

(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
2017 年 1 月 19 日 (19.01.2017)



(10) 国际公布号
WO 2017/008687 A1

- (51) 国际专利分类号:
G06F 9/54 (2006.01)
- (21) 国际申请号: PCT/CN2016/089308
- (22) 国际申请日: 2016 年 7 月 8 日 (08.07.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510413284.1 2015 年 7 月 14 日 (14.07.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛英属开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 张俊 (ZHANG, Jun); 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)。何乐 (HE, Le); 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)。张杰 (ZHANG, Jie); 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: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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。
- (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)。

[见续页]

(54) Title: MESSAGE QUEUING SYSTEM AND METHOD OF REALIZING MESSAGE COMMUNICATION

(54) 发明名称: 消息队列系统以及实现消息通信的方法

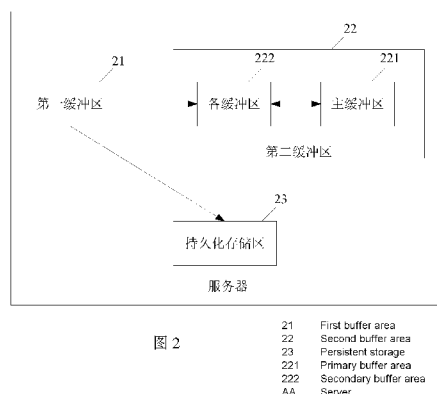


图 2

(57) Abstract: Provided are a message queuing system and a method of realizing message communication. The system comprises: a first buffer area (21), a second buffer area (22) and a persistent storage (23). The first buffer area (21) and the second buffer area (22) are located in a memory of a server in which a message producer and a message consumer are located. The second buffer area (22) comprises a primary buffer area (221) and a secondary buffer area (222). The first buffer area (21) is configured to buffer a message pushed by the message producer, store the message in the persistent storage (23), and send the message stored in the persistent storage (23) to the second buffer area (22). The persistent storage (23) is configured to persistently store the message sent by the first buffer area (21). The secondary buffer area (222) is configured to buffer the message sent by the first buffer area (21). The primary buffer area (221) is configured to deliver the message therein to the message consumer, and after all of the messages is delivered to the message consumer, perform primary-secondary switching with the secondary buffer area (222). The system improves the real-time performance of message transmission in the message queuing system.

(57) 摘要:

[见续页]



WO 2017/008687 A1



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种消息队列系统和实现消息通信的方法，所述系统包括：第一缓冲区（21）、第二缓冲区（22）和持久化存储区（23）；第一缓冲区（21）和第二缓冲区（22）位于消息生产者和消息消费者所在的服务器的内存中，第二缓冲区（22）包括主缓冲区（221）和备缓冲区（222）；第一缓冲区（21）用于缓存消息生产者推送的消息并存储在持久化存储区（23）中，将已存储到持久化存储区（23）的消息发送给第二缓冲区（22）；持久化存储区（23）用于对第一缓冲区（21）发送的消息进行持久化存储；备缓冲区（222）用于缓存第一缓冲区（21）发送的消息；主缓冲区（221）用于将自身的消息传递给消息消费者，在将自身的消息全部传递给消息消费者之后，与备缓冲区（222）进行主备切换。所述系统提高了消息在消息队列系统中传输的实时性。

消息队列系统以及实现消息通信的方法

本申请要求 2015 年 07 月 14 日递交的申请号为 201510413284.1、发明名称为“消息队列系统以及实现消息通信的方法”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。

5

技术领域

本申请涉及计算机领域，特别是涉及消息队列系统以及实现消息通信的方法。

背景技术

10 在计算机系统中，消息队列是实现不同进程间通信或者同一进程中的不同线程间通信的一种重要方式。也就是说，不同的进程之间或不同的线程之间在消息队列中完成消息传输。

在现有的相关技术中，已经存在多种关于消息队列系统的实现方案。其中一个方案为，消息队列系统由进程（或线程）所在的服务器上的缓存区以及本地磁盘构成。其中，
15 先将消息生产者推送的消息存储在该缓存区中，当该缓存区中的消息个数达到配置值或者消息的发布时间超过阈值时，再将该缓存区中的消息存储到本地磁盘中，以保证消息的可靠性。并且只有存储到本地磁盘中的消息才可以被进一步传递给消息消费者。另外，由于本地磁盘的存储空间有限，存储在本地磁盘中的消息无论是否被传递给消息消费者，其在保留一定时间后就会被删除，以便释放本地磁盘的存储空间。所谓“可靠性”就是
20 指，消息生产者推送到消息队列系统中的消息不丢失，即使消息队列系统故障也可以被恢复；并且，消息在未被传递给消息消费者之前可以一直保存在消息队列系统中，直到被传递给消息消费者为止。

在第一个方案中，只有位于本地磁盘的消息才可以被传递给消息消费者，而消息在被存储到本地磁盘之前，还必须先被存储在缓存区中，因此，消息传输的实时性会比较
25 差。并且，消息是从本地磁盘传递给消息消费者的，受到本地磁盘的 IO 操作速度的影响，消息传输的实时性就会更差。另外，由于本地磁盘中的消息会被定期删除，因此消息的可靠性也非常差。

另一个方案为：消息队列系统由进程（或线程）所在的服务器上的内存以及位于该服务器外部的数据库构成。其中，先将消息生产者推送的消息存储在内存中，再将消息

从内存中传递给消息消费者。并且，为了保证消息的可靠性，还将内存中的消息存储到数据库中。显然，在第二个方案中，消息直接从内存传递到消息消费者，并且，内存的IO操作速度比本地磁盘的IO操作速度快，因此，相较于第一个方案，消息传输的实时性更好。另外，由于数据库的存储空间比较大，存储在数据库中的消息不会被定期删除，

5 因此，相较于第一个方案，消息的可靠性也更好。

但是，在实现本申请的过程中，本申请的发明人发现现有技术中至少存在如下问题：虽然在第二个方案中，内存的IO操作速度比较快，使其消息在消息队列系统中传输的实时性相较于第一个方案好。但是，由于内存不支持同时执行读/写操作，因此，在将消息生产者推送的消息存储在内存中的同时，就不能将消息从内存中传递给消息消费者，这

10 样会很大程度地影响消息传输的实时性。

另外，在第二个方案中，如果消息消费者无法正常地从消息队列系统接收消息或者只能以很慢的速度接收消息，并且消息生产者生产消息的速度又非常快，这样就会有越来越多的消息占满内存，严重时会发生内存溢出。当发生内存溢出时，溢出的消息会被丢弃掉，以至于溢出的消息无法正常传递给消息消费者。

发明内容

为了解决上述技术问题，本申请实施例提供了消息队列系统、以及实现消息通信的方法，以解决现有技术中由于内存不支持同时执行读/写操作而影响消息在消息队列系统中传输的实时性的问题。另外，进一步解决现有技术中由于消息队列系统中的内存发生

20 内存溢出而导致溢出的消息无法正常地被传递给消息消费者的问题。

本申请实施例公开了如下技术方案：

一种消息队列系统，包括：第一缓冲区、第二缓冲区以及持久化存储区；其中，所述第一缓冲区和第二缓冲区位于消息生产者和消息消费者所在的服务器的内存中，所述第二缓冲区包括主缓冲区和备缓冲区；

