

(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
2014年8月14日 (14.08.2014)

(10) 国际公布号
WO 2014/121588 A1

- (51) 国际专利分类号:
H04L 29/08 (2006.01) G06F 12/08 (2006.01)
G06F 12/02 (2006.01)
- (21) 国际申请号: PCT/CN2013/079841
- (22) 国际申请日: 2013年7月23日 (23.07.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201310046412.4 2013年2月6日 (06.02.2013) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 莫良伟 (MO, Liangwei); 中国广东省深圳市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,

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, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, 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))。

(54) Title: DATA ACCESS SYSTEM, MEMORY SHARING DEVICE, AND DATA ACCESS METHOD

(54) 发明名称: 数据存取系统、内存共享设备及数据读取方法

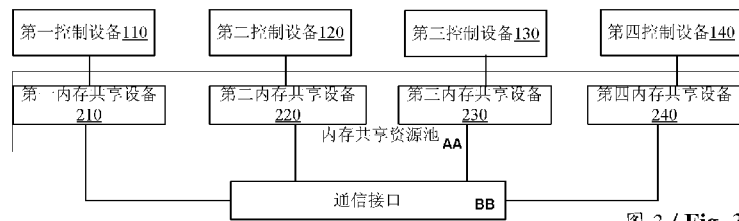
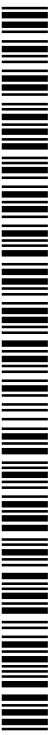


图3 / Fig. 3

- 110 FIRST CONTROL DEVICE
- 120 SECOND CONTROL DEVICE
- 130 THIRD CONTROL DEVICE
- 140 FOURTH CONTROL DEVICE
- 210 FIRST MEMORY SHARING DEVICE
- 220 SECOND MEMORY SHARING DEVICE
- 230 THIRD MEMORY SHARING DEVICE
- 240 FOURTH MEMORY SHARING DEVICE
- AA MEMORY SHARING RESOURCE POOL
- BB COMMUNICATIONS INTERFACE

(57) Abstract: The present invention relates to a data access system, comprising: a memory sharing resource pool formed by more than two memory sharing devices, and more than two control devices corresponding to all the memory sharing devices. Any memory sharing device in the more than two memory sharing devices has a uniformly addressed memory unit used for data storage, any memory sharing device obtains, under the control of a control device in the more than two control devices and corresponding to the memory sharing device, data in a memory unit in any memory sharing device in the memory sharing resource pool, and any control device in the more than two control devices is connected to the more than two memory sharing devices by using a communications interface, and obtains, by using uniformly addressed address information and a memory sharing device in the memory sharing resource pool and corresponding to any control device, data in a memory unit in any one or more memory sharing devices in the memory sharing resource pool.

(57) 摘要: 本发明涉及一种数据存取系统, 其包括: 由两个以上内存共享设备组成的内存共享资源池, 以及与每个内存共享设备对应的两个以上控制设备; 两个以上内存共享设备中的任一内存共享设备, 具有统一编址的用于存储数据的内存单元, 任一内存共享设备在两个以上控制设备中的与该内存共享设备对应的一个控制设备的控制下获取内存共享资源池中任一内存共享设备中的内存单元中的数据; 两个以上控制设备中的任一控制设备, 通过通信接口与两个以上内存共享设备连接, 并通过统一编址后的地址信息, 通过内存资源池中与任一控制设备对应的内存共享设备获取内存共享资源池中的任意的至少一个内存共享设备中的内存单元中的数据。



WO 2014/121588 A1

数据存取系统、内存共享设备及数据读取方法

技术领域

本发明涉及互联网领域，具体涉及到一种数据存取系统、内存共享设备
5 及数据读取方法。

背景技术

目前，随着互联网的发展与普及，分布式系统的应用越来越广泛。分布式文件系统（Distributed file system, DFS）是指建立在网络之上的软件系统。
10 其特点是具有高度的内聚性和透明性。内聚性是指系统中每一个数据库分布节点高度自治；透明性是指每一个数据库分布节点对用户的应用来说都是透明的。但是，分布式系统之间的信息共享的实时性影响了分布式系统的整体性能。

在现有技术中，为了解决分布式系统的信息共享问题，通常采用如图 1
15 所示的基于交换网络的信息共享方式。在图 1 中，作为获取方控制系统 A 中的 CPU 希望获取作为提供方控制系统 B 内存单元中的数据，过程为：1）控制系统 A 中的 CPU 向控制系统 B 的 CPU 发送请求消息；2）控制系统 B 中的 CPU 根据请求消息从本机内存单元中读取数据；3）控制系统 B 中的 CPU 将携带数据的执行结果返回给控制系统 A 的 CPU；4）控制系统 A 中的 CPU
20 从执行结果中提取数据，并交给上层系统进行处理。在上述过程中，所有数据的流动都需要经过控制系统 A 和控制系统 B 的中央处理器（Central Processing Unit, CPU）参与处理。

以现有的基于交换网络的信息共享方式为例，如图 2 中所示流程，控制系统 A 中 CPU 希望获取控制系统 B 内存单元中的数据，则控制系统 A 作
25 为获取方，控制系统 B 作为提供方，具体过程为，如图 2 中实线流程，控制系统 A 中 CPU 在执行某一程序时，需要获取控制系统 B 内存中的数据，

控制系统 A 中 CPU 通过交叉开关网络向控制系统 B 中 CPU 发送请求消息；
控制系统 B 中 CPU 对请求消息解析后，明确控制系统 A 中 CPU 需要获取自身内存中的数据，则控制系统 B 中 CPU 根据请求消息访问本机内存中的内存单元，并从中获取数据，控制系统 B 中 CPU 在从内存单元中获取数据后，
5 要将获取的数据反馈至控制系统 A 中 CPU 内，如图 2 中虚线流程；控制系统 B 中 CPU 将携带数据的执行结果通过交叉开关网络返回给控制系统中 A 中 CPU；控制系统 A 中 CPU 从执行结果中获取数据，并交给上层系统进行处理。因此，现有技术中的基于交换网络的信息共享方式，控制系统 A 在获取其他控制系统中的数据均需要经过控制系统 A 和其他控制系统 B 中 CPU
10 参与处理，尤其是在多个控制系统相互获取其他控制系统中的数据时，多个控制系统的 CPU 均会参与处理，使得在整个信息共享的过程中，过多消耗作为提供方控制系统中 CPU 的性能，造成 CPU 的浪费，再者，在任一作为提供方控制系统中 CPU 运行软件故障时，其所拥有的数据将不能被其它控制系统访问，降低整个系统的可靠性。

15 因此，在现有技术中存在以下缺点：1) 控制系统 A 与控制系统 B 的 CPU 均需要参与到信息的共享控制，使得在整个信息共享的过程中，过多消耗作为提供方控制系统中 CPU 的性能，造成 CPU 的浪费；2) 作为提供方控制系统 B 中的 CPU 的运行软件故障后，其所拥有的数据将不能被其它终端访问，降低整个系统的可靠性。

20

发明内容

本发明的目的是提供一种数据存取系统，以解决现有技术作为提供方的控制系统过多消耗 CPU 的性能，造成 CPU 的浪费及降低整个系统的可靠性的问题，提供了一种信息处理的方法、装置与系统。

25 第一方面，本发明实施例提供了一种数据存取系统，所述系统包括：由两个以上内存共享设备组成的内存共享资源池，以及两个以上与所述内存资

源池中的每个内存共享设备对应的控制设备;

所述任一内存共享设备具有统一编址的用于存储数据的内存单元, 所述任一内存共享设备中的第一内存共享设备在接收到与第一内存共享设备对应的第一控制设备发送的带有地址信息的内存访问请求时:

- 5 如果所述地址信息对应的内存单元为所述第一内存共享设备中的内存单元, 则在第一内存共享设备中的内存单元中读取数据, 并将所述数据反馈给所述第一控制设备, 如果所述地址信息对应的内存单元为所述内存共享资源池中的第二内存共享设备中的内存单元, 则第一内存共享设备将所述内存访问请求向所述第二内存共享设备转发, 并接收第二内存共享设备反
- 10 馈的所述第二内存共享设备的内存单元中的数据;

所述任一内存共享设备中的第一内存共享设备在接收到第二内存共享设备转发的带有地址信息的内存访问请求时, 读取数据并将读取到的数据反馈给所述第二内存共享设备。

- 基于第一方面, 在第一种可能的实施方式中, 所述内存共享设备包括
- 15 存储模块、处理模块以及通信单元:

存储模块, 与所述处理模块连接, 所述存储模块包括至少一个在所述内存共享资源池中统一编址的内存单元, 所述内存单元用于存储数据;

- 处理模块, 通过通信接口与一个控制设备相连, 通过所述通信单元与
- 20 所述内存共享资源池中的其他内存共享设备连接, 通过所述通信接口接收与该处理模块连接的所述控制设备发送的内存访问请求; 通过所述通信单元接收其他内存共享设备转发的内存访问请求, 和/或向所述内存共享资源池中的其他内存共享设备转发所述控制设备发送的内存访问请求; 所述其他内存共享设备转发的内存访问请求为该其他内存共享设备从对应的控制设备处接收到的, 所述内存访问请求中包括所述内存共享资源池中通过统
- 25 一编址后多个内存共享设备中的内存单元的地址信息, 所述地址信息用于获取所述内存共享资源池中的任意一个内存共享设备中的内存单元中的数

据。

基于第一方面的第一种可能的实施方式，在第一方面的第二种可能的实施方式中，所述内存共享设备通过所述通信单元向所述内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中统一编址后的内存单元的编址数据，并将获取到的所述编址数据保存在所述处理模块中。

基于第一方面，在第一方面的第三种可能的实施方式中，所述两个控制设备中的任一控制设备包括通信模块，所述任一控制设备的所述通信模块通过连接一可插拔的内存共享模块接入所述数据存取系统。

