

(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 1 月 9 日 (09.01.2014)



(10) 国际公布号
WO 2014/005261 A1

- (51) 国际专利分类号:
G06F 13/00 (2006.01)
- (21) 国际申请号: PCT/CN2012/078032
- (22) 国际申请日: 2012 年 7 月 2 日 (02.07.2012)
- (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): 杭州华为数字技术有限公司 (HANGZHOU HUAWEI DIGITAL TECHNOLOGIES CO., LTD.) [CN/CN]; 中国浙江省杭州市滨江区滨兴路 301 号 3 幢 A 楼 301 室, Zhejiang 310053 (CN)。
- (72) 发明人: 及
- (75) 发明人/申请人 (仅对美国): 常胜 (CHANG, Sheng) [CN/CN]; 中国浙江省杭州市滨江区滨兴路 301 号 3 幢 A 楼 301 室, Zhejiang 310053 (CN)。王工艺 (WANG, Gongyi) [CN/CN]; 中国浙江省杭州市滨江区滨兴路 301 号 3 幢 A 楼 301 室, Zhejiang 310053 (CN)。李涛 (LI, Tao) [CN/CN]; 中国浙江省杭州市滨江区滨兴路 301 号 3 幢 A 楼 301 室, Zhejiang 310053 (CN)。
- (74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国北京市海淀区西直门北大街 32 号枫蓝国际 A 座 8F-6, Beijing 100082 (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, 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: STORAGE EXPANSION DEVICE AND SERVER

(54) 发明名称: 存储扩展装置及服务器

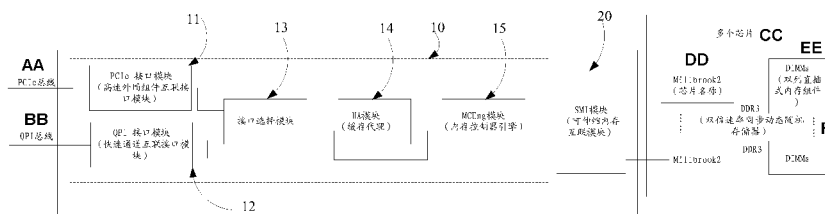


图 1 / FIG.1

11 PCIe INTERFACE MODULE (PERIPHERAL COMPONENT INTERCONNECT EXPRESS INTERFACE MODULE)
12 QPI INTERFACE MODULE (QUICK PATH INTERCONNECT INTERFACE MODULE)
13 INTERFACE SELECTION MODULE
14 HA MODULE (CACHING PROXY)
15 MCENG MODULE (INTERNAL MEMORY CONTROLLER ENGINE)
20 SERVER
AA PCIe BUS
BB QPI BUS
CC A PLURALITY OF CPUs
DD A PLURALITY OF CHIPS
EE DIMMS (DOUBLE IN-LINE MEMORY MODULES)
FF DDR3 (DOUBLE DATA RATE SYNCHRONOUS DYNAMIC RANDOM ACCESS MEMORY)

(57) Abstract: Provided are a storage expansion device (10) and server, the storage expansion device (10) comprising: a quick path interconnect (QPI) interface module (12) in communication with a central processing unit (CPU) via a QPI bus; a peripheral component interconnect express (PCIe) interface module (11) in communication with the CPU via a PCIe bus; an interface selection module (13) connected with the QPI interface module (12) and the PCIe interface module (11) respectively; a caching proxy HA module (14) connected with the interface selection module (13); and an internal memory controller engine (MCEng) module (15) connected with the HA module (14) and the interface selection module (13) respectively. The storage expansion device can be used as a CPU internal memory capacity expansion device, and can also be used as the storage expansion hardware of a storage IO.

(57) 摘要: 提供了一种存储扩展装置 (10) 及服务器, 其中, 该存储扩展装置 (10) 包括: 快速通道互联 QPI 接口模块 (12), 通过 QPI 总线与中央处理器 CPU 通信; 高速外围组件互联 PCIe 接口模块 (11), 通过 PCIe 总线与所述 CPU 通信; 接口选择模块 (13), 分别与所述 QPI 接口模块 (12)、所述 PCIe 接口模块 (11) 连接; 缓存代理 HA 模块 (14), 与所述接口选择模块 (13) 连接; 内存控制器引擎 MCEng 模块 (15), 分别与所述 HA 模块 (14)、所述接口选择模块 (13) 连接。上述存储扩展装置可以作为 CPU 内存容量的扩展设备, 还可以作为存储 IO 的存储扩展硬件。



本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

存储扩展装置及服务器

技术领域

5 本发明实施例涉及通信技术，尤其涉及一种存储扩展装置及服务器。

背景技术

在服务器中，计算资源与存储资源是整机性能的关键因素。随着微电子技术的飞速发展，处理器的性能在成倍的提高。虽然处理器技术在不断进步，但内存容量仍然受到限制，内存容量的限制造成了服务器的利用率无法随着处理器的性能提高而提高。其原因如下：

第一，现在处理器性能的提高更多是通过内核数量和缓存，而非主频；第二，内核数、缓存和线程数随中央处理器（Central Processing Unit，简称 CPU）工作频率的降低而增加，内存增长不足，导致失衡；第三，15 64 位的应用和操作系统需要更多内存。

从整个服务器的利用率的角度来看，服务器的利用率的瓶颈不再受限于 CPU 的性能，而主要受限于内存的容量与内存数据的读写速率。

当前，内存数据的读写速率受制造工业技术的限制，存储器的读写速度提升也相当困难。这样，就只能从扩展内存容量的方向考虑。如何扩展20 内存容量，可以达到既适用于多种接口多种配置的内存扩展设备，同时还可以切换为存储扩展硬件，从而提升服务器性能成为当前需要解决的技术问题。

发明内容

25 有鉴于此，本发明实施例提供一种存储扩展装置及服务器，能够作为 CPU 内存容量的扩展设备，还能够切换为存储输入输出（Input & Output，简称 IO）设备的存储扩展硬件。

一方面，本发明实施例提供的存储扩展装置，包括：

快速通道互连 QPI 接口模块，通过 QPI 总线与中央处理器 CPU 通信；
30 高速外围组件互联 PCIe 接口模块，通过 PCIe 总线与所述 CPU 通信；

接口选择模块,分别与所述 QPI 接口模块、所述 PCIe 接口模块连接;
缓存代理 HA 模块,与所述接口选择模块连接;

内存控制器引擎 MCEng 模块,分别与所述 HA 模块、所述接口选择模块连接;

5 当所述存储扩展装置作为存储 IO 接口设备时,所述 PCIe 接口模块用于完成所述 PCIe 总线中 PCIe 数据的接收,并对所述 PCIe 数据中 PCIe 协议存储器事务进行处理,以及将处理的 PCIe 协议存储器事务通过所述接口选择模块发送至所述 MCEng 模块,所述 MCEng 模块完成读写驱动功能并提供接口与存储 IO 设备相连;或者

10 当所述存储扩展装置作为内存扩展设备时,所述 QPI 接口模块用于完成所述 QPI 总线中 QPI 数据的接收,获取所述 QPI 数据中的 QPI 协议事务并发送给所述接口选择模块,进而通过所述接口选择模块将所述 QPI 数据中的 QPI 协议事务输出至所述 HA 模块,以及所述 HA 模块对所述 QPI 数据中的 QPI 协议事务进行处理,并将处理完成的 QPI 协议事务发
15 送至所述 MCEng 模块,所述 MCEng 模块完成读写驱动功能并提供接口与扩展内存相连;或者,