25 所述第一缓冲区，用于缓存消息生产者推送的消息，将消息生产者推送的消息存储到所述持久化存储区中，并将已存储到所述持久化存储区的消息发送给所述第二缓冲区；

所述持久化存储区，用于通过持久化存储介质对所述第一缓冲区发送的消息进行持久化存储；

所述备缓冲区，用于缓存所述第一缓冲区发送的消息；

30 所述主缓冲区，用于将位于所述主缓冲区的消息传递给消息消费者，在将位于所述

主缓冲区的消息全部传递给消息消费者之后，与所述备缓冲区进行主备切换。

优选的，所述备缓冲区还用于，在缓存所述第一缓冲区发送的消息之前，检查自身是否存在内存溢出，如果存在内存溢出，丢弃所述第一缓冲区发送的消息，并将自身的当前状态标记为非正常状态，如果不存在内存溢出，继续缓存所述第一缓冲区发送的消息；

所述主缓冲区还用于，在与所述备缓冲区进行主备切换之前，检查所述备缓冲区的当前状态是否为正常状态，如果为非正常状态，不与所述备缓冲区进行主备切换，而从所述持久化存储区中读取消息，并将读取的消息传递给消息消费者，如果为正常状态，继续与所述备缓冲区进行主备切换。

优选的，所述主缓冲区还用于，在从所述持久化存储区中读取消息之后，判断在所述备缓冲区中是否存在下一个读取的消息，如果存在，将所述备缓冲区的当前状态标记为正常状态，如果不存在，维持所述备缓冲区的当前状态。

优选的，所述持久化存储区位于所述服务器外部的分布式文件系统中。

优选的，所述主缓冲区和备缓冲区为同构的缓冲区。

优选的，所述消息生产者和消息消费者为运行在所述服务器上的两个进程或者为运行在所述服务器上的一个进程中的两个线程。

一种实现消息通信的方法，预先在消息生产者和消息消费者所在的服务器的内存中建立第一缓冲区和第二缓冲区，所述第二缓冲区包括主缓冲区和备缓冲区，所述方法包括：

第一缓冲区缓存消息生产者推送的消息，将消息生产者推送的消息存储到持久化存储区中，以便将所述消息持久化地存储在所述持久化存储区中；

第一缓冲区将已存储到所述持久化存储区的消息发送给第二缓冲区；

备缓冲区缓存第一缓冲区发送的消息；

主缓冲区将位于主缓冲区的消息传递给消息消费者，在将位于所述主缓冲区的消息全部传递给消息消费者之后，与备缓冲区进行主备切换。

优选的，所述方法还包括：

在缓存第一缓冲区发送的消息之前，备缓冲区检查自身是否存在内存溢出，如果存在内存溢出，丢弃所述第一缓冲区发送的消息，并将自身的当前状态标记为非正常状态；

在与备缓冲区进行主备切换之前，主缓冲区检查备缓冲区的当前状态是否为正常状态，如果为非正常状态，不与备缓冲区进行主备切换，而从持久化存储区中读取消息，

以便将读取的消息传递给消息消费者；

则所述备缓冲区缓存第一缓冲区发送的消息为：

如果不存在内存溢出，备缓冲区缓存第一缓冲区发送的消息；

则所述主缓冲区与备缓冲区进行主备切换为：

5 如果为正常状态，主缓冲区与备缓冲区进行主备切换。

优选的，所述方法还包括：

在从持久化存储区中读取消息之后，主缓冲区判断在备缓冲区中是否存在下一个读取的消息，如果存在，将所述备缓冲区的当前状态标记为正常状态，如果不存在，维持所述备缓冲区的当前状态。

10 优选的，所述持久化存储区位于所述服务器外部的分布式文件系统中。

优选的，所述主缓冲区和备缓冲区为同构的缓冲区。

优选的，所述消息生产者和消息消费者为运行在所述服务器上的两个进程或者为运行在所述服务器上的一个进程中的两个线程。

由上述实施例可以看出，与现有技术相比，本申请的优点在于：

15 在消息生产者和消息消费者所在的服务器的内存中设置用于缓存消息的缓冲区，在正常情况下，通过内存访问完成消息传输，保证了消息在消息队列系统中传输的实时性。并且，通过主、备缓冲区的双结构设计，可以实现在将消息生产者推送的消息存储在备缓冲区的同时，将消息从主缓冲区中传递给消息消费者，更进一步提高了消息在消息队列系统中传输的实时性。另外，还将消息生产者推送的消息存储在持久化存储区中，保
20 证了消息的可靠性。

此外，当消息队列系统中的备缓冲区出现内存溢出问题时，即，在非正常情况下，也可以从持久化存储区中取回内存中丢失的消息，避免了由于消息队列系统中的内存发生内存溢出而导致溢出的消息无法正常地被传递给消息消费者的问题。

25 附图说明

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。

30 图 1 示意性地示出了本发明的实施方式可以在其中实施的示例性应用场景；

图 2 示意性地示出了根据本申请实施例的一种消息队列系统的结构示意图；

图 3 示意性地示出了根据本申请实施例的另一种消息队列系统的结构示意图；

图 4 示意性地示出了根据本申请实施例的另一种消息队列系统的结构示意图；

图 5 示意性地示出了根据本申请实施例的主缓冲区读取消息的工作示意图；

5 图 6 示意性地示出了根据本申请实施例的一种实现消息通信的方法的流程图；

图 7 示意性地示出了根据本申请实施例的另一种实现消息通信的方法的流程图；

图 8 示意性地示出了根据本申请实施例的另一种实现消息通信的方法的流程图。

具体实施方式

10 参阅图 1 所示，图 1 示意性地示出了本发明的实施方式可以在其中实施的示例性应用场景。其中，服务器上存在一个消息生产者 10 和一个消息消费者 20，当消息生产者 10 与消息消费者 20 之间进行通信时，消息生产者 10 将消息 21 推送到消息队列系统 30 中，消息队列系统 30 负责将消息 21 传递给消息消费者 20。另外，为了保证消息 21 的可靠性，消息队列系统 30 还需要将消息生产者 10 推送的消息 21 存储到持久化存储介质
15 中，如，服务器上的本地磁盘。其中，“传递”可以表现为“推（push）”和“拉（pull）”两种模式，“推”模式是由消息队列系统 30 主动将消息 21 推送给消息消费者 20，“拉”模式是消息消费者 20 主动通过轮询的方式从消息队列系统 30 中拉取消息 21。消息生产者 10 和消息消费者 20 可以为服务器上的两个进程，也可以为服务器上同一个进程中的两个线程。本领域技术人员可以理解，图 1 所示的示意图仅是本发明的实施方式可以在
20 其中得以实现的一个示例。本发明实施方式的应用范围不受到该框架任何方面的限制。例如，服务器上存在多个消息生产者和多个消息消费者，而不是如图 1 所示的仅存在一个消息生产者和一个消息消费者。

