

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 13/40 (2006.01)

G06F 9/46 (2006.01)

H04L 12/24 (2006.01)



# [12] 发明专利说明书

专利号 ZL 200510103551.1

[45] 授权公告日 2007 年 6 月 27 日

[11] 授权公告号 CN 1323362C

[22] 申请日 2005.9.21

[21] 申请号 200510103551.1

[73] 专利权人 杭州华为三康技术有限公司

地址 310053 浙江省杭州市高新技术产业  
开发区之江科技工业园六和路东华  
为 3Com 总部

[72] 发明人 李玉峰

[56] 参考文献

US5737542A 1998.4.7

CN1601498A 2005.3.30

CN1435974A 2003.8.13

审查员 陈颖

[74] 专利代理机构 北京集佳知识产权代理有限公司  
代理人 逯长明

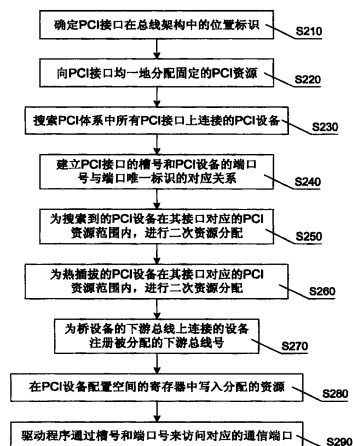
权利要求书 2 页 说明书 9 页 附图 3 页

[54] 发明名称

网络设备及其外围器件互连资源的分配方法

[57] 摘要

本发明公开了一种网络设备中 PCI 资源的分配方法，包括：确定 PCI 接口在总线架构中的位置标识；向 PCI 接口均一地分配 PCI 资源，使位置标识与被分配的 PCI 资源具有固定的对应关系。本发明还公开了一种应用上述方法的网络设备。本发明不仅能够避免初始化，尤其是热插拔造成的地址资源浮动，简化了驱动程序的管理；同时，地址空间的均一分配使得扩展卡在各个槽位能够得到同样的资源分配，避免了存储空间浪费。



1. 一种网络设备中外围器件互连 PCI 资源的分配方法, 该网络设备具有至少一条 PCI 总线和至少一个 PCI 接口, 其特征在于, 包括以下步骤:

确定 PCI 接口在总线架构中的位置标识;

向 PCI 接口均一地分配 PCI 资源, 使位置标识与被分配的 PCI 资源具有固定的对应关系。

2. 按照权利要求 1 所述 PCI 资源的分配方法, 其特征在于, 在向 PCI 接口均一分配资源之后还包括:

搜索 PCI 接口上连接的 PCI 设备;

在 PCI 接口的位置标识对应的 PCI 资源范围内, 为搜索到的 PCI 设备二次进行资源分配。

3. 按照权利要求 2 所述 PCI 资源的分配方法, 其特征在于, 所述方法还包括: 当在 PCI 接口中插入热插拔 PCI 设备时, 在该接口位置标识对应的 PCI 资源范围内, 为插入的设备进行二次资源分配。

4. 按照权利要求 3 所述 PCI 资源的分配方法, 其特征在于: 所述 PCI 资源包括总线号、中断号和物理地址空间; 所述物理地址空间包括配置空间、内存 Memory 空间、输入输出 I/O 空间。

5. 按照权利要求 4 所述 PCI 资源的分配方法, 其特征在于, 所述方法还包括: 当进行资源分配的 PCI 设备为桥设备时, 为该桥设备中下游总线上连接的设备注册被分配的下游总线号。

6. 按照权利要求 4 所述 PCI 资源的分配方法, 其特征在于, 为搜索到的 PCI 设备分配资源后还包括:

将分配给 PCI 设备的配置空间、Memory 空间和 I/O 空间映射为虚拟地址空间;

在该 PCI 设备配置空间的寄存器中写入为其分配的资源。

7. 按照权利要求 4 至 6 任意一项所述 PCI 资源的分配方法, 其特征在于: 所述 PCI 接口为 PCI 插槽, 所述 PCI 设备为一个通信端口;

所述方法还包括:

建立 PCI 插槽的槽号和通信端口的端口号的组合与端口唯一标识的对应关系, 以及端口唯一标识与该通信端口所连接 PCI 接口位置标识的对应关系; 所述端口唯一标识用来唯一地识别通信端口;

驱动程序通过提供槽号和端口号来访问具有相应端口唯一标识的通信端口。

8. 一种网络设备, 具有至少一条 PCI 总线和至少一个 PCI 接口, 其特征在于, 包括位置资源模块、设备搜索模块和资源分配模块, 其中:

位置资源模块用来存储均一分配给各个 PCI 接口的 PCI 资源范围, 均一分配根据 PCI 接口在总线架构中的位置进行;

设备搜索模块用来搜索 PCI 接口上连接的 PCI 设备;

资源分配模块查找位置资源模块中存储的与所述 PCI 设备连接的 PCI 接口的 PCI 资源范围, 在该范围内为所述 PCI 设备进行二次资源分配。

9. 按照权利要求 8 所述的网络设备, 其特征在于: 所述 PCI 资源包括总线号、中断号和物理地址空间; 所述物理地址空间包括配置空间、内存 Memory 空间、输入输出 I/O 空间。

10. 按照权利要求 9 所述的网络设备, 其特征在于, 所述网络设备还包括地址映射模块, 用来将 PCI 设备的配置空间、Memory 空间和 I/O 空间映射为虚拟地址空间。

## 网络设备及其外围器件互连资源的分配方法

### 技术领域

本发明涉及用于网络设备中的处理器系统，尤其涉及一种网络设备中外围器件互连资源的分配方法和应用该方法的网络设备。

### 背景技术

网络的发展对网络设备的功能提出了更高的要求，CPU（Central Process Unit，中央处理器）系统的应用越来越广泛。在网络设备的CPU系统中，PCI（Peripheral Component Interconnect，外围器件互连）是普遍采用的总线结构。PCI总线是一种具有多路地址线 and 数据线的高性能32/64位总线，在高度集成的外围控制器件、外围插件板和处理器/存储器之间作为互连结构应用。

PCI设备正常工作的前提是得到正确的初始化，在初始化过程中系统最为重要的工作是为PCI设备分配资源，包括总线号资源、地址资源和中断资源。PCI设备的地址资源主要是指PCI设备的配置空间、Memory（内存）空间和I/O（输入输出）空间，供CPU和该设备的驱动程序来控制设备工作。

现有技术中，PCI资源通常是按需分配的。系统启动后对PCI总线进行轮询，为查找到的设备依次分配PCI资源。其中PCI地址资源是CPU系统内部真实存在的地址空间范围，对每一个找到的PCI设备，根据由该设备得到的Memory空间和I/O空间的大小从地址空间范围中划分给该设备相应的空间。换言之，PCI设备被分配的地址空间的位置与系统搜索到该设备的顺序相关，当系统中插入、拔出了其他设备、或某个设备的Memory空间和I/O空间的大小发生了变化，则搜索顺序在其后的PCI设备地址空间的位置会相应变化，对热插拔PCI设备的地址分配也是如此。

PCI设备的配置空间、Memory空间和I/O空间位于CPU系统的物理地址

空间中，而出于安全性的考虑，绝大部分操作系统都不允许设备的驱动程序直接访问物理地址空间。操作系统提供了内存映射机制，将物理内存映射为虚拟内存后，设备驱动程序可以访问虚拟地址空间。

内存映射是根据地址来进行的，物理地址空间发生变化后需要重新进行内存映射，此时驱动程序必须对其访问的虚拟地址空间进行相应修改才能正常工作。因此，现有技术中的资源分配方法导致了地址空间处于浮动状态，尤其是对于支持热插拔的系统，造成了驱动程序内存管理的很大不便。