当所述存储扩展装置作为内存扩展设备时,所述 PCIe 接口模块用于完成所述 PCIe 总线中 PCIe 数据的接收,获取所述 PCIe 数据中的 QPI 协议事务并发送给所述接口选择模块,进而通过所述接口选择模块将所述
20 PCIe 数据中的 QPI 协议事务输出至所述 HA 模块,以及所述 HA 模块对所述 PCIe 数据中的 QPI 协议事务进行处理,并将处理完成的所述 PCIe 数据中的 QPI 协议事务发送至所述 MCEng 模块,所述 MCEng 模块完成读写驱动功能并提供接口与扩展内存相连。

在一种可能的实现方式中,上述的存储扩展装置,还包括:

25 SMbus 选择模块,与所述接口选择模块连接,用于向所述接口选择模块发送选择指令,以使所述接口选择模块根据所述选择指令,选择接收所述 QPI 接口模块发送的所述 QPI 接口模块获取的 QPI 协议事务;或者选择接收所述 PCIe 接口模块发送的所述 PCIe 接口模块获取的 QPI 协议事务,或者选择接收所述 PCIe 接口模块发送的所述处理的 PCIe 协议存储
30 器事务。

在另一种可能的实现方式中，上述的存储扩展装置，还包括：

选择指令切换模块，与所述 CPU 连接，并分别与所述 QPI 接口模块和所述 PCIe 接口模块连接，用于使所述 CPU 输出的数据通过所述 QPI 总线传输至所述 QPI 接口模块，以及向所述 SMBus 选择模块发送第一内存扩展指令，所述 SMBus 选择模块根据所述第一内存扩展指令向所述接口选择模块发送所述选择指令；

或者，用于使所述 CPU 输出的数据，经过接口适配后，通过所述 PCIe 总线传输至所述 PCIe 接口模块，以及向所述 SMBus 选择模块发送第二内存扩展指令，所述 SMBus 选择模块根据所述第二内存扩展指令向所述接口选择模块发送所述选择指令；

或者，用于使所述 CPU 输出的数据，通过所述 PCIe 总线传输至所述 PCIe 接口模块，以及向所述 SMBus 选择模块发送存储指令，所述 SMBus 选择模块根据所述存储指令向所述接口选择模块发送所述选择指令。

另一方面，本发明实施例提供的服务器，包括：CPU，和本发明任一所述的存储扩展装置，所述 CPU 通过 QPI 总线与所述存储扩展装置中的 QPI 接口模块连接，以及所述 CPU 通过 PCIe 总线与所述存储扩展装置中的 PCIe 接口模块连接。

由上述技术方案可知，本发明实施例的存储扩展装置及服务器，通过 QPI 接口模块和 PCIe 接口模块与 CPU 通信，通过接口选择模块选择接收 QPI 接口模块或 PCIe 接口模块输出的数据，若存储扩展装置作为存储 IO 设备，接口选择模块输出的数据可以直接输出至 MCEng 模块，若存储扩展装置作为内存扩展设备，接口选择模块输出的数据经过 HA 模块处理之后输出至 MCEng 模块，实现了上述的存储扩展装置可作为 CPU 内存容量的 QPI 接口和 PCIe 接口扩展设备，还可切换为存储 IO 的存储扩展硬件，从而提升服务器性能。

附图说明

为了更清楚地说明本发明的技术方案，下面将对实施例中所需要使用的附图作一简单地介绍，显而易见地：下面附图只是本发明的一些实施例的附图，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同样能实现本发明技术方案的其它附图。

图 1 为本发明一实施例提供的存储扩展装置的结构示意图；

图 2A 至图 2C 为本发明一实施例提供的存储扩展装置的部分结构示意图；

5 图 3 为本发明另一实施例提供的存储扩展装置的部分结构示意图；

图 4 为本发明另一实施例提供的存储扩展装置的部分结构示意图；

图 5A 和图 5B 为本发明另一实施例提供的存储扩展装置的部分结构示意图；

10 图 6 至图 8 为本发明另一实施例提供的存储扩展装置的实际应用的场景图；

图 9 为本发明另一实施例提供的存储扩展装置的使用方法示意图。

具体实施方式

15 为使本发明的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明的技术方案进行清楚、完整地描述。显然，下述的各个实施例都只是本发明一部分的实施例。基于本发明下述的各个实施例，本领域普通技术人员即使没有作出创造性劳动，也可以通过等效变换部分甚至全部的技术特征，而获得能够解决本发明技术问题，实现本发明技术效果的其它实施例，而这些变换而来的各个实施例显然并不脱离本发明所公开的范围。

20 本发明实施例主要是提供一种可以适用于多种接口多种配置的存储扩展装置，用以方便灵活地提升服务器的内存/存储容量，从而达到提升服务器的性能。

25 结合图 1 至图 2C 所示，图 1 示出了本发明一实施例提供的存储扩展装置的结构示意图，图 2A 至图 2C 示出了本发明一实施例提供的存储扩展装置的部分结构示意图，其中，为方便查看图 1 中所示的存储扩展装置，采用 2A、图 2B 和图 2C 分别示出存储扩展装置的部分结构，如图所示，图 2A 示出了图 1 中的部分结构示意图，图 2B 示出了存储扩展装置 10 中的 CPU 通过 PCIe 总线连接 PCIe 接口模块 11 的示意图，图 2C 示出了
30 存储扩展装置 10 中的 CPU 通过 QPI 总线连接 QPI 接口模块 12 的示意图。

本实施例的存储扩展装置 10（如图 1 中的虚线框中的结构）包括：
快速通道互联（Quick Path Interconnect，简称 QPI）接口模块 12 通过
QPI 总线与 CPU 通信，具体的 QPI 接口模块 12 可以与 CPU 的 QPI 接口
（图 1 中未示出）连接，进而实现与 CPU 通信；高速外围组件互联
5 （Peripheral Component Interconnect Express，简称 PCIe）接口模块
11 通过 PCIe 总线与 CPU 通信，具体的 PCIe 接口模块 11 可以与 CPU
的 QPI 接口连接，进而实现与 CPU 通信；接口选择模块 13 分别与 QPI
接口模块 12、PCIe 接口模块 11 连接；缓存代理（Home Agent，简称
HA）模块 14 与接口选择模块 13 连接；内存控制器引擎（Memory Controller
10 Engine，简称 MCEng）模块 15 分别与 HA 模块 14、接口选择模块 13 连
接；

当存储扩展装置 10 作为存储 IO 设备接口时，PCIe 接口模块 11 用于
完成 PCIe 总线中 PCIe 数据的接收，并对 PCIe 数据中 PCIe 协议存储器
事务进行处理，以及将处理后的 PCIe 协议存储器事务通过接口选择模块
15 13 发送至 MCEng 模块 15，MCEng 模块完成读写驱动功能并提供接口与
存储 IO 设备相连；

当存储扩展装置 10 作为内存扩展设备时，QPI 接口模块 12 用于完成
QPI 总线中 QPI 数据的接收，获取 QPI 数据中的 QPI 协议事务并发送给
接口选择模块 13，进而通过接口选择模块 13 将 QPI 接口模块获取的 QPI
20 协议事务输出至 HA 模块 14，以及 HA 模块 14 对 QPI 接口模块获取的
QPI 协议事务进行处理，并将处理完成的 QPI 协议事务发送至 MCEng 模
块 15，MCEng 模块完成读写驱动功能并提供接口与扩展内存相连；

或者，当存储扩展装置 10 作为内存扩展设备时，PCIe 接口模块 11
用于完成 PCIe 总线中 PCIe 数据的接收，获取携带在 PCIe 数据中的 QPI
25 协议事务并发送给接口选择模块 13，进而通过接口选择模块 13 将 PCIe
接口模块 11 获取的 QPI 协议事务输出至 HA 模块 14，以及 HA 模块 14
对 PCIe 接口模块获取的 QPI 协议事务进行处理，并将处理完成的 PCIe
数据中的 QPI 协议事务发送至 MCEng 模块 15，所述 MCEng 模块完成读
写驱动功能并提供接口与扩展内存相连。