为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图对本申请实施例进行详细描述。

25

装置实施例

本申请实施例还提供了一种消息队列系统。请参阅图 2，图 2 示意性地示出了根据本申请实施例的一种消息队列系统的结构示意图。其中，该消息队列系统 20 包括：第一缓冲区 21、第二缓冲区 22 以及持久化存储区 23，第一缓冲区 21 和第二缓冲区 22 位于
30 消息生产者和消息消费者所在的服务器的内存中，第二缓冲区 22 包括主缓冲区 221 和备

缓冲区 222。下面结合该装置的工作原理进一步介绍其内部结构以及连接关系。

第一缓冲区 21，用于缓存消息生产者推送的消息，将消息生产者推送的消息存储在持久化存储区 23 中，并将已存储到持久化存储区 23 的消息发送给第二缓冲区 22。

其中，在消息生产者生产的消息到达第一缓冲区 21 之后，会被首先持久化地存储到持久化存储区 23 中，以保证该消息的可靠性。如果存储失败，持久化存储区 23 告知第一缓冲区 21 消息存储失败，以便第一缓冲区 21 重新将消息持久化存储到持久化存储区 23 中直到存储成功为止，如果存储成功，持久化存储区 23 告知第一缓冲区 21 消息存储成功。

持久化存储区 23，用于通过持久化存储介质对第一缓冲区 21 发送的消息进行持久化存储。

在本申请的一个可选实施方式中，持久化存储区 23 位于消息生产者和消息消费者所在的服务器内部的磁盘中，如图 2 所示，也可以位于该服务器外部的数据库中。

在本申请的另一个可选实施方式中，持久化存储区 23 位于该服务器外部的分布式文件系统中，如图 3 所示。将持久化存储区 23 布置在分布式文件系统中，可以很可靠地实现对海量消息的持久化存储的需求。另外，在分布式文件系统中，通过 LogFile 追加方式记录消息，也能够避免因为第一缓冲区 21 远程操作分布式文件系统而带来的延迟问题。

备缓冲区 222，用于缓存第一缓冲区 21 发送的消息。

主缓冲区 221，用于将位于主缓冲区 221 的消息传递给消息消费者，在将位于主缓冲区 221 的消息全部传递给消息消费者之后，与备缓冲区 222 进行主备切换。

也就是说，备缓冲区 222 专用于缓存第一缓冲区发送的消息，而主缓冲区 221 专用于将位于自身的消息传递给消息消费者。一旦主缓冲区 221 将自身的消息全部传递出去之后，就与备缓冲区互换身份，转换主备关系，即，原来的主缓冲区成为备缓冲区，开始缓存来自第一缓冲区 21 的消息，而原来的备缓冲区成为主缓冲区，将其缓存的消息传递给消息消费者。在以上的正常场景下，整个消息队列系统不需要读取在持久化存储区 23 中备份的消息，所有的逻辑在三个缓冲区（即，第一缓冲区 21、主缓冲区 221 和备缓冲区 222）中完成。

可以理解的，当在服务器的内存中设置用于缓存消息的缓冲区时，通过主、备缓冲区的双结构设计，可以实现在将消息生产者推送的消息存储在备缓冲区的同时，将消息从主缓冲区中传递给消息消费者，更好地保证了消息在消息队列系统中传输的实时性。

在本申请的一个可选实施方式中，主缓冲区 221 和备缓冲区 222 可以为同构的缓冲区，即，两个缓冲区的存储介质以及存储空间大小都相同。

考虑到消息消费者有时可能无法正常从消息队列系统中接收消息，或者只能以很慢的速度接收消息，并且消息生产者生产消息的速度又非常快，在这种情况下，主缓冲区 221 传递消息的进度与备缓冲区 222 缓存消息的进度会出现不匹配，前者快于后者，因此，在主缓冲区 221 将自身的消息全部传递出去之前，备缓冲区 222 就会因为第一缓冲区 21 转发过多的消息而发生内存溢出，当发生内存溢出时，备缓冲区 222 就会主动丢掉溢出的消息，而导致溢出的消息无法正常被传递给消息消费者。因此，为了解决由于备缓冲区 222 发生内存溢出而导致溢出的消息无法正常地被传递给消息消费者的问题，

在本申请的另一个可选实施方式中，如图 4 所示，

备缓冲区 222 还用于，在缓存第一缓冲区 21 发送的消息之前，检查自身是否存在内存溢出，如果存在内存溢出，丢弃第一缓冲区 21 发送的消息，并将自身的当前状态标记为非正常状态，如果不存在内存溢出，继续缓存第一缓冲区 21 发送的消息。

主缓冲区 221 还用于，在与备缓冲区 222 进行主备切换之前，检查备缓冲区 222 的当前状态是否为正常状态，如果为非正常状态，不与备缓冲区 222 进行主备切换，而是从持久化存储区 23 中读取消息，并将读取的消息传递给消息消费者，如果为正常状态，继续与备缓冲区进行主备切换。

也就是说，当备缓冲区 222 发生内存溢出时，其被标记为非正常状态，表明当前的备缓冲区 222 所缓存的消息不完整，由于某些原因备缓冲区 222 已经丢弃了部分消息。因此，在这种非正常场景下，主缓冲区 221 不与备缓冲区 222 进行主备切换，而是直接访问持久化存储区 23（如，底层的分布式文件系统），将持久化存储区 23 中保存的消息读取出来后继续传递给消息消费者。其中，主缓冲区 221 可以根据自己的存储空间大小从持久化存储区 23 中读取相应数量的消息，然后将读取的消息逐个地传递给消息消费者。

这种方式可以一直持续到主缓冲区 221 发现自己传递消息的进度可以与备缓冲区 222 缓存消息的进度匹配，此时主缓冲区 221 则不再需要从持久化存储区 23 中读取消息，而是回到正常场景下，通过主备切换方式传递消息。因此，在本申请的另一个可选实施方式中：

主缓冲区 221 还用于，在从持久化存储区 23 中读取消息之后，判断在备缓冲区 222 中是否存在下一个读取的消息，如果存在，说明主缓冲区 221 传递消息的进度可以与备

缓冲区 222 缓存消息的进度匹配，主缓冲区 221 将备缓冲区 222 的当前状态标记为正常状态，如果不存在，维持备缓冲区 222 的当前状态。

例如，如图 5 所示，备缓冲区 222 缓存了标识为 30-50 的消息，而标识为 1-29 的消息由于内存溢出而被备缓冲区 222 丢弃掉了，主缓冲区 221 先从底层的分布式文件系统中读取消息，假设主缓冲区 221 一次性只能缓存 20 个消息，主缓冲区 221 从分布式文件系统中读取标识为 1-20 的消息，并逐个地传递给消息消费者。在读取完之后，主缓冲区 221 判断其下一个读取的消息（即标识为 21 的消息）是否在备缓冲区 222 中，经过判断发现，其下一个读取的标识为 21 的消息不在备缓冲区 222 中，主缓冲区 221 继续维持备缓冲区 222 的当前状态标记，即，非正常状态。当主缓冲区 221 将自身的消息全部传递出去后，即刻检查备缓冲区 222 的当前状态，由于检查出的备缓冲区 222 的当前状态仍然为非正常状态，因此继续从底层的分布式文件系统中读取消息，即读取标识为 21-40 的消息，并逐个地传递给消息消费者。在读取完之后，主缓冲区 221 继续判断其下一个读取的消息（即标识为 41 的消息）是否在备缓冲区 222 中，经过判断发现，其下一个读取的标识为 41 的消息在备缓冲区 222 中，主缓冲区 221 将备缓冲区 222 的当前状态标记为正常状态。当主缓冲区 221 将自身的消息再一次全部传递出去后，由于检查出的备缓冲区 222 的当前状态为正常状态，因此可以正常地与备缓冲区 222 进行主备切换。

