



(12) 发明专利

(10) 授权公告号 CN 114500413 B

(45) 授权公告日 2024.04.16

(21) 申请号 202111554375.9

CN 109891394 A, 2019.06.14

(22) 申请日 2021.12.17

CN 111159093 A, 2020.05.15

(65) 同一申请的已公布的文献号

CN 113452731 A, 2021.09.28

申请公布号 CN 114500413 A

US 10325343 B1, 2019.06.18

US 8891543 B1, 2014.11.18

(43) 申请公布日 2022.05.13

CN 110708185 A, 2020.01.17

(73) 专利权人 阿里巴巴(中国)有限公司

CN 111796932 A, 2020.10.20

地址 310052 浙江省杭州市滨江区长河街

CN 112416672 A, 2021.02.26

道网商路699号4号楼5楼508室

CN 113127160 A, 2021.07.16

CN 113377520 A, 2021.09.10

(72) 发明人 张岱 付斌章

US 2014059265 A1, 2014.02.27

(74) 专利代理机构 北京智信禾专利代理有限公

US 2019220311 A1, 2019.07.18

司 11637

黄家明;李冠广;乔英良;班华堂;梁锐.融合架构云服务器体系结构和关键技术.计算机工程与应用.2018,(第14期),全文.

专利代理师 赵杰

徐华;李京.面向带宽保障的云中虚拟集群调度算法.中国科学技术大学学报.2018,(第06期),全文.

(51) Int. Cl.

H04L 49/25 (2022.01)

(56) 对比文件

CN 107239346 A, 2017.10.10

CN 108845970 A, 2018.11.20

审查员 赵潜

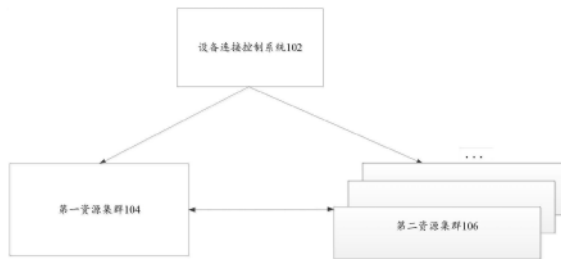
权利要求书3页 说明书19页 附图9页

(54) 发明名称

设备连接方法及装置、设备连接芯片

(57) 摘要

本说明书实施例提供设备连接方法及装置、设备连接芯片,其中设备连接方法应用于设备连接控制系统,包括:接收用户发送的针对目标处理设备的设备连接请求,其中,设备连接请求携带待连接外部设备的需求数量;在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备;向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令。



1. 一种设备连接方法,应用于设备连接控制系统,包括:

接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;

在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;

向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;

向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

2. 根据权利要求1所述的设备连接方法,所述确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量,包括:

确定第一资源集群中与第一目标交换机相连接的初始外部设备数量;

确定第一资源集群中与至少一个第二目标交换机相连接的备选外部设备数量;

在确定所述初始外部设备数量与所述备选外部设备数量之和小于所述需求数量的情况下,确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量。

3. 根据权利要求1或2所述的设备连接方法,所述在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,包括:

在至少一个第二资源集群中,确定与所述至少一个第二资源集群中的目标交换机相连接的待连接外部设备;

基于所述待连接外部设备确定满足所述需求数量的外部设备为第二目标外部设备。

4. 根据权利要求3所述的设备连接方法,所述向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,包括:

基于第一目标外部设备以及第二目标外部设备确定拓扑链路配置规则;

基于所述拓扑链路配置规则向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令。

5. 根据权利要求4所述的设备连接方法,所述接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路,包括:

接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令解析的所述拓扑链路配置规则;

接收所述管理模块基于所述拓扑链路配置规则,在预设拓扑链路图的目标芯片组中调整的处理设备与外部设备之间的更新拓扑链路图;

接收所述管理模块基于所述更新拓扑链路图确定的所述目标处理设备与所述第一目标外部设备之间的第一目标拓扑链路,以及所述目标处理设备和所述第二目标外部设备之间的第二目标拓扑链路。

6. 根据权利要求5所述的设备连接方法,所述向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令之后,还包括:

控制所述目标处理设备响应于所述连接指令,基于所述第一目标拓扑链路在所述第一资源集群中与所述第一目标外部设备进行连接;

控制所述目标处理设备响应于所述连接指令,基于所述第二目标拓扑链路在所述第二资源集群中与所述第二目标外部设备进行连接。

7. 根据权利要求6所述的设备连接方法,所述控制所述目标处理设备响应于所述连接指令,基于所述第二目标拓扑链路在所述第二资源集群中与所述第二目标外部设备进行连接,包括:

控制所述目标处理设备响应于所述连接指令,基于所述连接指令中携带的第二目标拓扑链路确定与所述目标处理设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述目标处理设备相对应的目标交换插槽中;

通过所述拓展线卡控制所述目标处理设备与所述第二资源集群中的第二目标外部设备进行连接。

8. 根据权利要求7所述的设备连接方法,所述通过所述拓展线卡控制所述目标处理设备与所述第二资源集群中的第二目标外部设备进行连接,包括:

通过所述拓展线卡控制所述目标处理设备的设备连接协议的报文转换为交换网络协议的报文;

控制所述目标处理设备基于所述交换网络协议的报文,通过网络交换机与所述第二资源集群中的第二目标外部设备进行连接。

9. 根据权利要求8所述的设备连接方法,所述控制所述目标处理设备基于所述交换网络协议的报文,通过网络交换机与所述第二资源集群中的第二目标外部设备进行连接,包括:

控制所述目标处理设备通过网络交换机将接收到的所述交换网络协议的报文发送至所述第二资源集群中的第二目标外部设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述第二目标外部设备相对应的目标交换插槽中;

控制所述目标处理设备基于所述拓展线卡将所述交换网络协议的报文进行转换,获得设备连接协议的报文;

控制所述目标处理设备基于所述设备连接协议的报文与所述第二资源集群中的第二目标外部设备建立连接。

10. 根据权利要求1所述的设备连接方法,所述接收用户发送的针对目标处理设备的设备连接请求之后,还包括:

在确定第一资源集群中待连接的第一目标外部设备满足所述需求数量的情况下,向所述第一目标外部设备的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备的管理模块响应于所述拓扑链路配置指令确定的目标拓扑链路;

向所述目标处理设备与所述第一目标外部设备分别发送连接指令,其中,所述连接指令携带有目标拓扑链路。

11. 一种设备连接芯片,包括:至少一个交换插槽、至少一个上行端口,所述上行端口用于与第一资源集群中处理设备相连接,所述芯片通过所述交换插槽与第二资源集群中的外部设备相连接,其中,所述第一资源集群与所述第二资源集群不同,所述第一资源集群中的处理设备、所述第二资源集群中的外部设备为权利要求1-10任一项所述的设备连接方法中

的第一目标外部设备、第二目标外部设备。

12. 一种设备连接装置,应用于设备连接控制系统,包括:

请求接收模块,被配置为接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;

设备确定模块,被配置为在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;

拓扑链路确定模块,被配置为向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;

指令发送模块,被配置为向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

13. 一种计算设备,包括:

存储器和处理器;

所述存储器用于存储计算机可执行指令,所述处理器用于执行所述计算机可执行指令,该计算机可执行指令被处理器执行时实现权利要求1-10任意一项所述设备连接方法的步骤。

14. 一种计算机可读存储介质,其存储有计算机可执行指令,该计算机可执行指令被处理器执行时实现权利要求1-10任意一项所述设备连接方法的步骤。

## 设备连接方法及装置、设备连接芯片

### 技术领域

[0001] 本说明书实施例涉及计算机技术领域,特别涉及一种设备连接方法。

### 背景技术

[0002] 随着人工智能的兴起,云服务器设备应运而生,本质上,云服务器设备就是一台高配置的服务器,内部可配置多个子设备,通过虚拟化技术,将这些子设备分成不同的实例,并将子设备分配给不同的实例中使用。

[0003] 但是在实际应用中,往往根据用户对于实例的需求,在具体分配子设备时,无法实现对子设备资源的充分利用,甚至在目标处理设备(与子设备相连接的设备)所在的本地中,也无法满足的用户需求,进而对分配实例的过程拓展性较差、灵活性较低。

### 发明内容

[0004] 有鉴于此,本说明书实施例提供了一种设备连接方法。本说明书一个或者多个实施例同时涉及一种设备连接芯片,一种设备连接装置,一种计算设备,一种计算机可读存储介质,以解决现有技术中存在的技术缺陷。

[0005] 根据本说明书实施例的第一方面,提供了一种设备连接方法,应用于设备连接控制系统,包括:

[0006] 接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;

[0007] 在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;

[0008] 向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;

[0009] 向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

[0010] 根据本说明书实施例的第二方面,提供了一种设备连接芯片,包括:至少一个交换插槽、至少一个上行端口,所述上行端口用于与第一资源集群中处理设备相连接,所述芯片通过所述交换插槽与第二资源集群中的外部设备相连接,其中,所述第一资源集群与所述第二资源集群不同。

[0011] 根据本说明书实施例的第三方面,提供了一种设备连接装置,应用于设备连接控制系统,包括:

[0012] 请求接收模块,被配置为接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;

[0013] 设备确定模块,被配置为在确定第一资源集群中待连接的第一目标外部设备不满

足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;

[0014] 拓扑链路确定模块,被配置为向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;

[0015] 指令发送模块,被配置为向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

[0016] 根据本说明书实施例的第四方面,提供了一种计算设备,包括:

[0017] 存储器和处理器;

[0018] 所述存储器用于存储计算机可执行指令,所述处理器用于执行所述计算机可执行指令,其中,所述处理器执行所述计算机可执行指令时实现所述设备连接方法的步骤。

[0019] 根据本说明书实施例的第五方面,提供了一种计算机可读存储介质,其存储有计算机可执行指令,该指令被处理器执行时实现任意一项所述设备连接方法的步骤。

[0020] 本说明书一个实施例应用于设备连接控制系统,通过接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