30 需要说明的是，在实际应用中，CPU 还需读取存储扩展装置 10 中缓

存/存储的相关数据。此时，上述各数据链路中数据的流向则按照原有数据链路的反向流入 CPU 中。例如，在存储扩展装置作为存储 IO 接口设备时，CPU 若需要读取存储在存储扩展装置中的数据，则 MCEng 模块 15 根据 CPU 的读取指令将相对应的数据发送至接口选择模块 13，接口选择模块 13 将相对应的数据发送至 PCIe 接口模块 11，该 PCIe 接口模块 11 通过 PCIe 总线将相对应的数据经过适配后传输至 CPU。

需要说明的是，在实际应用中，当存储扩展装置 10 作为内存扩展设备时，则 PCIe 接口模块 11 无法直接和 CPU 的 QPI 接口相连，需要通过一个 PCIe 协议与 QPI 协议的转换模块之后才能相连，用于处理 QPI 协议的相关内容。则上述的 PCIe 接口模块 11 通过 PCIe 总线与 CPU 的 QPI 接口连接，实际上是 PCIe 接口模块 11 经过接口适配后通过 PCIe 总线与 CPU 的 QPI 接口连接。

相应地，PCIe 接口模块 11 用于完成 PCIe 总线中 PCIe 数据的接收与发送；QPI 接口模块 12 也用于完成 QPI 总线中 QPI 数据的接收与发送。上述图 1 中所示的各模块的功能包括现有的各模块所能实现的功能，本实施例对各模块的功能仅为举例说明。

在实际应用中，MCEng 模块提供接口与扩展的内存或 IO 设备相连，具体可以为通过数据接口模块连接至少一个芯片，该数据接口模块可以为图 1 中所示的可伸缩内存互联(Scalable Memory Interconnect, 简称 SMI) 模块 20，或者为其他现有的数据接口模块。图 1 中所示的 SMI 模块可以是一种总线，用于增加双倍速率同步动态随机存储器(Double Data Rate RAM 3 ，简称 DDR3) 的驱动。在其他实施例中，数据接口模块可根据实际需要设置。

需要说明的是，前述的 MCEng 模块 15 主要是对内存进行控制，例如，对内存设备进行初始化、根据内存的状态向内存发送各种命令以对内存的运行状态进行控制、实现芯片对内存的访问等。

由上述实施例可知，本实施例的存储扩展装置，通过 QPI 接口模块和 PCIe 接口模块与 CPU 通信，进而通过接口选择模块选择接收 QPI 接口模块或 PCIe 接口模块输出的数据，若存储扩展装置作为存储 IO 设备，接口选择模块输出的数据可以直接输出至 MCEng 模块，若存储扩展装置作为

内存扩展设备，接口选择模块输出的数据经过 HA 模块处理之后输出至 MCEng 模块，实现了上述的存储扩展装置可作为 CPU 内存容量的 QPI 接口和 PCIe 接口扩展设备，即多种接口的内存扩展设备，还可切换为存储 IO 的存储扩展硬件，从而提升服务器性能。

5 如图 3 所示，图 3 示出了本发明另一实施例提供的存储扩展装置的部分结构示意图，图 3 所示的存储扩展装置与图 1 所示的存储扩展装置 10 不同的是，本实施例中的存储扩展装置还包括：系统管理总线（System Management Bus，简称 SMbus）选择模块 16，该 SMbus 选择模块 16 与接口选择模块 13 连接，用于向接口选择模块 13 发送选择指令，以使接
10 口选择模块 13 根据选择指令，选择接收 QPI 接口模块 12 发送的 QPI 接口模块获取的 QPI 协议事务，或选择接收 PCIe 接口模块 11 发送 PCIe 接口模块获取的 QPI 协议事务，或者选择接收 PCIe 接口模块 11 发送的 PCIe 协议存储器事务。需要注意的是，图 3 中未示出存储扩展装置中的 HA 模块和 MCEng 模块。该处 HA 模块和 MCEng 模块的结构和连接关系
15 与图 1 中所示的基本一致。

也就是说，在图 3 中所示的存储扩展装置作为存储 IO 接口设备时，SMbus 选择模块 16，向接口选择模块 13 发送选择指令，以使接口选择模块 13 根据选择指令，选择接收 PCIe 接口模块 11 发送的所处理的 PCIe 协议存储器事务。

20 在图 3 中所示的存储扩展装置作为内存扩展设备时，若此时 CPU 的 QPI 接口通过 PCIe 总线连接 PCIe 接口模块 11，则 SMbus 选择模块 16，向接口选择模块 13 发送另一选择指令，以使接口选择模块 13 根据另一选择指令，选择接收 PCIe 接口模块 11 发送的 PCIe 接口模块 11 获取的 QPI 协议事务；

25 在图 3 中所示的存储扩展装置作为内存扩展设备时，若此时 CPU 的 QPI 接口通过 QPI 总线连接 QPI 接口模块 12，则 SMbus 选择模块 16，向接口选择模块 13 发送再一选择指令，以使接口选择模块 13 根据再一选择指令，选择接收 QPI 接口模块 12 发送的 QPI 协议事务。

由上述实施例可知，本实施例的存储扩展装置，通过 SMbus 选择模
30 块使接口选择模块选择接收 QPI 接口模块 12 或者 PCIe 接口模块的数据，

以便实现上述的存储扩展装置可作为 CPU 内存容量的扩展设备，或者作为存储 IO 的存储扩展硬件。

如图 4 所示，图 4 示出了本发明另一实施例提供的存储扩展装置的部分结构示意图，图 4 所示的存储扩展装置与图 3 所示的存储扩展装置不同的是，本实施例中的存储扩展装置还包括：选择指令切换模块 17；

选择指令切换模块 17 与 CPU 连接，并分别与 QPI 接口模块和 PCIe 接口模块连接，用于使述 CPU 输出的数据通过 QPI 总线传输至所述 QPI 接口模块，以及向 SMBus 选择模块发送第一内存扩展指令，SMBus 选择模块根据第一内存扩展指令向所述接口选择模块发送选择指令；或者，用于使 CPU 输出的数据，经过接口适配后，通过 PCIe 总线传输至 PCIe 接口模块，以及向 SMBus 选择模块发送第二内存扩展指令，SMBus 选择模块根据第二内存扩展指令向接口选择模块发送选择指令；或者，用于使 CPU 输出的数据，通过 PCIe 总线传输至 PCIe 接口模块，以及向 SMBus 选择模块发送存储指令，SMBus 选择模块根据存储指令向接口选择模块发送选择指令。

在具体应用中，该选择指令切换模块 17 与 CPU 的 QPI 接口连接，并分别与 QPI 接口模块和 PCIe 接口模块连接，用于使 CPU 的 QPI 接口输出的数据通过 QPI 总线传输至 QPI 接口模块 12，以及向 SMBus 选择模块 16 发送第一内存扩展指令；相应地，SMBus 选择模块 16 根据第一内存扩展指令向接口选择模块 13 发送选择指令，使得接口选择模块 13 根据选择指令，选择接收 QPI 接口模块 12 发送的 QPI 协议事务。

或者，

该选择指令切换模块 17 与 CPU 的 QPI 接口连接，并分别与 QPI 接口模块和 PCIe 接口模块连接，用于使 CPU 的 QPI 接口输出的数据，经过接口适配后，通过 PCIe 总线传输至 PCIe 接口模块 11，以及向 SMBus 选择模块发送第二内存扩展指令；所述 SMBus 选择模块根据第二内存扩展指令向接口选择模块发送另一选择指令，使得接口选择模块 13 根据另一选择指令，选择接收 PCIe 接口模块 11 发送的 PCIe 接口模块 11 中获取的 QPI 协议事务，后续数据的流向如图 5B 所示；