中国专利 CN1601498A 中公开了一种向 PCI 适配器分配存储器地址的方法来解决上述问题，根据与各个 PCI 适配器所在槽位的标识符关联的配置信息，向上述 PCI 适配器非均一地分配存储器地址范围。用这种方法为 PCI 适配器分配地址适用于计算机系统，由于计算机系统可以扩展的 PCI 适配器种类很多，不同种类的适配器所需的存储器空间会有很大的差别，采用非均一分配可以在很大程度上节省为 PCI 适配器预留的存储器空间。并且，计算机系统扩展槽位要留待用户决定其用途，根据用户的自身应用环境选择各种不同厂商和型号的适配器，这些相同类型的适配器所需要的存储器空间也经常会有很大的不同，此时非均一的空间分配也可以解决计算机系统用途不同导致的资源分配有所侧重的问题。

但对于具有扩展 PCI 槽位的网络设备，这种存储器空间分配方法则并不适用。与计算机系统中不同，网络设备中的 PCI 槽位基本上都用来扩展通信端口，而每个槽位具有的通信能力主要依赖于设备的背板，实际应用中同一设备所支持的扩展卡上端口的总带宽基本相同，同时在网络设备中通常要求扩展卡在各个槽位都能够正常工作，在这种情况下分配给每个槽位的地址空间应当能够满足所有扩展卡的需要，因而地址空间的非均一分配只可能增加存储空间的浪费。

## 发明内容

本发明要解决的是现有技术中 PCI 设备浮动的地址空间造成的不便于管理和向 PCI 设备非均一分配存储器空间造成的存储空间浪费的问题。

本发明所述的网络设备具有至少一条 PCI 总线和至少一个 PCI 接口，其外围器件互连 PCI 资源的分配方法包括以下步骤：

确定 PCI 接口在总线架构中的位置标识；

向 PCI 接口均一地分配 PCI 资源，使位置标识与被分配的 PCI 资源具有固定的对应关系。

优选地，在向 PCI 接口均一分配资源之后还包括：

搜索 PCI 接口上连接的 PCI 设备；

在 PCI 接口的位置标识对应的 PCI 资源范围内，为搜索到的 PCI 设备二次进行资源分配。

优选地，所述方法还包括：当在 PCI 接口中插入热插拔 PCI 设备时，在该接口位置标识对应的 PCI 资源范围内，为插入的设备进行二次资源分配。

优选地，所述 PCI 资源包括总线号、中断号和物理地址空间；所述物理地址空间包括配置空间、内存 Memory 空间、输入输出 I/O 空间。

优选地，所述方法还包括：当进行资源分配的 PCI 设备为桥设备时，为该桥设备中下游总线上连接的设备注册被分配的下游总线号。

优选地，为搜索到的 PCI 设备分配资源后还包括：

将分配给 PCI 设备的配置空间、Memory 空间和 I/O 空间映射为虚拟地址空间；

在该 PCI 设备配置空间的寄存器中写入为其分配的资源。

优选地，所述 PCI 接口为 PCI 插槽，所述 PCI 设备为一个通信端口；

所述方法还包括：

建立 PCI 插槽的槽号和通信端口的端口号的组合与端口唯一标识的对应关系，以及端口唯一标识与该通信端口所连接 PCI 接口位置标识的对应关系；所述端口唯一标识用来唯一地识别通信端口；

驱动程序通过提供槽号和端口号来访问具有相应端口唯一标识的通信端口。

本发明还提供了一种具有至少一条 PCI 总线和至少一个 PCI 接口的网络设备，包括位置资源模块、设备搜索模块和资源分配模块，其中：

位置资源模块用来存储均一分配给各个 PCI 接口的 PCI 资源范围，均一分配根据 PCI 接口在总线架构中的位置进行；

设备搜索模块用来搜索 PCI 接口上连接的 PCI 设备；

资源分配模块查找位置资源模块中存储的与所述 PCI 设备连接的 PCI 接口的 PCI 资源范围，在该范围内为所述 PCI 设备进行二次资源分配。

优选地，所述 PCI 资源包括总线号、中断号和物理地址空间；所述物理地址空间包括配置空间、内存 Memory 空间、输入输出 I/O 空间。

优选地，所述网络设备还包括地址映射模块，用来将 PCI 设备的配置空间、Memory 空间和 I/O 空间映射为虚拟地址空间。

本发明根据 PCI 接口在总线架构中的位置为其分配均一的资源，令连接在某个 PCI 接口上的设备始终得到固定的 PCI 资源范围，并且各个 PCI 接口具有同等大小的资源范围，从而使得连接在同一接口上的 PCI 设备能够分配到固定的地址空间，简化了驱动程序的内存管理；同时地址空间的均一分配使得扩展卡在各个槽位能够得到同样的资源分配，避免了存储空间的浪费；

