

(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



(10) 国际公布号

WO 2022/174705 A1

(43) 国际公布日

2022 年 8 月 25 日 (25.08.2022)

(51) 国际专利分类号:

G06F 9/50 (2006.01)

G06F 16/27 (2019.01)

UNIVERSITY) [CN/CN]; 中国北京市海淀区清华大学, Beijing 100084 (CN)。

(21) 国际申请号:

PCT/CN2022/072422

(22) 国际申请日:

2022 年 1 月 17 日 (17.01.2022)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

202110197578.0

2021 年 2 月 22 日 (22.02.2021) CN

(71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。清华大学(TSINGHUA

Guangdong 518129 (CN)。

清华大学(TSINGHUA

Guangdong 518129 (CN)。

(72) 发明人: 舒继武(SHU, Jiwu); 中国北京市海淀区清华大学, Beijing 100084 (CN)。汪庆(WANG, Qing); 中国北京市海淀区清华大学, Beijing 100084 (CN)。陈游旻(CHEN, Youmin); 中国北京市海淀区清华大学, Beijing 100084 (CN)。陆游游(LU, Youyou); 中国北京市海淀区清华大学, Beijing 100084 (CN)。崔文林(CUI, Wenlin); 中国广东省深圳龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。钟刊(ZHONG, Kan); 中国广东省深圳龙岗区坂田华为总部办公楼, Guangdong 518129 (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,

(54) Title: DATA PROCESSING METHOD, COMPUTER SYSTEM AND INTERMEDIATE DEVICE

(54) 发明名称: 数据处理方法、计算机系统和中间设备

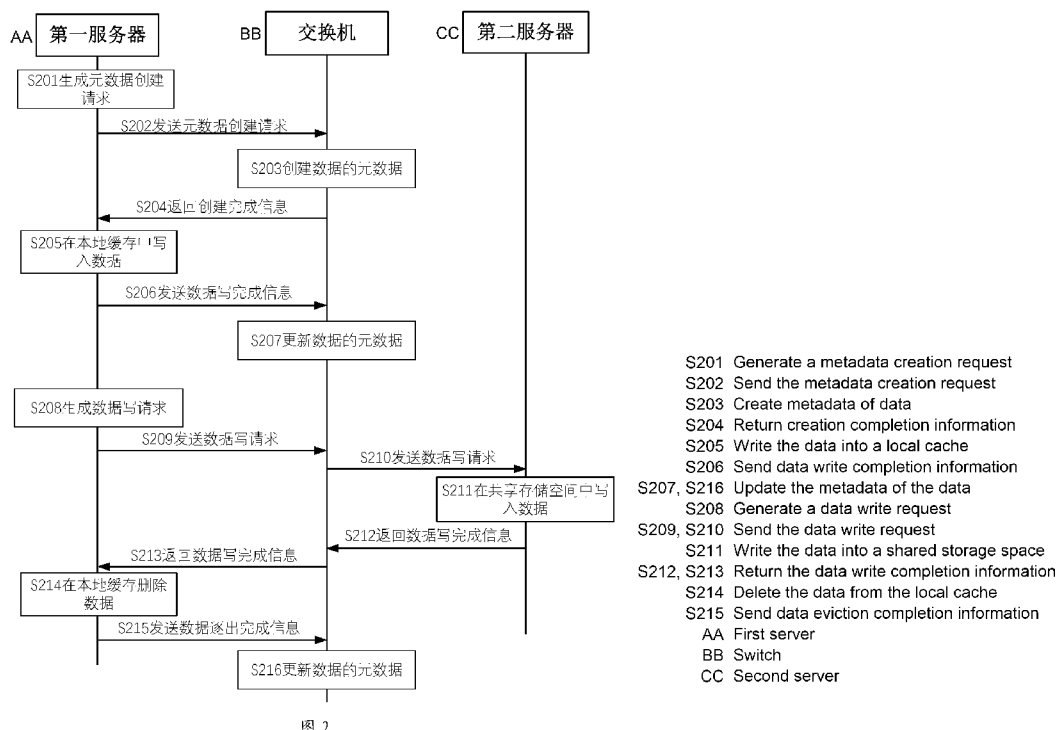


图 2

(57) Abstract: Provided are a data processing method, a computer system and an intermediate device. The computer system comprises an intermediate device, a first server and a second server, wherein the first server and the second server are connected by means of the intermediate device, the second server stores a first version of data, the first version is the latest version of the data, the intermediate device records information of the server which stores the latest version of the data, and the first server is used for acquiring the first version of the data from the second server by means of the intermediate device. Therefore, the burden on a server is reduced, thereby

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, IT, 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, 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, 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))。

improving the system performance.

(57) 摘要: 提供一种数据处理方法、计算机系统和中间设备, 计算机系统包括中间设备、第一服务器和第二服务器, 第一服务器和第二服务器通过中间设备连接, 第二服务器中存储有数据的第一版本, 第一版本为数据的最新版本, 中间设备中记录有存储数据的最新版本的服务器的信息, 第一服务器用于通过中间设备从第二服务器获取数据的第一版本。从而减少了服务器的负担, 提高了系统性能。

数据处理方法、计算机系统和中间设备

技术领域

本申请涉及存储技术领域，具体涉及一种数据处理方法、计算机系统和中间设备。

背景技术

在共享存储系统中，例如在具有可以被多个服务器共享的全局内存的存储系统或者分布式共享文件系统中，需要保证多个服务器中所缓存数据的一致性，即保证每个服务器都可以访问到最新版本的数据。在一种保证缓存一致性的方法中，对于共享存储系统中的每份数据，在该数据的归属服务器中记录该数据的元数据，元数据中包括缓存该数据最新版本的服务器的信息。在数据被归属服务器之外的其他服务器修改之后，归属服务器需要更新元数据，另外，当其他服务器访问数据的最新版本时，需要由归属服务器转发数据访问请求，从而增加了归属服务器的负担。

发明内容

本申请实施例旨在提供一种数据处理方案，通过由连接多个服务器的中间设备存储数据的最新版本所在的服务器的信息，通过中间设备获取数据的最新版本，从而减少了服务器的负担，提高了系统性能。

为实现上述目的，本申请第一方面提供一种计算机系统，包括中间设备、第一服务器和第二服务器，所述第一服务器和第二服务器通过所述中间设备连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述第一服务器用于通过所述中间设备从所述第二服务器获取所述数据的第一版本。

通过由中间设备记录存储数据最新版本的第二服务器的信息，其它服务器可通过中间设备转发数据获取请求和数据最新版本，从而可从第二服务器获取数据的最新版本，减少了该数据的归属服务器的负担，提高了系统性能。

在本申请第一方面的一种可能的实现方式中，所述第二服务器还用于在获取所述数据的第一版本后，指示所述中间设备记录所述第二服务器为存储所述数据的最新版本的服务器的信息。通过由中间设备在服务器获取数据的最新版本之后记录该服务器的信息，从而其它服务器可通过中间设备根据记录的信息获取数据的最新版本。其中，服务器可通过以下任一方式获取数据的最新版本：新写数据、从其他服务器或本地获取数据之后改写数据、从其他服务器读取数据。

在本申请第一方面的一种可能的实现方式中，所述中间设备存储有所述数据的元数据，所述中间设备用于在所述数据的元数据中记录所述第二服务器的信息。

在本申请第一方面的一种可能的实现方式中，所述第一服务器在通过所述中间设备从所述第二服务器获取所述数据的第一版本时，具体用于向所述中间设备发送对所述数据的获取请求，所述中间设备还用于：从所述第一服务器接收对所述数据的获取请求，根据存储所述数据的最新版本的服务器的信息将所述获取请求转发给所述第二服务器；从所述第二服务器接收所述数据的第一版本并转发给所述第一服务器。

通过由中间设备基于所记录的存储最新版本的服务器的信息将从第一服务器接收的数据获取请求转发给第二服务器，并将从第二服务器接收的数据的最新版本转发给第一服务器，减少了数据的归属服务器的负担。

在本申请第一方面的一种可能的实现方式中，所述第一服务器用于在接收到对所述数据的改写请求，并在本地没有命中所述数据的第一版本时，发送所述获取请求至所述中间设备；所述第一服务器在获取所述数据的第一版本之后，还用于将所述数据的第一版本改写为第二版本。

在本申请第一方面的一种可能的实现方式中，所述第一服务器还用于在将所述数据的第一版本改写为第二版本之后，向所述中间设备发送改写完成信息；所述中间设备还用于从所述第一服务器接收所述改写完成信息，基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

通过由中间设备在数据被改写之后接收改写完成信息并修改存储最新版本的服务器的信息，减少了数据的归属服务器的负担。

在本申请第一方面的一种可能的实现方式中，所述获取请求中携带有改写标识，在中间设备将所述获取请求转发给所述第二服务器之后，所述第二服务器用于根据所述获取请求中的改写标识在获取所述数据的第一版本之后删除所存储的所述数据的第一版本。

通过由第二服务器在获取数据的最新版本之后删除本地存储的数据的最新版本，保证了数据的缓存一致性，同时也减轻了服务器的负担。

在本申请第一方面的一种可能的实现方式中，所述中间设备中记录有对所述数据的写锁信息，所述中间设备还用于：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述元数据的写锁信息以对所述数据解锁。

通过由中间设备记录写锁信息并在接收到指示改写的获取请求之后通过写锁信息对数据进行锁定，保证了数据的缓存一致性，同时减轻了数据的归属服务器的负担。

在本申请第一方面的一种可能的实现方式中，所述数据为热点数据。通过使得中间设备只存储热点数据的元数据，减少了对中间设备的存储空间的需求。

在本申请第一方面的一种可能的实现方式中，所述系统还包括第三服务器，所述第三服务器与所述中间设备连接，所述第三服务器中也存储有所述数据的第一版本，所述第二服务器在接收到对所述数据的改写请求，并在本地命中所述数据的第一版本后，还用于向所述中间设备发送对所述数据的删除请求，所述中间设备还用于：从所述第二服务器接收对所述数据的删除请求，根据所述存储所述数据的最新版本的服务器的信息，将所述删除请求转发给所述第三服务器，指示所述第三服务器删除所述数据的第一版本；所述第二服务器还用于将所述数据的第一版本改写为第二版本。

在本申请第一方面的一种可能的实现方式中，所述中间设备为网络交换设备。

在本申请第一方面的一种可能的实现方式中，所述中间设备包括相互连接的第一中间设备和第二中间设备，其中，所述第一中间设备中存储有归属于所述第一服务器的数据的元数据，所述第二中间设备中存储有归属于所述第二服务器的数据的元数据。

在本申请第一方面的一种可能的实现方式中，所述计算机系统为全局内存系统或者分布式共享文件系统。

本申请第二方面提供一种数据处理方法，其特征在于，由中间设备执行，所述中间设备分别与第一服务器和第二服务器连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述方法包括：从所述第一服务器接收对所述数据的获取请求；根据所述存储所述数据的最新版本的服务器的信息，将所述获取请求转发给所述第二服务器；从所述第二服务器接收所述数据的第一版本；将所述数据的第一版本转发给所述第一服务器。

在本申请第二方面的一种可能的实现方式中，所述方法还包括：从所述第一服务器接收对所述数据的改写完成信息；基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

在本申请第二方面的一种可能的实现方式中，所述获取请求中携带有改写标识，所述方法还包括，在将所述获取请求转发给所述第二服务器之后，从所述第二服务器接收所述数据的第一版本被删除的信息。

在本申请第二方面的一种可能的实现方式中，所述中间设备中记录有对所述数据的写锁信息，所述方法还包括：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述写锁信息以对所述数据解锁。

在本申请第二方面的一种可能的实现方式中，所述数据为热点数据。

本申请第三方面提供一种中间设备，所述中间设备分别与第一服务器和第二服务器连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述中间设备包括：接收单元，用于从所述第一服务器接收对所述数据的获取请求；转发单元，用于根据所述存储所述数据的最新版本的服务器的信息，将所述获取请求转发给所述第二服务器；所述接收单元还用于从所述第二服务器接收所述数据的第一版本；所述转发单元还用于将所述数据的第一版本转发给所述第一服务器。

在本申请第三方面的一种可能的实现方式中，所述接收单元还用于从所述第一服务器接收对所述数据的改写完成信息；所述中间设备还包括修改单元，用于基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

在本申请第三方面的一种可能的实现方式中，所述获取请求中携带有改写标识，所述接收单元还用于在将所述获取请求转发给所述第二服务器之后，从所述第二服务器接收所述数据的第一版本被删除的信息。

在本申请第三方面的一种可能的实现方式中，所述中间设备中记录有对所述数据的写锁信息，所述修改单元还用于：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述写锁信息以对所述数据解锁。