10 基于第一方面，或第一方面的第一种至第三种可能的实施方式，在第一方面的第四种可能的实施方式中，所述通信接口为系统高速总线、交换机或以太网接口。

第二方面，本发明实施例提供了一种内存共享设备，其包括存储模块、处理模块以及通信单元，其中，

15 存储模块，与处理模块连接，用于存储数据；

存储模块，与处理模块连接，用于存储数据；

处理模块，通过通信接口与一个控制设备相连，通过所述通信单元与所述内存共享设备所在的内存共享资源池中的其他内存共享设备连接，通过所述通信接口接收与该处理模块连接的控制设备发送的内存访问请求；
20 通过所述通信单元接收所述内存共享资源池中其他内存共享设备转发的内存访问请求，和/或向所述内存共享资源池中的其他内存共享设备转发所述控制设备发送的内存访问请求；所述其他内存共享设备转发的内存访问请求为该其他内存共享设备从对应的控制设备处接收到的，所述内存访问请求中包括所述内存共享资源池中通过统一编址后多个内存共享设备中的内存单元的地址信息，用于获取所述内存共享资源池中的任意一个内存共享设备中的内存单元中的数据。
25

基于第二方面，在第二方面的第一种可能的实施方式中，所述处理模块通过所述通信单元接收到所述内存共享资源池中的所述其他内存共享设备转发的内存访问请求后，在所述内存访问请求中携带的地址信息对应的内存单元中获取数据，并将结果通过所述通信单元返回给所述其他内存共享设备，由所述其他内存共享设备通过所述通信接口向与所述其他内存共享设备发送内存访问请求的控制设备返回所述结果。

基于第二方面，在第二方面的第二种可能的实施方式中，所述处理模块进一步包括判断单元，如果所述判断单元判定所述地址信息对应的内存单元中的第一部分内存单元在本端内存共享设备，第二部分在第二内存共享设备，则所述处理模块在本端内存共享设备中的所述存储模块中读取第一部分内存单元中的数据，并将所述内存访问请求中的地址信息中第一部分内存单元的地址删除，通过所述通信单元向所述其他内存共享设备转发只包含第二部分内存单元地址的内存访问请求。

基于第二方面，或第二方面的第一种或第二种可能的实施方式，在第二方面的第三种可能的实施方式中，所述内存共享设备还包括电源模块，用于当不存在外部电源供电时，为所述存储提供电能，藉由该电源模块所述内存共享设备形成可插拔的非易失性随机访问存储器 NVRAM。

基于第二方面，或第二方面的第一种至第三种可能的实施方式，在第二方面的第四种可能的实施方式中，所述内存共享设备，还包括管理模块，用于调试所述存储模块中的应用程序及对所述通信单元中的流量进行控制。

基于第二方面，或第二方面的第一种至第四种可能的实施方式，在第二方面的第五种可能的实施方式中，所述处理模块为现场可编程门阵列 FPGA。

第三方面，本发明实施例提供了一种数据读取方法，所述方法包括：
内存共享资源池中的第一内存共享设备接收控制该第一内存共享设备

的第一控制设备发送的内存访问请求，所述内存访问请求中包括通过统一编址后的地址信息；其中，所述的地址信息是通过统一编址后的内存共享资源池中的两个以上内存共享设备中的用于存储数据的一个或多个内存单元的地址信息，所述两个以上的内存共享设备中的任意一个内存共享设备在两个以上控制设备中的一个控制设备的控制下进行数据的读取，所述第一内存共享设备为所述两个以上内存共享设备中的一个；

第一内存共享设备根据所述地址信息，确定所述地址信息对应的内存单元所在的内存共享设备；

如果所述地址信息对应的内存单元所在的内存共享设备为所述第一内存共享设备，则所述第一内存共享设备从第一内存共享设备自身中所述地址信息对应的内存单元中获取数据，

如果所述地址信息对应的内存单元所在的内存共享设备为第二内存共享设备，则所述第一内存共享设备向所述第二内存共享设备转发所述内存访问请求，并接收所述第二内存共享设备返回的数据；

所述第一内存共享设备将所述数据读取反馈给所述第一控制设备。

基于第三方面，在第一种可能的实施方式中，在内存共享资源池中的第一内存共享设备接收控制该第一内存共享设备的第一控制设备发送的内存访问请求之前，还包括：

内存共享资源池中的第一内存共享设备向所述内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中全部内存共享设备中内存单元的编址数据。

通过应用本发明实施例提供的数据存取系统，两个以上内存共享设备组成的内存共享资源池，以及两个以上与所述内存资源池中的每个内存共享设备对应的控制设备，任一内存共享设备具有统一编址的用于存储数据的内存单元，所述内存共享设备根据其对应的控制设备从所述内存单元中读取数据。当通过内存资源池中的其他内存设备读取数据，简化了现有技术

术中作为提供方控制设备的工作量；提高了作为提供方控制设备的利用率，同时，在内存共享资源池中存储了各控制系统所需的数据，各控制系统均可共享其他控制系统的数据库，提高了分布式系统存储数据的可靠性。

5 附图说明

为了更清楚地说明本发明实施例中的技术方案，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发明的一些实施例，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。

- 10 图 1 为现有技术提供的信息共享方式结构图；
图 2 为现有技术提供的信息共享方式流程示意图；
图 3 为本发明实施例提供的信息的数据存取系统示意图；
图 4 为本发明实施例提供的数据存取系统中内存共享设备的示意图；
图 5A 为本发明一实施方式内存共享设备工作过程示意图；
15 图 5B 为本发明另一实施方式内存共享设备工作过程示意图；
图 5C 为本发明另一实施方式内存共享设备工作过程示意图；
图 6 为本发明实施例数据存取系统的另一种实施方式的架构图；
图 7 为本发明实施例数据存取系统另一实施例的结构图；
图 8 为本发明实施例数据读取方法的流程图；
20 图 9 为本发明实施例数据读取方法另一实施例的流程图；
图 10 为本发明提供的内存共享设备的一种实施方式的结构示意图；
图 11 为本发明提供的内存共享设备另一种实施方式的结构示意图。

具体实施方式

- 25 下面通过附图和实施例，对本发明的技术方案做进一步的详细描述。
本发明实施例提供一种包括两个以上内存共享设备和通过通信接口与

上述的内存共享设备连接的两个以上控制设备的数据存取系统，其中，两个以上的内存共享设备组成所述数据存取系统中的内存共享资源池，每个控制设备与每个内存共享设备对应，当两个以上控制设备中的一个控制设备想要在内存共享资源池中读取数据时，通过与此控制设备对应的内存共享设备，在所述的内存共享资源池中读取所需的数据。

如图3所示，本发明实施例提供了一种数据存取系统，所述系统包括两个以上内存共享设备组成的内存共享资源池，以及通过通信接口与多个控制设备连接的多个内存共享设备。其中，通信接口包括但不限于系统高速总线、交换机或以太网接口等。所述的多个内存共享设备组成内存共享资源池，其中的每个内存共享设备都与多个控制设备中的一个对应，建立控制关系，如果是通过系统高速总线连接，则与每个内存共享设备连接的控制设备，就控制此内存共享设备，如果控制设备通过交换机或以太网接口通过网络连接内存共享设备，则内存共享设备和控制设备之间可以通过端口配置或者IP分配等方式建立对应关系，达到一个控制设备控制一个内存共享设备的目的，或者一个控制设备控制内存共享资源池中的任意一个内存共享设备的目的。

所述内存共享资源池中的任一内存共享设备具有统一编址的用于存储数据的内存单元，所述任一内存共享设备中的第一内存共享设备在接收到与第一内存共享设备对应的第一控制设备发送的带有地址信息的内存访问请求时：

如果所述地址信息对应的内存单元为所述第一内存共享设备中的内存单元，则在第一内存共享设备中的内存单元中读取数据，并将所述数据反馈给所述第一控制设备，如果所述地址信息对应的内存单元为所述内存共享资源池中的第二内存共享设备中的内存单元，则第一内存共享设备将所述内存访问请求向所述第二内存共享设备转发，并接收第二内存共享设备反馈的所述第二内存共享设备的内存单元中的数据；

任一内存共享设备中的第一内存共享设备在接收到第二内存共享设备转发的带有地址信息的内存访问请求时，在本地读取数据，并将读取到的数据反馈给所述第二内存共享设备。

如图 4 所示，每个内存共享设备都包括处理模块 501、存储模块 502 以及通信单元 503。通过所述的通信单元 503，所述内存共享设备与其他内存共享设备之间能够进行通信，这些内存共享设备同属于一个内存共享资源池；所述的存储模块 502 包括多个的内存单元，每个内存单元都在整个内存共享资源池中统一编址，每个内存单元都用于存储数据；所述的处理模块 501 用于对从控制设备中接收到的内存访问请求进行处理，并在存储模块 502 中读取数据。在每个内存共享设备的处理模块 501 可以保存有一个地址表，此地址表记录所述内存共享资源池中所有内存单元的地址。

内存共享设备中的处理模块 501 通过所述通信接口与一个控制设备相连，通过所述通信接口接收与该处理模块 501 连接的所述控制设备发送的内存访问请求；通过所述通信单元 503 与所述内存共享资源池中的其他内存共享设备连接，通过所述通信单元 503 接收其他内存共享设备转发的内存访问请求；此外，处理模块 501 还能够在接收到与该处理模块 501 连接的所述控制设备发送的内存访问请求后，当该内存访问请求所要访问的内存单元在其它内存共享设备时，通过所述通信单元 503 向所述要访问的内存单元所在的内存共享设备转发请求。所述内存访问请求中包括地址信息，处理模块 501 根据地址信息确定控制设备想要访问的内存单元所在的内存共享设备，并获取所述内存共享资源池中的任意一个内存共享设备中的数据，这些数据位于内存共享设备中的存储模块 502 中的内存单元中。