进而，本发明根据 PCI 接口上连接的通信端口进行二次资源分配，通过建立槽号和端口号与 PCI 接口位置标识的对应关系，驱动程序可以通过槽号和端口号来定位资源，使得现有的驱动程序可以不经修改而适用于本发明。

## 附图说明

图 1 为本发明所述包括具有位置标识的 PCI 接口的 PCI 架构;

图 2 为本发明所述 PCI 资源分配方法的流程图;

图 3 为本发明所述网络设备的结构示意图。

## 具体实施方式

PCI 总线架构是一种层次式的体系结构,如图 1 所示,CPU 通过 Host-PCI 桥接器连接到主 PCI 总线——总线 0 上。除位置 M 外,位置 1 至位置 N 为总线 0 上的各个 PCI 接口,PCI 接口是总线上留待接入 PCI 设备的连接接口,通常是用于插入扩展卡的 PCI 插槽,也可能是其他符合 PCI 规范的连接接口。在位置 M 上,PCI 桥将总线 L 与总线 0 连接起来,此时总线 0 为 PCI 桥的上游总线,总线 L 为 PCI 桥的下游总线。位置 (N+1) 至位置 (N+K) 为总线 L 上的 PCI 接口。图中各个空闲的 PCI 接口都可以用来连接 PCI 设备,可以是单个的 PCI 设备,也可以是超过 1 个 PCI 设备的组合,还可以是 PCI 桥设备,即包括 PCI 桥并且在下游总线上连接有 PCI 设备。

不难理解,图 1 中的 PCI 架构是一种基本的 PCI 体系结构,在具体应用中可以对图 1 中的架构进行修改或增删,以得到任何符合需要的 PCI 架构。本发明中采用图 1 的架构对 PCI 接口的位置标识进行说明,不应将图 1 的架构作为对本发明的限定。

图 1 中用于连接上游和下游 PCI 总线的桥还可以是 HT (Hyper Transport, 超级传输) 桥,或者其他符合 PCI 规范并具有同样功能的桥接器。

本发明中,位置标识用来唯一标定 PCI 接口在 PCI 总线架构中的位置,并且 PCI 资源分配根据位置标识来进行。通过一个 PCI 设备的位置标识,可以得知其连接的 PCI 接口,以及该 PCI 接口在总线架构中的位置。

图 2 所示为本发明所述资源分配方法的流程。在步骤 S210,根据 PCI 接



口在总线架构中的位置为其设置位置标识。

在步骤 S220，向各个 PCI 接口均一分配 PCI 资源，并且使得各个 PCI 接口的位置标识与其分得的 PCI 资源具有固定的对应关系。即 PCI 接口、PCI 接口的位置标识、某个固定的 PCI 资源范围三者之间互为一一对应的关系。

实际应用中 PCI 接口通常是 PCI 插槽，而设备驱动程序通常是根据插槽的槽号来访问该插槽中的 PCI 设备的。本发明中之所以采用位置标识而不采用 PCI 插槽的槽号进行资源分配，是因为槽号可以由用户自由定义，这样将无法通过槽号建立插槽与 PCI 资源范围之间的固定对应关系。

PCI 资源包括总线号、中断号和物理地址空间，其中，总线号用来给 PCI 接口上连接的桥分配下游总线号；中断号供连接的 PCI 设备向系统申请中断时使用。物理地址空间包括配置空间、Memory 空间、I/O 空间，其中配置空间包括 PCI 设备的信息、状态和配置寄存器，系统通过这些寄存器来识别 PCI 设备、控制设备的工作状态、配置设备所使用的中断号、物理地址空间的基址等，对桥设备还需要配置下游总线号。

