

(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
2012 年 12 月 13 日 (13.12.2012)



WIPO | PCT



(10) 国际公布号
WO 2012/167531 A1

- (51) 国际专利分类号:
H04L 29/06 (2006.01) H04N 21/238 (2011.01)
- (21) 国际申请号: PCT/CN2011/081366
- (22) 国际申请日: 2011 年 10 月 27 日 (27.10.2011)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人 (对除美国外的所有指定国): 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: 及
- (75) 发明人/申请人 (仅对美国): 方帆 (FANG, Fan) [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)。 陈克平 (CHEN, Keping) [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)。
- (74) 代理人: 深圳市深佳知识产权代理事务所 (普通合伙) (SHENPAT INTELLECTUAL PROPERTY

AGENCY); 中国广东省深圳市国贸大厦 15 楼西座 1521 室, Guangdong 518014 (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, 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, KM, 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, 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, MD, 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, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: DATA-FAST-DISTRIBUTION METHOD AND DEVICE

(54) 发明名称: 一种数据快速分发方法和装置

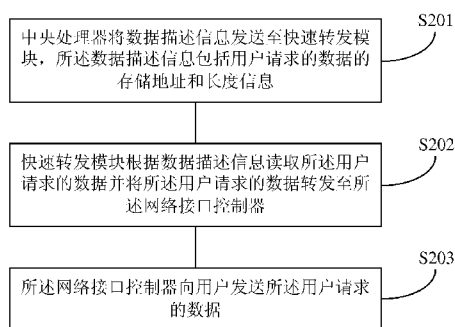


图 2 / Fig. 2

S201 THE CPU TRANSMITS DATA DESCRIPTION INFORMATION TO A FAST FORWARDING MODULE, AND THE DATA DESCRIPTION INFORMATION INCLUDES STORAGE ADDRESS AND LENGTH INFORMATION OF DATA REQUESTED BY A USER
S202 THE FAST FORWARDING MODULE READS THE DATA REQUESTED BY THE USER ACCORDING TO THE DATA DESCRIPTION INFORMATION AND THEN TRANSMITS IT TO A NETWORK INTERFACE CONTROLLER
S203 THE NETWORK INTERFACE CONTROLLER TRANSMITS THE DATA REQUESTED BY THE USER TO THE USER

(57) Abstract: Provided are a data-fast-distribution method and device to reduce the joint of a Central Processing Unit (CPU) and a memory, and improve the system performance. The method includes: the CPU transmits data description information to a fast forwarding module, and the data description information includes storage address and length information of data requested by a user; the fast forwarding module reads the data requested by the user according to the data description information and then transmits it to a Network Interface Controller (NIC); the NIC transmits the data requested by the user to the user. Compared with the prior art, after the service is increased (e.g. present network card bandwidth and storage capacity need to be enlarged), the method provided by the present embodiments only needs to increase an apparatus with an NIC and a storage apparatus without increasing the cost of the memory and the CPU, therefore one side the usage of the memory and the CPU is reduced and then the performance of the whole system is improved, on the other side the problem of memory walls is reduced.

(57) 摘要: 本发明提供一种数据快速分发方法和装置, 以减少中央处理器和内存的参与, 提高系统性能。所述方法包括: 中央处理器将数据描述信息发送至快速转发模块, 所述数据描述信息包括用户请求的数据的存储地址和长度信息; 所述快速转发模块根据所述数据描述信息读取所述用户请求的数据后转发至所述网络接口控制器; 所述网络接口控制器向用户发送所述用户请求的数据。与现有技术相比, 本发明实施例提供的方法在业务增加后 (例如, 需要扩大现有的网卡带宽和存储容量), 只需要增加网络接口控制器和存储设备的设备, 而不需要增加内存和中央处理器的成本, 一来可以减少对内存和中央处理器的使用, 提高整个系统的性能, 二来也减少了内存墙的问题。

**本国际公布:**

— 包括国际检索报告(条约第 21 条(3))。

— 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布(细则 48.2(h))。

— 根据申请人的请求, 在条约第 21 条(2)(a)所规定的期限届满之前进行。

一种数据快速分发方法和装置

技术领域

本发明实施例涉及通信领域，尤其涉及一种数据快速分发方法和装置。

5 背景技术

互联网应用现在已经进入Web2.0时代，用户的实际感受已经是用户选择业务的第一选择。为了解决时延导致用户感受的问题，提供大量的快速访问设计，满足用户良好的感受体验，几乎所有的互联网公司都在实施加速设计，其中，内容分发网络（Content Delivery Network，CDN）中的高速缓冲（Cache）设计尤其突出。在CDN中，每一个Cache都是一个中间设备，提供高速缓冲功能。

以CDN产品业务中的视频点播业务为例，现有技术提供的视频点播业务加速解决方案如附图1（带箭头的线条表示信息或数据流动的方向，线条旁的序号表示顺序关系）所示：用户的请求从网口，即网络接口控制器（Network Interface Controller，NIC）进入，中央处理单元（Central Processing Unit，CPU）收到请求后，读取存储设备（Storage）的数据到内存（Memory），并发送给NIC，通过NIC发送至客户端（Client）。在这种方案中，设备提供快速的闪存（Flash）和网络，针对热点的数据，提供高速缓冲（Cache）。如果存储设备中暂时没有用户请求的视频数据，则会通过NIC发送给其它可以提供用户请求的视频数据的设备，通过NIC将用户请求的视频数据取回，再发送给客户端（Client）；为了便于下次满足用户的请求，在将从其他设备取回的视频数据发送给客户端的同时或之后，将该视频数据存储到本地存储设备（Storage）。

上述现有技术的缺陷在于：所有的视频数据都必须过一次内存和CPU，CPU的工作就是搬移数据，进行分发，无法做到数据面和控制面的分离；对于CDN流，在高并发情况下，CPU使用率非常高，会带来明显的性能瓶颈。

发明内容

本发明实施例提供一种数据快速分发方法和装置,以减少中央处理器和内存的参与,提高系统性能。

本发明实施例提供一种数据快速分发方法包括:

- 5 中央处理器将数据描述信息发送至快速转发模块,所述数据描述信息包括用户请求的数据的存储地址和长度信息;