[0021] 具体的,设备连接控制系统在控制与目标处理设备相连接外部设备的资源选择的过程中,先通过在第一资源集群中确定第一外部设备能否满足用户的需求,在该第一外部设备不满足用户的需求时,可在至少一个第二资源集群确定外部设备,作为第二目标外部设备,同时,再对第一目标外部设备和第二目标外部设备的管理模型下发对拓扑链路的配置指令,进而确定目标处理设备与第一目标外部设备和第二目标外部设备的具体链接的拓扑链路,最后再下发连接指令,便于后续目标处理设备、第一目标外部设备、第二目标外部设备响应于该连接指令,实现目标处理设备和第一目标外部设备连接,或者是目标处理设备和第二目标外部设备连接,该种对外部设备资源的选择方式不仅能够实现在本地资源集群(第一资源集群)中选择与目标处理设备相连接的外部设备,还能够实现跨域资源集群(第二资源集群)中选择与目标处理设备相连接的外部设备,体现了连接设备的拓展性和资源选择的灵活性,进而能够进一步解决资源池化问题。

## 附图说明

[0022] 图1是本说明书一个实施例提供的一种设备连接方法应用于设备连接控制系统系统架构图;

[0023] 图2是本说明书一个实施例提供的一种设备连接方法的流程图;

- [0024] 图3是本说明书一个实施例提供的一种设备连接方法中资源集群的结构示意图；
- [0025] 图4是本说明书一个实施例提供的一种设备连接方法中配置管理模块的资源集群结构示意图；
- [0026] 图5是本说明书一个实施例提供的一种设备连接方法配置拓扑链路的示意图；
- [0027] 图6是本说明书一个实施例提供的一种设备连接方法跨域设备连接处理过程的结构示意图；
- [0028] 图7是本说明书一个实施例提供的一种设备连接芯片的结构示意图；
- [0029] 图8是本说明书一个实施例提供的一种设备连接方法的处理过程流程图；
- [0030] 图9是本说明书一个实施例提供的一种设备连接装置的结构示意图；
- [0031] 图10是本说明书一个实施例提供的一种计算设备的结构框图。

### 具体实施方式

[0032] 在下面的描述中阐述了很多具体细节以便于充分理解本说明书。但是本说明书能够以很多不同于在此描述的其它方式来实施,本领域技术人员可以在不违背本说明书内涵的情况下做类似推广,因此本说明书不受下面公开的具体实施的限制。

[0033] 在本说明书一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本说明书一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0034] 应当理解,尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本说明书一个或多个实施例范围的情况下,第一也可以被称为第二,类似地,第二也可以被称为第一。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0035] 首先,对本说明书一个或多个实施例涉及的名词术语进行解释。

[0036] host:指服务器的计算单元,即CPU。

[0037] PCIe(Peripheral Component Interconnect Express,外围组件互联总线):是计算机内部最常用的外部总线,采用的是树形拓扑结构,它的体系架构一般由RC,bridge,switch,EP等类型的PCIe设备组成。

[0038] PCIe Lane:通道,PCIe是串行总线,通过使用差分信号传输,这样一对差分信号组成一个PCIe Lane,也叫做x1通道。

[0039] RC(root complex,根桥设备):与CPU直接相连,主要负责PCIe报文的解析和生成。

[0040] PCIe switch (PCIe交换机):PCIe的转接器设备,目的是扩展PCIe总线,包含多个port。

[0041] PCIe bridge (PCIe桥):PCIe的转接器设备,可以理解为只包含2个port的pcie switch。

[0042] Port:端口,把n组Lane绑定在一起后的接口,Lane数越多,带宽越大。

[0043] Up port(上行端口):PCIe switch或Pcie bridge的上行口,在拓扑图中更接近

host。

[0044] Down port(下行端口):PCIe switch或Pcie bridge的下行口,在拓扑图中更远离host。

[0045] EP(end point,PCIe终端设备):是PCIe树形结构的叶子节点,比如网卡,NVME卡,显卡都是PCIe EP设备。

[0046] CPU domain(Cpu域):指的是CPU的RC下面挂载的所有PCIE设备,包括PCIE bridge、PCIE switch和EP,这些设备为该CPU独占。

[0047] GPU(graphics processing unit,图形处理器):又称显示核心、视觉处理器、显示芯片,是一种专门在个人电脑、工作站、游戏机和一些移动设备(如平板电脑、智能手机等)上图像运算工作的微处理器。

[0048] FPGA(Field Programmable Gate Array,现场可编程门阵列):FPGA实现了跨以太网络的PCIE switch功能。

[0049] 随着人工智能的兴起,云服务器应运而生,其中对于GPU的使用也越来越普及。以GPU的使用为例对目前资源分配问题进行说明。本质上,GPU云服务器就是一台高配置的服务器,其服务器内部配置多块GPU卡,并通过虚拟化技术,将CPU切分成不同的实例,并将GPU卡分配给其中这些实例。目前,GPU服务器可分为机头和机尾两部分:机头是一台标准的服务器;机尾是内置有8张GPU卡的盒子,盒子内部并没有CPU、内存等。机头和机尾之间通过PCIE电缆相连。将GPU服务器设计为机头和机尾两部分,主要是为了满足CPU和GPU独立维修和升级的需求,两者的连接关系是绑定的。