本发明实施例中虚拟的内存共享资源池可以通过如下方式实现：各内存共享设备通过通信单元组成一个数据交换域，将所述交换域中包含的内存单元进行统一地址编址，并将编址后的地址信息发送给所述交换域中的内存共享设备。

如图 3 所示, 作为获取方的第一控制设备 110 需要从内存共享资源池中获取所需的数据时, 向与该第一控制设备 110 对应的第一内存共享设备 210 发送内存访问请求, 所述内存访问请求携带第一控制设备 110 想要获取的数据的地址信息, 第一内存共享设备 210 设备根据所述地址信息和处理
5 模块 501 中的地址表进行查询, 以确定第一控制设备 110 所需数据的存储位置。当第一控制设备 110 所需的数据在第一内存共享设备 210 的存储模块 502 中的内存单元时, 则直接访问所述存储模块 502 中的内存单元中的数据, 并反馈给所述第一控制设备 110; 当所述第一控制设备 110 所需的数据不在第一内存共享设备 210 的存储模块 502 内时, 则通过所述第一内存
10 共享设备 210 中的通信单元 503, 将所述访问请求发送到内存访问请求中的地址信息对应的内存共享设备 (例如第二内存共享设备 220), 以获取第一控制设备所需的数据。在第一控制设备 110 所需的数据不在第一内存共享设备 210 内时, 整个获取数据的过程也不需要第二控制设备 120 参与, 不需要耗费第二控制设备 120 中的处理模块 501 中的资源。在本发明实施例中, 具体的访问、获取内存单元中数据的过程由内存共享设备完成, 进而
15 降低了现有技术中作为提供方控制设备中处理器的工作量; 提高了作为提供方控制设备中处理器利用率, 同时, 在内存共享资源池中存储了各控制设备所需的数据, 各控制设备均可共享其他控制设备的数据, 进而解决分布式系统信息共享的问题, 提高分布式系统之间信息共享的实时性。

20 以下结合附图, 对上述系统的几个具体实施例, 对上述系统做进一步详细叙述。

如图 5A 所示, 图 5A 为本发明实施例提供的内存共享设备中内存单元地址编址信息示意图。在图 5A 中, 以 4 个内存共享设备形成内存共享资源池为例进行说明, 实际应用中, 并不限制于 4 个内存共享设备形成内存共
25 享资源池。4 个内存共享设备分别与控制设备通过系统高速总线连接, 且每个内存共享设备中的处理模块 501 均完成自身的存储模块 502 中内存单元

的编址，明确自身的存储空间，并且第一内存共享设备 210、第二内存共享设备 220、第三内存共享设备 230 和第四内存共享设备 240 形成内存共享资源池，每个内存共享设备的地址编址信息如图 5A 所示，第一内存共享设备 210 的第一地址编址信息为 0-99；第二内存共享设备 220 的第二地址编址信息为 100-199；第三内存共享设备 230 的第三地址编址信息为 200-299；第四内存共享设备 240 的第四地址编址信息为 300-399。

当内存共享资源池形成后，各内存共享设备中的处理模块明确自身存储模块 502 中内存单元的地址编址信息，并利用通信单元 503 向内存共享资源池中的其他 3 个内存共享设备发送询问消息，以获取其他内存共享设备其内部的内存单元的地址编址信息。

每个内存共享设备中的处理模块 501 根据自身内存单元的地址编址信息，以及其他内存共享设备其内部的内存单元的地址编址信息，生成地址信息表，所述地址信息表用于内存共享设备快速查找内存单元的位置。

在第一个优选的实施例中，如图 5B 所示，第一控制设备 110 希望访问偏移 58，长度为 10 的内存单元中的数据，第一控制设备 110 通过系统高速总线向第一内存共享设备 210 的发送内存访问请求，所述内存访问请求包括内存地址信息，第一内存共享设备 210 的处理模块 502 接收到内存访问请求后，对内存访问请求进行解析，并提取内存访问请求中的地址信息，明确第一控制设备 110 希望访问偏移 58，长度为 10 的内存单元的数据，也即是地址信息为 58 至 67 的内存单元的数据；处理模块 501 根据地址信息表，查找并确定 58 至 67 的内存单元的位置，处理模块 501 根据地址信息表确定编址信息为 58 至 67 的内存单元存在于本地的存储模块 502 中，则处理模块 501 访问编址信息为 58 至 67 的内存单元，在完成访问后，生成结果消息，所述结果消息携带 58 至 67 内存单元中存储的数据。处理模块 501 通过系统高速总线向第一控制设备 110 发送获取到的数据。

在第二个优选的实施例中，如图 5C 所示，第二控制设备 120 希望访问

偏移 58, 长度为 10 的内存单元中的数据, 第二控制设备 120 通过系统高速总线向第二内存共享设备 220 的处理模块 501 发送内存访问请求, 所述内存访问请求包括内存地址信息, 第二内存共享设备 220 的处理模块 501 接收到内存访问请求后, 对内存访问请求进行解析, 并提取内存访问请求中的信息, 确定第二控制设备 120 希望访问偏移 58, 长度为 10 的内存单元的数据, 即地址信息为 58 至 67 的内存单元的数据; 第二内存共享设备 220 的处理模块 501 根据地址信息表, 查找并确定 58 至 67 的内存单元的位置, 第二内存共享设备 220 的处理模块 501 根据地址信息表确定编址信息为 58 至 67 的内存单元未存在于本地的第二内存共享设备 220 的存储模块 502 中, 而存在于第一内存共享设备 210 的存储模块 502 中, 则第二内存共享设备 220 的处理模块 501 将该内存访问请求通过通信单元将该内存访问请求向第一内存共享设备 210 转发。第一内存共享设备 210 接收到内存访问请求后, 将该消息传输至第一内存共享设备 210 的处理模块 501 中, 由第一内存共享设备 210 的处理模块 501 进行相应的处理, 第一内存共享设备 210 的处理模块 501 根据地址信息表, 查找并确定内存访问请求包括的内存地址信息存在于本地的存储模块 502 后, 则访问编址信息为 58 至 67 的内存单元, 在完成访问后, 生成结果消息。第一内存共享设备 210 的处理模块 501 通过通信单元 503 将结果消息传输至第二内存共享设备的处理模块 501, 再由第二内存共享设备 220 的处理模块 501 将结果消息通过系统高速总线向第二控制设备 120 发送, 所述结果消息携带 58 至 67 内存单元中存储的数据。在该实施例中, 当第二控制设备 120 需要从内存共享资源池中获取数据时, 即使第二控制设备 120 所需的数据不在第二控制设备 120 控制的第二内存共享设备 220 中, 也只需要第二内存共享设备 220 转发第二内存共享设备 220 收到的内存访问请求, 从而从其他的内存共享设备中获取第二控制设备 120 所需的数据, 而不需要其他的控制设备参与, 从而节省了其他控制设备的损耗。

在第三个优选的实施例中，如图 5C 所示，第二控制设备 120 希望访问偏移 58，长度为 100 的内存单元中的数据，第二控制设备 120 通过系统高速总线向第二内存共享设备 220 发送内存访问请求，所述内存访问请求包括内存地址信息，第二内存共享设备 220 的处理模块 501 对内存访问请求进行解析，并提取内存访问请求中的信息，确定第二控制设备 120 希望访问偏移 58，长度为 100 的内存单元的数据，即，地址信息为 58 至 157 的内存单元的数据；第二内存共享设备 220 的处理模块 501 根据地址信息表，查找并确定 58 至 157 的内存单元的位置，第二内存共享设备 220 的处理模块 501 根据地址信息表确定编址信息为 58 至 157 的内存单元未全部存在于本地的第二内存共享设备 220 的存储模块 502 中，而是部分存在于第一内存共享设备 210 的存储模块 501 中（58-99），部分存在与第二内存共享设备 220 的存储模块 502 中（100-157），则第二内存共享设备 220 的处理模块 501 将该内存访问请求转发给第一内存共享设备，所述内存访问请求包括内存地址信息。第二内存共享设备 220 的处理模块明确编址信息为 100 至 157 的内存单元存在于本地的第二内存共享设备 220 的存储模块 502 中，则第二内存共享设备 220 的处理模块 501 访问编址信息为 100 至 157 的内存单元，在完成访问后，生成第一结果消息，所述第一结果消息携带 100 至 157 内存单元中存储的数据。

第一内存共享设备 210 接收到第二内存共享设备 220 转发的内存访问请求后，由第一内存共享设备 210 的处理模块 501 进行相应的处理，查找并确定内存访问请求包括的内存地址信息存在于本地的第一内存共享设备 210 的存储模块 502 后，则访问编址信息为 58 至 99 的内存单元，在完成访问后，生成第二结果消息，所述第二结果消息携带 58 至 99 内存单元中存储的数据。第一内存共享设备 210 的处理模块 501 将第二结果消息传输至第二内存共享设备 220 的处理模块 501。

由于第二内存共享设备 220 的处理模块 501 将自身生成的携带第二控

制设备 220 所需数据的第一结果消息和接收的第二结果消息通过系统高速总线向第二控制设备 120 发送。通过第三实施例，当第二控制设备 120 需
要从内存共享资源池中获取数据时，即使第二控制设备 120 所需的数据并
不是只存储在第二控制设备 120 控制的第二内存共享设备 220 中，而是分
散存储的情况下，第二内存共享设备 220 除了在本地读取存储在第二内存
共享设备 220 中的部分数据外，也只需要第二内存共享设备 220 将第二内
存共享设备 220 收到的内存访问请求进行重新处理后，向其他内存共享设
备转发只携带剩余数据的存储地址的内存访问请求，从而从其他的内存共
享设备中获取第二控制设备 120 所需的数据，而不需要其他的控制设备参
与，从而节省了其他控制设备的损耗。