在步骤 S230，搜索 PCI 体系中所有 PCI 接口上连接的 PCI 设备。对 PCI 设备的搜索在系统上电后初始化的过程中进行，本发明中可以采用现有技术的方法来搜索 PCI 设备，此处不再赘述。

在步骤 S240，如果 PCI 接口为 PCI 插槽，且 PCI 插槽连接的 PCI 设备是通信端口，则采用端口唯一标识来在整个系统中唯一地识别该通信端口，即使得端口唯一标识与该通信端口所连接的插槽槽号和该通信端口在该槽中的端口号具有对应关系。其中，PCI 插槽槽号为用户可以自己定义的槽标识符，而端口号为该通信端口在其所连接插槽上的索引号，在该插槽上所有通信端口中唯一地标识该通信端口。

这种对应关系应该满足两个条件：其一是通过槽号和端口号可以唯一地获得该端口唯一标识；其二是通过端口唯一标识也可以唯一地获得槽号和端

口号。

同时，在本步骤中还需要建立端口唯一标识与其连接的 PCI 接口的位置标识的对应关系。

在步骤 S250，当在 PCI 接口上搜索到 PCI 设备时，在该 PCI 接口的位置标识对应的 PCI 资源范围内，即在步骤 S220 中为该 PCI 接口分配的 PCI 资源范围内，为搜索到的 PCI 设备进行二次资源分配。

由于本发明中 PCI 接口所连接的设备包括单个 PCI 设备、超过一个 PCI 设备的组合、或 PCI 桥设备，而设备驱动程序要针对每个 PCI 设备自身进行操作，所以需要 PCI 接口连接的 PCI 设备进行资源的二次分配。

在二次分配时，可以根据所连接设备的数量将该 PCI 接口对应的 PCI 资源范围平分给连接的设备，也可以按照各个设备具体需要的资源数量为其分配，或者采用其他由用户预定的能够满足设备所需的分配方法。由于该 PCI 接口对应的资源范围固定，只要这些设备仍连接在该 PCI 接口上，不论初始化时其他 PCI 接口上的设备情况发生什么变化，在重新分配资源时只要仍采用相同的二次分配方法，这些设备仍能得到与以前相同的 PCI 资源。

在步骤 S260，如果系统检测到某个 PCI 接口上接入了热插拔的 PCI 设备，则为插入的 PCI 设备在该 PCI 接口的位置标识对应的 PCI 资源范围内，进行资源的二次分配。本步骤中资源的二次分配方法与步骤 S250 中相同，不再重复。

支持热插拔的系统在运行过程中插入 PCI 设备后，会收到上报的中断，从而触发 PCI 资源的二次分配过程。

在步骤 S270，如果步骤 S250 中搜索到的 PCI 设备、或步骤 S260 中插入的热插拔 PCI 设备为桥设备，则连接在该桥设备中桥的下游总线上的 PCI 设备也被二次分配了 PCI 资源。此时应为该下游总线分配总线号，并且为连接在该下游总线上的设备注册下游总线号，以便于后续对该设备的访问。

在步骤 S280, 将分配给 PCI 设备的配置空间、Memory 空间和 I/O 空间映射为虚拟地址空间, 并且对 PCI 设备配置空间的寄存器进行写操作, 记录为该设备分配的资源, 例如中断号、Memory 空间的基址、I/O 空间地址的基址等, 以及为桥设备配置的下游总线号。其中, 将 Memory 空间的基址、I/O 空间地址的基址写入配置空间的基址寄存器, 将中断号写入中断线寄存器, 如果是桥设备则将下游总线号写入总线号寄存器。

PCI 设备的驱动程序通常不能直接访问位于物理地址空间中的 Memory 空间和 I/O 空间, 而需要通过建立内存映射后的虚拟地址空间来间接访问物理地址空间。

在步骤 S290, 如果在步骤 S240 中为通信端口建立的端口唯一标识, 则驱动程序可以通过 PCI 插槽的槽号和通信端口的端口号来访问某个通信端口。从槽号和端口号可以得到该通信端口的端口唯一标识, 由端口唯一标识可以得知该通信端口位于哪个 PCI 接口以及在该 PCI 接口上的索引, 从而找到该通信端口在 PCI 架构中的具体位置, 之后驱动程序可以通过对其虚拟地址空间的读写控制该设备工作。