[0050] 由于CPU和GPU的从属关系是固定的,这就造成了GPU资源的浪费。具体可参见下述两种情况:

[0051] 第一种情况:租户A购买了一个4GPU的实例,租户B购买了一个2GPU的实例,两者共用了一台服务器的6张GPU卡,剩余2张。这时候,这台服务器只能再提供一台2GPU的实例或两台1GPU的实例。如果后续的需求都是4GPU实例或8GPU实例,那该服务器的两张GPU卡就只能被闲置了。

[0052] 第二种情况:租户A购买了一个8GPU的实例,服务器在开出实例后,仅占用一部分CPU内核,并未全部占用,还剩余了多个CPU内核,但这些CPU核心最后会以CPU实例的形式全部售出。而后租户A释放了之前购买的8GPU实例,因此8张GPU卡空闲了出来。紧接着租户B购买了一个1GPU实例,占用了剩余的CPU内核。这时候,该服务器虽然还有7张GPU卡空闲,但是已经无法再开出GPU实例了,因为租户A购买时占用的一部分CPU内核资源已耗尽。

[0053] 基于此,每个租户开启/释放GPU实例的时刻是随机的,这就导致了云上的GPU卡的碎片化分布越来越严重,多GPU实例资源消耗殆尽。为了解决GPU碎片化的问题,可采取的方法可以为资源的池化操作。进而,本说明书实施例提供的设备连接方法可通过解除GPU与CPU之间的物理的绑定关系,将多个GPU设备资源池化,CPU可随时获取/释放GPU资源。

[0054] 因此,本说明书实施例提供的设备连接方法,不仅可在本地资源集群中选择对应的外部设备与目标处理设备连接,还可在跨域集群中选择外部设备与目标处理设备相连,以满足用户对连接外部设备的数量需求,并通过管控系统下发指令实现目标处理设备与外部设备连接。为了便于理解,本说明书实施例提供的设备方法以目标处理设备为CPU设备,外部设备为GPU为例对设备连接方法的具体步骤进行详细地描述,需要说明的是,目标处理



设备并不限于CPU设备,可以为服务器内任何一个待连接的处理设备;外部设备也并不限于GPU设备,可以为任意一个具有扩展处理功能的外部设备。

[0055] 在本说明书中,提供了一种设备连接方法,本说明书同时涉及一种设备连接芯片,一种设备连接装置,一种计算设备,一种计算机可读存储介质,在下面的实施例中逐一进行详细说明。

[0056] 图1示出了根据本说明书一个实施例提供的一种设备连接方法应用于设备连接控制系统系统架构图。

[0057] 图1中包括设备连接控制系统102、第一资源集群104、至少一个第二资源集群106,需要说明的是,设备连接控制系统102可以为区别于第一资源集群104以及第二资源集群106的单独的一套系统,是能够得知整个资源集群中资源占用情况。比如,第一资源集群104以及第二资源集群106中,可包括至少一个CPU设备、至少一个GPU设备,且CPU设备与GPU设备之间通过交换机连接总线相连接。

[0058] 实际应用中,设备连接控制系统102可用于接收用户发送的目标CPU设备的设备连接请求,且该目标CPU设备若在第一资源集群104时,可以理解为第一资源集群104相对于该目标CPU设备为本地资源集群,第二资源集群106相对于该目标CPU设备可以理解为跨区域资源集群,基于此,在设备连接控制系统102确定在第一资源集群104中可用的GPU数量不满足用户发送的设备连接请求中携带的外部设备需求数量时,然后设备连接控制系统102可控制在第二资源集群106中选择外部设备,以满足用户的外部设备需求数量,进而,在确定外部设备之后,再具体地确定连接的拓扑链路,设备连接控制系统102可继续通过下发控制指令,控制目标CPU设备与选取的GPU实现设备连接。

[0059] 需要说明的是,在第一资源集群中选择的外部设备数量有三种情况,第一种是在第一资源集群中即可选择出与用户设备需求数量相同的外部设备,即无需在第二资源集群中再次选择;第二种是在第一资源集群中只能选择出一部分数量的外部设备与目标CPU设备相连,还不满足用户设备需求数量,即可在第二资源集群中再选择外部设备,而满足用户设备需求数量,也即使说,在第一资源集群中选择外部设备数量与第二资源集群中选择外部设备数量的和大于等于用户设备需求数量;第三种是在第一资源集群中无法选择一个GPU设备与目标CPU设备相连,即可全部在第二资源集群中选择外部设备与目标CPU设备相连接。

[0060] 本说明书实施例提供的设备连接方法,在本地资源集群中待连接的外部设备数量不满足用户发送的设备需求数量的情况下,设备连接控制系统可在跨区域资源集群中选择可与目标CPU设备相连接的外部设备,以满足用户的设备需求数量,同时可提高外部设备在资源集群中的利用率,以实现资源池化。

[0061] 参见图2,图2示出了根据本说明书一个实施例提供的一种设备连接方法的流程图,具体包括以下步骤:

[0062] 需要说明的是,本说明书实施例提供的一种设备连接方法可应用于设备连接控制系统,该设备连接系统可以为区别与资源集群的单独的系统,也可以配置在任意一个资源集群中,本说明书实施例对此不作具体限定,为了便于理解,本说明书实施例中将设备连接控制系统作为独立的控制系统实现设备连接。

[0063] 步骤202:接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连

接请求携带待连接外部设备的需求数量。

[0064] 其中,目标处理设备可以理解为在服务器内需要连接外部设备的处理设备,比如CPU设备,即服务器内的CPU设备需要挂载多个GPU进行任务处理的情况。

[0065] 外部设备可以理解为与服务器内部的目标处理设备相连接的,区别与服务器之外的外部设备,比如,GPU设备。

[0066] 实际应用中,设备连接控制系统可接收到用户针对目标CPU设备的设备连接请求,即用户可通过触发指令发送对目标CPU设备的设备连接请求,其中,目标CPU设备可以由设备连接控制系统指定当前可用的CPU设备,同时,该设备连接请求中还需携带有待连接到目标CPU设备的GPU的需求数量,比如,用户A需要5个GPU实例,那么,用户A可向设备连接控制系统发送设备连接请求,其中,设备连接请求中携带的待连接的GPU设备的需求数量为5。

[0067] 步骤204:在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备。

[0068] 其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量。

[0069] 其中,第一资源集群可以理解为目标CPU设备所在的资源集群,第二资源集群可以理解为不存储目标CPU设备的集群,且第二资源集群与第一资源集群的结构相同,均具有多个CPU设备、多个GPU设备,且CPU设备与GPU设备之间通过交换机的连接总线进行连接。

[0070] 第一目标外部设备可以理解为在第一资源集群中所挂载的多个外部设备,第二目标外部设备可以理解为任意一个第二资源集群中所挂在的多个外部设备。

[0071] 可参见图3,图3示出了本说明书实施例提供的一种设备连接方法中资源集群的结构示意图。

[0072] 图3的结构示意图可以理解为第一资源集群的结构示意图,也可以理解为第二资源集群的结果示意图,需要说明的是,该资源集群中的具有两个交换机,分别为第一目标交换机、第二目标交换机,且具有8个机头,分别为机头0、机头1、机头2、机头3、机头4、机头5、机头6、机头7,每个机头中包含两个CPU设备(C0、C1),此外,该资源集群结构中还包括4个机尾,分别为机尾0、机尾1、机尾2、机尾3,其中,每个机尾中包含4个GPU设备(G0、G1、G2、G3),且在每个目标交换机中包括上下两层芯片组(分别为sw u、sw d),一共包含8个芯片组,且通过PCIE连接总线实现任意一个CPU设备可与该资源集群中任意一个GPU设备相连接,其中,机头为一台双CPU服务器,一共有4个x16的port接口。机尾内部除了有8张GPU卡,还可以有四个PCIE芯片。GPU每两个为一组,分属于四个PCIE芯片。每个PCIE暴露出一个x16的port接口,因此池化的最小颗粒度就是2GPU。

[0073] 需要说明的是,本说明书实施例提供的设备连接方法以两个目标交换机实现CPU设备与GPU设备相连接为例进行说明,但并不限定于每个资源集群中仅有两个目标交换机。

[0074] 进一步地,本说明书实施例提供的设备连接方法,为了能够实现CPU设备跨域与其他资源集群中的GPU设备相连接,在每个资源集群中的交换机的芯片内,可配置一个插槽,即参见图3中第一目标交换机与第二目标交换机的sw d层的芯片,每个芯片可配置两个插槽,一个为实现上行流逻辑的插槽,一个为实现下行流逻辑的插槽;实际应用中,可根据该芯片中的插槽实现跨域连接其他资源集群中的GPU设备,具体的实现方式可参见下述实施例对插槽应用的描述。

[0075] 实际应用中,在明确了各个资源集群中的各个设备的拓扑连接链路之后,本说明书实施例提供的设备连接可基于该资源集群确定具体选取外部设备的思路。在设备连接控制系统确定用户需求数量为5时,可在第一资源集群、第二资源集群中选择5个GPU设备,分配给用户使用;例如,若目标CPU设备是在第一资源集群中的与第一目标交换机相连接的机头0,那么,设备连接控制系统可先在第一资源集群中选择待连接的第一目标外部设备,在确定第一目标外部设备的数量不能够满足用户需求数量时,可在至少一个第二资源集群中继续选择第二目标外部设备,并且,第一资源集群与第二资源集群不同,第一目标外部设备与第二目标外部设备的数量之和不小于用户需求数量,也即是说,在用户需求数量为5时,第一目标外部设备的数量与第二目标外部设备的数量之和则不小于5,这种情况就可以有多种组合方式,比如,第一种:第一目标外部设备的数量可以为0,第二目标外部设备的数量则为5或者可大于等于5;第二种:第一目标外部设备的数量可以为2,第二目标外部设备的数量则可为3或者大于3;第三种:第一目标外部设备的数量可以为5,第二目标外部设备的数量则可为0或者大于0。需要说明的是,在第一目标外部设备的数量为0时,即可以理解为在第一资源集群中并没有空闲的GPU可分配给用户使用,则可在第二资源集群中继续寻找可用的GPU设备。