上述多个实施例详细描述了控制设备与内存共享设备之间的相互通信
及内存共享设备之间的相互通信，在实际应用中，不限于上述实施例所描
述的两个内存共享设备的通信，通过上述多个实施例，可以简单、方便地
实现作为获取方的控制设备共享内存共享设备的数据。

在上述实施例的一种应用方式中，如图 6 所示，每个控制设备或称为
CPU 子系统与其对应的内存共享设备构成一个分布式节点，对应于硬件中，
每个分布式节点可以是一个机架，每个机架服务器就是一个控制设备，插
上一个内存共享设备后，多个机架上通过交换机互联的内存共享设备，就
构成了一个内存共享资源池，每个机架服务器都能够访问内存共享资源池
中的任意一个内存共享设备中的数据。

在另外一种可能的实施方式中，如图 7 所示，将每个控制设备和内存
共享资源池中的内存共享设备采用通信接口连接，所述通信接口例如是交
换机、以太网接口等。如果未接入所述数据存取系统的控制设备，需要接
入数据存取系统，只需要在此控制设备上通过插槽、排线连接一个具有匹
配的插槽、排线接口的内存共享设备，之后，内存共享设备重新广播询问
消息，获取内存共享资源池中全部内存共享设备中的内存单元的编址，即

可接入整个数据存取系统，不多赘述，通过这种可以扩展的内存共享资源池，可以达到 PB 级以上的内存空间。

在本发明实施例中，控制设备在获取内存共享资源池中的内存共享设备中内存单元中数据的过程由内存共享设备完成，进而替代了现有技术中提供方控制设备中的 CPU，简化了现有技术中作为提供方控制设备中 CPU 的工作量；提高了作为提供方控制设备中 CPU 的利用率，同时，在内存共享资源池中存储了各控制设备所需的数据，各控制设备均可共享其他控制设备的数据，进而解决分布式系统信息共享的问题，提高分布式系统之间信息共享的实时性。

相应的，本发明实施例提供一种数据读取方法，所述数据读取方法的执行主体为前述实施例中的内存共享设备，由图 8 可见，所述方法包括：

801，内存共享资源池中的第一内存共享设备接收控制该第一内存共享设备的第一控制设备发送的内存访问请求，所述内存访问请求中包括通过统一编址后的地址信息通过统一编址后的地址信息；

其中，所述的地址信息是通过统一编址后的多个内存单元的地址信息，所述多个内存单元用于存储数据，分别位于内存共享资源池中的两个以上内存共享设备中。所述两个以上的内存共享设备中的任意一个内存共享设备，在两个以上控制设备中的一个控制设备的控制下进行数据的读取，所述第一内存共享设备为所述两个以上内存共享设备中的一个。

内存共享资源池中的第一内存共享设备，在接收第一控制设备发送的内存访问请求之前，还可以包括：

第一内存共享设备向内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中全部内存共享设备中内存单元的编址数据，之后第一内存共享设备可以采用地址信息表的形式保存所述编址数据，方便第一内存共享设备查找地址信息对应的内存单元。

802，第一内存共享设备根据所述地址信息，确定所述内存共享资源池

中与地址信息对应的内存空间所在的内存共享设备;

具体而言, 第一内存共享设备根据所述地址信息, 可以在自身的处理模块中的地址信息表中, 通过查找的方式确定所述地址信息对应的内存单元所在的具体位置。

5 如果所述地址信息对应的内存单元所在的内存共享设备为所述第一内存共享设备, 则第一内存共享设备在第一内存共享设备中的内存单元中读取数据;

如果所述地址信息对应的内存单元所在的内存共享设备为内存共享资源池中的其他内存共享设备, 例如第二内存共享设备, 则第一内存共享设备可以向所述第二内存共享设备转发所述内存访问请求, 之后, 等待所述
10 第二内存共享设备根据所述内存访问请求中的地址信息, 在第二内存共享设备的内存单元中进行数据读取后, 反馈给第一内存共享设备的数据读取结果。

在根据所述地址信息, 确定所述地址信息对应的内存单元所在的内存
15 共享设备之后,

如果所述地址信息对应的内存单元中, 第一部分内存单元在第一内存共享设备, 第二部分在第二内存共享设备, 则, 第一内存共享设备在第一内存共享设备中的内存单元中读取第一部分内存单元中的数据;

将所述内存访问请求中的地址信息中第一部分内存单元的地址删除,
20 向第二内存共享设备转发只包含第二部分内存单元地址的内存访问请求;

之后, 接收所述第二内存共享设备返回的数据读取结果。

803, 第一内存共享设备将所述数据读取结果反馈给所述第一控制设备。

具体而言, 第一内存共享设备在第一内存共享设备自身的内存单元中
25 读取到第一控制设备所需的数据, 或者, 在接收到内存共享资源池中的第二内存共享设备反馈的第一内存共享设备所需的数据, 或者, 将本地读取

到的数据和从第二内存共享设备中读取到的数据，通过所述通信接口向第一控制设备发送数据读取结果。

通过上述的方法，在数据存取系统中的某个控制设备需要在内存共享资源池中读取数据时，即使所需的数据不在该控制设备所控制的内存共享设备上，也不需要其他控制设备参与，从而降低了对数据存取系统中控制设备的 CPU 消耗。

相应的，从控制设备的角度，数据读取的过程如图 9 所示，所述方法包括：

901，第一控制设备向内存共享资源池中的第一内存共享设备发送内存访问请求，所述内存访问请求所述控制设备需要访问的内存单元的地址信息，

具体而言，所述地址信息为对整个内存共享资源池中的全部内存共享设备中的内存单元通过统一编址的地址信息，内存资源池中的任意一个内存共享设备在两个以上控制设备中的一个控制设备的控制下进行数据的读取，所述第一内存共享设备为所述两个以上内存共享设备中的一个；

902，第一控制设备接收所述第一内存共享设备反馈的数据读取结果。

具体而言，如果所述地址信息对应的内存空间所在的内存共享设备为所述第一内存共享设备，则，所述第一内存共享设备反馈的数据读取结果为所述第一内存共享设备在所述第一内存共享设备的内存单元中进行数据读取后生成的。

如果，所述地址信息对应的内存单元所在的内存共享设备为第二内存共享设备，则，所述第一内存共享设备反馈的数据读取结果为所述第一内存共享设备接收的所述第二内存共享设备发送的数据读取结果。

如果所述地址信息对应的内存单元中的第一部分内存单元在第一内存共享设备，第二部分在第二内存共享设备，则，

所述第一内存共享设备反馈的数据读取结果为：第一内存共享设备中

的第一部分内存单元中的数据，以及所述第二内存共享设备中第二部分内存单元中的数据。

通过上述实施例提供的数据读取方法，对于内存共享资源池中任意一个内存共享设备中的数据，都能够通过一个控制设备就能够读取，而不需要其他控制设备中的 CPU 的参与。

相应的，本发明实施例提供一种内存共享设备，如图 10 所示，所述内存共享设备包括：存储模块 1101、处理模块 1102 以及通信单元 1103：

存储模块 1101，与处理模块 1102 连接，用于存储数据；

处理模块 1102，通过通信接口与一个控制设备相连，通过所述通信单元 1103 与所述内存共享设备所在的内存共享资源池中的其他内存共享设备连接，通过所述通信接口接收与该处理模块 1102 连接的控制设备发送的内存访问请求；通过所述通信单元 1103 接收所述内存共享资源池中其他内存共享设备转发的内存访问请求，和/或向所述内存共享资源池中的其他内存共享设备转发所述控制设备发送的内存访问请求；所述其他内存共享设备转发的内存访问请求为该其他内存共享设备从对应的控制设备处接收到的，所述内存访问请求中包括所述内存共享资源池中通过统一编址后多个内存共享设备中的内存单元的地址信息，用于获取所述内存共享资源池中的任意一个内存共享设备中的内存单元中的数据。

更具体的，与所述处理模块 1102 连接的存储模块 1101 包括多个内存单元，用于存储数据，所述内存单元在内存共享资源池中的多个内存共享设备中的统一编址。

优选地，处理模块 1102 可以通过现场可编程门阵列（Field - Programmable Gate Array, FPGA）或中央处理器芯片实现。

具体的，例如处理模块 1102 可以包括询问单元，该询问单元通过通信单元 1103 向内存共享资源池中的全部内存共享设备广播询问消息，获取内存共享资源池中全部内存共享设备中的内存单元编址信息，之后以地址信

息表的形式将所述编址信息保存。

所述处理模块 1102 从通信接口（例如系统高速总线、以太网接口或交换机）中接收到的内存访问请求，根据所述内存访问请求中的地址信息和处理模块 1102 保存的地址信息表，确定所述地址信息对应的内存空间所在
5 的内存共享设备。

如果所述地址信息对应的内存空间所在的内存共享设备为与本端的内存共享设备，则所述处理模块 1102，在所述存储模块 1101 中的所述地址信息对应的内存单元中读取数据。

如果所述地址信息对应的内存空间所在的内存共享设备为与所述通信
10 单元 1103 连接的其它内存共享设备，则所述处理模块 1102 通过所述通信单元将所述内存访问请求向所述其它内存共享设备转发。

如果所述地址信息对应的内存单元中的第一部分内存单元在本端内存共享设备，第二部分内存单元在其它内存共享设备，则所述处理模块 1102 在本端内存共享设备中的所述存储模块 1101 中读取第一部分内存单元中的
15 数据，并将所述内存访问请求中的地址信息中第一部分内存单元的地址删除，通过通信单元 1103 向第二内存共享设备转发只包含第二部分内存单元地址的内存访问请求。

内存共享设备首先与其他内存共享设备形成虚拟的内存共享资源池，所述虚拟的内存共享资源池的形成过程具体为：各内存共享设备中的处理
20 模块 1102 为自身的存储模块 1101 中的内存单元进行编址，明确自身的存储空间，将多个内存共享设备通过同一个通信接口连接在一起，则多个内存共享设备同属于一个交换域，为交换域中的多个内存共享设备中存储模块 1101 的内存单元进行统一地址编址划分，并将划分后的地址编址信息告知各内存共享设备中的处理模块 1102，进而形成虚拟的内存共享资源池。