所述快速转发模块根据所述数据描述信息读取所述用户请求的数据后转发至所述网络接口控制器;

所述网络接口控制器向用户发送所述用户请求的数据。

- 10 本发明另一实施例提供一种数据快速分发方法包括:

接收中央处理器发送的数据描述信息,所述数据描述信息包括用户请求的数据的存储地址和长度信息;

- 15 根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器,由所述网络接口控制器向用户发送所述用户所请求的数据。

本发明实施例提供一种数据快速分发系统,所述系统包括中央处理器、网络接口控制器和快速转发模块;

所述中央处理器,用于将数据描述信息发送至快速转发模块,所述数据描述信息包括用户请求的数据的存储地址和长度信息;

- 20 所述快速转发模块,用于根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器;

所述网络接口控制器,用于向用户发送所述用户请求的数据。

本发明实施例提供一种数据快速分发装置,所述装置包括接收模块和数据读取模块;

所述接收模块,用于接收中央处理器发送的数据描述信息,所述数据描述信息包括用户请求的数据的存储地址和长度信息;

所述数据读取模块,用于根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器,以由所述网络接口控制器向
5 用户发送所述用户所请求的数据。

从上述本发明实施例可知,由于网络接口控制器是将存储地址和长度信息发送给快速转发模块,快速转发模块在读取所述用户请求的数据后转发至所述网络接口控制器,由所述网络接口控制器发送所述用户请求的数据,整个数据的传输过程通过网络接口控制器和存储设备之间的独立、直接通道完成,没有
10 经过中央处理器和内存,只需要中央处理器配置传输描述符,不需要中央处理器去拷贝数据。因此,与现有技术相比,本发明实施例提供的方法在业务增加后(例如,需要扩大现有的网卡带宽和存储容量),只需要增加网络接口控制器和存储设备的设备,而不需要增加内存和中央处理器的成本,一来可以减少对内存和中央处理器的使用,提高整个系统的性能,二来也减少了内存墙的问题。
15 题。

附图说明

为了更清楚地说明本发明实施例的技术方案,下面将对现有技术或实施例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,对于本领域技术人员来讲,还可以如这些附图获得其
20 他的附图。

图1是现有技术提供的视频点播业务加速解决方案示意图;

图2是本发明实施例提供的快速分发方法流程图;

图3是本发明实施例提供的传输描述符结构示意图;

图4是本发明实施例提供的快速分发系统结构示意图;

图 5 是本发明另一实施例提供的数据快速分发系统结构示意图；

图 6 是本发明另一实施例提供的数据快速分发系统结构示意图；

图 7 是本发明另一实施例提供的数据快速分发系统结构示意图；

图 8 是本发明另一实施例提供的数据快速分发系统结构示意图；

5 图 9 是本发明另一实施例提供的数据快速分发系统结构示意图；

图 10 是本发明实施例提供的数据快速分发装置结构示意图；

图 11 是本发明另一实施例提供的数据快速分发装置结构示意图；

图 12 是本发明另一实施例提供的数据快速分发装置结构示意图。

具体实施方式

10 本发明实施例提供一种数据快速分发方法和装置，以减少中央处理器和内存的参与，提高系统性能。

请参阅附图2，是本发明实施例提供的数据快速分发方法流程示意图，主要包括步骤：

15 S201，中央处理器将数据描述信息发送至快速转发模块，所述数据描述信息包括用户请求的数据的存储地址和长度信息。

在本发明提供的实施例中，快速转发模块是嵌入到PCI-Switch或者嵌入到网络接口控制器NIC中的一个功能模块，提供直接发送和接收功能。该功能模块通过CPU配置后，可以去读取存储设备或内存中的数据，发送到网络接口控制器NIC上，完成直接内存访问（Direct Memory Access，DMA）发送，即，
20 快速转发模块是以DMA方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

一般情况下，网络接口控制器NIC在内存中有一块对应的数据描述信息。中央处理器如果需要发送数据的时候，会通过配置数据描述信息，例如，传输描述符（Transmit Descriptor）。此时网络接口控制器NIC获取到需要发送数据，

则通知内部的DMA模块，根据传输描述符（Transmit Descriptor）中的缓存地址（Buffer Address）字段和长度（Length）字段，去相应的地址取数据，进行发送。每一个发送的数据都需要配置传输描述符（Transmit Descriptor），配置一个发一个。

- 5 如果发送的数据是一块连续的数据或者是中央处理器可预知的数据这种固定的流，中央处理器通过配置快速转发模块，网络接口控制器NIC直接把数据发送到存储设备（Storage）进行存储；也可以直接从存储设备（Storage）把数据直接发到网络接口控制器NIC出去，完成网络接口控制器NIC到存储设备（Storage）和存储设备（Storage）到网络接口控制器NIC的DMA操作，并且
- 10 不占用PCI-Switch和中央处理器之间的带宽。

因此，在本发明实施例中，数据的传输不需要中央处理器去干预，只需要中央处理器配置数据描述信息；也不需要中央处理器去拷贝数据和因为拷贝数据导致对内存资源的消耗。

- 一个用户链接建立时就被分配了一个打包模式标识（Package ID）给该用
- 15 户，并根据用户的回应的报文格式将打包模式标识下发至快速转发模块。在本发明实施例中，Package ID与打包模式对应，位于快速转发模块的打包列表（packagelist）中，这个表项目主要是一个报文格式定义，即IP头格式的内容。当后续从存储设备取到数据时，按照报文格式定义的格式将数据填充进去就可以直接发送使用。

- 20 当用户请求数据时，中央处理器通过网络接口控制器NIC收到用户发送的请求，这些请求中携带了用户请求的数据的存储地址和长度等信息。中央处理器将用户请求的数据的存储地址和长度等信息发送至网络接口控制器NIC，以便触发网络接口控制器NIC的DMA模块取用户请求的数据。

作为本发明一个实施例，中央处理器可以配置数据描述信息，例如，传输

描述符 (Transmit Descriptor), 将所述用户请求的数据的存储地址和长度信息携带在所述数据描述信息中; 除此之外, 中央处理器在配置数据描述信息时, 还可以将打包模式标识和打包选择标识携带在所述数据描述信息中, 所述打包模式标识与所述请求数据的用户对应, 所述打包选择标识用于标识是否需要对所
5 述用户请求的数据进行打包。中央处理器将所述数据描述信息发送至网络接口控制器, 由所述网络接口控制器将所述数据描述信息转发至所述快速转发模块。