在本申请的一个可选实施方式中，消息生产者和消息消费者为运行在服务器上的两个进程或者为运行在服务器上的一个进程中的两个线程。

由上述实施例可以看出，与现有技术相比，本申请的优点在于：

在消息生产者和消息消费者所在的服务器的内存中设置用于缓存消息的缓冲区，在正常情况下，通过内存访问完成消息传输，保证了消息在消息队列系统中传输的实时性。并且，通过主、备缓冲区的双结构设计，可以实现在将消息生产者推送的消息存储在备缓冲区的同时，将消息从主缓冲区中传递给消息消费者，更进一步提高了消息在消息队列系统中传输的实时性。另外，还将消息生产者推送的消息存储在持久化存储区中，保证了消息的可靠性。

此外，当消息队列系统中的备缓冲区出现内存溢出问题时，即，在非正常情况下，也可以从持久化存储区中取回内存中丢失的消息，避免了由于消息队列系统中的内存发生内存溢出而导致溢出的消息无法正常地被传递给消息消费者的问题。

请参阅图 6，图 6 示意性地示出了根据本申请实施例的一种实现消息通信的方法的流程图，例如，该方法可以由消息队列系统 20 执行，该方法例如可以包括以下步骤：

步骤 600：预先在消息生产者和消息消费者所在的服务器的内存中建立第一缓冲区和第二缓冲区，所述第二缓冲区包括主缓冲区和备缓冲区。

5 步骤 601：第一缓冲区缓存消息生产者推送的消息，将消息生产者推送的消息存储到持久化存储区中，以便将所述消息持久化地存储在所述持久化存储区中。

步骤 602：第一缓冲区将已存储到所述持久化存储区的消息发送给第二缓冲区。

步骤 603：备缓冲区缓存第一缓冲区发送的消息。

步骤 604：主缓冲区将位于主缓冲区的消息传递给消息消费者。

10 步骤 605：主缓冲区在将位于所述主缓冲区的消息全部传递给消息消费者之后，与备缓冲区进行主备切换。

如图 7 所示，在本申请的一个可选实施方式中，该方法例如可以包括以下步骤：

步骤 600：预先在消息生产者和消息消费者所在的服务器的内存中建立第一缓冲区和第二缓冲区，所述第二缓冲区包括主缓冲区和备缓冲区。

15 步骤 601：第一缓冲区缓存消息生产者推送的消息，将消息生产者推送的消息存储到持久化存储区中，以便将所述消息持久化地存储在所述持久化存储区中。

步骤 602：第一缓冲区将已存储到所述持久化存储区的消息发送给第二缓冲区。

步骤 603a：备缓冲区检查自身是否存在内存溢出，如果存在内存溢出，进入步骤 606，如果不存在内存溢出，进入步骤 603。

20 步骤 603：备缓冲区缓存第一缓冲区发送的消息。

步骤 604：主缓冲区将位于主缓冲区的消息传递给消息消费者。

步骤 605a：在将位于所述主缓冲区的消息全部传递给消息消费者之后，主缓冲区检查备缓冲区的当前状态是否为正常状态，如果为非正常状态，进入步骤 607，如果为正常状态，进入步骤 605。

25 步骤 605：主缓冲区与备缓冲区进行主备切换，结束流程。

步骤 606：备缓冲区丢弃所述第一缓冲区发送的消息，并将自身的当前状态标记为非正常状态，结束流程。

步骤 607：主缓冲区不与备缓冲区进行主备切换，而从持久化存储区中读取消息，以便将读取的消息传递给消息消费者，结束流程。

30 如图 8 所示，在本申请的另一个可选实施方式中，该方法例如可以包括以下步骤：

步骤 600: 预先在消息生产者和消息消费者所在的服务器的内存中建立第一缓冲区和第二缓冲区, 所述第二缓冲区包括主缓冲区和备缓冲区。

步骤 601: 第一缓冲区缓存消息生产者推送的消息, 将消息生产者推送的消息存储到持久化存储区中, 以便将所述消息持久化地存储在所述持久化存储区中。

5 步骤 602: 第一缓冲区将已存储到所述持久化存储区的消息发送给第二缓冲区。

步骤 603a: 备缓冲区检查自身是否存在内存溢出, 如果存在内存溢出, 进入步骤 606, 如果不存在内存溢出, 进入步骤 603。

步骤 603: 备缓冲区缓存第一缓冲区发送的消息。

步骤 604: 主缓冲区将位于主缓冲区的消息传递给消息消费者。

10 步骤 605a: 在将位于所述主缓冲区的消息全部传递给消息消费者之后, 主缓冲区检查备缓冲区的当前状态是否为正常状态, 如果为非正常状态, 进入步骤 607, 如果为正常状态, 进入步骤 605。

步骤 605: 主缓冲区与备缓冲区进行主备切换, 结束流程。

15 步骤 606: 备缓冲区丢弃所述第一缓冲区发送的消息, 并将自身的当前状态标记为非正常状态, 结束流程。

步骤 607: 主缓冲区不与备缓冲区进行主备切换, 而从持久化存储区中读取消息, 以便将读取的消息传递给消息消费者。

步骤 608: 主缓冲区判断在备缓冲区中是否存在下一个读取的消息, 如果存在, 进入步骤 609, 如果不存在, 进入步骤 610。

20 步骤 609: 将所述备缓冲区的当前状态标记为正常状态, 结束流程。

步骤 610: 维持所述备缓冲区的当前状态, 结束流程。

在本申请的另一个可选实施方式中, 所述持久化存储区位于所述服务器外部的分布式文件系统中。

在本申请的另一个可选实施方式中, 所述主缓冲区和备缓冲区为同构的缓冲区。

25 在本申请的另一个可选实施方式中, 所述消息生产者和消息消费者为运行在所述服务器上的两个进程或者为运行在所述服务器上的一个进程中的两个线程。

由上述实施例可以看出, 与现有技术相比, 本申请的优点在于:

在消息生产者和消息消费者所在的服务器的内存中设置用于缓存消息的缓冲区, 在正常情况下, 通过内存访问完成消息传输, 保证了消息在消息队列系统中传输的实时性。

30 并且, 通过主、备缓冲区的双结构设计, 可以实现在将消息生产者推送的消息存储在备

缓冲区的同时，将消息从主缓冲区中传递给消息消费者，更进一步提高了消息在消息队列系统中传输的实时性。另外，还将消息生产者推送的消息存储在持久化存储区中，保证了消息的可靠性。

此外，当消息队列系统中的备缓冲区出现内存溢出问题时，即，在非正常情况下，也可以从持久化存储区中取回内存中丢失的消息，避免了由于消息队列系统中的内存发生内存溢出而导致溢出的消息无法正常地被传递给消息消费者的问题。