[0076] 进一步地,设备连接控制系统在第一资源集群中,可在单设备交换机的场景下确认是否有可用的GPU设备,也可在双设备交换机的场景下确认是否有可用的GPU设备,进而在第一资源集群中,确定选取的GPU设备的数量是否满足用户需求的GPU设备的数量;具体的,所述确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量,包括:

[0077] 确定第一资源集群中与第一目标交换机相连接的初始外部设备数量;

[0078] 确定第一资源集群中与至少一个第二目标交换机相连接的备选外部设备数量;

[0079] 在确定所述初始外部设备数量与所述备选外部设备数量之和小于所述需求数量的情况下,确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量。

[0080] 其中,第一目标交换机可以理解为目标CPU设备所直接连接的交换机,第二目标交换机可以理解为在第一资源集群中与目标CPU设备没有直接相连的交换机,需要说明的是,本说明书实施例提供的设备连接方法的资源集群中的交换机的数量在此不做过多限定,对于未与目标CPU设备相连接的交换机均可描述为第二目标交换机。

[0081] 初始外部设备数量可以理解为设备连接控制系统在第一资源集群中的与第一目标交换机相连接的外部设备的数量;备选外部设备数量可以理解为设备连接控制系统在第一资源集群中与至少一个第二目标交换机相连接的外部设备的数量。

[0082] 第一目标外部设备可以理解为在初始外部设备与备选外部设备的统称,即在第一资源集群中确定的外部设备。

[0083] 实际应用中,设备连接控制系统在确定第一资源集群中是否有可用的GPU设备时,可将与第一目标交换机相连接的可用GPU设备确定为初始外部设备,再将与至少一个第二目标交换机相连接的可用GPU设备确定为备选外部设备,并统计初始外部设备与备选外部设备的数量之和,若该数量之和小于用户预先需求数量的情况下,设备连接控制系统可确定第一资源集群中待连接的第一目标外部设备并未满足用户预先需求数量。例如,若确定用户预先需求数量为5时,设备连接控制系统可确定初始外部设备数量为1,备选外部设备数量为2,则可确定初始外部设备数量与备选外部设备数量之和小于5,则设备连接控制系

统可确定第一资源集群中的第一目标外部设备的数量为3 (1+2), 因此, 可确定在第一资源集群中待连接的第一目标外部设备不满足用户预先需求数量。

[0084] 本说明书实施例提供的设备连接方法, 设备连接控制系统可确定与第一目标交换机相连接的外部设备的数量以及确定与第二目标交换机相连接的外部设备数量, 进而能够实现单设备交换机场景下确定第一外部设备数量, 或者是在双设备交换机场景下确定第一外部设备数量, 最后确定第一资源集群中待连接的第一目标外部设备是否满足用户需求数量, 便于后续在第二资源集群中继续选取外部设备与目标CPU设备相连接。

[0085] 更进一步地, 设备连接控制系统在第一资源集群中无法选取到满足用户需求数量的第一目标外部设备之后, 设备连接控制系统可通过跨域选择的方式, 在第二资源集群中选取第二目标外部设备, 以实现第一目标外部设备和第二目标外部设备的数量之和满足用户需求数量; 具体的, 所述在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备, 包括:

[0086] 在至少一个第二资源集群中, 确定与所述至少一个第二资源集群中的目标交换机相连接的待连接外部设备;

[0087] 基于所述待连接外部设备确定满足所述需求数量的外部设备为第二目标外部设备。

[0088] 实际应用中, 在除了第一资源集群以外的其他资源集群中, 可能会有至少一个第二资源集群, 在实现跨域连接时, 可能一个第二资源集群确定的第二目标外部设备的数量也并不一定满足用户需求数量, 则还可能在多个第二资源集群中查询, 因此, 设备连接控制系统在至少一个第二资源集群中, 继续确定可用的外部设备作为第二目标外部设备, 但基于上述实施例的描述, 第二资源集群的设备连接结构与第一资源集群相同, 因此, 在第二资源集群中, 与交换机相连接的多个GPU设备中, 是否具有可用的GPU设备, 可将可用的GPU设备作为待连接外部设备, 由此, 与待连接外部设备相连接的交换机则为目标交换机; 其中, 第二资源集群中的交换机的数量与第一资源集群中的数量要求相一致, 不作具体限定。最后, 在第二资源集群中确定的待连接外部设备中选取满足了用户需求数量的外部设备作为第二目标外部设备; 例如, 在第二资源集群中确定的待连接外部设备(可用外部设备)的数量为5, 由于在第一资源集群中已经确定好的第一目标外部设备的数量为2, 用户需求数量为5, 那么, 设备连接控制系统在第二资源集群中可从5个待连接外部设备中可选取3个、4个或者5个外部设备作为第二目标外部设备, 需要强调的是, 用户需求数量为5, 那么设备连接控制系统返回给用户的外部设备数量的最低要求为5, 也可以返回给用户外部设备数量比5个需求数量大, 对于5个以外的外部设备可作为实际应用中的备用的外部设备, 若有任意一个外部设备损坏或者宕机之后, 随时可用备用的外部设备参与实际应用, 本说明书实施例对此不作具体限定, 具体的方式根据不同的应用场景可确定不同的方式。