本发明实施例提供的传输描述符 (Transmit Descriptor) 的格式如附图3所示, 其中:

10 内存选择 (Memory Select) 标识字段表示是从内存 (Memory) 还是存储设备 (Storage) 读取存放的数据, 可用1位 (bit) 二进制表示, 例如, 该字段是“1”表示是从内存 (Memory) 读取存放的数据, 该字段是“0”表示是从存储设备 (Storage) 读取存放的数据, 或者反之, 该字段是“0”表示是从内存 (Memory) 读取存放的数据, 该字段是“1”表示是从存储设备 (Storage)
15 读取存放的数据;

打包选择标识 (Package Select) 字段表示是否需要快速转发模块对读取到的数据进行打包或封装;

打包模式标识 (Package ID) 字段表示使用何种打包或封装模式, 该字段也与请求数据的用户对应;

20 缓存地址 (Buffer Address) 字段表示用户请求的数据的存储地址, 用户请求的数据的存储地址由用户请求数据时发送的请求提供;

长度 (Length) 字段表示用户请求的数据的长度;

传输描述符 (Transmit Descriptor) 还包括其他控制字段, 不做赘述。

S202, 快速转发模块根据数据描述信息读取所述用户请求的数据并将所述

用户请求的数据转发至所述网络接口控制器。

快速转发模块根据网络接口控制器的DMA模块传送过来的用户请求的数据的缓存地址、长度信息、打包选择标识和内存选择标识等数据描述信息（例如，传输描述符）读取所述用户请求的数据

- 5 具体地，快速转发模块可以根据所述数据描述信息包含的信息，例如，缓存地址、长度信息等，读取所述用户请求的数据后，进一步通过传输描述符（Transmit Descriptor）中的打包选择标识（Package Select）字段判断是否需要
- 10 对所述用户请求的数据进行打包，若需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。例如，快速转发模块可以根据传输描述符（Transmit Descriptor）中的打包模式标识对应的打包模式，对所述用户请求的数据进行打包后转发至所述网络接口控制器。

- 再如，快速转发模块也可以通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；若判断是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。
- 15

S203，所述网络接口控制器向用户发送所述用户请求的数据。

- 从上述本发明实施例可知，由于中央处理器是将存储地址和长度信息发送给快速转发模块，快速转发模块在读取所述用户请求的数据后转发至所述网络接口控制器，由所述网络接口控制器发送所述用户请求的数据，整个数据的传输过程通过网络接口控制器和存储设备之间的独立、直接通道完成，没有经过中央处理器和内存，只需要中央处理器配置数据描述信息，不需要中央处理器去拷贝数据。因此，与现有技术相比，本发明实施例提供的方法在业务增加后（例如，需要扩大现有的网卡带宽和存储容量），只需要增加网络接口控制器
- 20

和存储设备的设备，而不需要增加内存和中央处理器的成本，一来可以减少对内存和中央处理器的使用，提高整个系统的性能，二来也减少了内存墙的问题。

请参阅附图4，是本发明实施例提供的数据快速分发方法流程示意图，本实施例的执行主体可以是附图2示例的数据快速分发方法中快速转发模块，附

5 图4示例主要包括步骤：

S401，接收中央处理器发送的数据描述信息，所述数据描述信息包括用户请求的数据的存储地址和长度信息。

当用户请求数据时，中央处理器通过网络接口控制器NIC收到用户发送的请求，这些请求中携带了用户请求的数据的存储地址和长度等信息。中央处理
10 器将用户请求的数据的存储地址和长度等信息发送至网络接口控制器NIC，以便触发网络接口控制器NIC的DMA模块取用户请求的数据。

作为本发明一个实施例，中央处理器可以配置数据描述信息，例如，传输描述符 (Transmit Descriptor)，将所述用户请求的数据的存储地址和长度信息携带在所述数据描述信息中；除此之外，中央处理器在配置数据描述信息时，还
15 可以将打包模式标识和打包选择标识携带在所述数据描述信息中，所述打包模式标识与所述请求数据的用户对应，所述打包选择标识用于标识是否需要对所
述用户请求的数据进行打包。传输描述符的具体格式如附图3所示，具体含义可参阅前述实施例的相关文字说明。

中央处理器将数据描述信息发送至网络接口控制器，快速转发模块接收由
20 所述网络接口控制器转发的所述数据描述信息。

S402，根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器，以便由所述网络接口控制器向用户发送所述用户所请求的数据。

快速转发模块根据网络接口控制器的DMA模块传送过来的用户请求的数

据的缓存地址、长度信息、打包选择标识和内存选择标识等数据描述信息（例如，传输描述符）读取所述用户请求的数据

具体地，快速转发模块可以根据所述数据描述信息，例如，缓存地址、长度信息，读取所述用户请求的数据后，进一步通过传输描述符（Transmit
5 Descriptor）中的打包选择标识（Package Select）字段判断是否需要所述用户请求的数据进行打包，若需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。例如，快速转发模块可以根据传输描述符（Transmit Descriptor）中的打包模式标识对应的打包模式，对
10 所述用户请求的数据进行打包后转发至所述网络接口控制器。

再如，快速转发模块也可以通过所述内存选择标识判断是从内存读取所述
15 用户请求的数据还是从存储设备读取所述用户请求的数据；若判断是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

需要说明的是，本实施例的快速转发模块是嵌入到PCI-Switch或者嵌入到
20 网络接口控制器NIC中的一个功能模块，提供直接发送和接收功能。该功能模块通过中央处理器配置后，可以去读取存储设备（Storage）或内存中的数据，发送到网络接口控制器NIC上，完成直接内存访问（Direct Memory Access，DMA）发送，即，快速转发模块是以DMA方式读取所述用户请求的数据并将
25 所述用户请求的数据转发至所述网络接口控制器。

请参阅附图5，本发明实施例提供的数据快速分发系统结构示意图。为了便于说明，仅仅示出了与本发明实施例相关的部分，其中包含的功能模块/单元可以是软件模块/单元、硬件模块/单元或软硬件相结合模块/单元（本说明书提供的各个实施例都可应用这一描述原则）。图5示例的数据快速分发系统包括