在本申请第三方面的一种可能的实现方式中，所述数据为热点数据。

本申请第四方面提供一种中间设备，包括处理单元和存储单元，所述存储单元中存储有可执行代码，所述处理单元执行所述可执行代码以实现本申请第二方面所述的方法。

本申请第五方面提供一种中间设备，包括：通信接口，用于与服务器进行数据传输；处理单元，用于对所述通信接口接收的数据进行处理，以执行本申请第二方面所述的方法。

附图说明

通过结合附图描述本申请实施例，可以使得本申请实施例更加清楚：

图 1 为本申请实施例提供的共享存储系统的架构图；

图 2 为本申请实施例提供的在共享存储系统中写数据的方法流程图；

图 3 为本申请实施例提供的在共享存储系统中改写数据的方法流程图；

图 4 为图 3 中的第一服务器中未缓存数据时所执行的流程 A 的流程图；

图 5 为在数据状态为未共享状态的情况下的数据改写过程示意图；

图 6 为在数据状态为共享状态的情况下的数据改写过程示意图；

图 7 为在数据状态为修改状态的情况下的数据改写过程示意图；

图 8 为数据的状态随对数据的操作的变化示意图；

图 9 为图 3 中的第一服务器中缓存有数据且数据的状态为共享状态时所执行的流程 B 的流程图；

图 10 为在数据访问请求为数据删除请求情况下的数据改写过程示意图；

图 11 为本申请实施例提供的在共享存储系统中读取数据的方法流程图；

图 12 为在数据状态为共享或修改状态的情况下的数据读取过程示意图；

图 13 为本申请实施例提供的共享磁盘系统的架构图；

图 14 为本申请实施例提供的一种中间设备的架构图；

图 15 为本申请实施例提供的一种中间设备的结构图；

图 16 为本申请实施例提供的一种中间设备的结构图。

具体实施方式

下面将结合附图，对本申请实施例中的技术方案进行描述。

图 1 为本申请实施例提供的共享存储系统的架构图，所述系统包括构成服务器集群的多个服务器和至少一台中间设备，其中每台中间设备与至少一个服务器连接，所述中间设备可以为可编程网络连接设备，例如交换机，下文将以交换机作为中间设备的示例进行描述。图 1 中示意示出了，交换机 1 与服务器 S1~服务器 Si 连接，交换机 2 与服务器 Sj~服务器 Sn 连接。如图 1 中所示，每个服务器将自己的内存（例如动态随机存取存储器（Dynamic Random Access Memory, DRAM））分为两部分：共享内存和私有的本地缓存（即回写缓存）。服务器集群中的全部服务器的共享内存构成了一个内存池（下文称为全局内存），服务器集群中的任一服务器可以访问所述全局内存。其中，存储在全局内存中的数据的地址可以用 16 位的节点编号加 48 位的节点内地址表示，全局内存中的数据可以以地址作为数据标识（或数据编号），节点编号即为在共享内存中存储该数据的服务器的编号，节点内地址即为数据存储在服务器所提供的共享内存中的地址偏移。在各个服务器中，在服务器读写数据时，会首先访问本地缓存，本地缓存中存储有服务器中的应用程序（Application, APP）热数据，例如应用程序新写入的数据或者经常被读取的数据，当本地缓存中的存储空间不够时，可将本地缓存中的数据淘汰至服务器的共享内存中。当在本地缓存没有命中所访问的数据时，则从全局内存中将数据读取至本地内存，再从本地内存中读取数据。所述服务器可以为物理服务器，或者也可以为虚拟机、容器等基于通用硬件资源抽象的虚拟实体。交换机连接各自对应的服务器，并且交换机之间相互连接，以相对于所述多个服务器进行数据的转发。

在图 1 所示的共享存储系统中，需要保证多个服务器中的本地缓存中所缓存的数据的一

致性。即，当一份数据被多个服务器缓存在本地缓存时，保证多个服务器中的本地缓存中所缓存的该份数据的一致性，具体地，当一个服务器对该份数据进行修改之后，保证其他服务器可以访问到该份数据的最新版本。例如，假设服务器 S1 和服务器 Sn 在其各自的本地缓存中都缓存有编号为 d1 的数据 A，服务器 S1 先对本地缓存中的数据 A 进行了改写操作，将数据 A 更新为数据 A1。假设此时服务器 Sn 运行的应用程序对编号 d1 对应的数据进行读取操作，服务器 Sn 将从本地缓存中读取数据 A，而不是最新版本的数据 A1，从而导致应用程序不能读取到最新的数据。

在相关技术中，为了保证数据 A 被服务器 S1 修改为数据 A1 后，其他服务器仍然能读到最新版本的数据 A1，需要在数据修改后，进行缓存一致性处理。具体是，对于全局内存中的数据，例如由在共享内存中存储该数据的服务器作为该数据的归属服务器，并由该归属服务器中生成并存储数据的元数据，该元数据中包括缓存该数据最新版本的服务器的信息。在数据被归属服务器之外的其他服务器修改之后，归属服务器需要更新元数据中的缓存该数据最新版本的服务器的信息，另外，当其他服务器访问数据的最新版本时，需要由归属服务器向缓存该数据的服务器转发数据访问请求，以获取数据的最新版本，从而增加了归属服务器的负担。

在本申请实施例提供的数据处理方案中，如图 1 所示，由交换机存储全局内存中的数据的元数据，在该元数据中会记录存储每个数据的最新版本的服务器的信息。这样，在服务器访问该数据时，服务器可将数据访问请求发送给交换机，通过交换机的基于数据的元数据的处理，从存储数据的最新版本的服务器获取数据的最新版本。通过如此，在数据被服务器修改之后，由交换机更新数据的元数据中的缓存该数据最新版本的服务器的信息，当其他服务器访问数据的最新版本时，通过交换机将数据访问请求转发给缓存该数据的服务器，从而减少了服务器的处理和网络流量，减少了服务器的负担。

如图 1 所示，在本申请实施例中，每台交换机可与服务器集群中的部分服务器连接，从而可由多台交换机协同处理各个服务器的数据访问请求。在一种实施方式中，可确定全局内存中的数据与各台交换机之间的对应关系。例如，可按照全局内存中的存储地址范围将全局内存中的数据划分为与多台交换机分别对应。之后，可根据各台服务器的共享内存中存储的数据与各台交换机的对应关系建立交换机与服务器之间的连接，具体是，使得每台交换机与在共享内存中存储与该台交换机对应的数据的服务器连接。每台交换机可以存储其对应的每个数据的元数据以对其对应的每个数据进行缓存一致性处理，或者，每台交换机可以仅存储其对应的数据中的热点数据的元数据，以仅对该热点数据进行缓存一致性处理。在后一情况中，交换机和服务器之间可根据数据的访问热度对数据的元数据进行迁移。例如，对于被频繁访问的数据，该数据的归属服务器可将该数据的元数据发送给它连接的交换机，从而由交换机进行对该数据的缓存一致性处理。反之，如果当前由交换机进行缓存一致性处理的数据的访问频次降低，交换机可将该数据的元数据发送给该数据的归属服务器，从而恢复由归属服务器进行对该数据的缓存一致性处理。如此使得交换机仅处理热点数据的缓存一致性处理，减少了对交换机的存储需求。

其中，数据的元数据中例如包括缓存该数据的服务器列表。服务器列表的格式例如可以是一个位图 (bitmap)，每个服务器对应于位图中的一位，若服务器中缓存了该数据，则数据的服务器列表中与该服务器对应的位为 1，以表示该服务器在该服务器列表中。

数据的元数据中还可以包括数据的状态。数据的状态包括以下三种状态：未共享 (unshared)，即在服务器集群中，没有服务器在本地缓存中存储有该数据；共享 (shared)，

即服务器集群中的至少一个服务器的本地缓存中存储有该数据,并且该至少一个服务器仅具有对该数据的读权限;修改(modified),即服务器集群中的一个服务器的本地缓存中存储有该数据,并且该服务器具有对该数据的写权限。

在交换机中存储了数据的元数据的情况中,服务器在以下三种情况下向相应的交换机发送数据访问请求:本地缓存中未存储希望读写的数据,即缓存不命中;将要对本本地缓存中存储的共享状态的数据进行改写,即写共享;删除本地缓存中存储的数据,即数据逐出。服务器根据数据在本地缓存中的存储情况可向交换机发送以下五种数据访问请求:当本地缓存中未缓存数据,并且服务器希望对该数据进行改写的情况中,服务器发送指示写缺失操作的数据获取请求;当本地缓存中未缓存数据,并且服务器希望对该数据进行读取的情况中,服务器发送指示读缺失操作的数据获取请求;当本地缓存中缓存了数据,在服务器中记录的该数据的状态为共享状态,并且该服务器希望对该数据进行改写的情况中,服务器发送指示写共享操作的数据删除请求;当本地缓存中缓存了数据,在服务器中记录的该数据的状态为修改状态,并且该服务器希望删除该数据的情况中,服务器发送指示逐出修改操作的数据写请求;当本地缓存中缓存了数据,在服务器中记录的该数据的状态为共享状态,并且该服务器希望删除该数据的情况中,服务器发送指示逐出共享操作的处理请求。其中,写缺失操作、读缺失操作、写共享操作、逐出修改操作和逐出共享操作为服务器对数据进行的五种操作类型。

交换机在从服务器接收到数据访问请求之后,可根据数据访问请求的类型、元数据中记录的缓存数据的服务器列表和数据的状态,进行相对应的处理,以用于保证各个服务器的本地缓存中所缓存的数据的一致性,这将在下文进行详细描述。

图2为本申请实施例提供的在共享存储系统中写数据的方法流程图。该方法由图1中的服务器集群中的服务器和交换机共同执行。其中,第一服务器可以为服务器集群中的任一服务器,下文中以第一服务器为服务器S1进行示例描述。

当服务器S1中的应用希望向全局内存新写入数据(例如数据A)时,执行步骤S201,生成元数据创建请求。具体是,服务器S1确定全局内存中写入数据的地址,如上文所述,该地址包括服务器节点编号和服务器的节点内地址,将该确定的地址作为数据的编号(下文表示为编号d1),并据此生成元数据创建请求。该元数据创建请求中例如包括数据编号d1等信息。该元数据创建请求中还可以包括请求服务器(即服务器S1)的编号。

服务器S1在生成元数据创建请求之后执行步骤S202,向交换机发送元数据创建请求。该交换机为编号为d1的数据A所对应的交换机,例如,如上文所述,该交换机为与数据的地址相对应的交换机。假设数据的地址指示该数据对应的交换机为图1中的交换机2,则服务器S1可通过交换机1将该元数据创建请求转发给交换机2。

交换机2在从服务器S1接收到元数据创建请求之后,创建数据的元数据,该元数据中例如包括数据A的编号d1、数据A的服务器列表等信息。其中,在创建该元数据之后,由于此时还没有任一服务器在其本地缓存中缓存有数据A,因此初始的服务器列表为空列表,即其中不包括任何服务器的信息。

所述元数据中还可以包括读写锁表。交换机2可通过元数据中的读锁或写锁,避免同时接收的对编号d1的数据的多个数据访问请求之间的冲突。该读写锁例如为16位锁,最高位为写锁,如果值为1,表示该数据被加上写锁,其他服务器将不能对该数据进行其他数据操作(如读操作、写操作、逐出操作等);读写锁的后15位为读锁,表示读该数据的服务器的数量,每加一次读锁,读写锁的后15位的值加1,每解一次读锁,读写锁的后15位的值减1,通过如此设置,如果读写锁的后15位的值不为0,表示该数据被加上读锁,其他服务器将不

能对该数据进行写操作，并且可控制读该数据的服务器的最大数目为 $2^{15} - 1$ 。

交换机 2 在创建数据的元数据之后，可通过该读写锁对编号 d1 的数据加写锁。如上文所述，在加写锁期间，其他服务器将不能通过交换机 2 对该数据进行其他任何操作。交换机 2 在后续接收到其他对编号 d1 的数据的数据访问请求时，如果接收的数据访问请求为上述指示读缺失操作的数据获取请求，交换机 2 在确定未加写锁的情况中，在元数据中加读锁，如果是上述五种数据访问请求中的其他请求，则交换机 2 在确定未加写锁的情况中在元数据中加写锁，如果交换机 2 确定数据已经加写锁或者读锁的数量已经为 $2^{15} - 1$ 次，则本次加锁失败，交换机 2 返回处理失败消息。

在步骤 S204，交换机 2 在创建好数据的元数据之后，对服务器 S1 返回创建完成信息。该创建完成信息中例如包括对数据加写锁的信息。

在步骤 S205，服务器 S1 在确定交换机 2 已经创建好数据 A 的元数据之后，在本地缓存中写入数据 A。服务器 S1 在写入数据 A 之后，可在本地记录编号为 d1 的数据的状态为修改状态。当服务器 S1 希望再次修改编号为 d1 的数据时，基于该修改状态可直接对本地缓存中的数据 A 进行修改，即该修改状态指示服务器 S1 具有对编号为 d1 的数据的修改权限。另外，服务器 S1 可记录编号 d1 与本地缓存中存储数据 A 的地址（虚拟地址或物理地址）的对应关系，从而可根据该对应关系从本地缓存中读取到数据 A。