或者，该选择指令切换模块 17 与 CPU 的 QPI 接口连接，并分别与

QPI 接口模块和 PCIe 接口模块连接，用于使 CPU 的 QPI 接口输出的数据，通过 PCIe 总线传输至 PCIe 接口模块 11，以及向 SMBus 选择模块发送存储指令，所述 SMBus 选择模块根据存储指令向接口选择模块发送再一选择指令，选择接收 PCIe 接口模块 11 发送的 PCIe 协议存储器事务，
5 后续数据的流向如图 5A 所示。

需要注意的是，图 4 中未示出存储扩展装置中的 HA 模块和 MCEng 模块。该处 HA 模块和 MCEng 模块的结构和连接关系与图 1 中所示的基本一致。

进一步地，在实际应用中，若存储扩展装置作为内存扩展设备在使用，
10 此时，用户感觉内存扩展设备不再需要，但需要较大的存储器设备时，用户可触发上述的选择指令切换模块 17 向 SMBus 选择模块 16 发送存储指令，用以实现存储扩展配置。

由上述实施例可知，本实施例的存储扩展装置，通过选择指令切换模块触发 SMBus 选择模块向接口选择模块发送相关的选择指令，进而使得
15 接口选择模块选择接收 QPI 接口模块 12 或者 PCIe 接口模块 11 的数据，以便实现上述的存储扩展装置可作为 CPU 内存容量的扩展设备，或者作为存储 IO 的存储扩展设备。进一步地，用户还可使用选择指令切换模块切换存储扩展装置的使用属性。

在实际应用中，前述的 QPI 接口模块 12 主要完成的是 QPI 总线输出的
20 的 QPI 数据的接收，以及对 QPI 数据中 QPI 协议事务的获取等。当然，QPI 接口模块 12 具体用于实现 QPI 协议数据的处理，完成 QPI 数据的接收、发送等，例如实现 QPI 协议的物理层、链路层和包解析的功能。实际的 QPI 接口模块 12 传输的数据的速率与 CPU 的 QPI 接口中传输的速率相同。

前述的 PCIe 接口模块 11 主要完成的是 PCIe 总线输出的 PCIe 数据的接收，以及对 PCIe 协议存储器事务的处理，进一步还用于获取加载到
25 PCIe 消息事务类型中的 QPI 协议事务。可以理解的是，PCIe 接口模块 11 用于实现标准的 PCIe 协议的传输功能。当然，为了尽量匹配 QPI 总线传输的速率，上述任一实施例中的 PCIe 接口模块 11 可采用具体的硬件为
30 PCIe3.0 2X8 端口，或 PCIe3.0 X16 端口。如图 4 所示，接口选择模块

13 可接收 SMBus 选择模块发送的选择指令，进而根据选择指令相适应的处理相对应的数据，如选择处理 PCIe 接口模块 11 传输的数据或者选择处理 QPI 接口模块 12 的传输的数据。

本实施例中的 HA 模块，用于实现高速缓存一致性（Cache Coherence，简称 CC）协议，如处理 QPI 协议中的 CC 事务操作，保持内存的一致性。

MCEng 模块 15 用于实现内存命令的转换和数据的加密分发等功能，例如目录高速缓存、数据高速缓存等，以及实现双列直插式内存组件（dual in-line memory module，简称 DIMM）的访问和基本稳定性、可靠性、可服务性（Reliability, Availability and Serviceability，简称 RAS）等特性。

由上述实施例可知，本发明实施例的存储扩展装置，在用户的选择下，该存储扩展装置可以作为扩展内存配置使用，还可以作为存储 IO 的存储扩展硬件使用。

另外，需要说明的是，在图 5A 中，虚线框中所示的存储扩展装置可作为存储 IO 功能的设备，如硬盘、固态硬盘（solid state disk，简称 SSD）等。此时，在 CPU 中，存储扩展装置可映射为 PCIe 设备。

在图 5B 中，虚线框中所示的存储扩展装置可作为内存扩展设备，此时，在 CPU 中，存储扩展装置可映射为 QPI 可见。

举例来说，当用户需配置上述的存储扩展装置适用于图 6 所示的应用场景（即作为存储 IO 设备）时，触发 SMBus 选择模块 16 根据选择指令驱动接口选择模块 13 接收 PCIe 接口模块输出的包含 PCIe 协议存储器事务的数据，并将该数据直接输入到 MCEng 模块 15，实现存储 IO 的配置使用，如图 5A 虚线框所示的数据链路。

举例来说，当用户需配置上述的存储扩展装置适用于图 7 所示的应用场景（即作为内存扩展设备）时，触发 SMBus 选择模块 16 根据选择指令驱动接口选择模块 13 接收 PCIe 接口模块输出的包含 QPI 协议事务的数据，并将该数据输入 HA 模块 14，进而 HA 模块 14 处理后的数据输入到 MCEng 模块 15，实现内存扩展的配置使用，如图 5B 虚线框所示的数据链路。

举例来说，当用户需配置上述的存储扩展装置适用于图 8 所示的应用

场景（即作为内存扩展设备）时，触发 SMBus 选择模块 16 根据选择指令驱动接口选择模块 13 接收 QPI 接口模块 12 的数据，并将该数据输入 HA 模块 14，进而 HA 模块 14 处理后的数据输入到 MCEng 模块 15，实现内存扩展的配置使用。

- 5 参见图 9 所示，图 9 示出了本发明另一实施例提供的存储扩展装置的使用方法示意图，具体地，存储扩展装置的使用方法如下所述。

101、用户将设备组件（如上所述的存储扩展装置）插入服务器中相对应的位置。

- 102、选择指令切换模块使 CPU 的 QPI 接口通过 PCIe 总线连接 PCIe 接口模块，或者进行接口适配后连接 PCIe 接口模块，或者通过 QPI 总线连接 QPI 接口模块，若 CPU 的 QPI 接口通过 PCIe 总线连接 PCIe 接口模块，则执行如下步骤 103，否则执行如下的步骤 113。

103、在步骤 102 中，若用户使 CPU 的 QPI 接口通过 PCIe 总线连接 PCIe 接口模块，则对服务器上电。

- 15 需要说明的是，在实际应用中，PCIe 接口模块实现的是内存扩展功能时，则 PCIe 接口模块无法直接和 CPU 的 QPI 接口相连，需要通过一个 PCIe 协议与 QPI 协议的转换模块之后才能相连，用于处理 QPI 协议的相关内容，进而上述是 CPU 的 QPI 接口通过 PCIe 总线进行接口适配后连接 PCIe 接口模块。

- 20 可以理解的是，该步骤 103 还可以是用户使 CPU 的 QPI 接口通过 PCIe 总线进行接口适配后连接 PCIe 接口模块，进而对服务器上电。

104、用户决定该存储扩展装置是否作为存储 IO 接口设备使用，若作为存储 IO 接口设备使用，则执行步骤 105，否则，执行步骤 125。

- 25 105、在步骤 104 中，用户确定存储扩展装置作为存储 IO 接口设备使用，则触发选择指令切换模块向 SMBus 选择模块发送存储指令，SMBus 选择模块根据存储指令向接口选择模块下发选择指令，使得接口选择模块根据选择指令选择接收 PCIe 接口模块的数据（如接收 PCIe 数据中获取的 PCIe 协议存储器事务），并将 PCIe 接口模块的数据（如 PCIe 协议存储器事务）直接传输至 MCEng 模块。