中央处理器501、快速转发模块502和网络接口控制器503，其中：

中央处理器501，用于将数据描述信息发送至快速转发模块502，所述数据描述信息包括用户请求的数据的存储地址和长度信息；

快速转发模块502，用于根据所述数据描述信息读取所述用户请求的数据
5 并将所述用户请求的数据转发至所述网络接口控制器503；

网络接口控制器503，用于向用户发送所述用户请求的数据。一个用户链接建立时就被分配了一个打包模式标识（Package ID）给该用户，并根据用户的回应的报文格式将打包模式标识下发至快速转发模块。在本发明实施例中，Package ID与打包模式对应，位于快速转发模块的打包列表（packagelist）中，
10 这个表项目主要是一个报文格式定义，即IP头格式的内容。当后续从存储设备取到数据时，按照报文格式定义的格式将数据填充进去就可以直接发送使用。

当用户请求数据时，中央处理器501通过网络接口控制器503收到用户发送的请求，这些请求中携带了用户请求的数据的存储地址和长度等信息。中央处理器501将用户请求的数据的存储地址和长度等信息发送至网络接口控制器
15 503，以便触发网络接口控制器503的DMA模块取用户请求的数据。

如附图6所示本发明另一实施例提供的快速分发系统，中央处理器501可以包括配置单元601和发送单元602，其中：

配置单元601用于配置数据描述信息，例如，传输描述符（Transmit Descriptor），如附图3所示，将所述用户请求的数据的存储地址和长度信息携带
20 在所述数据描述信息中；除此之外，配置单元601在配置数据描述信息时，还可以将打包模式标识、打包选择标识和内存选择标识携带在所述数据描述信息中，所述打包模式标识与所述请求数据的用户对应，所述打包选择标识用于标识是否需要对所述用户请求的数据进行打包，所述内存选择标识用于标识是从内存（Memory）还是存储设备（Storage）读取存放的数据。

发送单元602用于将所述数据描述信息发送至网络接口控制器503,由所述网络接口控制器503将所述数据描述信息转发至所述快速转发模块502。

在本发明提供的实施例中,快速转发模块502是嵌入到PCI-Switch或者嵌入到网络接口控制器503中的一个功能模块,提供直接发送和接收功能。该功能模块通过中央处理器501配置后,可以去读取存储设备或内存中的数据,发送到网络接口控制器NIC上,完成DMA发送,即,快速转发模块502是以DMA方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

一般情况下,网络接口控制器503在内存中有一块对应的传输描述符(Transmit Descriptor)链表。中央处理器501如果需要发送数据的时候,会通过配置传输描述符(Transmit Descriptor)。此时网络接口控制器503获取到需要发送数据,则通知内部的DMA模块,根据传输描述符(Transmit Descriptor)中的缓存地址(Buffer Address)字段和长度(Length)字段,去相应的地址取数据,进行发送。每一个发送的数据都需要配置传输描述符(Transmit Descriptor),配置一个发一个。

如果发送的数据是一块连续的数据或者是中央处理器501可预知的数据这种固定的流,中央处理器501通过配置快速转发模块,网络接口控制器503NIC直接将数据发送到存储设备(Storage)进行存储;也可以直接从存储设备(Storage)将数据直接发到网络接口控制器503出去,完成网络接口控制器503到存储设备(Storage)和存储设备(Storage)到网络接口控制器503的DMA操作,并且不占用PCI-Switch和中央处理器之间的带宽。

因此,在本发明实施例中,数据的传输不需要中央处理器去干预,只需要中央处理器配置数据描述信息;也不需要中央处理器去拷贝数据和因为拷贝数据导致对内存资源的消耗。

如附图7所示本发明另一实施例提供的快速转发模块502可以包括第一判断单元701和打包单元702，其中：

第一判断单元701，用于根据所述数据描述信息读取所述用户请求的数据后，通过所述打包选择标识判断是否需要所述用户请求的数据进行打包；

5 打包单元702，用于若所述第一判断单元701判断的结果为需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器503。

如附图8所示本发明另一实施例提供的快速转发模块502可以包括第二判断单元801和转发单元802，其中：

10 第二判断单元801，用于通过内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

15 转发单元802，用于若所述第二判断单元801判断的结果为是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器503，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器503。

附图5至附图8任一示例的中央处理器501还包括下发模块901，如附图9所示。下发模块901用于将打包模式标识下发至所述快速转发模块502，所述打包模式标识与打包模式对应。

20 请参阅附图10，本发明实施例提供的快速转发装置结构示意图。为了便于说明，仅仅示出了与本发明实施例相关的部分，其中包含的功能模块/单元可以是软件模块/单元、硬件模块/单元或软硬件相结合模块/单元（本说明书提供的各个实施例都可应用这一描述原则）。图5示例的快速转发装置可以是附图5至附图9任一示例的快速转发模块502，其包括接收模块1001和数据读取模块1002，其中：

接收模块1001, 用于接收中央处理器发送的数据描述信息, 所述数据描述信息可以是如附图3示例的输描述符 (Transmit Descriptor), 包括用户请求的数据的存储地址和长度信息;

数据读取模块1002, 用于根据所述数据描述信息读取所述用户请求的数据
5 并将所述用户请求的数据转发至网络接口控制器, 以由所述网络接口控制器向用户发送所述用户所请求的数据。具体地, 数据读取模块1002用于根据所述数据描述信息, 以直接内存访问方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

因此, 在本实施例中, 数据的传输不需要中央处理器去干预, 只需要中央
10 处理器配置数据描述信息, 也不需要中央处理器去拷贝数据和因为拷贝数据导致对内存资源的消耗。

需要说明的是, 以上数据快速分发装置的实施方式中, 各功能模块的划分仅是举例说明, 实际应用中可以根据需要, 例如相应硬件的配置要求或者软件的实现的便利考虑, 而将上述功能分配由不同的功能模块完成, 即将所述数据
15 快速分发装置的内部结构划分成不同的功能模块, 以完成以上描述的全部或者部分功能。而且, 实际应用中, 本实施例中的相应的功能模块可以由相应的硬件实现, 也可以由相应的硬件执行相应的软件完成, 例如, 前述的接收模块, 可以是具有执行前述接收中央处理器发送的数据描述信息的硬件, 例如接收器, 也可以是能够执行相应计算机程序从而完成前述功能的一般处理器或者其
20 他硬件设备 (本说明书提供的各个实施例都可应用上述描述原则)。