[0089] 本说明书实施例提供的设备连接方法, 通过在第二资源集群中选取第二目标外部设备, 以满足用户需求数量, 能够较大化地实现对外部资源的利用, 提高资源利用率, 实现资源池化。

[0090] 步骤206: 向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令, 并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路。

[0091] 其中,管理模块可以理解为配置在交换机中,负责配置和管理交换机中的芯片的模块,比如MCPU模块。

[0092] 拓扑链路配置指令可以理解为改变CPU设备和GPU设备相连接的拓扑关系指令。

[0093] 具体的,可参见图4,图4示出了本说明书实施例提供的一种设备连接方法中配置管理模块的资源集群结构示意图。

[0094] 需要说明的是,图4以单设备交换机为例,对管理模块(MCPU模块)所配置的位置进行说明,且强调的是,各个资源集群中的各个交换机中均可配置MCPU模块,本说明书实施例对此不作具体限定。图4中,一台交换机支持四个机头和两个机尾互联。最多可令16张GPU卡都挂载到一个CPU上。每个PCIE芯片都预留一个port用于两个交换机之间的拓展。同时下面的四个PCIE芯片还预留了8个PCIE插槽,可在插槽上安装FPGA卡,FPGA卡可提供进一步的拓展。MCPU模块是配置在交换机中,可负责接收设备连接控制系统下发的控制指令,同时基于该控制指令实现对每个交换机中的CPU设备和GPU设备之间的拓扑链路,以通过调整拓扑链路,提高拓扑链路的利用率,进而也提高对CPU设备和GPU设备的资源分配。

[0095] 实际应用中,设备连接控制系统向第一目标外部设备和第二目标外部设备的各自的MCPU模块发送拓扑链路配置指令,便于MCPU模块基于该拓扑链路配置指令实现对该MCPU模块所对应的交换机所连接的CPU设备和GPU设备之间的拓扑链路,进而,设备连接控制系统就可再次接收到MCPU模块响应于该拓扑链路配置指令确定的对应的拓扑链路,便于后续基于配置后的拓扑链路实现CPU设备和GPU设备之间的连接。

[0096] 进一步地,设备连接控制系统可预先根据资源集群中的资源配置情况,确定好拓扑链路配置规则,再根据该拓扑链路配置规则下发指令;具体的,所述向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,包括:

[0097] 基于第一目标外部设备以及第二目标外部设备确定拓扑链路配置规则;

[0098] 基于所述拓扑链路配置规则向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令。

[0099] 其中,拓扑链路配置规则可以理解为改变CPU设备和GPU设备相连接的拓扑关系的规则。

[0100] 实际应用中,设备连接控制系统可先根据当前资源使用状况确定第一目标外部设备对应的拓扑链路的配置规则,再确定第二目标外部设备对应的拓扑链路的配置规则,再根据该第一目标外部设备对应的拓扑链路的配置规则确定相应的拓扑链路配置指令,再将该拓扑链路配置指令发送至第一目标外部设备对应的MCPU中;同样地,再根据第二目标外部设备对应的拓扑链路的配置规则确定相应的拓扑链路的配置指令,在将该拓扑链路的配置指令发送至第二目标外部设备对应的MCPU中。

[0101] 本说明书实施例提供的设备连接方法,设备连接控制系统通过确定拓扑链路配置规则,进而确定拓扑链路配置规则对应的拓扑链路配置指令,并将该拓扑链路配置指令下发至第一目标外部设备和第二目标外部设备各自的管理模块,以实现第一目标外部设备与目标CPU设备之间,以及第二目标外部设备与目标CPU设备之间的拓扑链路的合理使用,进而提高资源利用率。

[0102] 相应地,设备连接控制系统还可接收到第一目标外部设备和所述第二目标外部设备各自的管理模块确定后的拓扑链路,便于后续基于确定后的拓扑链路实现高利用率的拓

扑链路连接;具体的,所述接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路,包括:

[0103] 接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令解析的所述拓扑链路配置规则;

[0104] 接收所述管理模块基于所述拓扑链路配置规则,在预设拓扑链路图的目标芯片组中调整的处理设备与外部设备之间的更新拓扑链路图;

[0105] 接收所述管理模块基于所述更新拓扑链路图确定的所述目标处理设备与所述第一目标外部设备之间的第一目标拓扑链路,以及所述目标处理设备和所述第二目标外部设备之间的第二目标拓扑链路。

[0106] 其中,更新拓扑链路图可以理解为MCPU模块对原拓扑链路图按照拓扑链路配置规则进行调整后的拓扑链路图。

[0107] 第一目标拓扑链路可以理解为由MCPU模块确定的目标CPU设备与第一目标外部设备相连接的拓扑链路;第二目标拓扑链路可以理解为由MCPU模块确定的目标CPU设备与第二目标外部设备相连接的拓扑链路。