25 由内存共享设备中的处理模块 1102 将数据写入存储模块 1101 的内存单元中，使得内存单元中存储着各控制系统所需的数据，在各控制设备写

入数据的过程中，各控制设备将该数据的存储位置一同告知内存共享设备中的处理模块 1102，使得内存共享设备中的处理模块 1102 根据自身内存单元的地址编址划分，准确地将数据写入内容单元。

5 当内存共享设备接收到作为获取方的一个第一控制设备需要从内存共享资源池中获取所需的数据的内存访问请求时，所述内存访问请求中所述访问请求包括所述第一控制设备想要获取的数据的地址信息；处理模块 1102 根据所述地址信息，确定内存共享资源池中与地址信息对应的内存单元所在的内存共享设备。

10 处理模块 1102 根据所述地址信息，可以由一个判断单元，在自身的地址信息表中通过查找的方式确定所述地址信息对应的内存单元所在的具体位置。

具体而言，如果所述地址信息对应的内存单元所在的内存共享设备为本端内存共享设备，则内存共享设备的处理模块 1102 在存储模块 1101 中的内存单元中读取数据；

15 如果所述地址信息对应的内存空间所在的内存共享设备为内存共享资源池中的其他内存共享设备，例如第二内存共享设备，则内存共享设备可以通过处理模块 1102 和通信单元 1103 向第二内存共享设备转发所述内存访问请求，之后，等待所述第二内存共享设备根据所述内存访问请求中的地址信息，在第二内存共享设备中的内存单元中进行数据读取后，反馈给
20 第一内存共享设备的数据读取结果。

内存共享设备在自身的存储模块中读取到第一控制设备所需的数据，或者在接收到内存共享资源池中的第二内存共享设备反馈的控制设备所需的数据后，通过通信接口向控制设备发送数据读取结果。

25 如图 11 所示，在一种较佳的实施方式中，所述内存共享设备还可以包括电源模块 1105，用于外部电源供电失效时，为所述内存共享设备提供电源，通过该电源模块 1105，所述内存共享设备可以形成非易失性随机访问

存储器(**Non-Volatile Random Access Memory, NVRAM**), 该电源模块 1105 可以是超级电容或电池备用单元 BBU (**Battery Backup Unit, BBU**)。

在一种较佳的实施方式中, 所述的内存共享设备还可以包括: 闪存模块 1104, 用于存储应用程序, 所述应用程序包括所述处理模块执行的指令。

5 在一种较佳的实施方式中, 所述的内存共享设备还可以包括管理模块 1106, 用于调试所述存储模块 1101 中的应用程序及对所述通信单元 1103 中的流量进行控制, 该管理模块 1106 可以通过 CPU IP 核实现。

通过应用本发明实施例提供的内存共享设备, 内存共享设备根据接收的内存访问请求, 查找并确定对应的内存单元的位置, 当本地的存储模块
10 中存在对应的内存单元时, 则直接访问内存单元, 从内存单元中获取数据, 并将获取的数据携带在结果消息中, 反馈至控制设备, 在本发明实施例中, 具体的访问、获取内存单元中数据的过程由第一内存共享设备完成, 进而替代了现有技术中提供方控制设备中 CPU, 简化了现有技术中作为提供方控制设备中 CPU 的工作量; 提高了作为提供方控制设备中 CPU 利用率。

15 专业人员应该还可以进一步意识到, 结合本文中所公开的实施例描述的各示例的单元及算法步骤, 能够以电子硬件、计算机软件或者二者的结合来实现, 为了清楚地说明硬件和软件的可互换性, 在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行, 取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能, 但是这种
20 实现不应认为超出本发明的范围。

结合本文中所公开的实施例描述的方法或算法的步骤可以用硬件、处理器执行的软件模块, 或者二者的结合来实施。软件模块可以置于随机存储器 (RAM)、内存、只读存储器 (ROM)、电可编程 ROM、电可擦除可编程 ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的
25 任意其它形式的存储介质中。

以上所述的具体实施方式，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施方式而已，并不用于限定本发明的保护范围，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

5

10

权利要求

1、一种数据存取系统，其特征在于，所述系统包括：由两个以上内存共享设备组成的内存共享资源池，以及两个以上与所述内存资源池中的每个内存共享设备对应的控制设备；

5 所述任一内存共享设备具有统一编址的用于存储数据的内存单元，所述任一内存共享设备中的第一内存共享设备在接收到与第一内存共享设备对应的第一控制设备发送的带有地址信息的内存访问请求时：

如果所述地址信息对应的内存单元为所述第一内存共享设备中的内存单元，则在第一内存共享设备中的内存单元中读取数据，并将所述数据反
10 馈给所述第一控制设备，如果所述地址信息对应的内存单元为所述内存共享资源池中的第二内存共享设备中的内存单元，则第一内存共享设备将所述内存访问请求向所述第二内存共享设备转发，并接收第二内存共享设备反馈的所述第二内存共享设备的内存单元中的数据；

所述任一内存共享设备中的第一内存共享设备在接收到第二内存共享
15 设备转发的带有地址信息的内存访问请求时，读取数据并将读取到的数据反馈给所述第二内存共享设备。

2、如权利要求1所述的数据存取系统，其特征在于，所述内存共享设备包括存储模块、处理模块以及通信单元：

存储模块，与所述处理模块连接，所述存储模块包括至少一个在所述
20 内存共享资源池中统一编址的内存单元，所述内存单元用于存储数据；

处理模块，通过通信接口与一个控制设备相连，通过所述通信单元与
所述内存共享资源池中的其他内存共享设备连接，通过所述通信接口接收
与该处理模块连接的所述控制设备发送的内存访问请求；通过所述通信单
元接收其他内存共享设备转发的内存访问请求，和/或向所述内存共享资源
25 池中的其他内存共享设备转发所述控制设备发送的内存访问请求；所述其
他内存共享设备转发的内存访问请求为该其他内存共享设备从对应的控制

设备处接收到的，所述内存访问请求中包括所述内存共享资源池中通过统一编址后多个内存共享设备中的内存单元的地址信息，所述地址信息用于获取所述内存共享资源池中的任意一个内存共享设备中的内存单元中的数据。

5 3、如权利要求 2 所述的数据存取系统，其特征在于，所述内存共享设备通过所述通信单元向所述内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中统一编址后的内存单元的编址数据，并将获取到的所述编址数据保存在所述处理模块中。

10 4、如权利要求 1 所述的数据存取系统，其特征在于，所述两个控制设备中的任一控制设备包括通信模块，所述任一控制设备的所述通信模块通过连接一可插拔的内存共享模块接入所述数据存取系统。

5、如权利要求 1-5 任一项所述的数据存取系统，其特征在于，所述通信接口为系统高速总线、交换机或以太网接口。

15 6、一种内存共享设备，其特征在于，包括：包括存储模块、处理模块以及通信单元，其中，

存储模块，与处理模块连接，用于存储数据；

处理模块，通过通信接口与一个控制设备相连，通过所述通信单元与所述内存共享设备所在的内存共享资源池中的其他内存共享设备连接，通过所述通信接口接收与该处理模块连接的控制设备发送的内存访问请求；
20 通过所述通信单元接收所述内存共享资源池中其他内存共享设备转发的内存访问请求，和/或向所述内存共享资源池中的其他内存共享设备转发所述控制设备发送的内存访问请求；所述其他内存共享设备转发的内存访问请求为该其他内存共享设备从对应的控制设备处接收到的，所述内存访问请求中包括所述内存共享资源池中通过统一编址后多个内存共享设备中的内存单元的地址信息，用于获取所述内存共享资源池中的任意一个内存共享
25 设备中的内存单元中的数据。

7、如权利要求 6 所述的内存共享设备，其特征在于，所述处理模块通过所述通信单元接收到所述内存共享资源池中的所述其他内存共享设备转发的内存访问请求后，在所述内存访问请求中携带的地址信息对应的内存单元中获取数据，并将结果通过所述通信单元返回给所述其他内存共享设备，由所述其他内存共享设备通过所述通信接口向与所述其他内存共享设备发送内存访问请求的控制设备返回所述结果。

8、如权利要求 6 所述的内存共享设备，其特征在于，所述处理模块进一步包括判断单元，如果所述判断单元判定所述地址信息对应的内存单元中的第一部分内存单元在本端内存共享设备，第二部分在第二内存共享设备，则所述处理模块在本端内存共享设备中的所述存储模块中读取第一部分内存单元中的数据，并将所述内存访问请求中的地址信息中第一部分内存单元的地址删除，通过所述通信单元向所述其他内存共享设备转发只包含第二部分内存单元地址的内存访问请求。

9、如权利要求 6-8 任一项所述的内存共享设备，其特征在于，所述处理模块还包括询问单元，用以通过所述通信单元向所述内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中统一编址的内存单元的编址数据，并将获取到的所述编址数据保存在所述处理模块中。

10、如权利要求 6-9 任一项所述的内存共享设备，其特征在于，所述内存共享设备还包括电源模块，用于当不存在外部电源供电时，为所述存储提供电能，藉由该电源模块所述内存共享设备形成可插拔的非易失性随机访问存储器 NVRAM。

11、如权利要求 6-10 所述的内存共享设备，其特征在于，还包括管理模块，用于调试所述存储模块中的应用程序及对所述通信单元中的流量进行控制。

12、如权利要求 6-11 任一项所述的内存共享设备，其特征在于，所述

处理模块为现场可编程门阵列 FPGA。

13、一种数据读取方法，其特征在于，包括：

内存共享资源池中的第一内存共享设备接收控制该第一内存共享设备的第一控制设备发送的内存访问请求，所述内存访问请求中包括通过统一编址后的地址信息；其中，所述的地址信息是通过统一编址后的内存共享资源池中的两个以上内存共享设备中的用于存储数据的的一个或多个内存单元的地址信息，所述两个以上的内存共享设备中的任意一个内存共享设备在两个以上控制设备中的一个控制设备的控制下进行数据的读取，所述第一内存共享设备为所述两个以上内存共享设备中的一个；