进一步地, 所述数据描述信息还可以包括打包模式标识和打包选择标识, 所述打包模式标识与所述请求数据的用户对应, 所述打包选择标识用于标识是否需要所述用户请求的数据进行打包, 附图10示例的数据读取模块1002可以包括第一判断单元1101和打包单元1102, 如附图11所示本发明另一实施例提供

的数据快速分发装置，其中：

第一判断单元1101，用于根据所述数据描述信息读取所述用户请求的数据后，通过所述打包选择标识判断是否需要与所述用户请求的数据进行打包；

打包单元1102，用于若第一判断单元1101判断的结果为需要与所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。

进一步地，所述数据描述信息还可以包括内存选择标识，所述内存选择标识用于标识是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据，附图10示例的数据读取模块1002可以包括第二判断单元1201和转发单元1202，如附图12所示本发明另一实施例提供的快速分发装置，其中：

第二判断单元1201，用于通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

转发单元1202，用于若第二判断单元1201判断的结果为是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

需要说明的是，上述装置各模块/单元之间的信息交互、执行过程等内容，由于与本发明方法实施例基于同一构思，其带来的技术效果与本发明方法实施例相同，具体内容可参见本发明方法实施例中的叙述，此处不再赘述。

本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成，比如以下各种方法的一种或多种或全部：

方法一：

中央处理器将数据描述信息发送至快速转发模块，所述数据描述信息包括

用户请求的数据的存储地址和长度信息;

快速转发模块根据所述数据描述信息读取所述用户请求的数据后转发至所述网络接口控制器;

所述网络接口控制器向用户发送所述用户请求的数据。

5 方法二:

接收中央处理器发送的数据描述信息,所述数据描述信息包括用户请求的数据的存储地址和长度信息;

根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器,以由所述网络接口控制器向用户发送所述用户所请求的数据。

该程序可以存储于一计算机可读存储介质中,存储介质可以包括:只读存储器 (ROM, Read Only Memory)、随机存取存储器 (RAM, Random Access Memory)、磁盘或光盘等。

以上对本发明实施例提供的快速分发方法、装置和系统进行了详细介绍,本文中应用了具体个例对本发明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发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,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发明的限制。

权 利 要 求

1、一种数据快速分发方法，其特征在于，所述方法包括：

中央处理器将数据描述信息发送至快速转发模块，所述数据描述信息包括用户请求的数据的存储地址和长度信息；

5 所述快速转发模块根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器；

所述网络接口控制器向用户发送所述用户请求的数据。

2、如权利要求1所述的方法，其特征在于，所述中央处理器将数据描述信息发送至快速转发模块包括：

10 所述中央处理器配置所述数据描述信息，将所述用户请求的数据的存储地址和长度信息携带在所述数据描述信息；

所述中央处理器将所述数据描述信息发送至网络接口控制器，由所述网络接口控制器将所述数据描述信息转发至所述快速转发模块。

3、如权利要求2所述的方法，其特征在于，所述中央处理器配置数据描述信息还包括：所述中央处理器将打包模式标识和打包选择标识携带在所述传输描述符，所述打包模式标识与所述请求数据的用户对应，所述打包选择标识用于标识是否需要对所述用户请求的数据进行打包；

所述快速转发模块根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器包括：

20 所述快速转发模块根据所述数据描述信息读取所述用户请求的数据后，通过所述打包选择标识判断是否需要对所述用户请求的数据进行打包；若需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。

4、如权利要求3所述的方法，其特征在于，所述对所述用户请求的数据进行打包后转发至所述网络接口控制器包括：

所述快速转发模块根据所述打包模式标识对应的打包模式，对所述用户请求的数据进行打包后转发至所述网络接口控制器。

5、如权利要求2所述的方法，其特征在于，所述中央处理器配置数据描述

信息还包括: 所述中央处理器将内存选择标识携带在所述数据描述信息, 所述内存选择标识用于标识是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据;

所述快速转发模块根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器包括: 所述快速转发模块通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据; 若判断是从存储设备读取所述用户请求的数据, 则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器, 否则, 从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

10 6、如权利要求1所述的方法, 其特征在于, 所述快速转发模块根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器包括: 所述快速转发模块根据所述数据描述信息, 以直接内存访问方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

15 7、如权利要求2至5任意一项所述的方法, 其特征在于, 所述中央处理器将用户请求的数据的存储地址和长度信息发送至网络接口控制器之前还包括:

所述中央处理器将打包模式标识下发至所述快速转发模块, 所述打包模式标识与打包模式对应。

8、一种数据快速分发方法, 其特征在于, 所述方法包括:

20 接收中央处理器发送的数据描述信息, 所述数据描述信息包括用户请求的数据的存储地址和长度信息;

根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器, 以由所述网络接口控制器向用户发送所述用户所请求的数据。

25 9、如权利要求8所述的方法, 其特征在于, 所述数据描述信息还包括打包模式标识和打包选择标识, 所述打包模式标识与所述请求数据的用户对应, 所述打包选择标识用于标识是否需要对所述用户请求的数据进行打包;

所述根据所述数据描述信息读取所述用户请求的数据并将所述用户请求

的数据转发至所述网络接口控制器包括: 根据所述数据描述信息读取所述用户请求的数据后, 通过所述打包选择标识判断是否需要所述用户请求的数据进行打包; 若需要对所述用户请求的数据进行打包, 则对所述用户请求的数据进行打包后转发至所述网络接口控制器。

- 5 10、如权利要求9所述的方法, 其特征在于, 所述对所述用户请求的数据进行打包后转发至所述网络接口控制器包括:

根据所述打包模式标识对应的打包模式, 对所述用户请求的数据进行打包后转发至所述网络接口控制器。

- 10 11、如权利要求8所述的方法, 其特征在于, 所述数据描述信息还包括内存选择标识, 所述内存选择标识用于标识是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据;