在步骤 S206，服务器 S1 向交换机 2 发送数据写完成信息。该数据写完成信息中例如包括数据 A 的状态（修改状态）。另外，该数据写完成信息中还可以包括服务器列表，该服务器列表中例如包括服务器 S1 的信息。

在步骤 S207，交换机 2 在从服务器 S1 接收到数据写完成信息之后，基于数据写完成信息更新编号为 d1 的数据的元数据。具体是，交换机 2 可将元数据中的服务器列表修改为包括服务器 S1 的信息，并在元数据中记录数据 A 的状态为修改状态。在进行该更新之后，交换机 2 可在元数据中对编号为 d1 的数据解写锁。

服务器 S1 在通过上述步骤在本地缓存中写入数据 A 之后，根据预定逐出规则，例如当数据 A 不再是应用的热数据时，或者当数据 A 写入超过预定时间之后，服务器 S1 将从本地缓存中逐出修改状态的数据 A。为此，服务器 S1 需要通过交换机 2 将数据的最新版本（即数据 A）存储到其归属服务器中。从而，服务器 S1 可执行步骤 S208，生成指示逐出修改的数据写请求，以将数据 A 存入归属服务器。由于此时数据 A 的状态为修改状态，该数据写请求具体为上述指示逐出修改操作的数据写请求。该数据写请求中例如包括请求类型（指示逐出修改操作的数据写请求类型）、数据编号 d1、数据 A 和请求服务器（即服务器 S1）的编号。

在步骤 S209，服务器 S1 将数据写请求发送给交换机 2。

在步骤 S210，交换机 2 在从服务器 S1 接收到数据写请求之后，根据编号 d1 对应的数据地址，将该数据写请求转发给数据 A 的归属服务器（图 2 中以第二服务器示出），例如该归属服务器为服务器 Sn。

在步骤 S211，服务器 Sn 在接收到数据写请求之后，将数据 A 写入共享内存中的编号 d1 对应的地址中。

在步骤 S212，服务器 Sn 在完成对数据 A 的写入之后，向交换机 2 返回数据写完成信息。交换机 2 在从服务器 Sn 接收到数据写完成信息之后，在步骤 S213 向服务器 S1 返回数据写完成信息。

在步骤 S214，服务器 S1 在接收到数据写完成信息之后，在本地缓存中删除数据 A，并在步骤 S215 向交换机 2 发送数据逐出完成信息。交换机 2 在接收到逐出完成信息之后，可更新

数据 A 的元数据,具体是,可将元数据中的服务器列表中的服务器 S1 的信息删除,并将数据状态修改为未共享状态。

在图 2 所示的流程中虽然示出了通过交换机创建数据的元数据,本申请实施例不限于此。在另一种实施方式中,交换机在接收到数据写请求之后,可根据数据编号中的数据归属服务器的编号,将该元数据创建请求发送给服务器 Sn,由服务器 Sn 创建数据的元数据,并在该数据的访问频次增大成为热点数据之后,由服务器 Sn 将该元数据移至交换机 2 中。

在图 2 所示流程中,通过服务器和交换机进行了创建数据元数据的过程,以及在共享内存中存入数据的过程。在创建了数据的元数据之后,交换机 2 可根据服务器 S1 对数据的写操作或逐出操作在元数据中记录缓存编号 d1 的数据的最新版本的服务器的信息,从而使得其他服务器可通过交换机 2 获取到数据的最新版本。在该过程中,由交换机进行对元数据的创建和更新,并由交换机进行对数据访问请求的转发,减轻了服务器的负担。

图 3 为本申请实施例提供的在共享存储系统中改写数据的方法流程图。

图 3 中的第一服务器为服务器集群中希望对数据(例如编号为 d1 的数据)进行改写的任意服务器,下文中以服务器 S1 作为第一服务器进行示例描述。

如图 3 所示,当服务器 S1 中的应用希望对编号为 d1 的数据进行改写时,首先在步骤 S301,确定本地缓存中是否存储有编号为 d1 的数据。服务器 S1 可确定在本地记录的数据编号与本地缓存中的地址的对应关系表中是否包括编号 d1,如果不包括编号 d1,则可确定本地缓存中没有存储编号为 d1 的数据,从而可执行流程 A。如果包括编号 d1,则可确定本地缓存中存储有编号 d1 的数据。

在确定本地缓存中存储有编号 d1 的数据的情况下,服务器 S1 可执行步骤 S302,基于本地记录的编号 d1 的数据状态确定该数据是否为修改状态,如果为修改状态,服务器 S1 可执行步骤 S303,直接对本地缓存中的编号 d1 的数据进行改写,如果不是修改状态,可确定该数据的状态为共享状态,则服务器 S1 执行流程 B。

图 4 为图 3 中的第一服务器中未缓存数据时所执行的流程 A 的流程图。

如图 4 所示,在步骤 S401,服务器 S1 生成数据获取请求。该数据获取请求为上述指示写缺失操作的数据获取请求。该数据获取请求中包括请求类型(指示写缺失操作的数据获取请求)、请求改写的的数据编号(例如编号 d1)和请求服务器的编号(例如服务器 S1 的编号)。

在步骤 S402,服务器 S1 将数据获取请求发送给交换机 2。

服务器 S1 在生成数据获取请求之后,根据数据的编号 d1 与交换机 2 的对应关系可确定该数据与交换机 2 相对应,并可通过交换机 1 将该数据获取请求发送给交换机 2。

交换机 2 可确定本地是否存储有编号 d1 的数据的元数据,在确定本地存储有数据的元数据的情况下,进行图 4 所示流程的后续步骤。在确定本地未存储数据的元数据的情况下,交换机 2 根据数据 d1 的数据编号中的服务器 Sn 的编号,将该数据获取请求转发给服务器 Sn,以由服务器 Sn 对该数据获取请求进行缓存一致性处理。假设交换机 2 中存储有编号为 d1 的数据的元数据,即交换机 2 在接收到上述数据获取请求之后进行对该请求的缓存一致性处理。

如上文所述,交换机 2 在接收到数据获取请求之后,可以在元数据中对数据加写锁。

交换机 2 在每次接收到数据访问请求之后,可以检查该数据访问请求的有效性,包括:检查数据访问请求与交换机元数据中的数据状态是否存在冲突;以及检查该数据访问请求中的请求服务器编号和元数据中该数据的服务器列表是否冲突。若发生冲突,交换机可向请求服务器返回处理失败消息。该处理失败消息中可包括失败的原因,请求服务器在接收到处理失败消息之后,可以重新尝试数据处理操作。其中,数据处理请求与交换机元数据中的数据

状态之间的冲突例如包括以下冲突：请求类型为指示写共享或者逐出共享的请求，数据的状态为不共享；请求类型为指示逐出修改的请求，数据的状态不为修改。请求服务器编号和服务器列表之间的冲突例如包括以下冲突：请求类型为指示读缺失或者写缺失的请求，但请求服务器的编号已经在缓存块的服务器列表中；请求类型为指示写共享、逐出共享或者逐出修改的请求时，请求服务器不在服务器列表中。

在步骤 S403，交换机 2 在接收到数据获取请求之后可根据数据的元数据将数据获取请求转发给第二服务器。

在一种情况中，元数据中记录的数据状态为未共享状态，和/或元数据中记录了服务器列表中不包括任何服务器的信息，也即，服务器集群中没有服务器在本地缓存中缓存了编号为 d1 的数据。在该情况中，交换机 2 根据编号 d1 确定数据的归属服务器为服务器 Sn，并将该数据获取请求转发给服务器 Sn，在该情况中，服务器 Sn 为第二服务器。

图 5 为在数据状态为未共享状态的情况下的数据改写过程示意图。如图 5 所示，服务器 S1 在将数据获取请求发送给交换机 2 之后，通过交换机 2 从服务器 Sn 的共享内存中拉取编号为 d1 的数据 A，服务器 S1 将数据 A 修改为数据 A1，并在本地缓存中存储修改后的数据 A1，最后交换机 2 在服务器 S1 修改完成之后更新编号为 d1 的数据的元数据，以指示服务器 S1 中存储有编号为 d1 的数据的最新版本。

在另一种情况中，元数据中记录的数据状态为共享状态，在该情况中，交换机根据元数据中的服务器列表，将该数据获取请求发送给服务器列表中记录的各个服务器。交换机在向各个服务器转发数据获取请求的同时，还可以指定其中一个服务器返回其缓存的数据。例如，交换机可向各个服务器发送数据包，该数据包中包括指示写缺失操作的数据获取请求和“是否返回数据”字段，通过该字段指示其中一个服务器（例如服务器 Si）返回其缓存的数据。此外，该数据包中还可以包括其他字段，如数据 d1 的状态、数据 d1 的服务器列表、以及元数据是否在交换机 2 中等字段。

图 6 为在数据状态为共享状态的情况下的数据改写过程示意图。如图 6 所示，服务器 S1 在将数据获取请求发送给交换机 2 之后，通过交换机 2 将数据获取请求转发给缓存数据的各个服务器（例如服务器 Si 和服务器 Sj），并从服务器 Sj 的本地缓存中拉取编号 d1 的数据 A，服务器 S1 将数据 A 修改为数据 A1，并在本地缓存中存储修改后的数据 A1，最后交换机 2 在服务器 S1 修改完成之后更新编号为 d1 的数据的元数据，以指示服务器 S1 中存储有编号为 d1 的数据的最新版本。也就是说，在该情况中，服务器 Sj 为图 4 中的第二服务器。

在又一种情况中，元数据中记录的数据的状态为修改状态，在该情况中，交换机根据元数据中的服务器列表，将该数据获取请求发送给服务器列表中记录的一个服务器（例如服务器 Si），即，在该情况中，服务器 Si 为图 4 中的第二服务器。

图 7 为在数据状态为修改状态的情况下的数据改写过程示意图。如图 7 所示，服务器 S1 在将数据获取请求发送给交换机 2 之后，通过交换机 2 从服务器 Si 的本地缓存中拉取编号 d1 的数据 A，服务器 S1 将数据 A 修改为数据 A1，并在本地缓存中存储修改后的数据 A1，最后交换机 2 在服务器 S1 修改完成之后更新编号为 d1 的数据的元数据，以指示服务器 S1 中存储有编号为 d1 的数据的最新版本。

在步骤 S404，第二服务器在接收到数据获取请求之后读取数据。

具体是，在图 5 所示的情况中，第二服务器为服务器 Sn，服务器 Sn 在接收到数据获取请求之后，从共享内存中读取数据 A。

在图 6 所示的情况中，第二服务器为服务器 Sj，服务器 Sj 在接收到包括数据获取请求

的数据包之后,根据数据包中的“是否返回数据”字段,从本地缓存中读取数据 A 以用于返回。在读取数据 A 之后,服务器 S_j 根据数据获取请求的请求类型,将本地缓存中的数据 A 删除,以用于保证数据的缓存一致性。同样地,图 6 中的服务器 S_i 在接收到交换机转发的数据获取请求之后,也将本地缓存中的数据 A 删除,以用于保证数据的缓存一致性。

在图 7 所示的情况中,第二服务器为服务器 S_i,服务器 S_i 在接收到数据获取请求之后,从本地缓存中读取编号为 d1 的数据的最新版本(假设为数据 A),并将本地缓存中的数据 A 删除,以保证数据的缓存一致性。

在步骤 S405,第二服务器在读取数据 A 之后,将读取的数据 A 返回给交换机 2。

在步骤 S406,交换机 2 在从第二服务器接收到数据 A 之后,将该数据 A 返回给服务器 S₁。交换机 2 在将数据 A 返回给服务器 S₁ 的同时,还可以把元数据中编号为 d1 的数据的数据状态和服务器列表中的至少一个发送给服务器 S₁。

在步骤 S407,服务器 S₁ 在接收到返回的数据 A 之后,将数据 A 改写为数据 A₁。

具体是,在图 5 所示的情况中,服务器 S₁ 在还从交换机 2 接收到编号为 d1 的数据的服务器列表的情况中,可基于该服务器列表确定只需要从服务器 S_n 接收回复信息,从而在确定接收到服务器 S_n 返回的数据 A 之后进行对数据 A 的改写。服务器 S₁ 在接收到数据的状态为未共享状态的情况中,可基于该状态确定只需要从数据的归属服务器(即服务器 S_n)接收回复信息。服务器 S₁ 在接收到数据 A 之后,将数据 A 修改为数据 A₁ 作为编号为 d1 的数据的最新版本,并将数据 A₁ 存储到本地缓存中,其中,对数据 A 的修改包括对数据 A 中的至少部分数据的修改。

此外,服务器 S₁ 可根据接收的数据 d1 的状态(即未共享状态)和对数据的操作类型确定数据 d1 的最新状态。图 8 为数据的状态随对数据的操作的变化示意图,图 8 中的节点为数据的当前状态,节点之间的边为数据操作类型。如图 8 所示,服务器 S₁ 根据数据 d1 的当前未共享状态和写缺失操作可确定数据 d1 的状态变为修改状态。同时,服务器 S₁ 还相应地修改其接收的服务器列表,更新后的服务器列表中仅包括服务器 S₁。