图 3 所示为应用上述 PCI 资源分配方法的网络设备的结构示意图, 其中网络设备具有一条 PCI 总线和至少一个 PCI 接口。设备搜索模块 320、资源分配模块 330 和地址映射模块 340 串接, 且分别通过总线连接到 PCI 接口; 位置资源模块 310 与资源分配模块 330 连接; 某个 PCI 接口上连接有 PCI 设备。

根据网络设备中 CPU 系统的 PCI 架构来进行资源分配, 按照 PCI 接口在 PCI 架构中的位置将 PCI 资源均一地分配给各个 PCI 接口, 并将 PCI 接口与所分得的资源范围的对应关系存储在位置资源模块 310 中。

PCI 资源包括总线号、中断号和物理地址空间; 其中物理地址空间包括配置空间、内存 Memory 空间和输入输出 I/O 空间。

设备搜索模块 320 在网络设备上电初始化时、或网络设备运行过程中热

插拔的 PCI 接口上接入 PCI 设备时对 PCI 接口连接的设备进行搜索。

对设备搜索模块 320 搜索到的 PCI 设备，资源分配模块 330 从位置资源模块 310 中查找到该设备所连接 PCI 接口对应的 PCI 资源范围，在该范围中为搜索到的设备进行二次资源分配。二次资源分配的方法与本发明所述资源分配方法中步骤 S250 中相同，不再重复。

设备搜索模块 320 对 PCI 设备的配置空间进行读写操作，将分配给该 PCI 设备的资源范围写入其配置空间的寄存器中，包括中断号、Memory 空间的基址和 I/O 空间的基址等，以及桥设备的下游总线号。

地址映射模块 340 将资源分配模块 330 分配给 PCI 设备的 Memory 空间和 I/O 空间映射为虚拟内存。

应用本发明后，在网络设备的 PCI 架构确定之后就安排好各个 PCI 接口的资源空间。即使这个 PCI 接口没有接入设备、即使这个 PCI 接口在目前产品设计中还没有表现为具体的硬件，但未来有可能扩展出来，都给其分配好固定的空间。并且，PCI 设备的物理地址空间和虚拟地址空间的映射也可以进行统一的规划，后续可以通过一个全局的数据结构进行访问。

因而，在应用本发明前，要注意对网络设备 PCI 架构的设计不仅要满足当前系统的需要，而且要预留足够的未来扩展能力，以适应用户增长的性能和数量要求。

可见，本发明不仅能够避免初始化，尤其是热插拔造成的地址资源浮动，简化了驱动程序的管理，而且通过建立槽号和端口号与 PCI 接口位置标识的对应关系，驱动程序可以通过槽号和端口号来定位资源，避免了对现有驱动程序的修改。