所述根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器包括: 通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据; 若判断是从存储设备读取所述用户请求的数据, 则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器, 否则, 从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

- 20 12、如权利要求8所述的方法, 其特征在于, 所述根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器包括: 根据所述数据描述信息, 以直接内存访问方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

13、一种数据快速分发系统, 其特征在于, 所述系统包括中央处理器、网络接口控制器和快速转发模块;

- 25 所述中央处理器, 用于将数据描述信息发送至快速转发模块, 所述数据描述信息包括用户请求的数据的存储地址和长度信息;

所述快速转发模块, 用于根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器;

所述网络接口控制器, 用于向用户发送所述用户请求的数据。

14、如权利要求13所述的系统，其特征在于，所述中央处理器包括：

配置单元，用于配置所述数据描述信息，将所述用户请求的数据的存储地址和长度信息携带在所述数据描述信息；

5 发送单元，用于将所述数据描述信息发送至网络接口控制器，由所述网络接口控制器将所述数据描述信息转发至所述快速转发模块。

15、如权利要求14所述的系统，其特征在于，所述配置单元还用于将打包模式标识和打包选择标识携带在所述数据描述信息，所述打包模式标识与所述请求数据的用户对应，所述打包选择标识用于标识是否需要对所述用户请求的数据进行打包；

10 所述快速转发模块包括第一判断单元和打包单元；

所述第一判断单元，用于根据所述数据描述信息读取所述用户请求的数据后，通过所述打包选择标识判断是否需要对所述用户请求的数据进行打包；

15 所述打包单元，用于若所述第一判断单元判断的结果为需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。

16、如权利要求14所述的系统，其特征在于，所述配置单元还用于将内存选择标识携带在所述数据描述信息，所述内存选择标识用于标识是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

所述快速转发模块包括第二判断单元和转发单元；

20 所述第二判断单元，用于通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

25 所述转发单元，用于若所述第二判断单元判断的结果为是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所述网络接口控制器，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

17、如权利要求14所述的系统，其特征在于，所述快速转发模块根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器具体为：所述快速转发模块根据所述数据描述信息，以直接内

存访问方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。

18、如权利要求15所述的系统，其特征在于，所述中央处理器还包括：

5 下发模块，用于将打包模式标识下发至所述快速转发模块，所述打包模式标识与打包模式对应。

19、一种数据快速分发装置，其特征在于，所述装置包括：

接收模块，用于接收中央处理器发送的数据描述信息，所述数据描述信息包括用户请求的数据的存储地址和长度信息；

10 数据读取模块，用于根据所述数据描述信息读取所述用户请求的数据并将所述用户请求的数据转发至网络接口控制器，以由所述网络接口控制器向用户发送所述用户所请求的数据。

20、如权利要求19所述的装置，其特征在于，所述数据描述信息还包括打包模式标识和打包选择标识，所述打包模式标识与所述请求数据的用户对应，所述打包选择标识用于标识是否需要对所述用户请求的数据进行打包；

15 所述数据读取模块包括第一判断单元和打包单元；

所述第一判断单元，用于根据所述数据描述信息读取所述用户请求的数据后，通过所述打包选择标识判断是否需要对所述用户请求的数据进行打包；

20 所述打包单元，用于若所述第一判断单元判断的结果为需要对所述用户请求的数据进行打包，则对所述用户请求的数据进行打包后转发至所述网络接口控制器。

21、如权利要求19所述的装置，其特征在于，所述数据描述信息还包括内存选择标识，所述内存选择标识用于标识是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

所述数据读取模块包括第二判断单元和转发单元；

25 所述第二判断单元，用于通过所述内存选择标识判断是从内存读取所述用户请求的数据还是从存储设备读取所述用户请求的数据；

所述转发单元，用于若所述第二判断单元判断的结果为是从存储设备读取所述用户请求的数据，则从所述存储设备读取所述用户请求的数据后转发至所

述网络接口控制器，否则，从所述内存读取所述用户请求的数据后转发至所述网络接口控制器。

- 22、如权利要求19所述的装置，其特征在于，所述数据读取模块具体用于根据所述数据描述信息，以直接内存访问方式读取所述用户请求的数据并将所述用户请求的数据转发至所述网络接口控制器。
- 5