在图 6 所示的情况中,服务器 S₁ 在还从交换机 2 接收到编号为 d1 的数据的服务器列表的情况中,假设该服务器列表中包括服务器 S_i 和服务器 S_j 的信息,服务器 S₁ 可基于该服务器列表确定需要从服务器 S_i 和 S_j 接收回复信息,从而在确定接收到服务器 S_i 和 S_j 的回复信息之后,进行对数据 A 的改写,其中,服务器 S_i 的回复信息中包括删除数据 A 的信息,服务器 S_j 的回复信息中包括数据 A 和删除数据 A 的信息。服务器 S₁ 根据图 8 可确定,在对共享状态的数据进行写缺失操作之后,数据的状态变为修改状态,并将该更新的数据状态发送给交换机 2。

在图 7 所示的情况中,与图 5 所示情况类似地,服务器 S₁ 根据服务器列表确定在接收到服务器 S_i 发送的数据 A 之后,可进行对数据 A 的改写。服务器 S₁ 根据图 8 可确定,在对修改状态的数据进行写缺失操作之后,数据的状态仍为修改状态,并可将该数据状态发送给交换机 2。

在步骤 S408,服务器 S₁ 向交换机 2 发送数据改写完成信息。

该数据改写完成信息例如包括在进行写缺失操作之后的数据 d1 的状态和数据 d1 的服务器列表。在一种实施方式中,在交换机 2 在接收到指示写缺失操作的数据获取请求之后对数据 d1 上写锁的情况中,在该数据改写完成信息中还包括写锁解锁信息。

步骤 S409,交换机 2 在接收到数据改写完成信息之后,根据数据改写完成信息更新数据的元数据。

交换机 2 在接收到上述数据改写完成信息之后,根据数据改写完成信息中的数据 d1 的状态和数据 d1 的服务器列表更新编号为 d1 的数据的元数据。具体是,参考图 5-图 7,在该对元数据的更新之后,元数据中记录的数据 d1 的状态为修改状态,并且,元数据中的服务器列表中包括服务器 S1。通过在元数据中记录数据的状态为修改状态以及数据的服务器列表中包括服务器 S1,表示服务器 S1 可在本地缓存中对数据 d1 进行修改,并指示其他服务器:服务器 S1 的本地缓存中存储有编号为 d1 的数据的最新版本。从而,当其他服务器希望读写编号为 d1 的数据时,可通过交换机 2 从服务器 S1 读取编号为 d1 的数据的最新版本,从而保证了数据 d1 在各个服务器的缓存一致性。

在交换机 2 已经根据指示写缺失操作的数据获取请求对数据 d1 上写锁的情况中,在对元数据更新完成之后,交换机 2 根据数据处理完成信息中的写锁解锁信息对数据 d1 的写锁解锁,即,如上文所述,将数据 d1 的读写锁的最高位的值恢复为 0。

图 9 为图 3 中的第一服务器中缓存有数据且数据的状态为共享状态时所执行的流程 B 的流程图。

如图 9 所示,假设第一服务器为服务器 S1,服务器 S1 在确定本地缓存中缓存有编号为 d1 的数据且本地记录了数据的状态为共享状态的情况中,服务器 S1 执行步骤 S901,生成指示写共享操作的数据删除请求,该数据删除请求中包括请求类型(指示写共享操作的数据删除请求类型)、请求改写的数据编号(编号 d1)和请求服务器的编号(服务器 S1 的编号)。

在步骤 S902,服务器 S1 将数据删除请求发送给交换机 2。该步骤可参考上文对步骤 S402 的描述,在此不再赘述。

在步骤 S903,交换机 2 根据数据的元数据向第二服务器转发数据删除请求。

交换机 2 可根据元数据中的服务器列表,确定除服务器 S1 之外的缓存数据的至少一个服务器,并将该数据删除请求转发给每个缓存该数据的服务器。也就是说,图 9 中的第二服务器可以是除服务器 S1 之外的缓存该数据的任一服务器,例如服务器 Si。

图 10 为在数据访问请求为数据删除请求情况下的数据改写过程示意图。在数据元数据的服务器列表中包括服务器 S1 和服务器 Si 的信息的情况中,交换机 2 将数据删除请求转发给服务器 Si,从而使得服务器 Si 删除本地缓存中的数据 A,以保证数据的缓存一致性。即,在该情况中,服务器 Si 为图 9 中的第二服务器。

在步骤 S904,服务器 Si 在接收到数据删除请求之后删除数据。

如图 10 所示,服务器 Si 在接收到上述数据删除请求之后,删除本地缓存中的数据 A。

在步骤 S905,服务器 Si 向交换机 2 返回删除完成信息。

在步骤 S906,交换机 2 将接收的删除完成信息转发给服务器 S1。交换机在转发删除完成信息的同时,还可以把元数据中数据的状态和数据的服务器列表一起发送给服务器 S1。

在步骤 S907,服务器 S1 在接收到来自服务器 Si 的删除完成信息之后,对本地缓存中的数据 A 进行改写。服务器 S1 根据从交换机 2 接收的数据的服务器列表可确定仅需要从服务器 Si 接收删除完成信息,并在确定已经从服务器 Si 接收到删除完成信息之后进行对数据 A 的改写。具体是,如图 10 所示,服务器 S1 从本地缓存中读取数据 A,将数据 A 修改为数据 A1 作为编号为 d1 的数据的最新版本,并将数据 A1 存储到本地缓存中。此外,参考图 8,服务器 S1 可根据从交换机 2 接收的数据 d1 的状态(即共享状态)和写共享操作类型将数据 d1 的状态更新为修改状态。同时,服务器 S1 还可以修改从交换机 2 接收的服务器列表,更新后的服务器列表中仅包括服务器 S1。

在步骤 S908,服务器 S1 向交换机 2 发送数据改写完成信息。

该数据改写完成信息例如包括更新的数据的状态和数据的服务列表。

步骤 S909, 交换机 2 根据改写完成信息更新数据的元数据。

交换机 2 在接收到上述改写完成信息之后, 根据改写完成信息中的数据的状态和数据的服务列表更新编号为 d1 的数据的元数据。具体是, 如图 10 所示, 在该对元数据的更新之后, 元数据中记录的数据 d1 的状态为修改状态, 并且, 元数据中的服务器列表仅包括服务器 S1。

图 11 为本申请实施例提供的在共享存储系统中读取数据的方法流程图。

图 11 中的第一服务器为服务器集群中希望读取数据的任一服务器, 下文中以服务器 S1 作为第一服务器进行示例描述。

如图 11 所示, 当服务器 S1 中的应用希望读取编号为 d1 的数据时, 首先在步骤 S1101, 确定本地缓存中是否缓存有编号为 d1 的数据。该确定的过程可参考上文对步骤 S301 的描述, 在此不再赘述。在确定本地缓存中缓存有数据的情况中, 可直接读取本地缓存中的数据返回给应用。

在确定本地缓存中未缓存数据的情况中, 服务器 S1 执行步骤 S1102, 生成指示读缺失操作的数据获取请求。该数据获取请求中包括请求类型(指示读缺失操作的数据获取请求类型)、请求获取的数据编号(编号 d1)和请求服务器的编号等信息。

在步骤 S1103, 服务器 S1 将生成的数据获取请求发送给数据对应的交换机 2。

在步骤 S1104, 交换机 2 在接收到数据获取请求之后可根据数据的元数据将数据获取请求转发给第二服务器。

在一种情况中, 元数据中记录的数据状态为未共享状态, 和/或元数据中记录了服务器列表中不包括任何服务器的信息, 也即, 服务器集群中没有服务器在本地缓存中缓存了编号为 d1 的数据。在该情况中, 交换机 2 根据编号 d1 确定数据的归属服务器为服务器 Sn, 并将该数据获取请求转发给服务器 Sn, 在该情况中, 服务器 Sn 为图 11 中的第二服务器。

在另一种情况中, 元数据中记录的数据状态为共享状态, 在该情况中, 交换机根据元数据中的服务器列表, 将该数据获取请求发送给服务器列表中记录的任一服务器(例如服务器 Si), 即, 在该情况中, 服务器 Si 为图 11 中的第二服务器。

在又一种情况中, 元数据中记录的数据的状态为修改状态, 在该情况中, 交换机根据元数据中的服务器列表, 将该数据获取请求发送给服务器列表中记录的一个服务器(例如服务器 Si), 即, 在该情况中, 服务器 Si 为图 11 中的第二服务器。

图 12 为在数据状态为共享或修改状态的情况下的数据读取过程示意图。如图 12 所示, 服务器 S1 通过交换机 2 将指示读缺失操作的数据获取请求发送给服务器 Si, 以通过交换机 2 从服务器 Si 的本地缓存中拉取数据 A, 存储到本地缓存中, 并返回给应用, 交换机 2 在服务器 S1 读取完成之后更新编号为 d1 的数据的元数据, 以指示服务器 S1 和服务器 Si 中都缓存有数据 A。

在步骤 S1105, 第二服务器读取数据。

交换机 2 可随数据获取请求一起将数据元数据中记录的数据状态一起发送给第二服务器。在上述第二服务器为服务器 Sn 的情况中, 如果服务器 Sn 接收的数据状态为未共享状态, 则服务器 Sn 可以根据数据的编号 d1, 从共享内存中读取数据 A。在上述第二服务器为服务器 Si 的情况中, 如果服务器 Si 接收的数据状态为共享状态, 或者服务器 Si 可从本地记录的数据 A 的状态中确定数据 A 为共享状态, 则服务器 Si 可从本地缓存中读取数据 A。如果服务器 Si 接收的数据状态为修改状态, 或者服务器 Si 可从本地记录的数据 A 的状态中确定数

据 A 为修改状态, 则服务器 S_i 可从本地缓存中读取数据 A, 并在该读取之后, 在本地记录数据 A 的状态为共享状态, 以指示当前不再具有对数据 A 的修改权限。

在步骤 S1106, 第二服务器在读取数据之后向交换机 2 返回读取的数据 A。

在步骤 S1107, 交换机 2 将从第二服务器接收的数据 A 返回给服务器 S₁。交换机 2 在向服务器 S₁ 返回数据 A 的同时, 还可以把元数据中数据的状态和数据的服务器列表一起发送给服务器 S₁。

在步骤 S1108, 服务器 S₁ 在从交换机 2 接收到数据 A 之后, 将数据 A 存储到本地缓存中, 并向应用返回数据 A。此外, 参考图 8, 服务器 S₁ 可根据从交换机 2 接收的数据 A 的状态和读缺失操作类型确定是否更新数据 A 的状态。例如, 如果数据 A 的初始状态为未共享状态, 根据图 8, 在经过读缺失操作之后, 数据的状态变为共享状态, 从而服务器 S₁ 将数据 A 的状态更改为共享状态。如果数据 A 的初始状态为修改状态, 参考图 8, 同样地, 服务器 S₁ 将数据 A 的状态更改为共享状态。如果数据 A 的初始状态为共享状态, 参考图 8, 在经过该读取之后, 数据 A 的状态仍然为共享状态, 不需要更改。同时, 服务器 S₁ 还相应地修改从交换机 2 接收的服务器列表, 更新后的服务器列表中添加了服务器 S₁。例如, 在数据 A 的初始状态为共享状态或修改状态的情况中, 在该读取操作之后, 更新的服务器列表中包括 S₁ 和 S_i。

在步骤 S1109, 服务器 S₁ 向交换机 2 发送数据读取完成信息。

该数据读取完成信息例如包括更新的数据的状态和数据的服务器列表。在交换机 2 在接收到指示读缺失操作的数据获取请求之后对数据 d₁ 上读锁的情况中, 该数据处理完成信息中还可以包括读锁解锁信息。

步骤 S1110, 交换机 2 根据数据读取完成信息更新数据的元数据。

交换机 2 在接收到上述数据读取完成信息之后, 根据数据读取完成信息中的数据的状态和数据的服务器列表更新编号为 d₁ 的数据的元数据。例如, 在数据初始状态为共享或修改状态的情况中, 如图 12 所示, 在该对元数据的更新之后, 元数据中记录的数据的状态为共享状态, 并且, 元数据中的服务器列表中包括服务器 S₁ 和服务器 S_i, 该元数据中记录的内容表示, 服务器 S₁ 和 S_i 中都缓存了数据 A, 其他服务器可从服务器 S₁ 或者 S_i 读取数据 A。另外, 交换机 2 在完成对编号为 d₁ 的数据的元数据的更新之后, 对数据的读写锁解读锁, 即将读写锁的后 15 位的值减 1。

服务器 S₁ 在通过上述步骤在本地缓存中写入数据 A 之后, 根据预定逐出规则, 例如当数据 A 不再是应用的热数据时, 或者当数据 A 写入超过预定时间之后, 服务器 S₁ 可执行步骤 S1111, 生成数据的指示逐出共享操作的处理请求, 以将数据 A 从本地缓存中逐出。该处理请求中例如包括请求类型 (指示逐出共享操作的处理请求类型)、数据编号 d₁、和请求服务器 (即服务器 S₁) 的编号。