- 30 106、PCIe 数据链路建立完成，设备组件作为 PCIe 存储 IO 接口设

备使用。

进一步地，还包括步骤 107，此时，用户判断是否需要改变存储扩展装置的使用属性，若需要改变，则执行步骤 108 之后，再执行前述的步骤 104，否则，结束。

- 5 步骤 108、在用户的触发下，操作系统（Operating System，简称 OS）将存储扩展装置停止使用，然后再执行前述的步骤 104。

特别注意的是，为防止数据的遗漏，在做存储 IO 接口设备与内存扩展设备切换应用时，一定要先由 OS 将该存储扩展装置停止使用。

- 10 125、在步骤 104 中，若用户决定该存储扩展装置作为内存扩展设备使用，则触发选择指令切换模块向 SMBus 选择模块发送第二内存扩展指令，SMBus 选择模块根据第二内存扩展指令向接口选择模块下发选择指令，使得接口选择模块根据选择指令选择接收 PCIe 接口模块的数据（如 QPI 协议事务），并将 PCIe 接口模块的数据进行处理后（如将 PCIe 接口模块获取的 QPI 协议事务）传输至 HA 模块，进而 HA 模块处理后的数据
- 15 数据传输至 MCEng 模块，实现内存扩展。

126、PCIe 数据链路建立完成，设备组件作为 PCIe 内存扩展设备使用。

113、在步骤 102 中，若用户使 CPU 的 QPI 接口通过 QPI 总线连接 QPI 接口模块，则对服务器上电。

- 20 114、触发选择指令切换模块向 SMBus 选择模块发送第一内存扩展指令，SMBus 选择模块根据第一内存扩展指令向接口选择模块下发选择指令，使得接口选择模块根据选择指令选择接收 QPI 接口模块的数据（如 QPI 协议事务），并将处理后的数据传输至 HA 模块，该 HA 模块将处理后的数据传输至 MCEng 模块，实现内存扩展。

- 25 115、QPI 数据链路建立完成，设备组件作为 QPI 内存扩展设备使用。

由上，存储扩展装置可以完成多种配置（服务器扩展内存、存储 IO）的不同需求应用，且不需要增加新的硬件设备，提高了存储扩展装置的利用效率。

- 30 在实际应用中，上述的存储扩展装置可位于现场可编程门阵列（Field - Programmable Gate Array，FPGA 简称）或者，专用集成电路

(Application Specific Integrated Circuit, 简称 ASIC) 中。

根据本发明的另一方面，本发明还提供一种服务器，包括： CPU，以及本发明任意实施例所述的存储扩展装置，CPU 通过 QPI 总线与所述存储扩展装置中的 QPI 接口模块连接，以及所述 CPU 通过 PCIe 总线与
5 所述存储扩展装置中的 PCIe 接口模块连接。

举例来说，上述服务器包括：具有 QPI 接口的 CPU，以及还包括：

本发明任意实施例所述的存储扩展装置，所述 CPU 的 QPI 接口通过 QPI 总线与所述存储扩展装置中的 QPI 接口模块连接，以及所述 QPI 接口通过 PCIe 总线与所述存储扩展装置中的 PCIe 接口模块连接。

10 最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。
15

权 利 要 求 书

1、一种存储扩展装置，其特征在于，包括：

快速通道互连 QPI 接口模块，通过 QPI 总线与中央处理器 CPU 通信；

高速外围组件互联 PCIe 接口模块，通过 PCIe 总线与所述 CPU 通信；

5 接口选择模块，分别与所述 QPI 接口模块、所述 PCIe 接口模块连接；

缓存代理 HA 模块，与所述接口选择模块连接；

内存控制器引擎 MCEng 模块，分别与所述 HA 模块、所述接口选择模块连接；

10 当所述存储扩展装置作为存储 IO 接口设备时，所述 PCIe 接口模块用于完成所述 PCIe 总线中 PCIe 数据的接收，并对所述 PCIe 数据中 PCIe 协议存储器事务进行处理，以及将处理的 PCIe 协议存储器事务通过所述接口选择模块发送至所述 MCEng 模块，所述 MCEng 模块完成读写驱动功能并提供接口与存储 IO 设备相连；或者

15 当所述存储扩展装置作为内存扩展设备时，所述 QPI 接口模块用于完成所述 QPI 总线中 QPI 数据的接收，获取所述 QPI 数据中的 QPI 协议事务并发送给所述接口选择模块，进而通过所述接口选择模块将所述 QPI 数据中的 QPI 协议事务输出至所述 HA 模块，以及所述 HA 模块对所述 QPI 数据中的 QPI 协议事务进行处理，并将处理完成的 QPI 协议事务发送至所述 MCEng 模块，所述 MCEng 模块完成读写驱动功能并提供接口
20 与扩展内存相连；或者，

当所述存储扩展装置作为内存扩展设备时，所述 PCIe 接口模块用于完成所述 PCIe 总线中 PCIe 数据的接收，获取所述 PCIe 数据中的 QPI 协议事务并发送给所述接口选择模块，进而通过所述接口选择模块将所述 PCIe 数据中的 QPI 协议事务输出至所述 HA 模块，以及所述 HA 模块对
25 所述 PCIe 数据中的 QPI 协议事务进行处理，并将处理完成的所述 PCIe 数据中的 QPI 协议事务发送至所述 MCEng 模块，所述 MCEng 模块完成读写驱动功能并提供接口与扩展内存相连。

2、根据权利要求 1 所述的装置，其特征在于，还包括：

30 系统管理总线 SMBus 选择模块，与所述接口选择模块连接，用于向所述接口选择模块发送选择指令，以使所述接口选择模块根据所述选择指

令，选择接收所述 QPI 接口模块发送的所述 QPI 接口模块获取的 QPI 协议事务；或者选择接收所述 PCIe 接口模块发送的所述 PCIe 接口模块获取的 QPI 协议事务；或者选择接收所述 PCIe 接口模块发送的所述处理的 PCIe 协议存储器事务。

5 3、根据权利要求 2 所述的装置，其特征在于，还包括：

选择指令切换模块，与所述 CPU 连接，并分别与所述 QPI 接口模块和所述 PCIe 接口模块连接，用于使所述 CPU 输出的数据通过所述 QPI 总线传输至所述 QPI 接口模块，以及向所述 SMBus 选择模块发送第一内存扩展指令，所述 SMBus 选择模块根据所述第一内存扩展指令向所述接口选择模块发送所述选择指令；或者，用于使所述 CPU 输出的数据，经过接口适配后，通过所述 PCIe 总线传输至所述 PCIe 接口模块，以及向所述 SMBus 选择模块发送第二内存扩展指令，所述 SMBus 选择模块根据所述第二内存扩展指令向所述接口选择模块发送所述选择指令；或者，用于使所述 CPU 输出的数据，通过所述 PCIe 总线传输至所述 PCIe 接口模块，以及向所述 SMBus 选择模块发送存储指令，所述 SMBus 选择模块根据所述存储指令向所述接口选择模块发送所述选择指令。

10

15

4、一种服务器，包括：CPU，其特征在于，还包括：

如权利要求 1 至 3 任一所述的存储扩展装置，所述 CPU 通过 QPI 总线与所述存储扩展装置中的 QPI 接口模块连接，以及所述 CPU 通过 PCIe 总线与所述存储扩展装置中的 PCIe 接口模块连接。