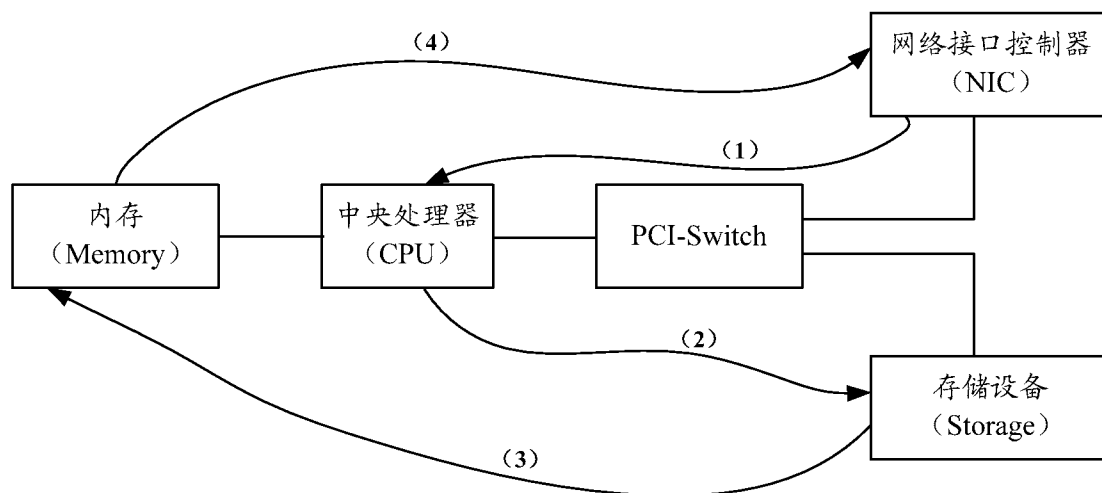


图 1

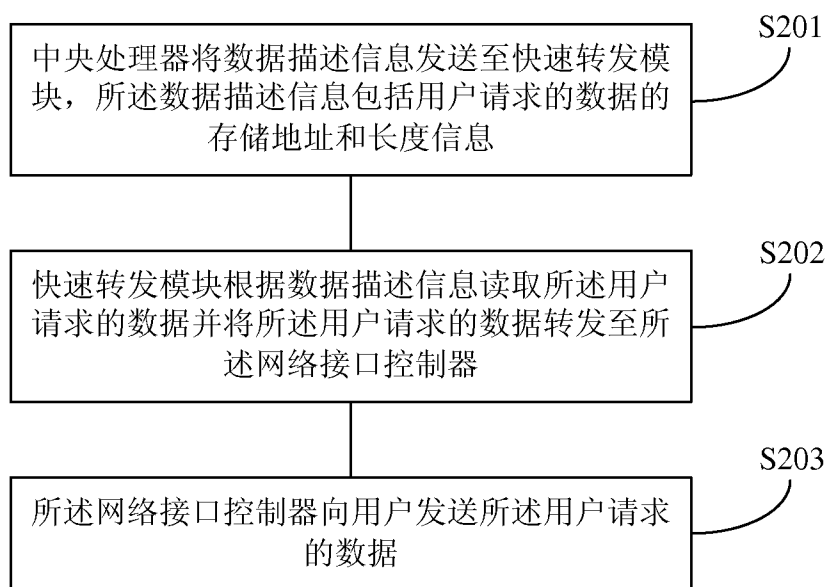


图 2

内存选择 (Memory Select)	打包选择标识 (Package Select)	打包模式标识 (Package ID)	缓存地址 (Buffer Address)
长度 (Length)	• • •		