在步骤 S1112, 服务器 S₁ 将处理请求发送给交换机 2。

在步骤 S1113, 交换机 2 返回对处理请求的回复信息。

交换机 2 在接收到处理请求之后可进行在元数据中对数据加写锁等操作。

由于该逐出共享操作不会影响其它服务器对数据的缓存一致性, 因此, 交换机 2 不需要将该指示逐出共享操作的处理请求转发给其它服务器, 从而可直接向服务器 S₁ 发送对指示逐出共享操作的处理请求的回复信息, 该回复信息中可包括编号为 d₁ 的数据的元数据中的状态和服务器列表。

在步骤 S1114, 服务器 S₁ 在接收到回复信息之后在本地缓存中删除数据 A。具体是, 服务器 S₁ 删除本地缓存中的数据 A, 根据数据 A 的服务器列表确定数据 A 的状态, 如果服务器

列表中只包括服务器 S1 的信息，则在该删除之后，数据 A 的状态应变为未共享状态，从而服务器 S1 修改数据 A 的状态，如果服务器列表中还包括其他服务器的信息，则在该删除之后，数据 A 的状态仍为共享状态。同时，在进行该删除之后，服务器 S1 修改接收的服务器列表，即，在服务器列表中删除服务器 S1 的信息。

在步骤 S1115，服务器 S1 在删除数据 A 之后将逐出共享完成信息发送给交换机 2，该逐出共享完成信息中可包括数据的状态和服务器列表。

在步骤 S1116，交换机 2 在接收到逐出共享完成信息之后，更新编号为 d1 的数据的元数据，具体是，在元数据的服务器列表中删除服务器 S1 的信息。交换机 2 例如可根据从服务器 S1 接收的数据的服务器列表更新元数据中数据的服务器列表，如果从服务器接收的数据的状态发生变化，则在元数据中更新数据的状态，并在完成更新之后释放写锁。

在上述参考图 2-图 12 描述的数据处理方案中，在数据被服务器修改之后，由交换机更新数据的元数据中的缓存该数据最新版本的服务器的信息，当其他服务器访问数据的最新版本时，通过交换机将数据访问请求转发给缓存该数据的服务器，从而减少了服务器的处理和网络流量，减少了服务器的负担。

可以理解，上文中虽然以全局内存为例描述基于共享存储系统的数据处理方案，本申请实施例提供的数据处理方案不限于用于包括全局内存的共享存储系统，而是还适用于其它带缓存的共享存储系统。图 13 示出本申请实施例提供的共享磁盘系统的架构图。与图 1 类似地，在该共享磁盘系统中包括多个服务器和至少一个交换机。每个服务器中自己的部分磁盘用作作为共享磁盘，全部服务器的共享磁盘构成全局磁盘，服务器集群中的任一服务器可以访问该全局磁盘中存储的数据。该共享磁盘系统例如可用作作为分布式共享文件系统，在分布式共享文件系统中，通过将共享文件存储在全局磁盘中，任一服务器都可以访问这些共享文件。全局磁盘中的数据的地址可通过服务器的节点编号和服务器所提供的共享磁盘中的地址偏移构成。在各个服务器中，还设置有本地缓存，在服务器读写数据时，会首先访问本地缓存，本地缓存中存储有服务器中的应用程序（Application，APP）热数据，例如应用程序新写入的数据或者经常被读取的数据，当本地缓存中的存储空间不够时，可将本地缓存中的数据淘汰至服务器的共享磁盘中。当在本地缓存没有命中所访问的数据时，则从全局磁盘中将数据读取至本地缓存，再从本地缓存中读取数据。

在图 13 所示的共享磁盘系统中，同样需要保证多个服务器中的本地缓存中所缓存的数据的一致性，本申请图 2-图 12 所示的数据处理方法同样适用于在共享磁盘系统中保证各个服务器的数据的缓存一致性。

图 14 为本申请实施例提供的一种中间设备的架构图，所述中间设备分别与第一服务器和第二服务器连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述中间设备包括：

接收单元 141，用于从所述第一服务器接收对所述数据的获取请求；

转发单元 142，用于根据所述存储所述数据的最新版本的服务器的信息，将所述获取请求转发给所述第二服务器；

所述接收单元 141 还用于从所述第二服务器接收所述数据的第一版本；

所述转发单元 142 还用于将所述数据的第一版本转发给所述第一服务器。

在一种可能的实现方式中，所述接收单元 141 还用于从所述第一服务器接收对所述数据

的改写完成信息；所述中间设备还包括修改单元 143，用于基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

在一种可能的实现方式中，所述获取请求中携带有改写标识，所述接收单元 141 还用于在将所述获取请求转发给所述第二服务器之后，从所述第二服务器接收所述数据的第一版本被删除的信息。

在一种可能的实现方式中，所述中间设备中记录有对所述数据的写锁信息，所述修改单元 143 还用于：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述写锁信息以对所述数据解锁。

在一种可能的实现方式中，所述数据为热点数据。

图 15 为本申请实施例提供的一种中间设备的结构图，包括处理单元 151 和存储单元 152，所述存储单元 152 中存储有可执行代码，所述处理单元 151 执行所述可执行代码以实现如图 2-图 12 所示的方法。

图 16 为本申请实施例提供的一种中间设备的结构图，包括：通信接口 161，用于与服务器进行数据传输；处理单元 162，用于对所述通信接口接收的数据进行处理，以执行如图 2-图 12 所示的方法。

需要理解，本文中的“第一”，“第二”等描述，仅仅为了描述的简单而对相似概念进行区分，并不具有其他限定作用。

本领域的技术人员可以清楚地了解到，本申请提供的各实施例的描述可以相互参照，为描述的方便和简洁，例如关于本申请实施例提供的各装置、设备的功能以及执行的步骤可以参照本申请方法实施例的相关描述，各方法实施例之间、各装置实施例之间也可以互相参照。

本领域技术人员可以理解：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全部或部分步骤；而前述的存储介质包括：只读存储器（read-only memory，ROM）、随机存取存储器（random-access memory，RAM）、磁盘或者光盘等各种可以存储程序代码的介质。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本发明实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一个计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质、或者半导体介质（例如固态硬盘（Solid State Disk，SSD）等。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，在没有超过本申请的范围内，可以通过其他方式实现。例如，以上所描述的实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例

如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

另外，所描述装置和方法以及不同实施例的示意图，在不超出本申请的范围内，可以与其它系统，模块，技术或方法结合或集成。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电子、机械或其它的形式。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应所述以权利要求的保护范围为准。

权 利 要 求 书

1. 一种计算机系统，其特征在于，包括中间设备、第一服务器和第二服务器，所述第一服务器和第二服务器通过所述中间设备连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述第一服务器用于通过所述中间设备从所述第二服务器获取所述数据的第一版本。

2. 根据权利要求 1 所述的系统，其特征在于，所述第二服务器还用于在获取所述数据的第一版本后，指示所述中间设备记录所述第二服务器为存储所述数据的最新版本的服务器。

3. 根据权利要求 2 所述的系统，其特征在于，所述中间设备存储有所述数据的元数据，所述中间设备用于在所述数据的元数据中记录所述第二服务器的信息。

4. 根据权利要求 1-3 任一项所述的系统，其特征在于，所述第一服务器在通过所述中间设备从所述第二服务器获取所述数据的第一版本时，具体用于向所述中间设备发送对所述数据的获取请求，

所述中间设备还用于：从所述第一服务器接收所述获取请求，根据存储所述数据的最新版本的服务器的信息将所述获取请求转发给所述第二服务器；从所述第二服务器接收所述数据的第一版本并转发给所述第一服务器。

5. 根据权利要求 4 所述的系统，其特征在于，所述第一服务器用于在接收到对所述数据的改写请求，并在本地没有命中所述数据的第一版本时，发送所述获取请求至所述中间设备；

所述第一服务器在获取所述数据的第一版本之后，还用于将所述数据的第一版本改写为第二版本。

6. 根据权利要求 5 所述的系统，所述第一服务器还用于在将所述数据的第一版本改写为第二版本之后，向所述中间设备发送改写完成信息；

所述中间设备还用于从所述第一服务器接收所述改写完成信息，基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

7. 根据权利要求 5 或 6 所述的系统，其特征在于，所述获取请求中携带有改写标识，在中间设备将所述获取请求转发给所述第二服务器之后，所述第二服务器用于根据所述获取请求中的改写标识在获取所述数据的第一版本之后删除所存储的所述数据的第一版本。

8. 根据权利要求 7 所述的系统，其特征在于，所述中间设备中记录有对所述数据的写锁信息，所述中间设备还用于：在从所述第一服务器接收到对所述获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述元数据的写锁信息以对所述数据解锁。

9. 根据权利要求 1-8 任一项所述的系统，其特征在于，所述数据为热点数据。

10. 根据权利要求 1-9 任一项所述的系统，其特征在于，还包括第三服务器，所述第三服务器与所述中间设备连接，所述第三服务器中也存储有所述数据的第一版本，所述第二服务器在接收到对所述数据的改写请求，并在本地命中所述数据的第一版本后，还用于向所述中

间设备发送对所述数据的删除请求，

所述中间设备还用于：从所述第二服务器接收对所述数据的删除请求，根据所述存储所述数据的最新版本的服务器的信息，将所述删除请求转发给所述第三服务器，指示所述第三服务器删除所述数据的第一版本；

所述第二服务器还用于将所述数据的第一版本改写为第二版本。

11. 根据权利要求 1-10 任一项所述的系统，其特征在于，所述中间设备为网络交换设备。

12. 根据权利要求 3-11 任一项所述的系统，其特征在于，所述中间设备包括相互连接的第一中间设备和第二中间设备，其中，所述第一中间设备中存储有归属于所述第一服务器的数据的元数据，所述第二中间设备中存储有归属于所述第二服务器的数据的元数据。

13. 根据权利要求 1-12 任一项所述的系统，其特征在于，所述计算机系统为全局内存系统或者分布式共享文件系统。

14. 一种数据处理方法，其特征在于，由中间设备执行，所述中间设备分别与第一服务器和第二服务器连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述方法包括：

从所述第一服务器接收对所述数据的获取请求；

根据所述存储所述数据的最新版本的服务器的信息，将所述获取请求转发给所述第二服务器；

从所述第二服务器接收所述数据的第一版本；

将所述数据的第一版本转发给所述第一服务器。

15. 根据权利要求 14 所述的方法，其特征在于，所述方法还包括：从所述第一服务器接收对所述数据的改写完成信息；

基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

16. 根据权利要求 15 所述的方法，其特征在于，所述获取请求中携带有改写标识，所述方法还包括，在将所述获取请求转发给所述第二服务器之后，从所述第二服务器接收所述数据的第一版本被删除的信息。

17. 根据权利要求 16 所述的方法，其特征在于，所述中间设备中记录有对所述数据的写锁信息，所述方法还包括：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述写锁信息以对所述数据解锁。

18. 根据权利要求 14-17 任一所述的方法，其特征在于，所述数据为热点数据。

19. 一种中间设备，所述中间设备分别与第一服务器和第二服务器连接，所述第二服务器中存储有数据的第一版本，所述第一版本为所述数据的最新版本，所述中间设备中记录有存储所述数据的最新版本的服务器的信息，所述中间设备包括：

接收单元，用于从所述第一服务器接收对所述数据的获取请求；

转发单元，用于根据所述存储所述数据的最新版本的服务器的信息，将所述获取请求转发给所述第二服务器；

所述接收单元还用于从所述第二服务器接收所述数据的第一版本；

所述转发单元还用于将所述数据的第一版本转发给所述第一服务器。

20. 根据权利要求 19 所述的中间设备，其特征在于，所述接收单元还用于从所述第一服务器接收对所述数据的改写完成信息；

所述中间设备还包括修改单元，用于基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息。

21. 根据权利要求 20 所述的中间设备，其特征在于，所述获取请求中携带有改写标识，所述接收单元还用于在将所述获取请求转发给所述第二服务器之后，从所述第二服务器接收所述数据的第一版本被删除的信息。

22. 根据权利要求 21 所述的中间设备，其特征在于，所述中间设备中记录有对所述数据的写锁信息，所述修改单元还用于：在从所述第一服务器接收到对所述数据的获取请求之后，根据所述获取请求中的改写标识修改所述写锁信息以锁定所述数据，在基于所述改写完成信息，将所述存储所述数据的最新版本的服务器的信息修改为所述第一服务器的信息之后，修改所述写锁信息以对所述数据解锁。

23. 根据权利要求 19-22 任一所述的中间设备，其特征在于，所述数据为热点数据。

24. 一种中间设备，其特征在于，包括处理单元和存储单元，所述存储单元中存储有可执行代码，所述处理单元执行所述可执行代码以实现权利要求 14-18 任意一项所述的方法。

