所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号；所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述
15 数据块的最新版本号。

具体地，响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，例如源端服务器的、所述数据对应的数据块的当前版本号为 17，
20 所述目的端服务器的、所述数据对应的数据块的当前版本号为 18，则将该数据块的版本号更新为 18，完成本次数据复制过程。如果所述目的端服务器的、所述数据对应的数据块的当前版本号为 16，则再次接收到数据复制请求，且经判断请求为有效后，还需要向所述源端服务器发送该数据复制请求的数据块标识、版本信息和空洞信息。

25 本申请实施例中，目的端服务器获得数据块标识、版本信息和空洞信息；所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息

对应的空洞，从而实现了随机读写方式下只对空洞进行数据复制，不需要进行数据的全量复制从而减少了待复制的数据量及复制时间，提高了数据复制效率。

如图 3 所示，为实施例二所提出另的一种数据更新的方法的流程示意图，该方法应用于包括包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，下面将在源端服务器侧对该方法进行详细说明。

该方法，具体包括以下步骤：

步骤 S201、源端服务器接收到目的端服务器发送的数据块标识、版本信息和空洞信息。

其中，所述空洞信息具体包括：数据偏移位置和数据偏移长度。

步骤 S202、所述源端服务器在判断所述源端服务器的、所述数据对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，获得所述数据块标识、版本信息和空洞信息对应的空洞所对应的数据。

源端服务器判断自身的、所述数据块标识对应的数据块的当前版本号是否大于等于所述版本信息，当判断结果为是，则查找到所述数据块标识、版本信息和空洞信息对应的数据，并将所述数据携带在响应请求中发送给所述目的端服务器，以使所述目的端服务器将该数据复制到对应的空洞；当判断结果为否，则源端服务器并不存在待复制的数据，则向所述目的端服务器发送失败响应。

其中，响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

本申请实施例中，目的端服务器获得数据块标识、版本信息和空洞信息；所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞，从而实现了随机读写方式下只对空洞进行数据复制，不需要进行数据的全量复制从而减少了待复制的数据量及复制时间，提高了数据复制效率。

基于上述实施例一同样的发明构思，本发明实施例三中还提供了一种目的端服务器，应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，如图4所示，该目的端服务器包括：

5 获得模块41，用于获得数据块标识、版本信息和空洞信息。

发送模块42，用于将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据。

10 复制模块43，用于接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到数据块标识对应的数据块。

所述获得模块，具体用于获得数据复制请求，所述数据复制请求中包括：数据块标识、版本信息和空洞信息。

所述数据复制请求还包括：数据块复制起始版本号。

所述版本信息具体为：数据复制终止版本号。

15 所述发送模块还用于：在判断所述数据块标识对应的数据块的数据复制终止版本号大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号一致时，将所述数据块标识、版本信息和空洞信息发送给源端服务器。

20 所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号。

所述复制模块还用于，在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块
25 的最新版本号。

所述获得模块还用于：根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

所述复制模块还用于，将所述数据块标识、版本信息和空洞信息删除。

所述空洞信息具体包括：数据偏移位置和数据偏移长度。

本申请实施例中，目的端服务器获得数据块标识、版本信息和空洞信息；所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，
5 并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞，从而实现了随机读写方式下只对空洞进行数据复制，不需要进行数据的全量复制从而减少了待复制的数据量及复制时间，提高了数据复制效率。

基于上述实施例二同样的发明构思，本发明实施例四中还提供了一种目的端服务器，应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，如图 5 所示，该源端服务器包括：
10

接收模块 51，用于接收目的端服务器发送的数据块标识、版本信息和空洞信息。

获得模块 52，用于在判断所述源端服务器的、所述数据对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，
15 获得所述数据块标识、版本信息和空洞信息对应空洞所对应的数据。

发送模块 53，用于向所述目的端服务器发送携带所述数据的响应消息，以使所述目的端服务器根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。
20

所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。所述空洞信息具体包括：数据偏移位置和数据偏移长度。
25