10 第一内存共享设备根据所述地址信息，确定所述地址信息对应的内存单元所在的内存共享设备；

如果所述地址信息对应的内存单元所在的内存共享设备为所述第一内存共享设备，则所述第一内存共享设备从第一内存共享设备自身中所述地址信息对应的内存单元中获取数据，

15 如果所述地址信息对应的内存单元所在的内存共享设备为第二内存共享设备，则所述第一内存共享设备向所述第二内存共享设备转发所述内存访问请求，并接收所述第二内存共享设备返回的数据；

所述第一内存共享设备将所述数据读取反馈给所述第一控制设备。

20 14、如权利要求 13 所述的方法，其特征在于，在内存共享资源池中的第一内存共享设备接收控制该第一内存共享设备的第一控制设备发送的内存访问请求之前，还包括：

内存共享资源池中的第一内存共享设备向所述内存共享资源池中的其它内存共享设备广播询问消息，以获取所述内存共享资源池中全部内存共享设备中内存单元的编址数据。

25

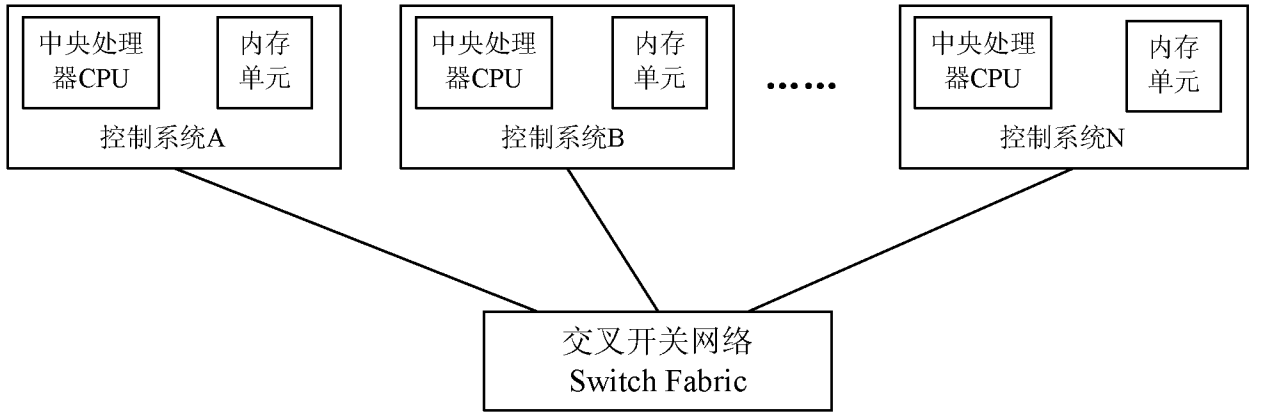


图 1

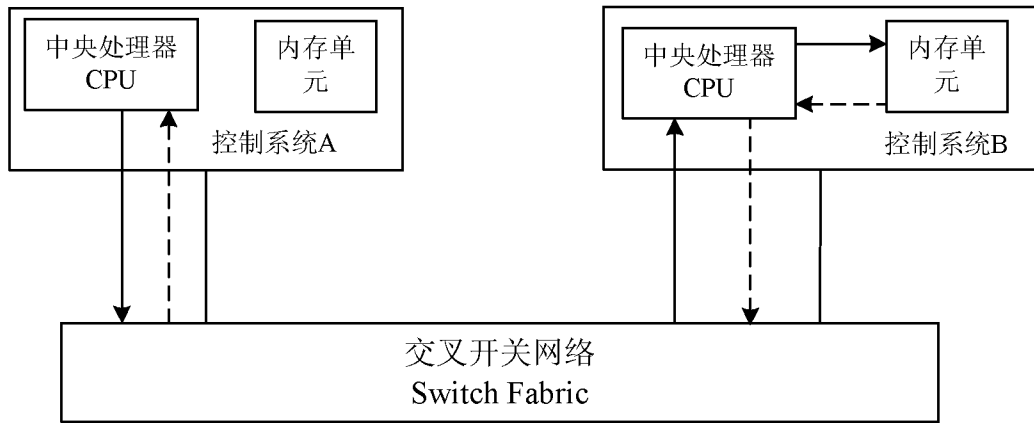


图 2

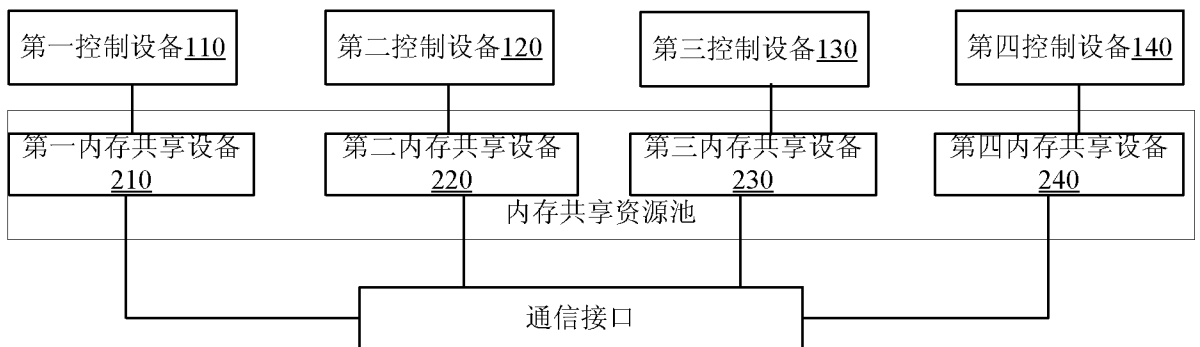


图 3

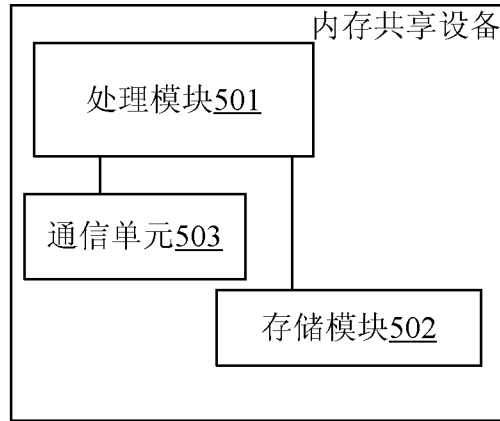


图 4

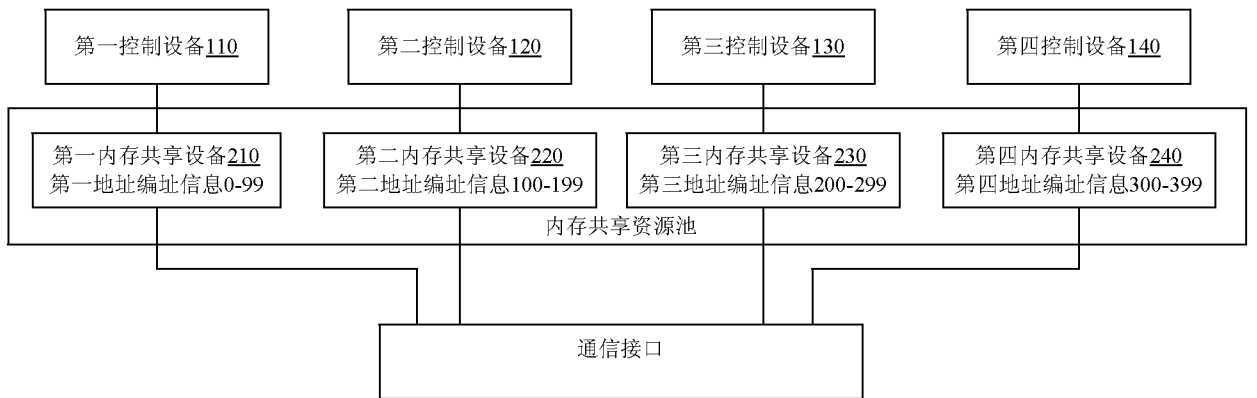


图 5A

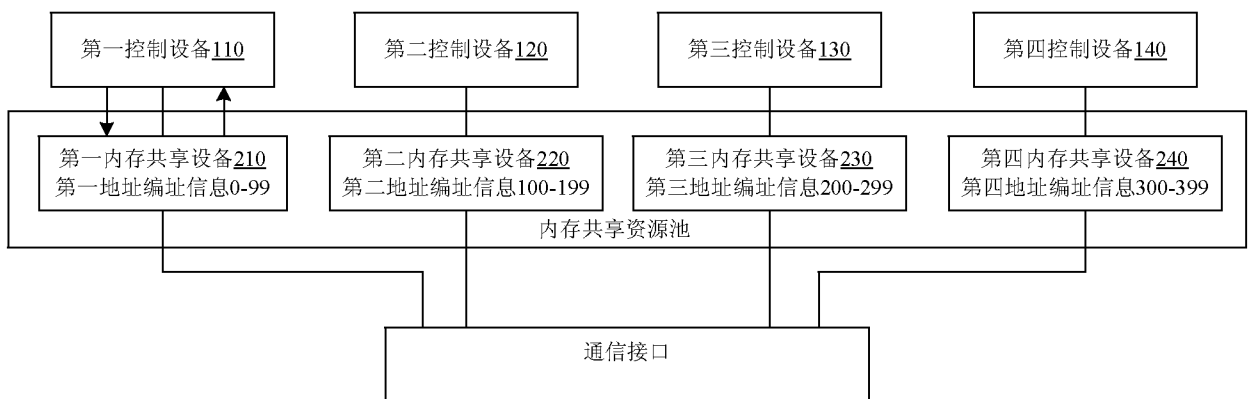


图 5B

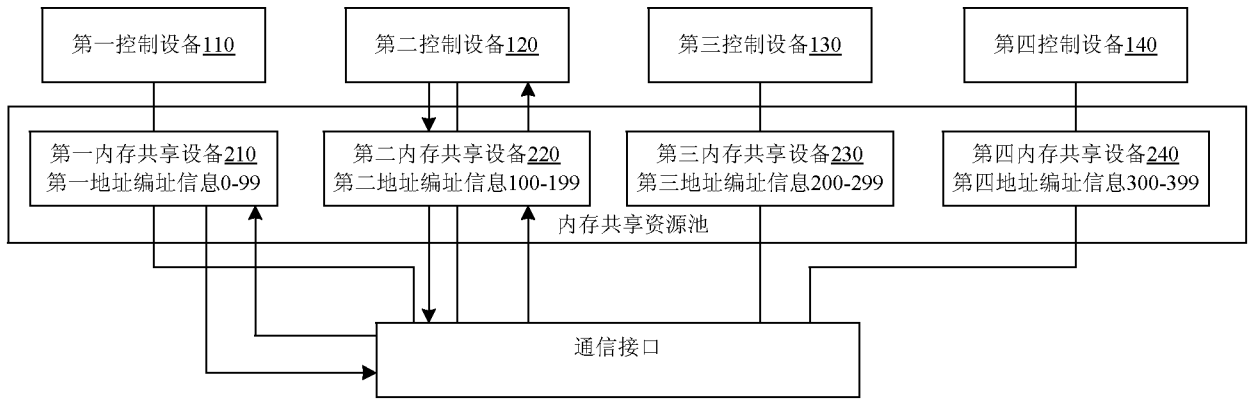


图 5C

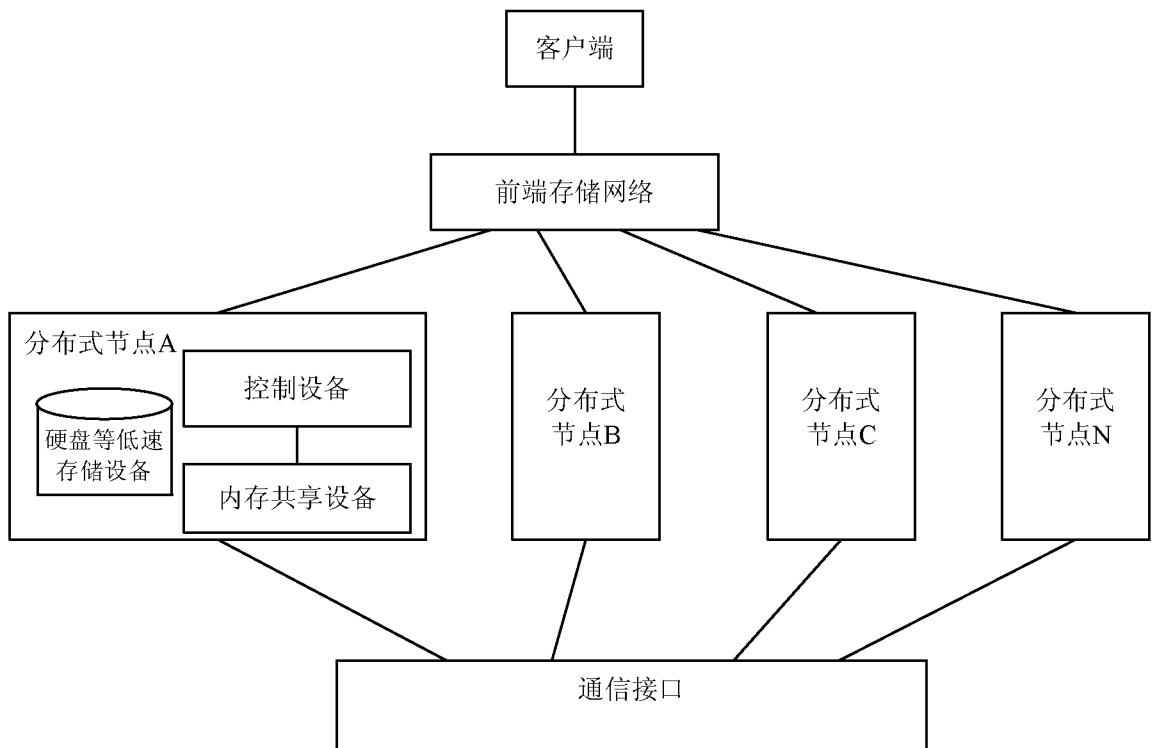


图 6

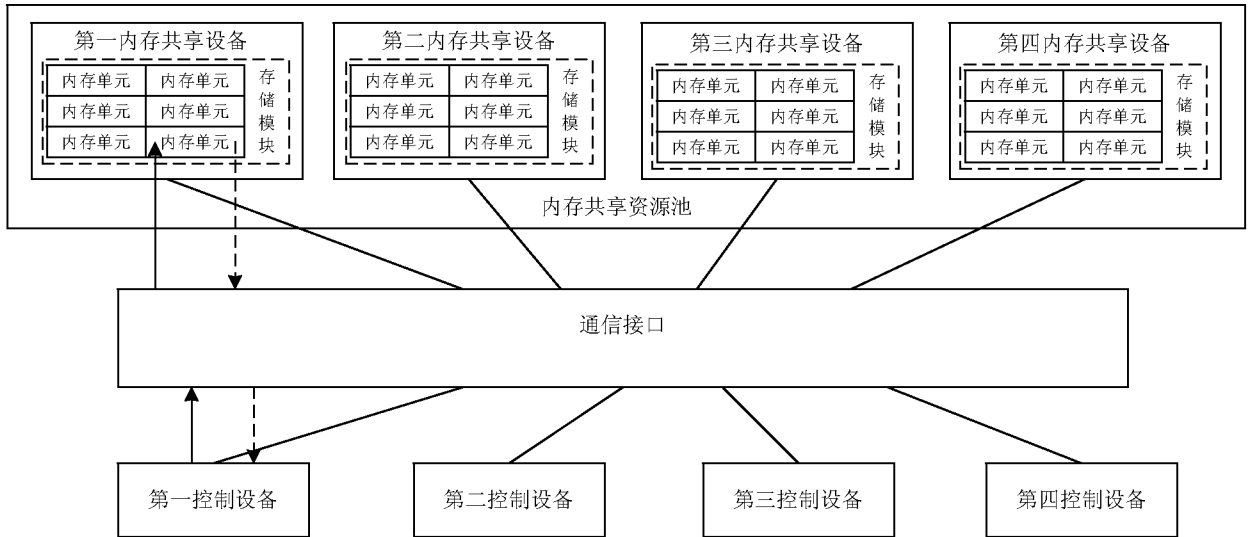


图 7

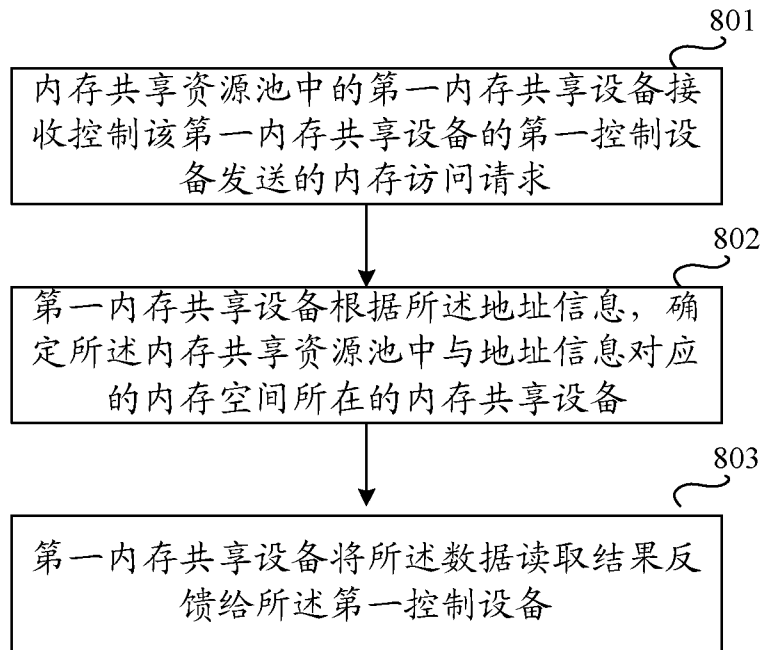


图 8

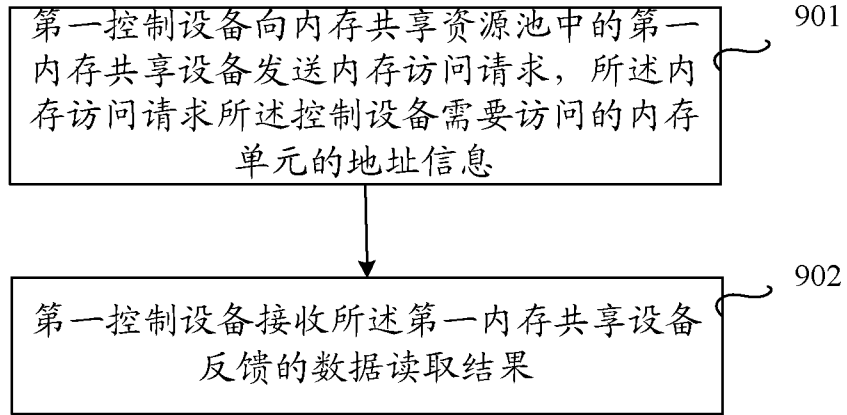


图 9

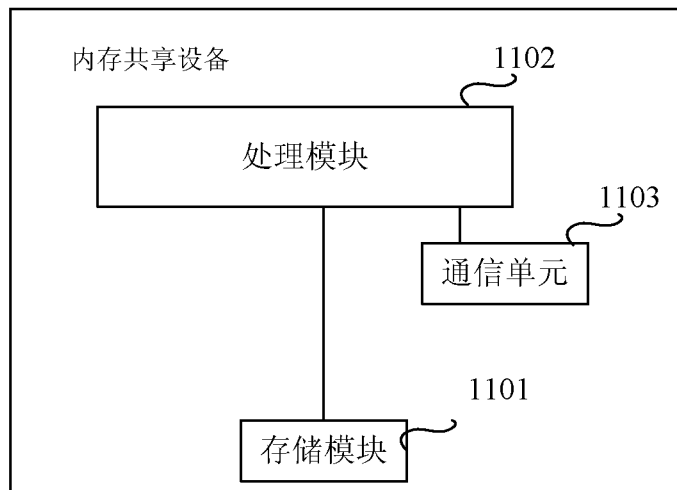


图 10

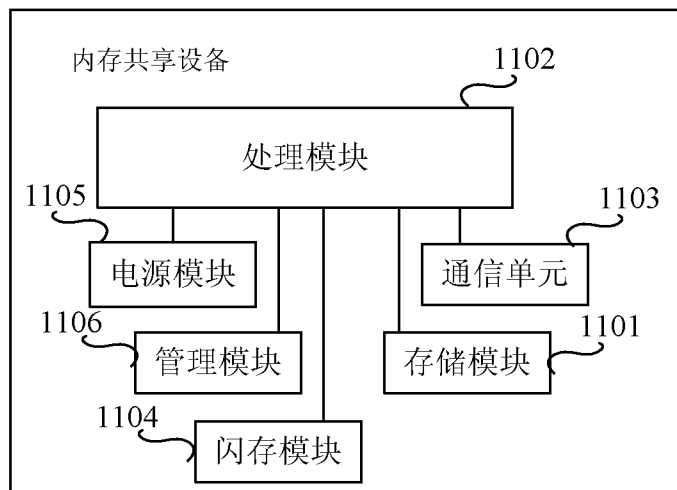


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/079841

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F, 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: controller?, read+, data, memory, shar+, storage, two, dual, first, second

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 103152394 A (HUAWEI TECHNOLOGIES CO., LTD.), 12 June 2013 (12.06.2013), claims 1-14	1-14
X	CN 101840309 A (UNITED INFORMATION TECHNOLOGY CO., LTD. et al.), 22 September 2010 (22.09.2010), description, paragraphs [0004] and [0045]-[0059], and figures 1 and 4	1-14
A	CN 1675625 A (MICRON TECHNOLOGY, INC.), 28 September 2005 (28.09.2005), the whole document	1-14
A	CN 1489058 A (INTERNATIONAL BUSINESS MACHINES CORPORATION), 14 April 2004 (14.04.2004), the whole document	1-14
A	CN 101135953 A (HANGZHOU H3C TECHNOLOGIES CO., LTD.), 05 March 2008 (05.03.2008), the whole document	1-14
A	CN 101158927 A (INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES), 09 April 2008 (09.04.2008), the whole document	1-14
A	US 2010/0228894 A1 (FREESCALE SEMICONDUCTOR, INC.), 09 September 2010 (09.09.2010), the whole document	1-14

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
25 October 2013 (25.10.2013)

Date of mailing of the international search report
21 November 2013 (21.11.2013)

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
ZHAO, Ting
Telephone No.: (86-10) **62414434**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/079841

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103152394 A	12.06.2013	None	
CN 101840309 A	22.09.2010	None	
CN 1675625 A	28.09.2005	US 2003229770 A1	11.12.2003
		WO 03104996 A1	18.12.2003
		AU 2003238961 A1	22.12.2003
		EP 1512077 A1	09.03.2005
		US 2007055817 A1	08.03.2007
		US 2011219196 A1	08.09.2011