25. 一种中间设备，其特征在于，包括：

通信接口，用于与服务器进行数据传输；

处理单元，用于对所述通信接口接收的数据进行处理，以执行权利要求 14-18 任意一项所述的方法。

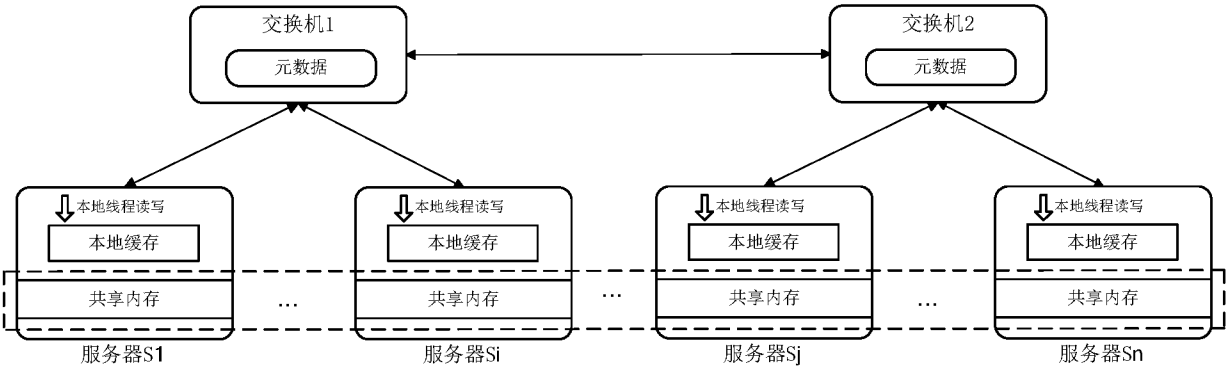


图 1

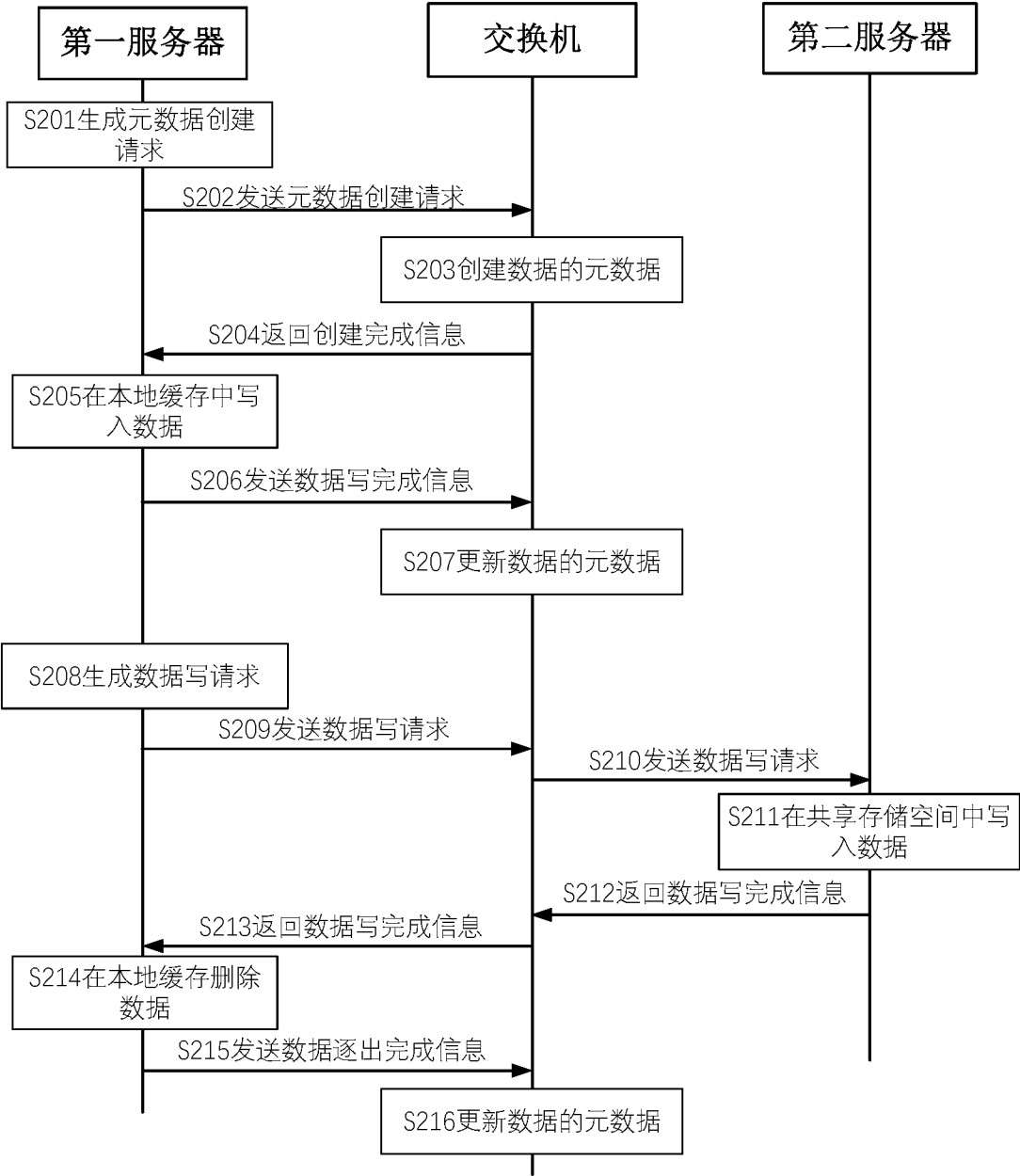


图 2

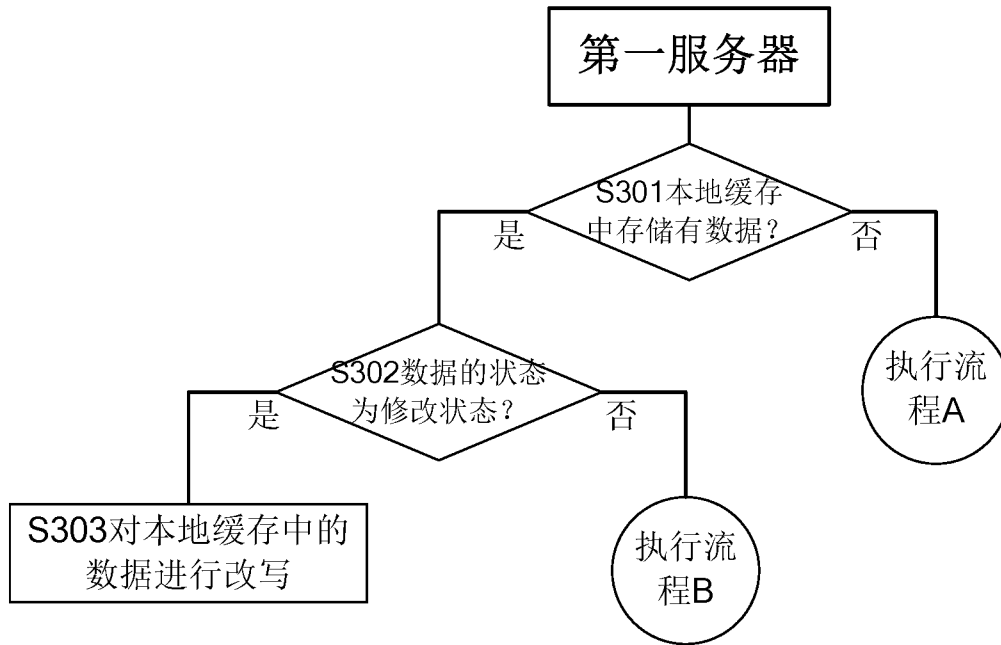


图 3

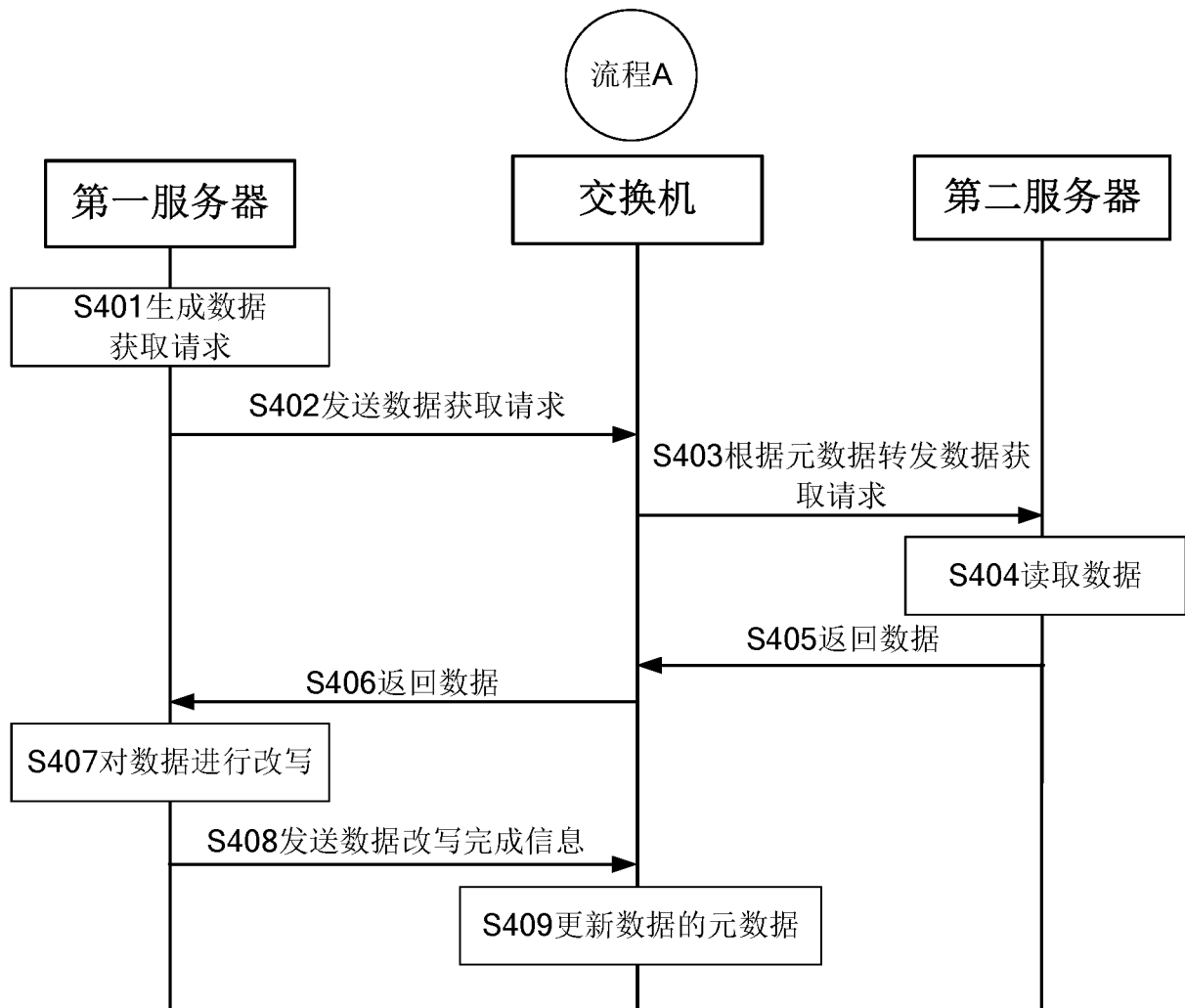


图 4

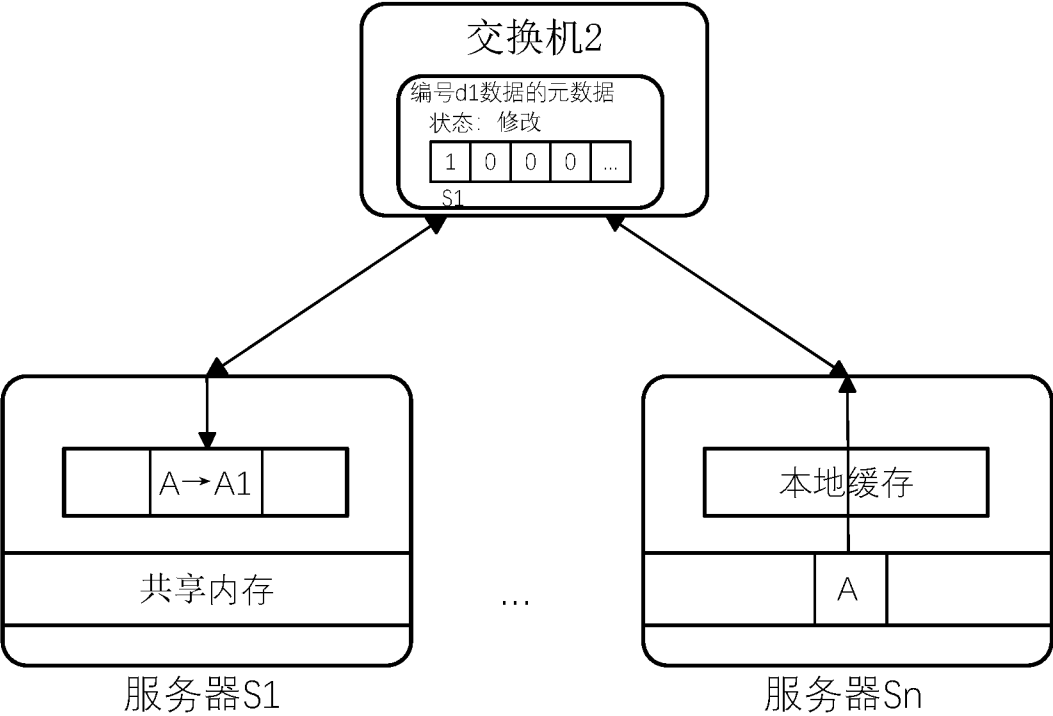


图 5

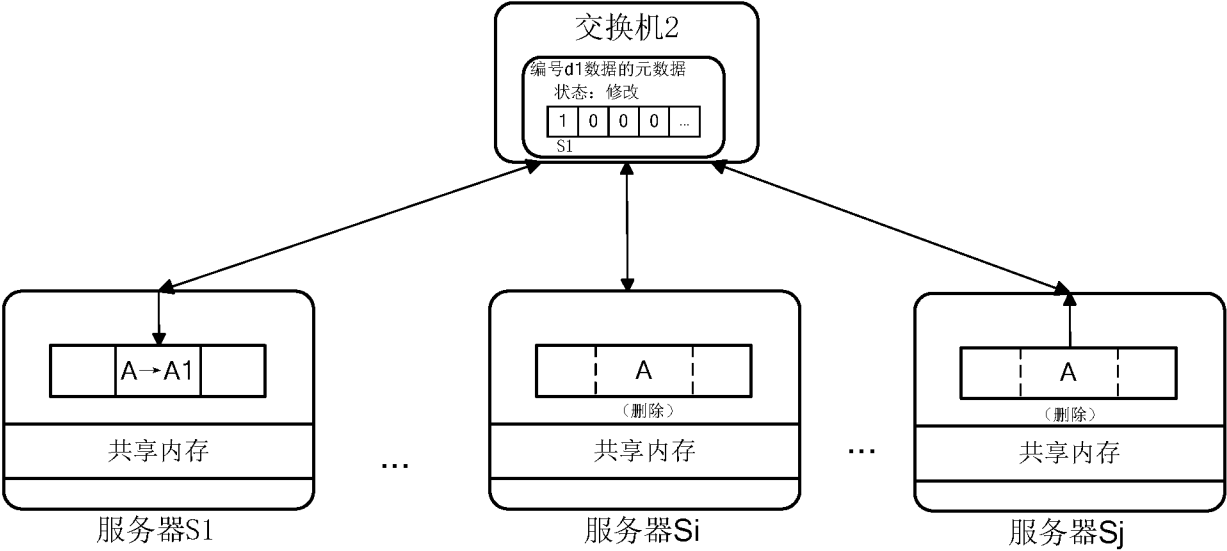


图 6

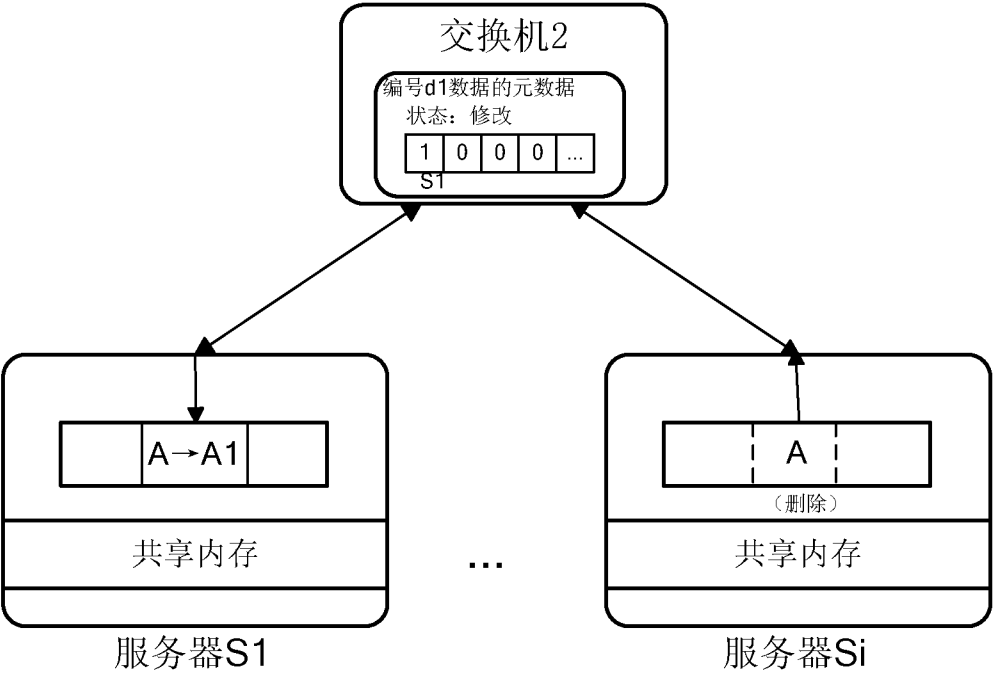


图 7

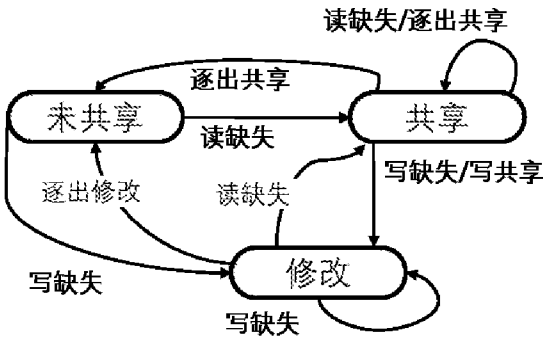


图 8

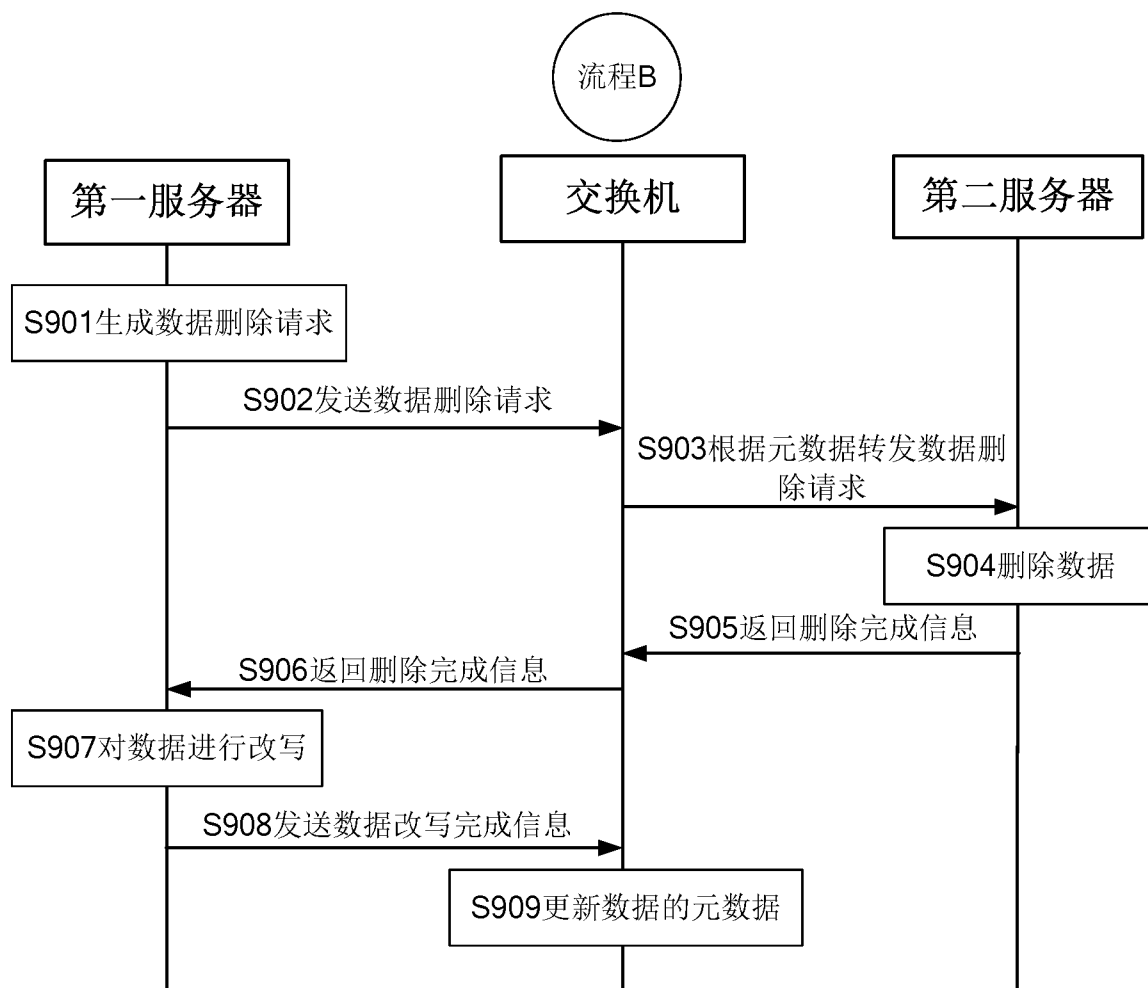


图 9

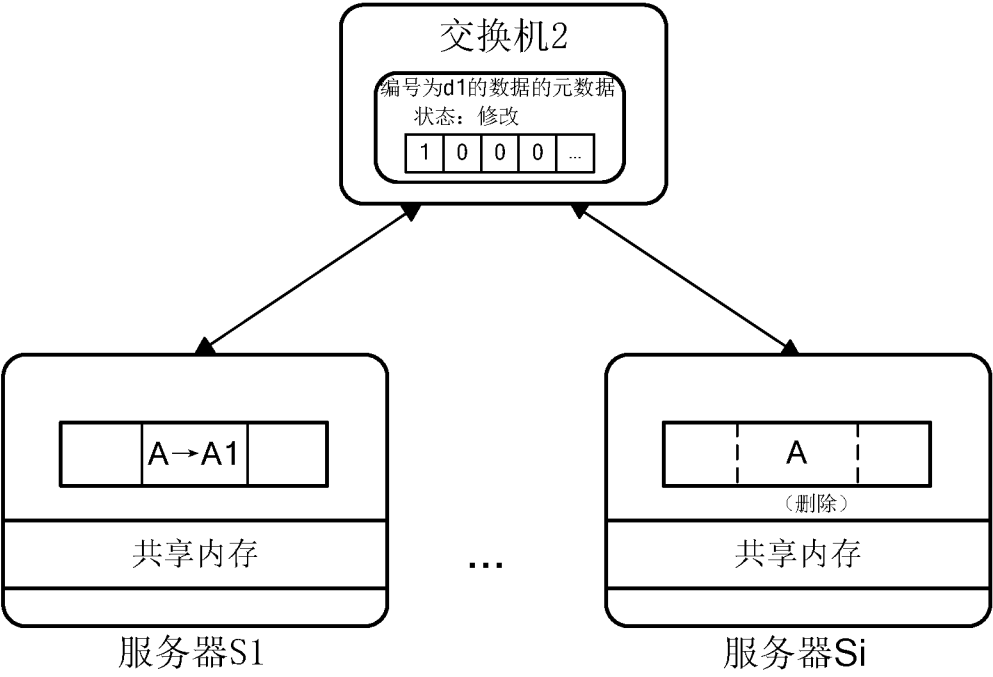


图 10

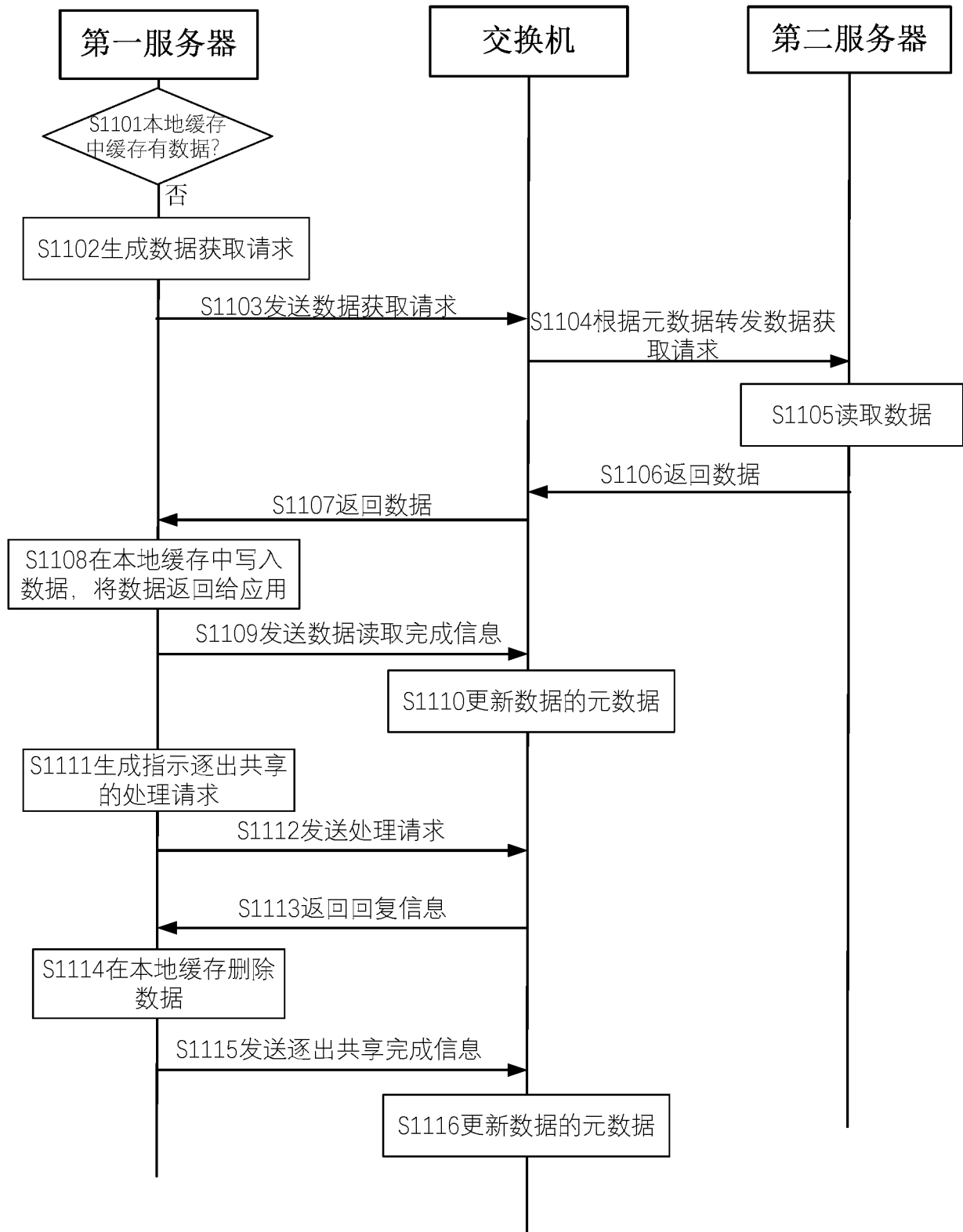


图 11

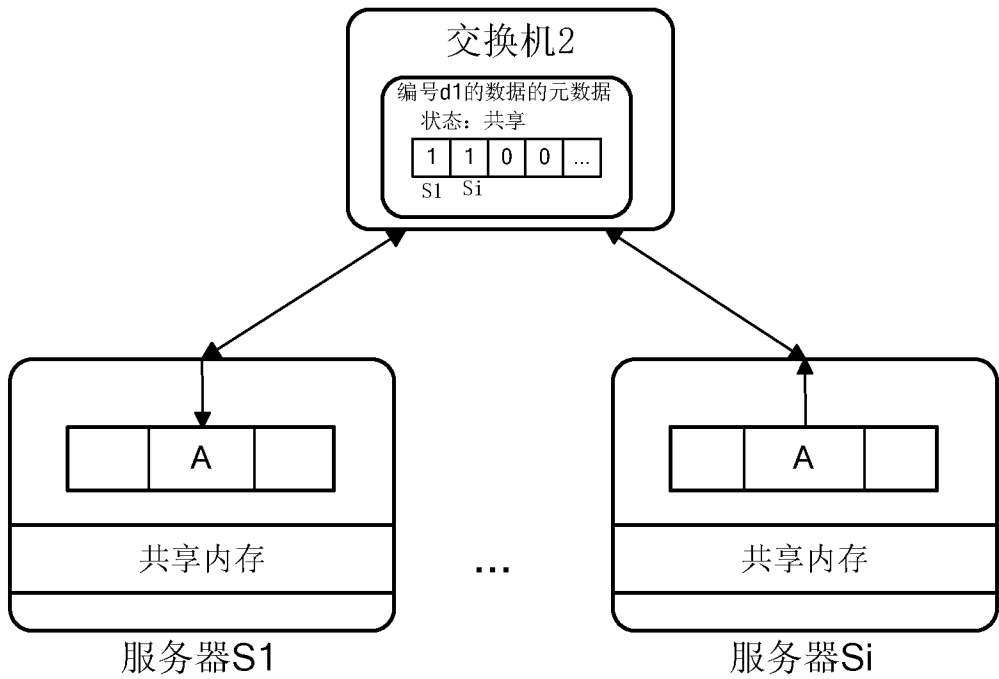


图 12

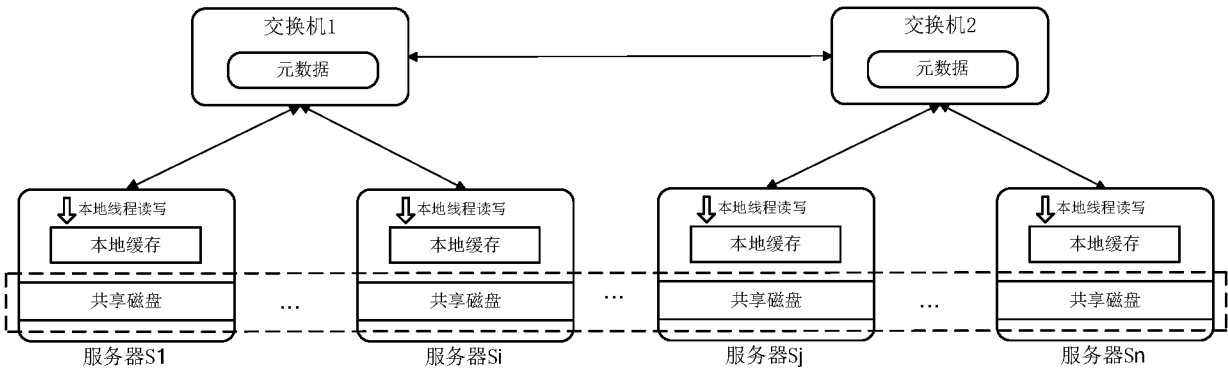


图 13

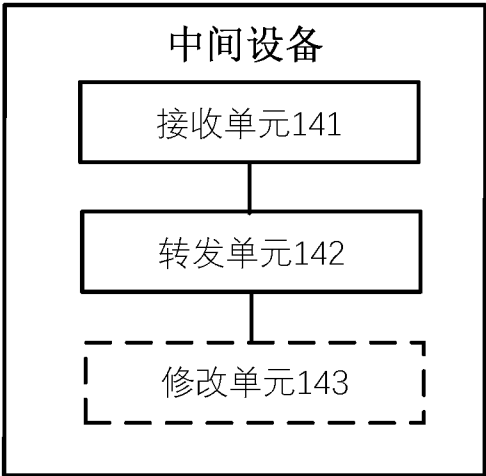


图 14

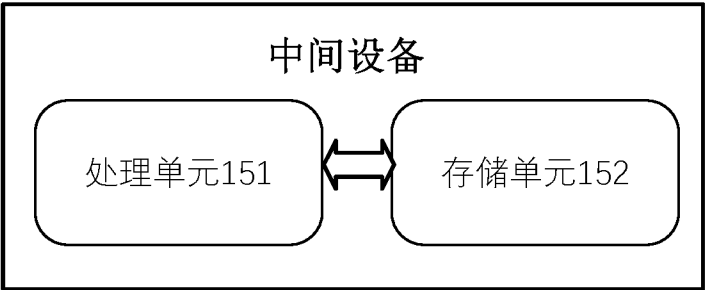


图 15

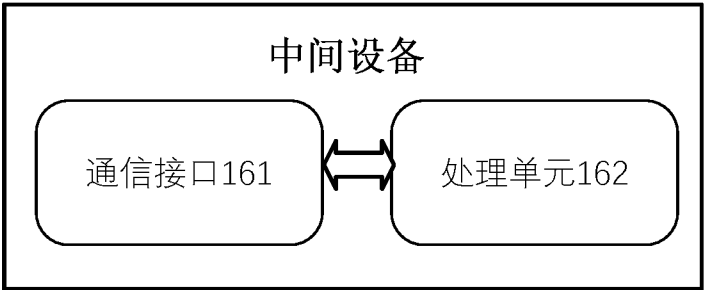


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/072422

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/50(2006.01)i; G06F 16/27(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)

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)

CNABS; CNTXT; VEN; ENTXT, CNKI: 交换机, 路由器, 中间, 设备, 装置, 服务器, 节点, 数据, 版本, 最新, 共享, 缓存, 保存, 记录, 存储, 地址, switch, router, intermediate, equipment, device, server, node, data, version, newest, latest, share, cache, buffer, save, record, store, adress, location, url

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112632029 A (HANGZHOU HIKVISION SYSTEM TECHNOLOGY CO., LTD.) 09 April 2021 (2021-04-09) description, paragraphs 121-206, and figures 1-12	1-25