本申请实施例中，目的端服务器获得数据块标识、版本信息和空洞信息；所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息

对应的空洞，从而实现了随机读写方式下只对空洞进行数据复制，不需要进行数据的全量复制从而减少了待复制的数据量及复制时间，提高了数据复制效率。

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本发明可以通过硬件实现，也可以借助软件加必要的通用硬件平台的方式来实现。基于这样的理解，本发明的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本发明各个实施场景所述的方法。

本领域技术人员可以理解附图只是一个优选实施场景的示意图，附图中的模块或流程并不一定是实施本发明所必须的。

本领域技术人员可以理解实施场景中的装置中的模块可以按照实施场景描述进行分布于实施场景的装置中，也可以进行相应变化位于不同于本实施场景的一个或多个装置中。上述实施场景的模块可以合并为一个模块，也可以进一步拆分成多个子模块。

上述本发明序号仅仅为了描述，不代表实施场景的优劣。

以上公开的仅为本发明的几个具体实施场景，但是，本发明并非局限于此，任何本领域的技术人员能思之的变化都应落入本发明的保护范围。

权 利 要 求 书

1、一种数据复制的方法，应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，其特征在于，该方法包括：

目的端服务器获得数据块标识、版本信息和空洞信息；

5 所述目的端服务器将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；

所述目的端服务器接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息
10 对应的空洞。

2、如权利要求 1 所述的方法，其特征在于，所述目的端服务器获得数据块标识、版本信息和空洞信息，具体包括：

所述目的端服务器获得数据复制请求，所述数据复制请求中包括：数据块标识、版本信息和空洞信息。

15 3、如权利要求 2 所述的方法，其特征在于，

所述数据复制请求还包括：数据块复制起始版本号；

所述版本信息具体为：数据复制终止版本号；

所述目的端服务器获得数据块标识、版本信息和空洞信息之后还包括：

所述目的端服务器在判断所述数据块标识对应的数据块的数据复制终止
20 版本号大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号一致时，执行将所述数据块标识、版本信息和空洞信息发送给源端服务器的步骤。

4、如权利要求 3 所述的方法，其特征在于，

25 所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号；

所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

- 5 5、如权利要求 2 所述的方法，其特征在于，所述目的端服务器获得数据块标识、版本信息和空洞信息之后，所述方法还包括：

所述目的端服务器根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

- 10 6、如权利要求 1 所述的方法，其特征在于，所述根据所述响应消息将所述数据复制到所述空洞信息对应的空洞之后，所述方法还包括：

所述目的端服务器将所述数据块标识、版本信息和空洞信息删除。

- 7、如权利要求 1-6 任一项所述的方法，其特征在于，

所述空洞信息具体包括：数据偏移位置和数据偏移长度。

- 15 8、一种数据更新的方法，应用于包含有源端服务器和目的端服务器的随机读写的分布式存储系统中，其特征在于，该方法包括：

源端服务器接收到目的端服务器发送的数据块标识、版本信息和空洞信息；

- 20 所述源端服务器在判断所述源端服务器的、所述数据块标识对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，获得所述数据块标识、版本信息和空洞信息对应的空洞所对应的数据；

所述源端服务器向所述目的端服务器发送携带所述数据的响应消息，以使所述目的端服务器根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。