所述领域的技术人员可以清楚地了解到，为了描述的方便和简洁，上述描述的系统、装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述到的装置实施例仅仅是示意性的，例如，所述单元的划分，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可以集成到另一个系统，或一些特征可以忽略，或不执行。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，可以是电性、机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以是物理上分开的，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单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，可以采用软件功能单元的形式实现。

需要说明的是，本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）或随机存储记忆体（Random Access Memory，RAM）等。

以上对本申请所提供的消息队列系统以及实现消息通信的方法进行了详细介绍，本文中应用了具体实施例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本

申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权利要求书

1、一种消息队列系统，其特征在于，包括：第一缓冲区、第二缓冲区以及持久化存储区；其中，所述第一缓冲区和第二缓冲区位于消息生产者和消息消费者所在的服务器的内存中，所述第二缓冲区包括主缓冲区和备缓冲区；

5 所述第一缓冲区，用于缓存消息生产者推送的消息，将消息生产者推送的消息存储到所述持久化存储区中，并将已存储到所述持久化存储区的消息发送给所述第二缓冲区；

所述持久化存储区，用于通过持久化存储介质对所述第一缓冲区发送的消息进行持久化存储；

10 所述备缓冲区，用于缓存所述第一缓冲区发送的消息；

所述主缓冲区，用于将位于所述主缓冲区的消息传递给消息消费者，在将位于所述主缓冲区的消息全部传递给消息消费者之后，与所述备缓冲区进行主备切换。

2、根据权利要求 1 所述的系统，其特征在于，

15 所述备缓冲区还用于，在缓存所述第一缓冲区发送的消息之前，检查自身是否存在内存溢出，如果存在内存溢出，丢弃所述第一缓冲区发送的消息，并将自身的当前状态标记为非正常状态，如果不存在内存溢出，继续缓存所述第一缓冲区发送的消息；

所述主缓冲区还用于，在与所述备缓冲区进行主备切换之前，检查所述备缓冲区的当前状态是否为正常状态，如果为非正常状态，不与所述备缓冲区进行主备切换，而从所述持久化存储区中读取消息，并将读取的消息传递给消息消费者，如果为正常状态，
20 继续与所述备缓冲区进行主备切换。

3、根据权利要求 2 所述的系统，其特征在于，所述主缓冲区还用于，在从所述持久化存储区中读取消息之后，判断在所述备缓冲区中是否存在下一个读取的消息，如果存在，将所述备缓冲区的当前状态标记为正常状态，如果不存在，维持所述备缓冲区的当前状态。

25 4、根据权利要求 1 至 3 中任意一项所述的系统，其特征在于，所述持久化存储区位于所述服务器外部的分布式文件系统中。

5、根据权利要求 1 至 4 中任意一项所述的系统，其特征在于，所述主缓冲区和备缓冲区为同构的缓冲区。

30 6、根据权利要求 1 至 4 中任意一项所述的系统，其特征在于，所述消息生产者和消息消费者为运行在所述服务器上的两个进程或者为运行在所述服务器上的一个进程

中的两个线程。

7、一种实现消息通信的方法，其特征在于，预先在消息生产者和消息消费者所在的服务器的内存中建立第一缓冲区和第二缓冲区，所述第二缓冲区包括主缓冲区和备缓冲区，所述方法包括：

5 第一缓冲区缓存消息生产者推送的消息，将消息生产者推送的消息存储到持久化存储区中，以便将所述消息持久化地存储在所述持久化存储区中；

 第一缓冲区将已存储到所述持久化存储区的消息发送给第二缓冲区；

 备缓冲区缓存第一缓冲区发送的消息；

 主缓冲区将位于主缓冲区的消息传递给消息消费者，在将位于所述主缓冲区的消息
10 全部传递给消息消费者之后，与备缓冲区进行主备切换。

8、根据权利要求7所述的方法，其特征在于，所述方法还包括：

 在缓存第一缓冲区发送的消息之前，备缓冲区检查自身是否存在内存溢出，如果存在内存溢出，丢弃所述第一缓冲区发送的消息，并将自身的当前状态标记为非正常状态；

 在与备缓冲区进行主备切换之前，主缓冲区检查备缓冲区的当前状态是否为正常状态，如果为非正常状态，不与备缓冲区进行主备切换，而从持久化存储区中读取消息，
15 以便将读取的消息传递给消息消费者；

 则所述备缓冲区缓存第一缓冲区发送的消息为：

 如果不存在内存溢出，备缓冲区缓存第一缓冲区发送的消息；

 则所述主缓冲区与备缓冲区进行主备切换为：

20 如果为正常状态，主缓冲区与备缓冲区进行主备切换。

9、根据权利要求8所述的方法，其特征在于，所述方法还包括：

 在从持久化存储区中读取消息之后，主缓冲区判断在备缓冲区中是否存在下一个读取的消息，如果存在，将所述备缓冲区的当前状态标记为正常状态，如果不存在，维持所述备缓冲区的当前状态。