以上所述的本发明实施方式，并不构成对本发明保护范围的限定。任何在本发明的精神和原则之内所作的任何修改、等同替换和改进等，均应包含在本发明的权利要求保护范围之内。

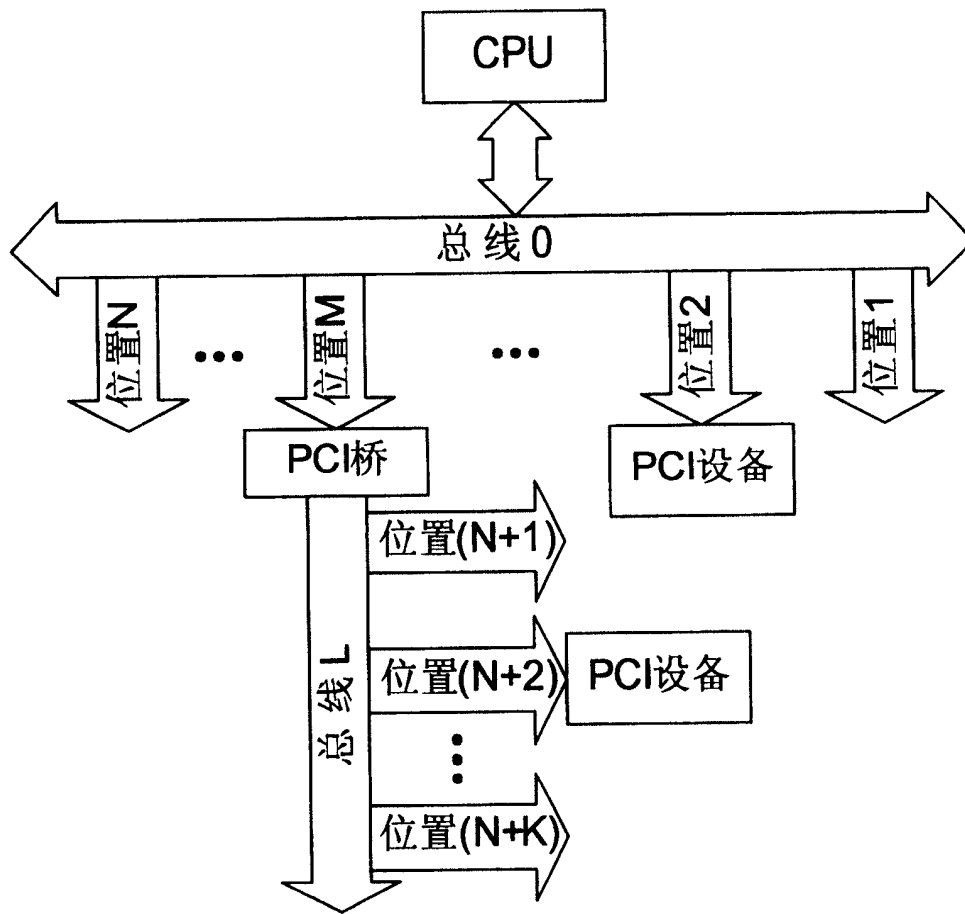