- 9、如权利要求 7 所述的方法，其特征在于，

- 25 所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本

号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

10、如权利要求 7-9 任一项所述的方法，其特征在于，

5 所述空洞信息具体包括：数据偏移位置和数据偏移长度。

11、一种目的端服务器，应用于包含目的端服务器和源端服务器的随机读写的分布式存储系统中，所述目的端服务器包括：

获得模块，用于获得数据块标识、版本信息和空洞信息；

10 发送模块，用于将所述数据块标识、版本信息和空洞信息发送给源端服务器，以使所述源端服务器根据所述数据块标识、版本信息和空洞信息查找到对应的数据；

复制模块，用于接收所述源端服务器发送的携带所述数据的响应消息，并根据所述响应消息将所述数据复制到数据块标识对应的数据块。

12、如权利要求 11 所述的所述目的端服务器，其特征在于，

15 所述获得模块，具体用于获得数据复制请求，所述数据复制请求中包括：数据块标识、版本信息和空洞信息。

13、如权利要求 12 所述的所述目的端服务器，其特征在于，

所述数据复制请求还包括：数据块复制起始版本号；

所述版本信息具体为：数据复制终止版本号；

20 所述发送模块还用于：在判断所述数据块标识对应的数据块的数据复制终止版本号大于所述数据块标识对应的数据块的数据块复制起始版本号，且所述数据块标识对应的数据块的数据块复制起始版本号与目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号一致时，将所述数据块标识、版本信息和空洞信息发送给源端服务器。

25 14、如权利要求 13 所述的方法，其特征在于，

所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号；

所述复制模块还用于，在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

15、如权利要求 12 所述的所述目的端服务器，其特征在于，

所述获得模块还用于：根据所述数据块标识、版本信息和空洞信息确定待复制数据的空洞，并在查询到与所述空洞有重合的数据复制请求后，将有重合的空洞整合成一个新的空洞。

16、如权利要求 11 所述的方法，其特征在于，

所述复制模块还用于，将所述数据块标识、版本信息和空洞信息删除。

17、如权利要求 11-16 任一项所述的所述目的端服务器，其特征在于，

所述空洞信息具体包括：数据偏移位置和数据偏移长度。

18、一种源端服务器，应用于包含目的端服务器和源端服务器的随机读写的分布式存储系统中，其特征在于，所述源端服务器包括：

接收模块，用于接收目的端服务器发送的数据块标识、版本信息和空洞信息；

获得模块，用于在判断所述源端服务器的、所述数据对应的数据块的当前版本号大于等于所述版本信息时，根据数据块标识、版本信息和空洞信息，获得所述数据块标识、版本信息和空洞信息对应空洞所对应的数据；

发送模块，用于向所述目的端服务器发送携带所述数据的响应消息，以使所述目的端服务器根据所述响应消息将所述数据复制到所述数据块标识、版本信息和空洞信息对应的空洞。

19、如权利要求 18 所述的源端服务器，其特征在于，

所述响应消息还包括：所述源端服务器的、所述数据对应的数据块的当前版本号，用于使所述目的端服务器在判断目的端服务器的、当前的数据块版本号大于等于所述源端服务器的、所述数据对应的数据块的当前版本号时，将目的端服务器的、所述数据块标识对应的数据块当前的数据块版本号更新为所述数据块的最新版本号。