图 3

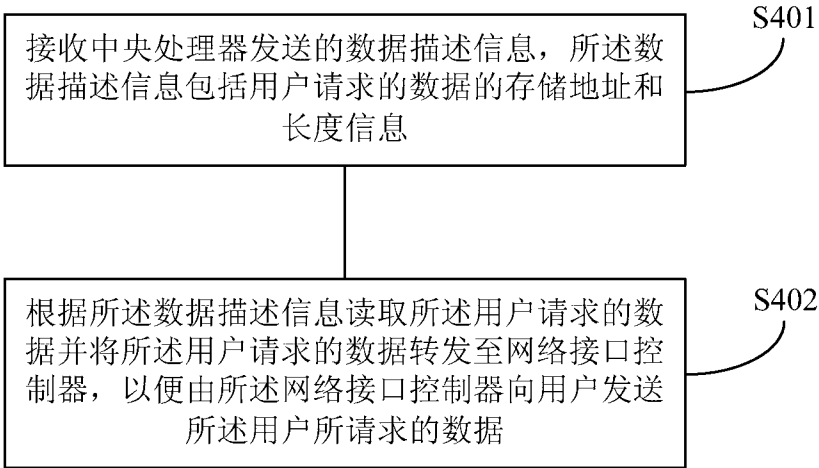


图 4

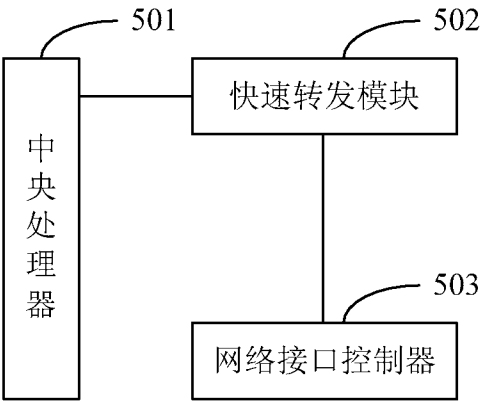


图 5

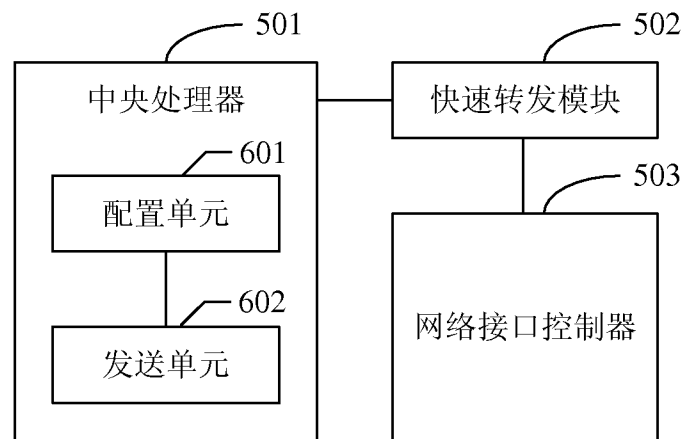


图 6

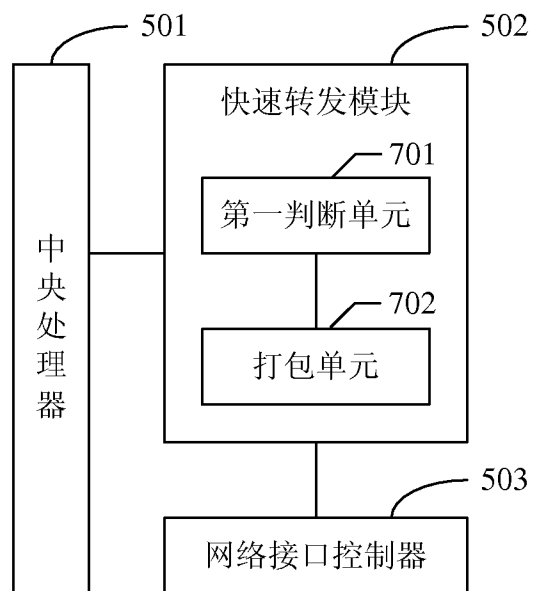


图 7

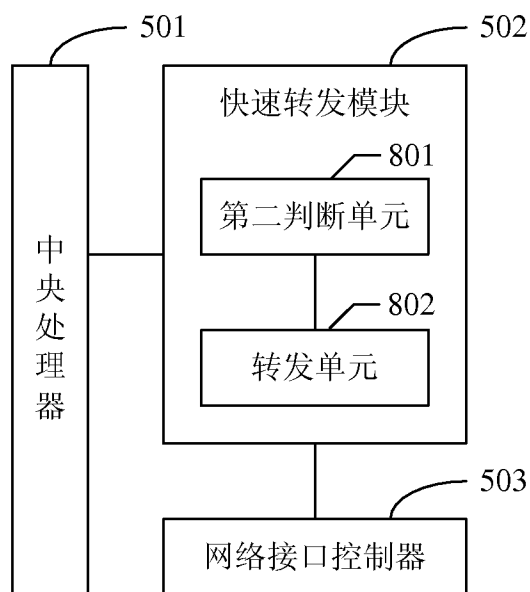


图 8

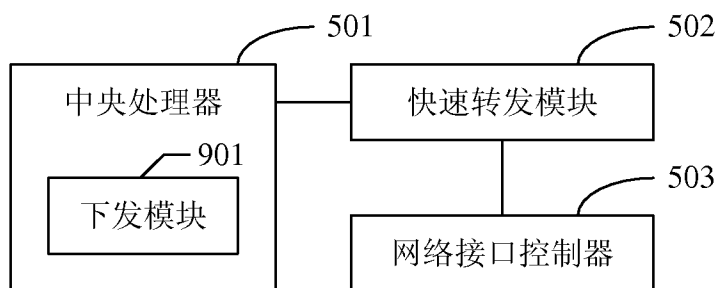


图 9

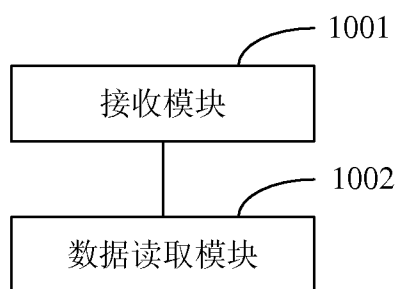


图 10

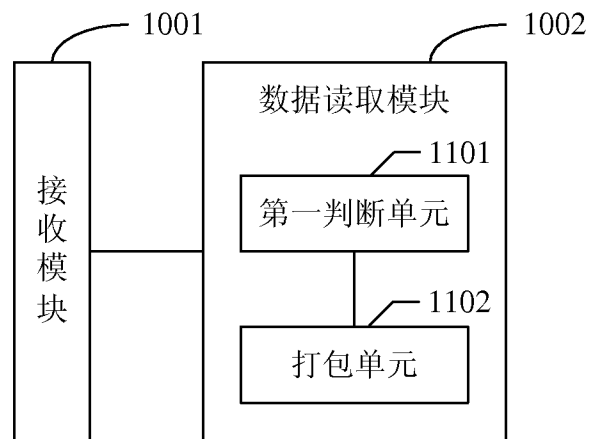


图 11

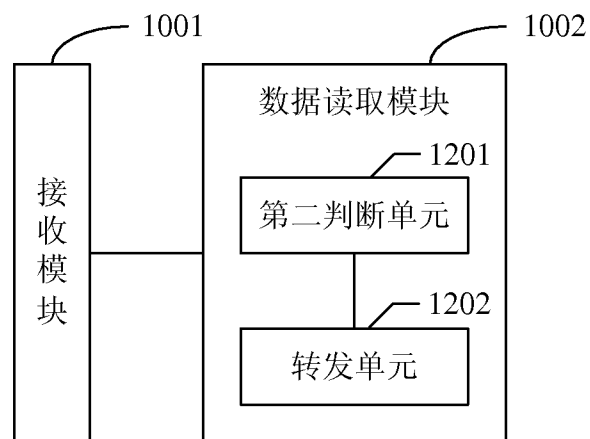


图 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2011/081366

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:H04L, H04N

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,CNXTX,CNKI,VEN:DMA, NIC, CDN, storage, data, request, content w delivery w network, description, network w interface

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Chen, Zesheng et al, Study and Realization of Interface Technology in ATM Network Adapter, Communications Technology, No.1, 2001 page 34	1-2,6,8,12-14,19,22
A	pages 34-36	3-5,7,9-11,15-18,20-21
A	CN1871587A(THE BOARD OF GOVERNORS FOR HIGHER EDUCATION STATE OF RHODE ISLAND: PROVIDENCE PLANTATION) 29 Nov. 2006(29.11.2006) the whole document	1-22

☐ 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
10 Jul. 2012(10.07.2012)

Date of mailing of the international search report
16 Aug. 2012(16.08.2012)

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10)62019451

Authorized officer
SUN, Zhiling
Telephone No. (86-10)62411252

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2011/081366

Patent Documents referred	Publication Date	Patent Family	Publication Date
CN1871587A	29.11.2006	US2005144223A1	30.06.2005
		CN100428185C	22.10.2008
		EP1690185A1	16.08.2006
		JP2007510978T	26.04.2007
		WO2005043395A1	12.05.2005
		EP1690185A4	04.04.2007

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/081366

CLASSIFICATION OF SUBJECT MATTER

H04L 29/06(2006.01)i

H04N 21/238(2011.01)i

国际检索报告

国际申请号
PCT/CN2011/081366

A. 主题的分类

参见附加页

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)

IPC:H04L,H04N

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,CNTEXT,CNKI,VEN:直接内存访问, DMA, 网络接口控制器, NIC, 数据, 存储, 描述, 请求, 内容
分发网络, CDN, storage, data, request, content w delivery w network, description, network w interface

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
X	陈则盛等, ATM 网络适配器接口技术研究与实现, 通信技术, 2001 年, 第 1 期,	1-2,6,8,12-14,19,22
A	第 34 页	3-5,7,9-11,15-18,20-21
A	第 34-36 页	
	CN1871587A(罗得岛及普罗维登斯属地高等教育管理委员会)	1-22
	29.11 月 2006 (29.11.2006) 全文	

☐ 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
10.7 月 2012 (10.07.2012)

国际检索报告邮寄日期
16.8 月 2012 (16.08.2012)

ISA/CN 的名称和邮寄地址:
中华人民共和国国家知识产权局
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

孙志玲
电话号码: (86-10) **62411252**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2011/081366

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN1871587A	29.11.2006	US2005144223A1	30.06.2005
		CN100428185C	22.10.2008
		EP1690185A1	16.08.2006
		JP2007510978T	26.04.2007
		WO2005043395A1	12.05.2005
		EP1690185A4	04.04.2007

主题的分类

H04L29/06(2006.01)i

H04N21/238(2011.01)i