25 10、根据权利要求7至9中任意一项所述的方法，其特征在于，所述持久化存储区位于所述服务器外部的分布式文件系统中。

 11、根据权利要求7至10中任意一项所述的方法，其特征在于，所述主缓冲区和备缓冲区为同构的缓冲区。

 12、根据权利要求7至10中任意一项所述的方法，其特征在于，所述消息生产者
30 和消息消费者为运行在所述服务器上的两个进程或者为运行在所述服务器上的一个进

程中的两个线程。

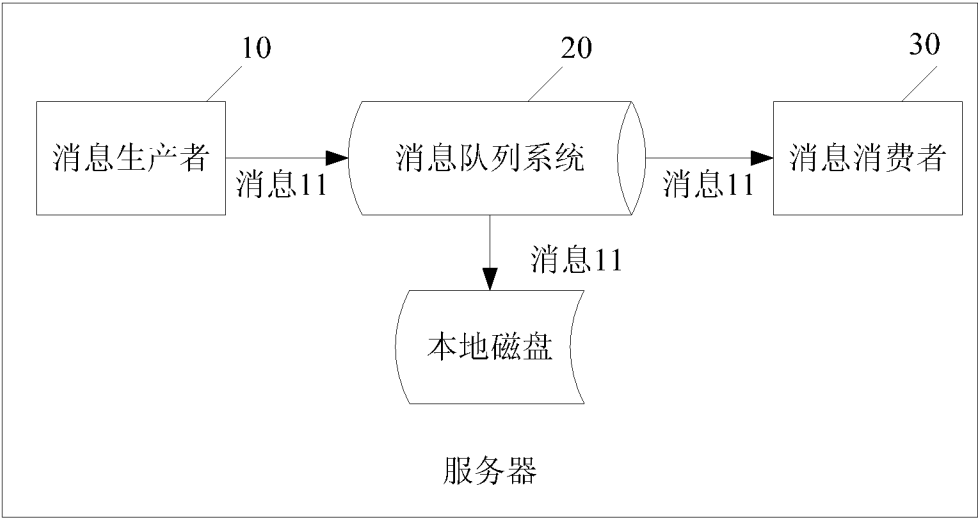


图 1

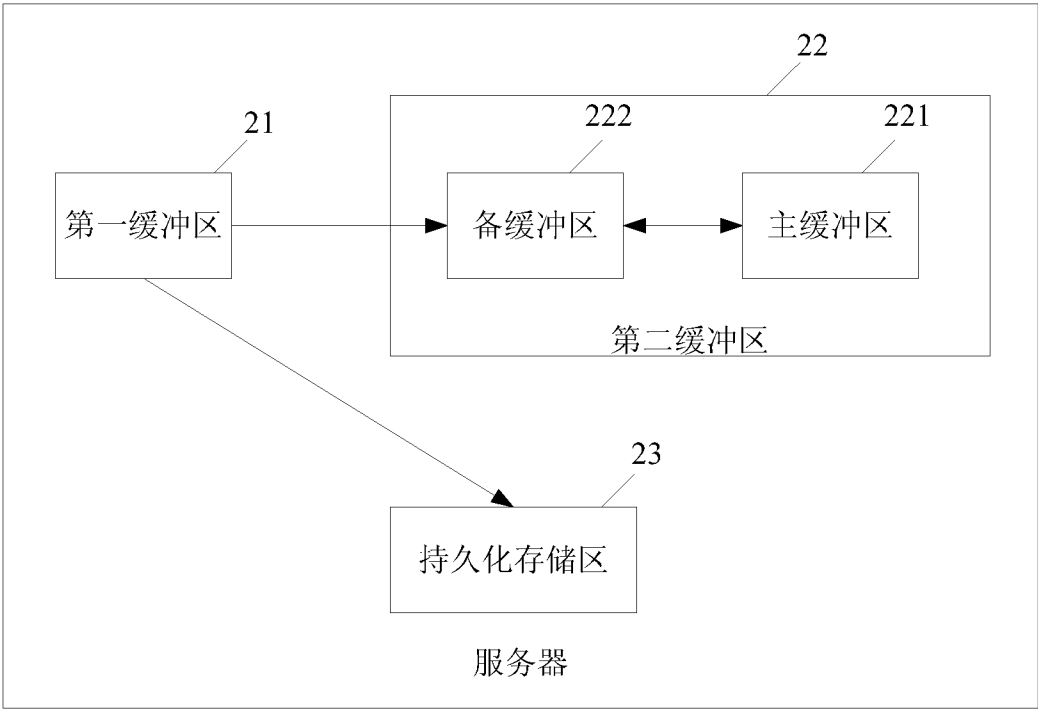


图 2

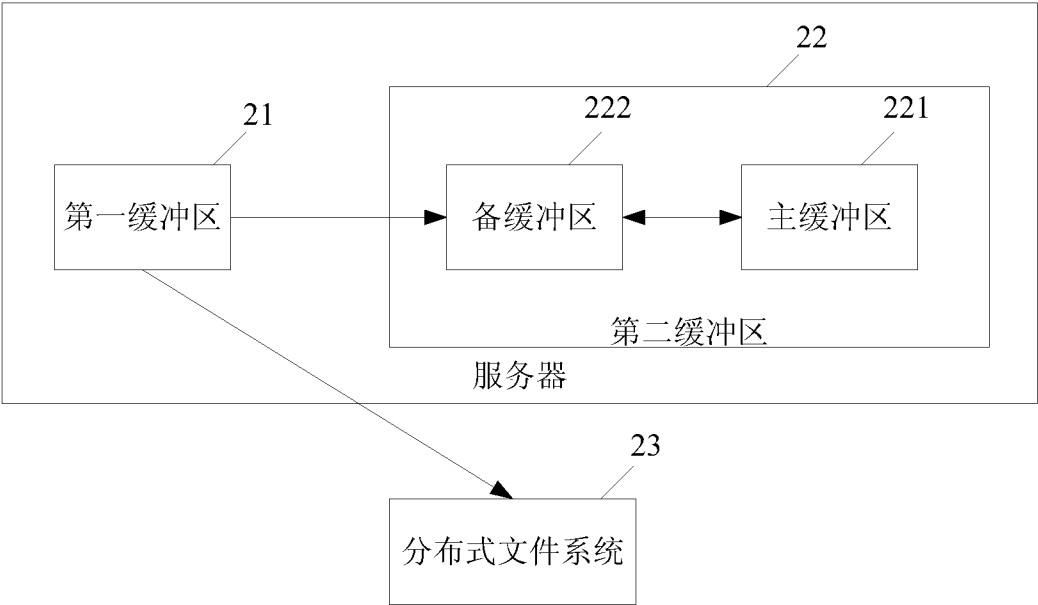


图 3

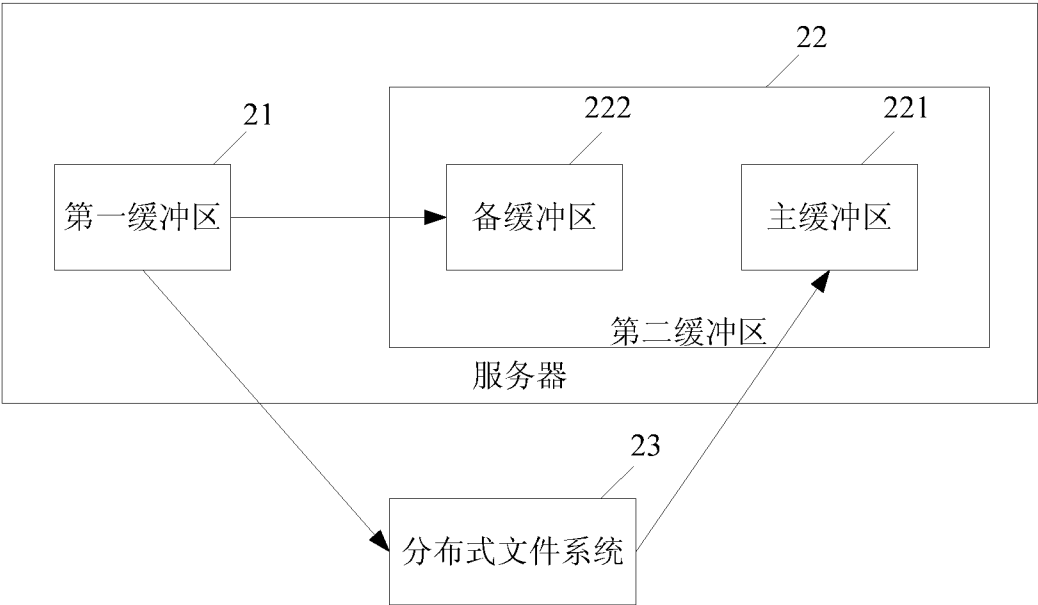


图 4

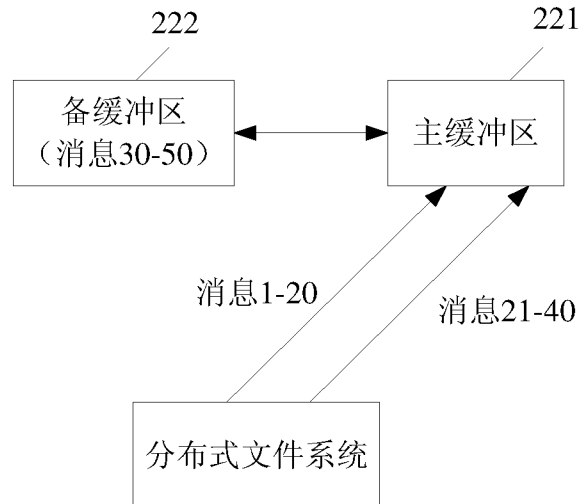


图 5

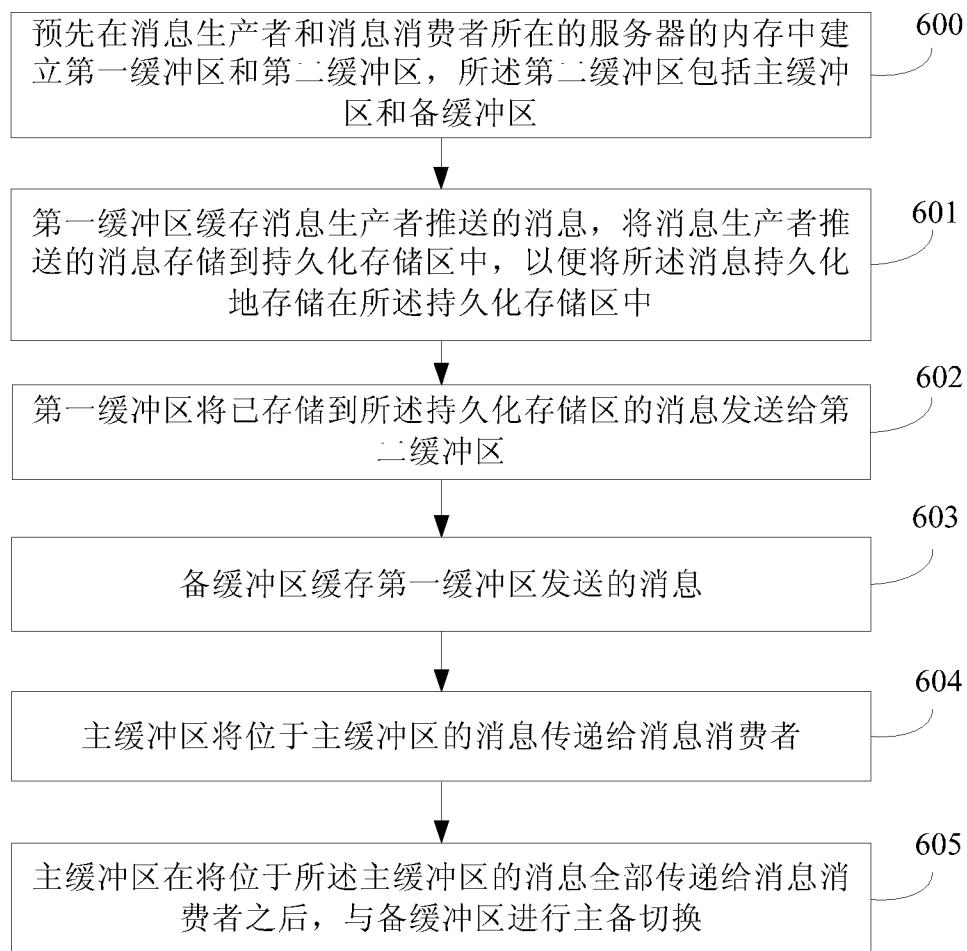


图 6

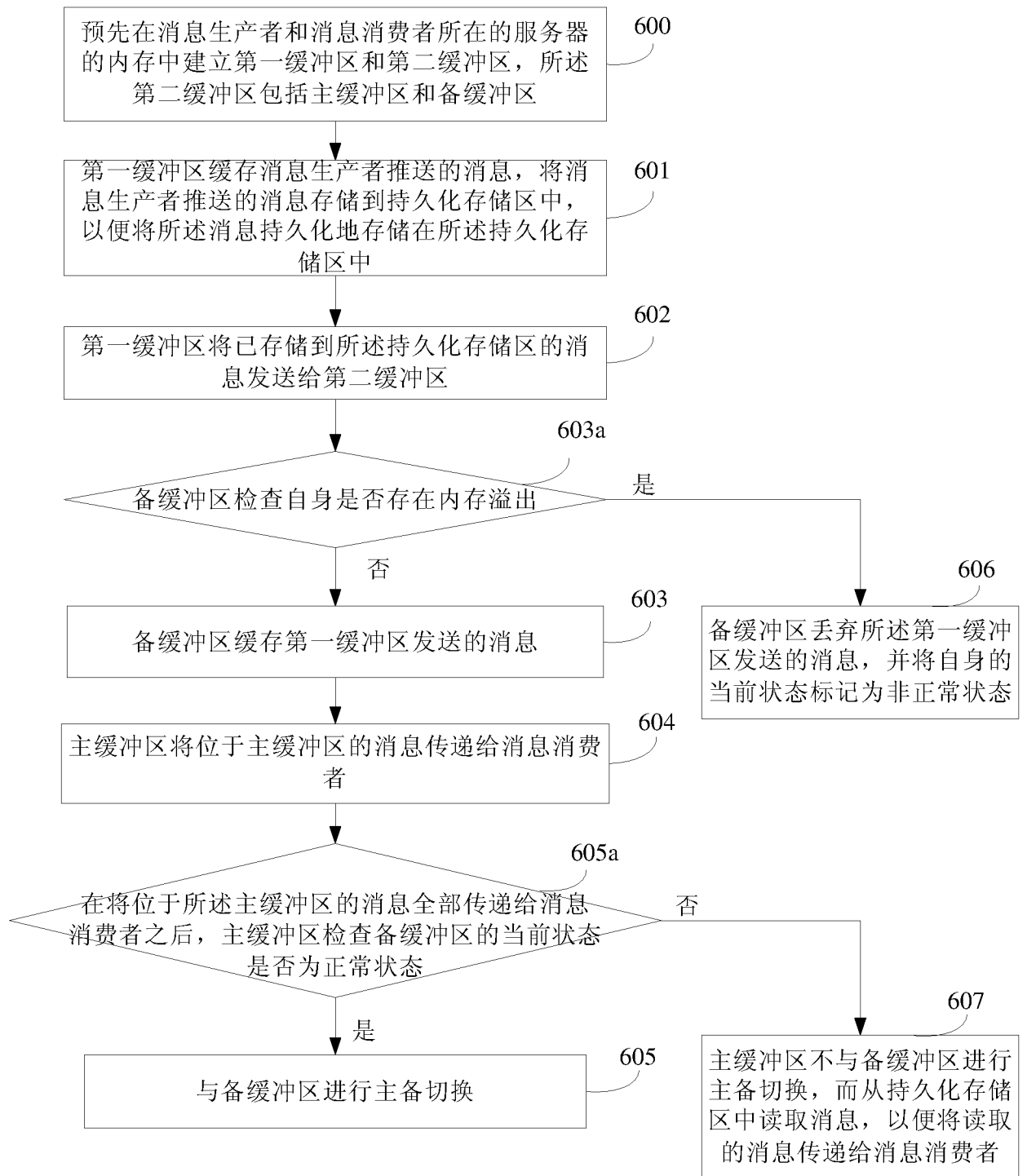


图 7

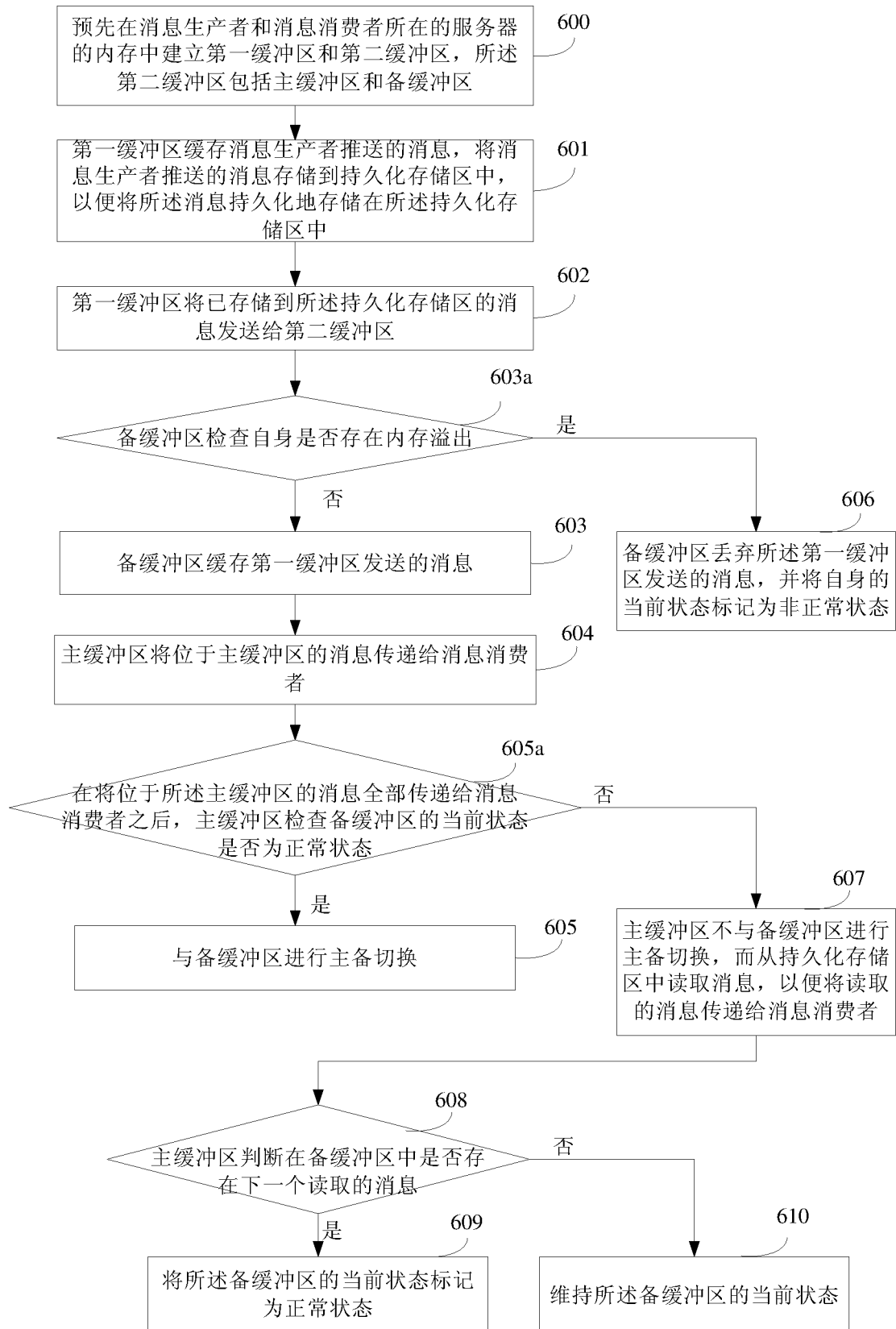


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/089308

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/54 (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)

G06F 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)

WPI, EPODOC, CNPAT, CNKI, IEEE: buffer, replace, message, queue, send, receive, cache, store, memory, switch, realtime

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101853238 A (HUAWEI TECHNOLOGIES CO., LTD.), 06 October 2010 (06.10.2010), description, paragraphs [0026]-[0027] and [0045]-[0048]	1-12
A	CN 1585378 A (ZTE CORP.), 23 February 2005 (23.02.2005), the whole document	1-12