X	CN 102314506 A (BEIJING BASESOFT INFORMATION TECHNOLOGIES INC.) 11 January 2012 (2012-01-11) description, paragraphs 9-40, and figures 1-2	1-3, 9, 10, 12, 13
Y	CN 102314506 A (BEIJING BASESOFT INFORMATION TECHNOLOGIES INC.) 11 January 2012 (2012-01-11) description, paragraphs 9-40, and figures 1-2	4-25
Y	US 6324584 B1 (IBM) 27 November 2001 (2001-11-27) description column 3, figure 8	4-25
A	CN 110059026 A (HUAWEI TECHNOLOGIES CO., LTD.) 26 July 2019 (2019-07-26) entire document	1-25
A	CN 110933137 A (BEIJING INSPUR DATA TECHNOLOGY CO., LTD.) 27 March 2020 (2020-03-27) entire document	1-25



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

22 February 2022

Date of mailing of the international search report

29 March 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (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/CN2022/072422

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2013009806 A1 (BANK OF AMERICA et al.) 17 January 2013 (2013-01-17) entire document	1-25

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/072422

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112632029	A	09 April 2021	None			
CN	102314506	A	11 January 2012	CN	102314506	B	09 September 2015
US	6324584	B1	27 November 2001	None			
CN	110059026	A	26 July 2019	WO	2019140885	A1	25 July 2019
				EP	3736701	A1	11 November 2020
				EP	3736701	A4	17 March 2021
				US	2020349078	A1	05 November 2020
CN	110933137	A	27 March 2020	None			
WO	2013009806	A1	17 January 2013	US	2013019013	A1	17 January 2013
				US	9015320	B2	21 April 2015

国际检索报告

国际申请号

PCT/CN2022/072422

A. 主题的分类