20、如权利要求 18-19 任一项所述的源端服务器，其特征在于，

所述空洞信息具体包括：数据偏移位置和数据偏移长度。

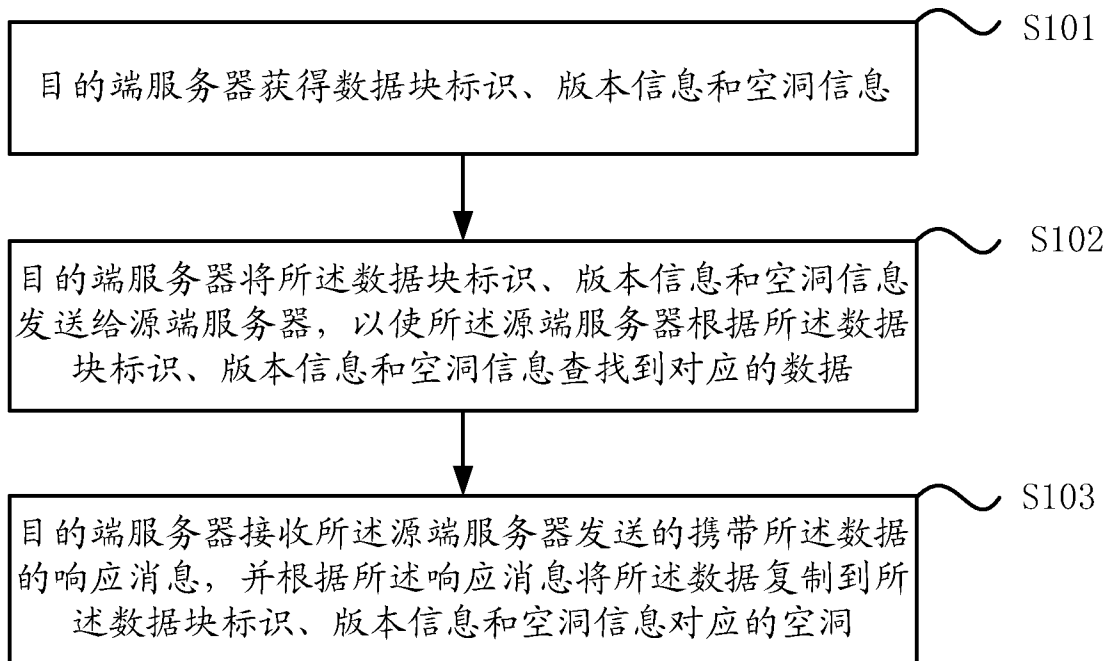


图 1

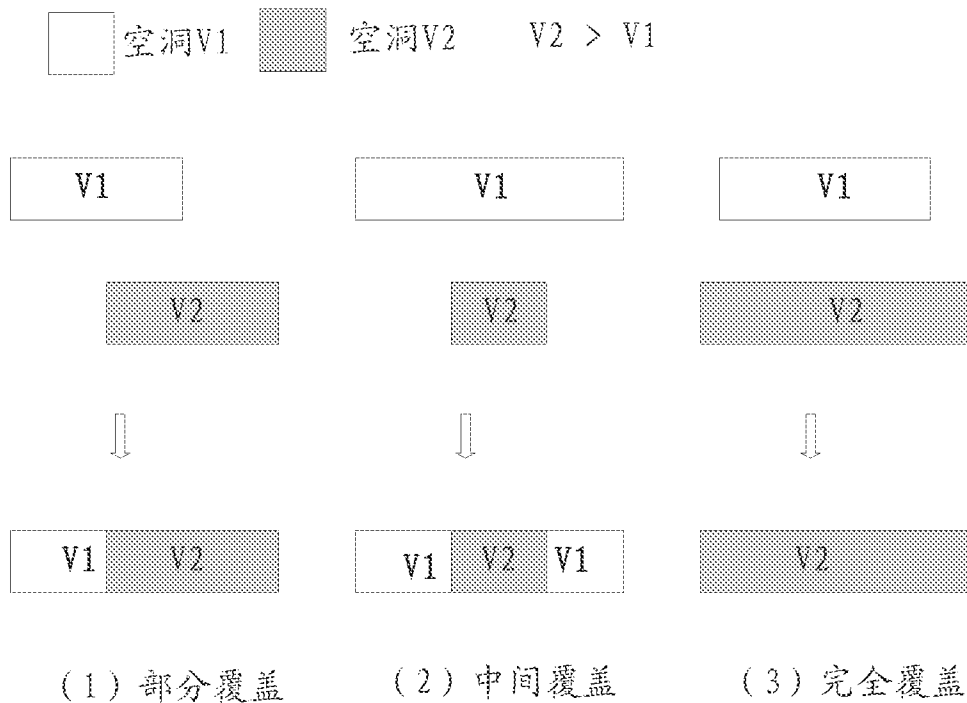


图 2

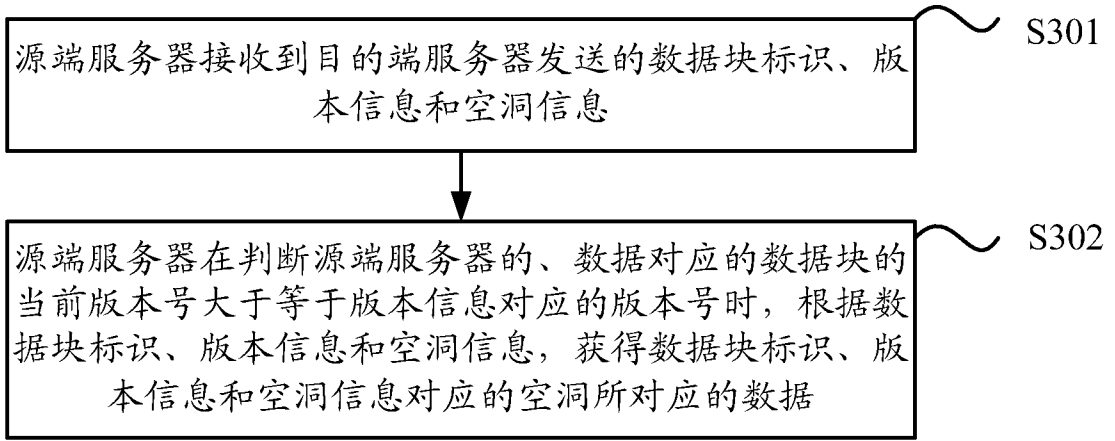


图 3

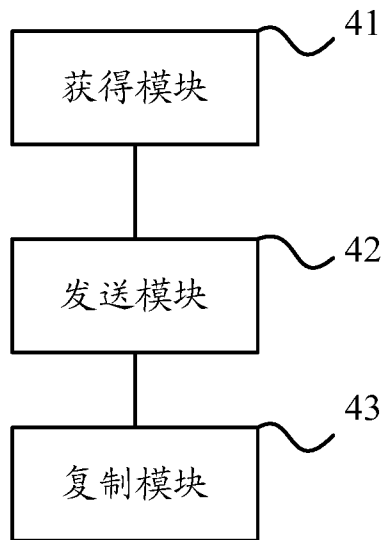


图 4

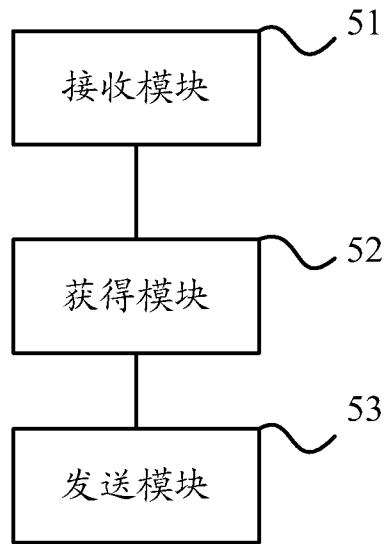


图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/079552

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (2006.01) 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)

H04L

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, VEN: duplicat???, synchronizat???, copy, incremental, destination, source, ID, sequence, version, edition, hole, void???, offset

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103795754 A (CHINA TELECOM CO., LTD.) 14 May 2014 (14.05.2014) description, paragraphs [0073] to [0100], claim 1	1-20
A	CN 102508835 A (UFIDA SOFTWARE CO., LTD.) 20 June 2012 (20.06.2012) the whole document	1-20
A	CN 102541461 A (ALIBABA GROUP HOLDING CO., LTD.) 04 July 2012 (04.07.2012) the whole document	1-20
A	US 2003158869 A1 (INT BUSINESS MACHINES CORP.) 21 August 2003 (21.08.2003) the whole document	1-20

Further documents are listed in the continuation of Box C. See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search 27 June 2016	Date of mailing of the international search report 11 July 2016
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LI, Fan Telephone No. (86-10) 62089572

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/079552

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103795754 A	14 May 2014	None	
CN 102508835 A	20 June 2012	CN 102508835 B	15 April 2015
CN 102541461 A	04 July 2012	CN 102541461 B	19 November 2014
		HK 1168177 A1	31 July 2015
		HK 1168177 A0	21 December 2012
US 2003158869 A1	21 August 2003	US 7747576 B2	29 June 2010