		US 2012239885 A1	20.09.2012
		US 2009125688 A1	14.05.2009
		KR 20050005553 A	13.01.2005
		JP 2005529407 A	29.09.2005
CN 1489058 A	14.04.2004	US 2004073742 A1	15.04.2004
CN 101135953 A	05.03.2008	CN 101131624 A	27.02.2008
CN 101158927 A	09.04.2008	None	
US 2010/0228894 A1	09.09.2010	WO 2007083197 A1	26.07.2007

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/079841

CONTINUATION: A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (2006.01) i

G06F 12/02 (2006.01) i

G06F 12/08 (2006.01) i

<p>A. 主题的分类</p> <p style="text-align: center;">(参见附加页)</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>IPC: G06F, 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>WPI, EPODOC, CNPAT, CNKI, IEEE: controller?, read+, data, memory, shar+, 控制器, 读, 数据, 内存, 共享, 存储, 储存, 两, 双, 第一, 第二</p>																																				
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN103152394A(华为技术有限公司) 12.6 月 2013(12.06.2013) 权利要求 1-14</td> <td>1-14</td> </tr> <tr> <td>X</td> <td>CN101840309A(创新科存储技术有限公司 等) 22.9 月 2010(22.09.2010) 说明书第[0004], [0045]-[0059]段、附图 1, 4</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN1675625A(米克伦技术公司) 28.9 月 2005(28.09.2005) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN1489058A(国际商业机器公司) 14.4 月 2004(14.04.2004) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN101135953A(杭州华三通信技术有限公司) 05.3 月 2008(05.03.2008) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN101158927A(中国科学院计算技术研究所) 09.4 月 2008(09.04.2008) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>US2010/0228894A1(FREESCALE SEMICONDUCTOR, INC.) 09.9 月 2010(09.09.2010) 全文</td> <td>1-14</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在 C 栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型:</p> <table border="0"> <tr> <td>“A” 认为不特别相关的表示了现有技术一般状态的文件</td> <td>“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</td> </tr> <tr> <td>“E” 在国际申请日的当天或之后公布的在先申请或专利</td> <td>“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</td> </tr> <tr> <td>“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)</td> <td>“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</td> </tr> <tr> <td>“O” 涉及口头公开、使用、展览或其他方式公开的文件</td> <td>“&” 同族专利的文件</td> </tr> <tr> <td>“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</td> <td></td> </tr> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN103152394A(华为技术有限公司) 12.6 月 2013(12.06.2013) 权利要求 1-14	1-14	X	CN101840309A(创新科存储技术有限公司 等) 22.9 月 2010(22.09.2010) 说明书第[0004], [0045]-[0059]段、附图 1, 4	1-14	A	CN1675625A(米克伦技术公司) 28.9 月 2005(28.09.2005) 全文	1-14	A	CN1489058A(国际商业机器公司) 14.4 月 2004(14.04.2004) 全文	1-14	A	CN101135953A(杭州华三通信技术有限公司) 05.3 月 2008(05.03.2008) 全文	1-14	A	CN101158927A(中国科学院计算技术研究所) 09.4 月 2008(09.04.2008) 全文	1-14	A	US2010/0228894A1(FREESCALE SEMICONDUCTOR, INC.) 09.9 月 2010(09.09.2010) 全文	1-14	“A” 认为不特别相关的表示了现有技术一般状态的文件	“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	“E” 在国际申请日的当天或之后公布的在先申请或专利	“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	“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)	“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	“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件	“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	