图 1

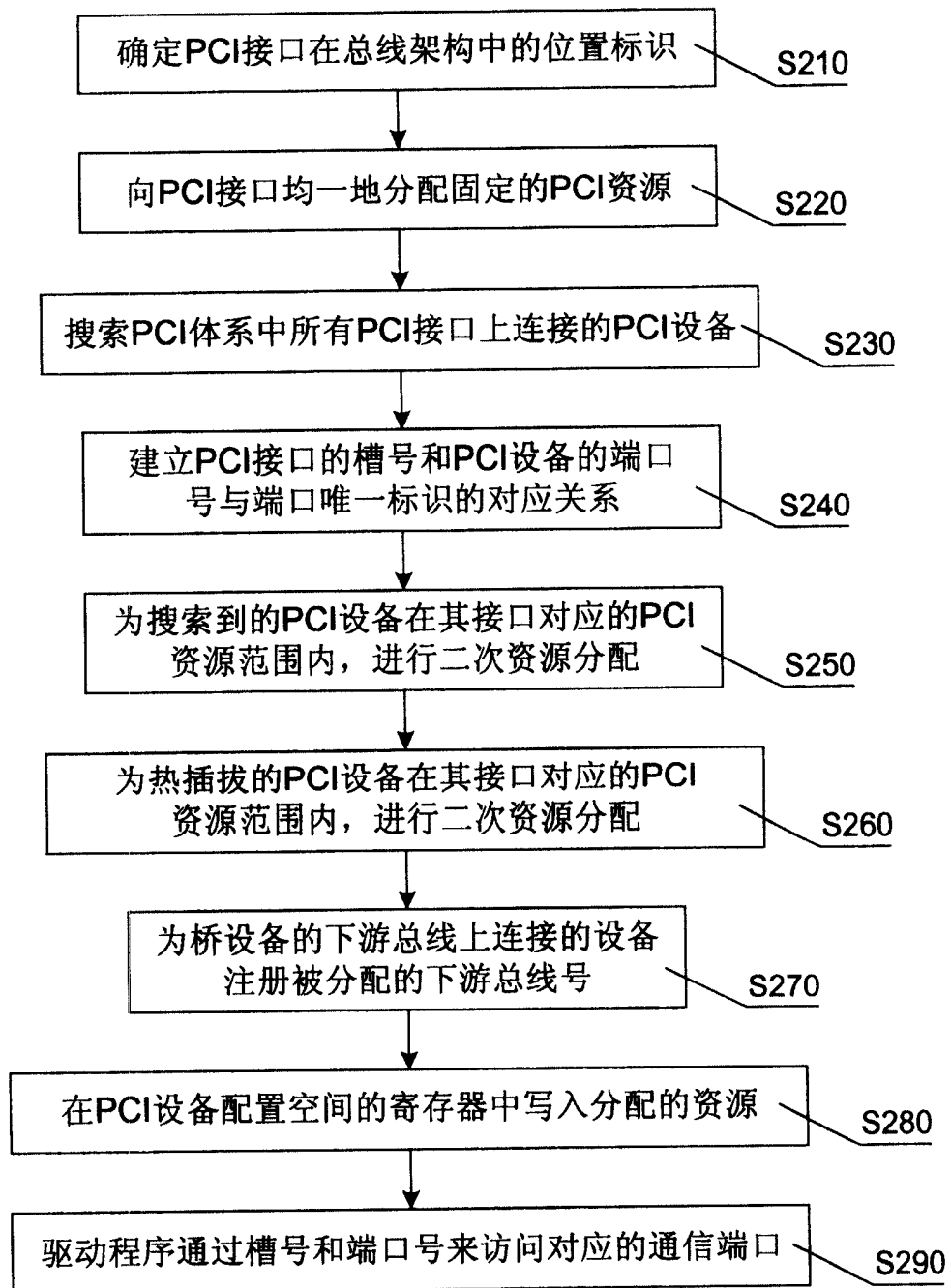


图 2

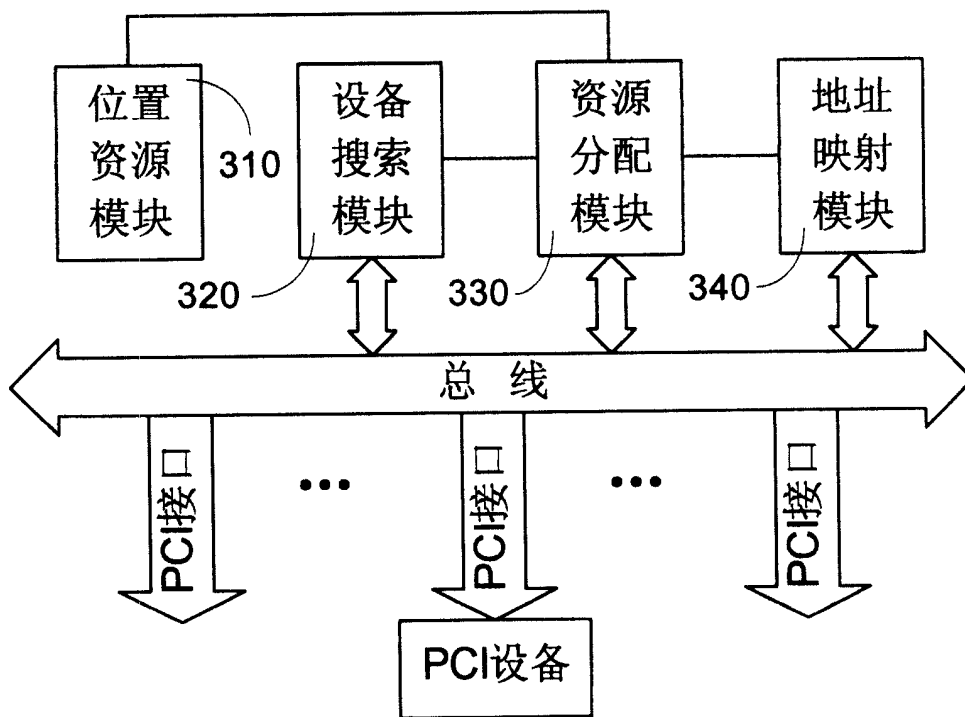


图 3