G06F 9/50(2006.01)i; G06F 16/27(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)

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))

CNABS;CNTXT;VEN;ENTXT,CNKI:交换机,路由器,中间,设备,装置,服务器,节点,数据,版本,最新,共享,缓存,保存,记录,存储,地址,switch,router,intermediate,equipment, device,server,node,data,version,newest,latest,share,cache,buffer,save,record,store,adress,location,url

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 112632029 A (杭州海康威视系统技术有限公司) 2021年4月9日 (2021 - 04 - 09) 说明书第121-206段, 图1-12	1-25
X	CN 102314506 A (北京人大金仓信息技术股份有限公司) 2012年1月11日 (2012 - 01 - 11) 说明书第9-40段, 图1-2	1-3、9、10、12、13
Y	CN 102314506 A (北京人大金仓信息技术股份有限公司) 2012年1月11日 (2012 - 01 - 11) 说明书第9-40段, 图1-2	4-25
Y	US 6324584 B1 (IBM) 2001年11月27日 (2001 - 11 - 27) 说明书第3栏, 图8	4-25
A	CN 110059026 A (华为技术有限公司) 2019年7月26日 (2019 - 07 - 26) 全文	1-25
A	CN 110933137 A (北京浪潮数据技术有限公司) 2020年3月27日 (2020 - 03 - 27) 全文	1-25
A	WO 2013009806 A1 (BANK OF AMERICA等) 2013年1月17日 (2013 - 01 - 17) 全文	1-25

☐ 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

2022年2月22日

国际检索报告邮寄日期

2022年3月29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

传真号 (86-10)62019451

受权官员

胡赢

电话号码 (86-27)59371827

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/072422

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	112632029	A	2021年4月9日	无			
CN	102314506	A	2012年1月11日	CN	102314506	B	2015年9月9日
US	6324584	B1	2001年11月27日	无			
CN	110059026	A	2019年7月26日	WO	2019140885	A1	2019年7月25日
				EP	3736701	A1	2020年11月11日
				EP	3736701	A4	2021年3月17日
				US	2020349078	A1	2020年11月5日
CN	110933137	A	2020年3月27日	无			
WO	2013009806	A1	2013年1月17日	US	2013019013	A1	2013年1月17日
				US	9015320	B2	2015年4月21日