<p>A. 主题的分类</p> <p>H04L 29/08 (2006. 01) i</p> <p>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</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>H04L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNABS, CNTXT, CNKI, VEN: 复制, 同步, 拷贝, 增量, 目的, 源, 标识, 序号, 版本, 空洞, 偏移, duplicat???, synchronizat???, copy, incremental, destination, source, ID, sequence, version, edition, hole, void???, offset</p>																	
<p>C. 相关文件</p> <table border="1" style="width:100%; border-collapse: collapse;"> <thead> <tr> <th style="width:10%;">类型*</th> <th style="width:70%;">引用文件, 必要时, 指明相关段落</th> <th style="width:20%;">相关的权利要求</th> </tr> </thead> <tbody> <tr> <td style="text-align:center;">X</td> <td>CN 103795754 A (中国电信股份有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0073-0100段, 权利要求1</td> <td style="text-align:center;">1-20</td> </tr> <tr> <td style="text-align:center;">A</td> <td>CN 102508835 A (用友软件股份有限公司) 2012年 6月 20日 (2012 - 06 - 20) 全文</td> <td style="text-align:center;">1-20</td> </tr> <tr> <td style="text-align:center;">A</td> <td>CN 102541461 A (阿里巴巴集团控股有限公司) 2012年 7月 4日 (2012 - 07 - 04) 全文</td> <td style="text-align:center;">1-20</td> </tr> <tr> <td style="text-align:center;">A</td> <td>US 2003158869 A1 (INT BUSINESS MACHINES CORP) 2003年 8月 21日 (2003 - 08 - 21) 全文</td> <td style="text-align:center;">1-20</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103795754 A (中国电信股份有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0073-0100段, 权利要求1	1-20	A	CN 102508835 A (用友软件股份有限公司) 2012年 6月 20日 (2012 - 06 - 20) 全文	1-20	A	CN 102541461 A (阿里巴巴集团控股有限公司) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-20	A	US 2003158869 A1 (INT BUSINESS MACHINES CORP) 2003年 8月 21日 (2003 - 08 - 21) 全文	1-20
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 103795754 A (中国电信股份有限公司) 2014年 5月 14日 (2014 - 05 - 14) 说明书第0073-0100段, 权利要求1	1-20															
A	CN 102508835 A (用友软件股份有限公司) 2012年 6月 20日 (2012 - 06 - 20) 全文	1-20															
A	CN 102541461 A (阿里巴巴集团控股有限公司) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-20															
A	US 2003158869 A1 (INT BUSINESS MACHINES CORP) 2003年 8月 21日 (2003 - 08 - 21) 全文	1-20															
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																	
<p>* 引用文件的具体类型:</p> <table style="width:100%;"> <tr> <td style="width:50%; vertical-align: top;"> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> </td> <td style="width:50%; vertical-align: top;"> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p> </td> </tr> </table>			<p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>													
<p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																
<p>国际检索实际完成的日期</p> <p style="text-align:center;">2016年 6月 27日</p>	<p>国际检索报告邮寄日期</p> <p style="text-align:center;">2016年 7月 11日</p>																
<p>ISA/CN的名称和邮寄地址</p> <p style="text-align:center;">中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10) 62019451</p>	<p>受权官员</p> <p style="text-align:center;">李凡</p> <p>电话号码 (86-10) 62089572</p>																

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/079552

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	103795754	A	2014年 5月 14日	无			
CN	102508835	A	2012年 6月 20日	CN	102508835	B	2015年 4月 15日
CN	102541461	A	2012年 7月 4日	CN	102541461	B	2014年 11月 19日
				HK	1168177	A1	2015年 7月 31日
				HK	1168177	A0	2012年 12月 21日
US	2003158869	A1	2003年 8月 21日	US	7747576	B2	2010年 6月 29日