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																																		
PX	CN103152394A(华为技术有限公司) 12.6 月 2013(12.06.2013) 权利要求 1-14	1-14																																		
X	CN101840309A(创新科存储技术有限公司 等) 22.9 月 2010(22.09.2010) 说明书第[0004], [0045]-[0059]段、附图 1, 4	1-14																																		
A	CN1675625A(米克伦技术公司) 28.9 月 2005(28.09.2005) 全文	1-14																																		
A	CN1489058A(国际商业机器公司) 14.4 月 2004(14.04.2004) 全文	1-14																																		
A	CN101135953A(杭州华三通信技术有限公司) 05.3 月 2008(05.03.2008) 全文	1-14																																		
A	CN101158927A(中国科学院计算技术研究所) 09.4 月 2008(09.04.2008) 全文	1-14																																		
A	US2010/0228894A1(FREESCALE SEMICONDUCTOR, INC.) 09.9 月 2010(09.09.2010) 全文	1-14																																		
“A” 认为不特别相关的表示了现有技术一般状态的文件	“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																																			
“E” 在国际申请日的当天或之后公布的在先申请或专利	“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																																			
“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)	“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																																			
“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件																																			
“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																																				
<p>国际检索实际完成的日期</p> <p style="text-align: center;">25.10 月 2013(25.10.2013)</p>	<p>国际检索报告邮寄日期</p> <p style="text-align: center;">21.11 月 2013 (21.11.2013)</p>																																			
<p>ISA/CN 的名称和邮寄地址:</p> <p>中华人民共和国国家知识产权局</p> <p>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;">赵婷</p> <p>电话号码: (86-10) 62414434</p>																																			

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/079841

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN103152394A	12.06.2013	无	
CN101840309A	22.09.2010	无	
CN1675625A	28.09.2005	US2003229770A1	11.12.2003
		WO03104996A1	18.12.2003
		AU2003238961A1	22.12.2003
		EP1512077A1	09.03.2005
		US2007055817A1	08.03.2007
		US2011219196A1	08.09.2011
		US2012239885A1	20.09.2012
		US2009125688A1	14.05.2009
		KR20050005553A	13.01.2005
		JP2005529407A	29.09.2005
CN1489058A	14.04.2004	US2004073742A1	15.04.2004
CN101135953A	05.03.2008	CN101131624A	27.02.2008
CN101158927A	09.04.2008	无	
US2010/0228894A1	09.09.2010	WO2007083197A1	26.07.2007

续: A. 主题的分类

H04L29/08(2006.01)i

G06F12/02(2006.01)i

G06F12/08(2006.01)i