20

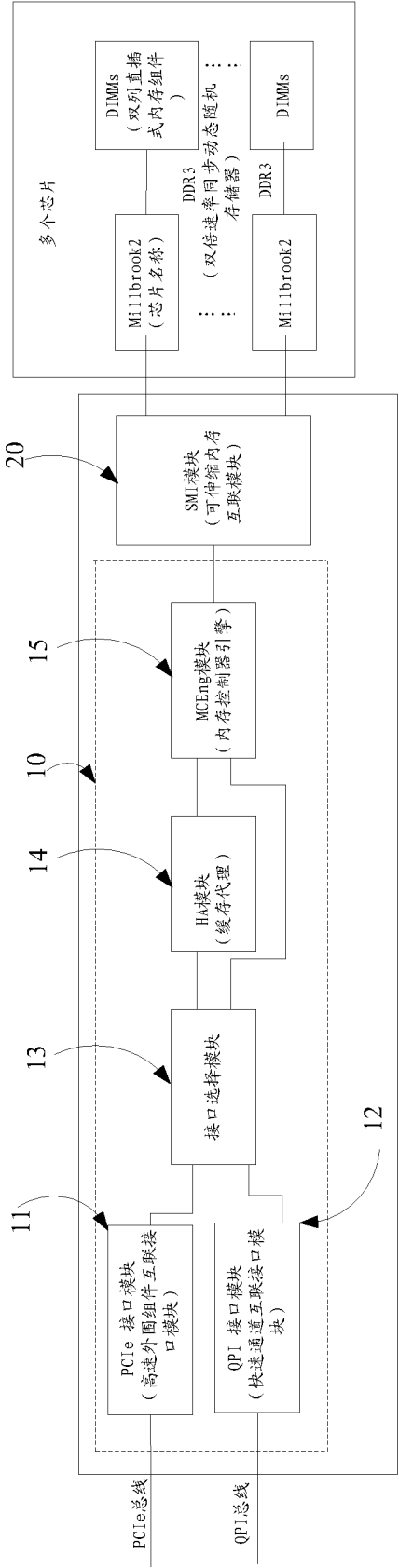


图 1

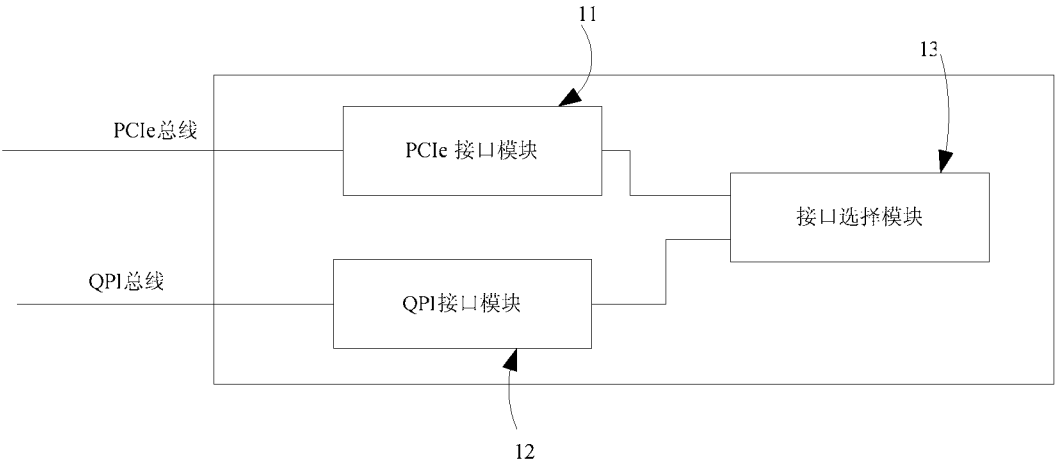


图 2A

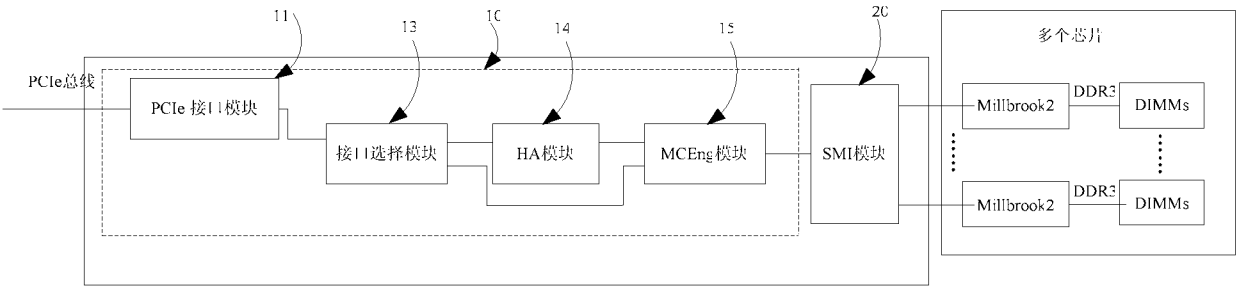


图 2B

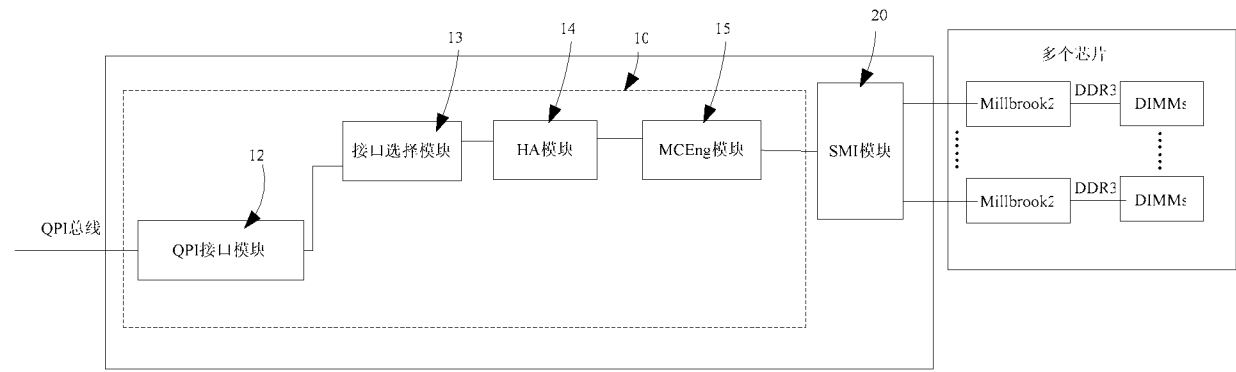


图 2C

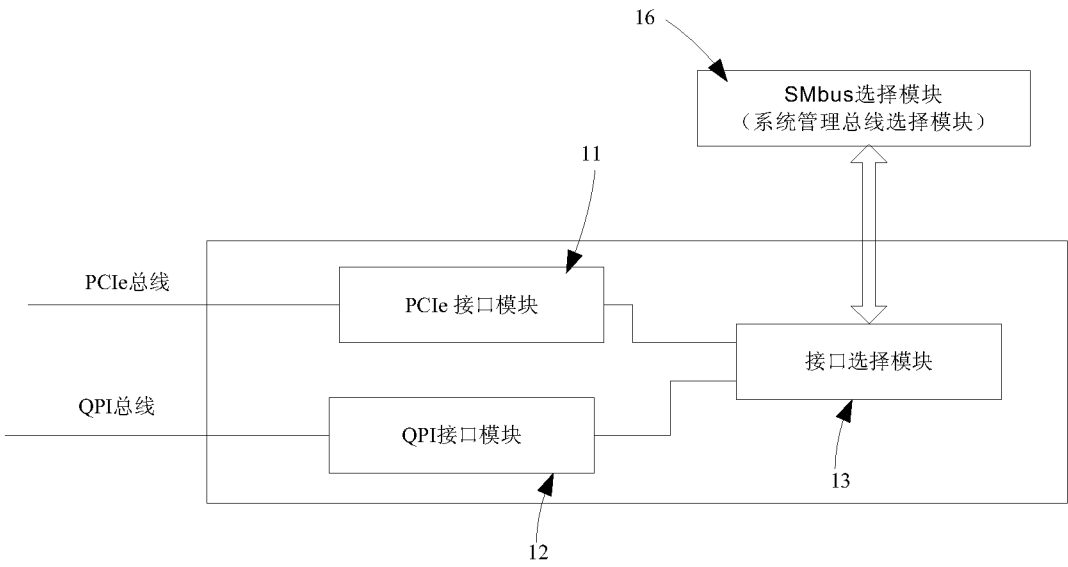


图 3

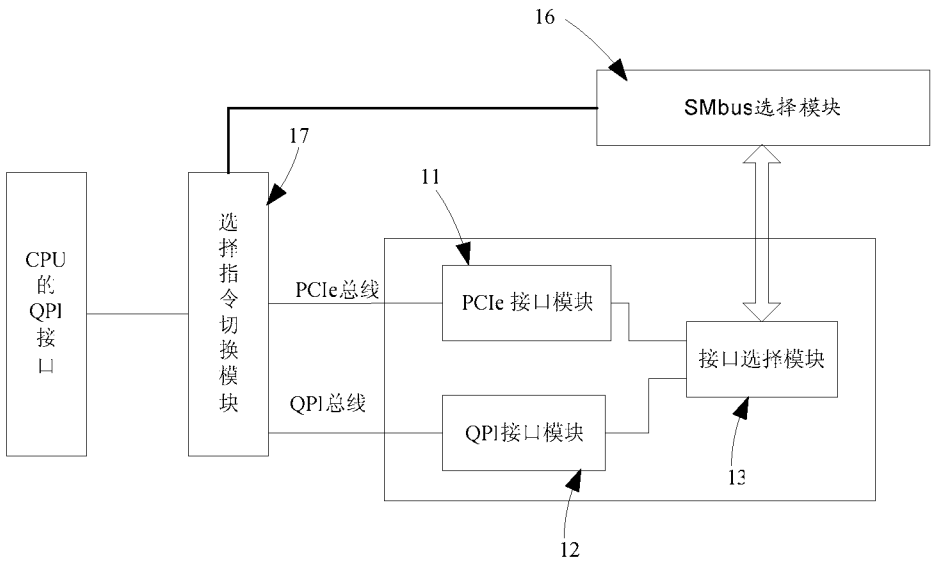


图 4

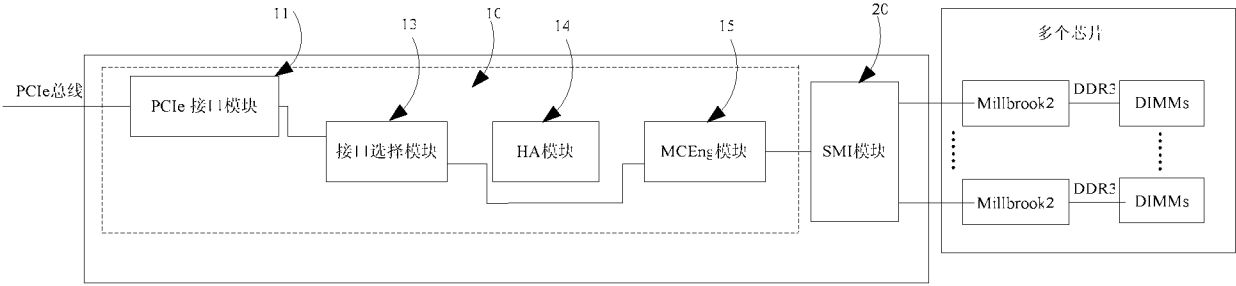


图 5A

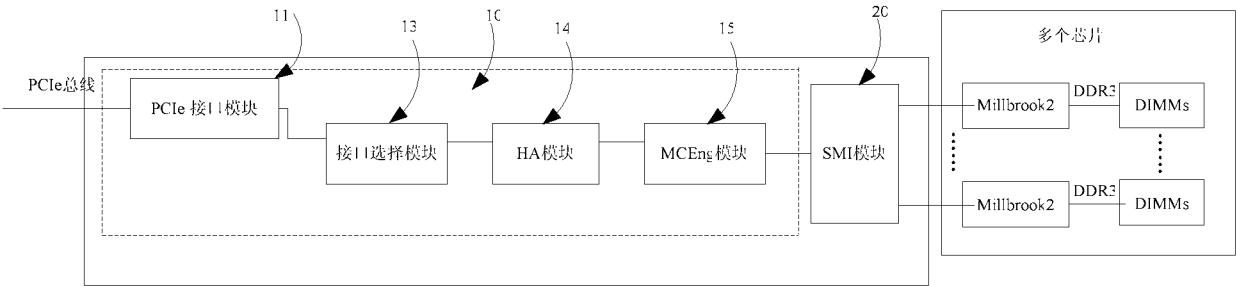


图 5B

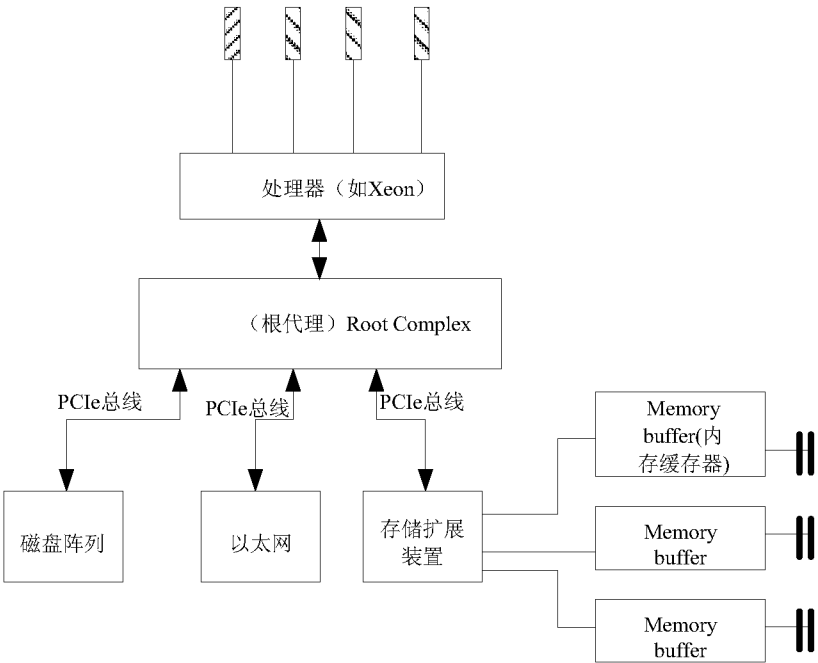


图 6

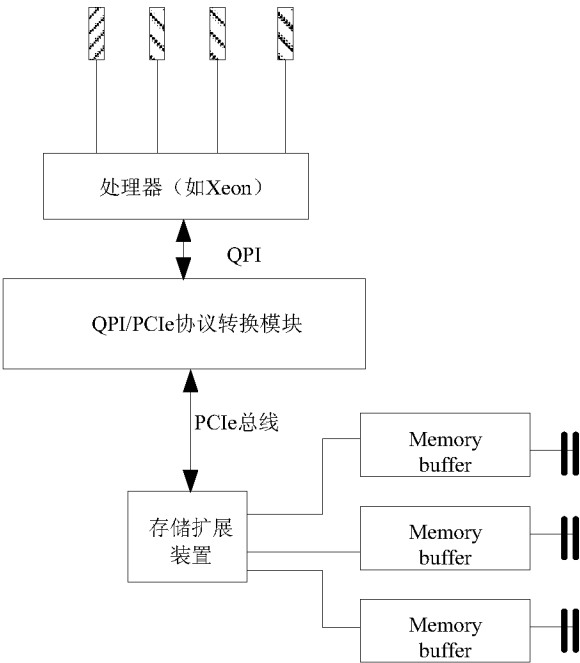


图 7

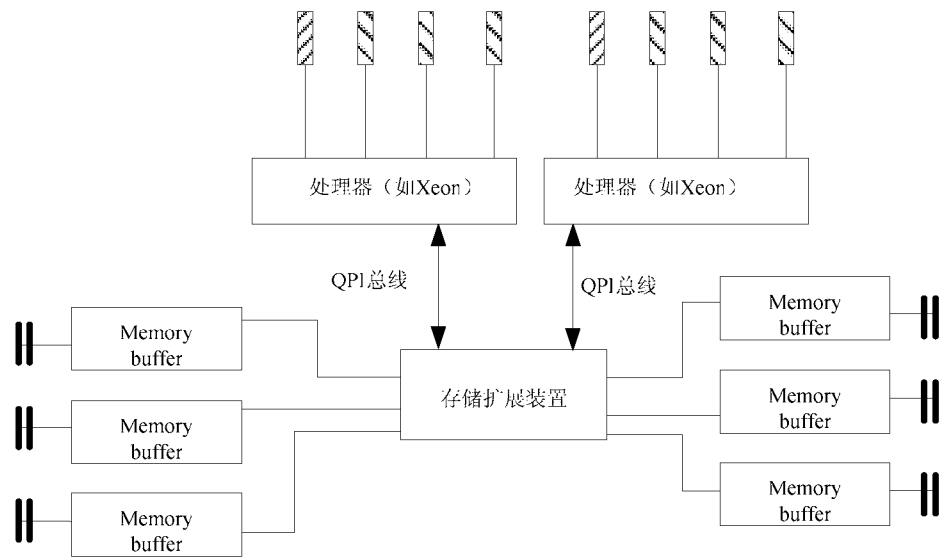


图 8

7/7

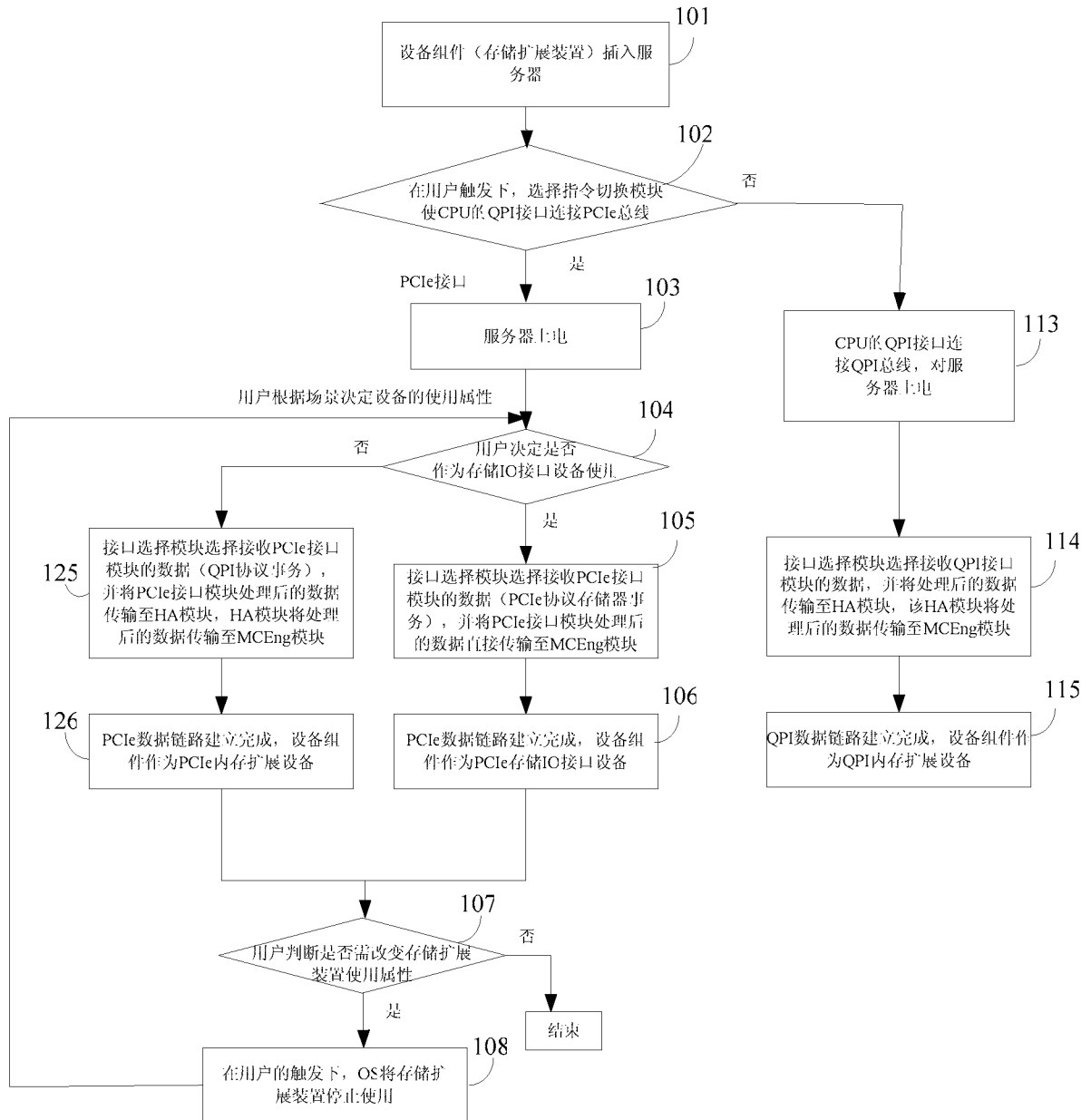


图 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2012/078032

A. CLASSIFICATION OF SUBJECT MATTER

G06F 13/00 (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)

IPC: G06F 13/-

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)

CNPAT; CNKI; WPI; EPODOC; CNTXT

memory, storage, device, expand+, extend+, spread+, processor, CPU, input+, output, IO, PCIe, QPI, interface, select+, option, switch+, performance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 102200952 A (HANGZHOU C SKY MICROSYSTEMS CO LTD) 28 September 2011 (28.09.2011) see the whole document	1-4