A	US 7277990 B2 (JAIN, S. et al.), 02 October 2007 (02.10.2007), the whole document	1-12
A	CN 103064749 A (SHANGHAI FEIXUN COMMUNICATION CO., LTD.), 24 April 2013 (24.04.2013), the whole document	1-12
A	CN 103390049 A (LINKAGE TECHNOLOGY CO., LTD.), 13 November 2013 (13.11.2013), the whole document	1-12
A	CN 103078921 A (TCL CORPORATION), 01 May 2013 (01.05.2013), the whole document	1-12

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

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 27 September 2016 (27.09.2016)	Date of mailing of the international search report 10 October 2016 (10.10.2016)
Name and mailing address of the ISA/CN: 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 REN, Xingchao Telephone No.: (86-10) 62413667

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2016/089308

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101853238 A	06 October 2010	None	
CN 1585378 A	23 February 2005	None	
US 7277990 B2	02 October 2007	None	
CN 103064749 A	24 April 2013	None	
CN 103390049 A	13 November 2013	None	
CN 103078921 A	01 May 2013	None	

国际检索报告

国际申请号

PCT/CN2016/089308

A. 主题的分类

G06F 9/54 (2006.01) 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 H04L

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))

WPI, EPDOC, CNPAT, CNKI, IEEE: 发送, 接收, 消息, 队列, 缓冲, 缓存, 存储, 储存, 替换, 更换, 实时, message, queue, send, receive, cache, store, memory, switch, realtime

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	CN 101853238 A (华为技术有限公司) 2010年 10月 6日 (2010 - 10 - 06) 说明书第[0026]-[0027]段, 第[0045]-[0048]段	1-12
A	CN 1585378 A (中兴通讯股份有限公司) 2005年 2月 23日 (2005 - 02 - 23) 全文	1-12
A	US 7277990 B2 (JAIN, SANJEEV 等) 2007年 10月 2日 (2007 - 10 - 02) 全文	1-12
A	CN 103064749 A (上海斐讯数据通信技术有限公司) 2013年 4月 24日 (2013 - 04 - 24) 全文	1-12
A	CN 103390049 A (南京联创科技集团股份有限公司) 2013年 11月 13日 (2013 - 11 - 13) 全文	1-12
A	CN 103078921 A (TCL集团股份有限公司) 2013年 5月 1日 (2013 - 05 - 01) 全文	1-12

☐ 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

2016年 9月 27日

国际检索报告邮寄日期

2016年 10月 10日

ISA/CN的名称和邮寄地址

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

传真号 (86-10) 62019451

受权官员

任兴超

电话号码 (86-10) 62413667

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/089308

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	101853238	A	2010年 10月 6日	无	
CN	1585378	A	2005年 2月 23日	无	
US	7277990	B2	2007年 10月 2日	无	
CN	103064749	A	2013年 4月 24日	无	
CN	103390049	A	2013年 11月 13日	无	
CN	103078921	A	2013年 5月 1日	无	