A	CN 1731369 A (ZTE CORP.) 08 February 2006 (08.02.2006) see the whole document	1-4
A	US 2005071534 A1 (DELL PRODUCTS LP) 31 March 2005 (31.03.2005) see the whole document	1-4
A	US 2010161924 A1 (FUJITSU LTD) 24 June 2010 (24.06.2010) see the whole document	1-4

☐ 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 22 March 2013 (22.03.2013)	Date of mailing of the international search report 04 April 2013 (04.04.2013)
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 CUI, Zhe Telephone No. (86-10) 62411849

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2012/078032

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102200952 A	28.09.2011	None	
CN 1731369 A	08.02.2006	CN 100517267 C	22.07.2009
US 2005071534 A1	31.03.2005	None	
US 2010161924 A1	24.06.2010	JP 2010146450 A	01.07.2010
		JP 4615595 B2	19.01.2011
		US 8180985 B2	15.05.2012

A. 主题的分类

G06F 13/00 (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)

IPC: G06F 13/-

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))

CNPAT; CNKI; WPI; EPODOC; CNTXT

存储, 内存, 设备, 扩展, 处理器, 输入, 输出, 接口, 选择, 切换, 性能; memory, storage, device, expand+, extend+, spread+, processor, CPU, input+, output, IO, PCIe, QPI, interface, select+, option, switch+, performance

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	CN102200952A (杭州中天微系统有限公司) 28.9 月 2011 (28.09.2011) 参见全文	1-4
A	CN1731369A (中兴通讯股份有限公司) 08.2 月 2006 (08.02.2006) 参见全文	1-4
A	US2005071534A1 (DELL PRODUCTS LP) 31.3 月 2005 (31.03.2005) 参见全文	1-4
A	US2010161924A1 (FUJITSU LTD) 24.6 月 2010 (24.06.2010) 参见全文	1-4

☐ 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
22.3 月 2013(22.03.2013)国际检索报告邮寄日期
04.4 月 2013 (04.04.2013)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授权官员
崔哲
电话号码: (86-10) 62411849

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2012/078032

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN102200952A	28.09.2011	无	
CN1731369A	08.02.2006	CN100517267C	22.07.2009
US2005071534A1	31.03.2005	无	
US2010161924A1	24.06.2010	JP2010146450A	01.07.2010
		JP4615595B2	19.01.2011
		US8180985B2	15.05.2012