[0108] 实际应用中,设备连接控制系统可接收到第一目标外部设备的MCPU模块以及第二目标外部设备的MCPU模块响应于拓扑链路配置指令解析的各自的拓扑链路配置规则,然后还可接收到第一目标外部设备以及第二目标外部设备各自的MCPU模块基于确定的各自的拓扑链路配置规则,在预设拓扑链路图的目标芯片组中调整的CPU设备和GPU设备之间的更新拓扑链路图,其中,预设拓扑链路图可以理解为上述实施例中资源集群中CPU设备和GPU设备通过交换机相连接的拓扑链路图,目标芯片组可以理解为与确定的第一目标外部设备以及第二目标外部设备相连接的,在其对应的交换机中的目标芯片组(sw);最后,设备连接控制系统还可接收到第一目标外部设备对应的MCPU模块基于更新拓扑链路图确定的目标CPU设备与第一目标外部设备(GPU设备)之间的第一目标拓扑链路,以及,接收到第二目标外部设备对应的MCPU模块基于更新拓扑链路图确定的目标CPU设备与第二目标外部设备(GPU设备)之间的第二目标拓扑链路。

[0109] 可参见图5,图5示出了本说明书实施例提供的一种设备连接方法配置拓扑链路的示意图。

[0110] 图5中的(a)为交换机中原拓扑链路图,图5中的(b)为MCPU模块进行调整后的更新拓扑链路图;其中,图5中的(a),CPU0通过RC,经过交换机中芯片组的拓扑链路,与3个外部设备相连接,CPU1通过RC,经过交换机中芯片组的拓扑链路,与1个外部设备(EP)相连接;在MCPU模块对图5中的(a)的拓扑链路进行调整后,可获得图5中的(b)的拓扑链路,其中,图5中的(b),CPU0通过RC,经过交换机中芯片组的拓扑链路,与2个外部设备相连接,CPU1通过RC,经过交换机中芯片组的拓扑链路,与2个外部设备(EP)相连接。

[0111] 实际应用中,图5中的(a)的CPU0挂载的EP个数有3个,调整为CPU0挂载的EP个数有2个,CPU1挂载的EP个数有1个,调整为CPU1挂载的EP个数有2个;根据PCIE芯片的特性,一个芯片支持多个port,每个port都可设置为up或down。一个up port可下挂多个down port,每个down port只能从属于一个up port。通过配置,up port和down port可自由连接,组成多个树状拓扑。如图5中的芯片,被“切分”成了两个PCIE芯片,而且通过更改芯片内部的拓扑结构,可将下游的EP改配给另一个CPU。单个PCIE芯片的port数量有限,所以需要多个芯片

级联,组成一个更大的交换网络,以支持更多的CPU和EP设备;需要说明的是,根据不同应用场景下的负载情况,按照拓扑链路配置规则对拓扑链路图进行调整,以满足外部设备利用的最大化。

[0112] 本说明书实施例提供的设备连接方法,设备连接控制系统可接收到第一目标外部设备以及第二目标外部设备各自的管理模块确定的拓扑链路,以实现第一目标外部设备按照确定的第一目标拓扑链路与目标CPU设备相连接,以及实现第二目标外部设备按照确定的第二目标拓扑链路与目标CPU设备相连接。

[0113] 步骤208:向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

[0114] 实际应用中,设备连接控制系统还可向目标CPU设备、第一目标外部设备、第二目标外部设备分别发送连接指令,便于后续实现目标CPU设备与第一目标外部设备之间建立连接,目标CPU设备与第二目标外部设备之间建立连接,其中,具体的连接链路可由各自发送的连接指令中携带的拓扑链路。

[0115] 此外,设备连接控制系统在向第一目标外部设备和第二目标外部设备发送连接指令时,可直接将连接指令发送至第一目标外部设备和第二目标外部设备的MCPU模块;具体的,所述向所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,包括:

[0116] 向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送连接指令。

[0117] 实际应用中,目标CPU设备与第一目标外部设备之间连接过程、目标CPU设备与第二目标外部设备之间连接过程,可由第一目标外部设备和第二目标外部设备各自的MCPU模块执行,设备连接控制系统是通过下发指令,控制第一目标外部设备和第二目标外部设备各自的MCPU模块按照指令进行执行。需要说明的是,设备连接控制系统对下发指令的对象以及方式可在本说明书实施例中不作具体限定。

[0118] 本说明书实施例提供的设备连接方法,通过向第一目标外部设备和第二目标外部设备各自的MCPU模块下发指令,便于后续控制MCPU模块即可完成目标CPU设备与第一目标外部设备相连接、目标CPU设备与第二目标外部设备相连接。

[0119] 进一步地,设备连接控制系统在下发连接指令之后,还可控制目标CPU设备与第一目标外部设备相连接、目标CPU设备与第二目标外部设备相连接;具体的,所述向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令之后,还包括:

[0120] 控制所述目标处理设备响应于所述连接指令,基于所述第一目标拓扑链路在所述第一资源集群中与所述第一目标外部设备进行连接;

[0121] 控制所述目标处理设备响应于所述连接指令,基于所述第二目标拓扑链路在所述第二资源集群中与所述第二目标外部设备进行连接。

[0122] 实际应用中,设备连接控制系统控制目标CPU设备响应于下发的连接指令,按照确定的第一目标拓扑链路在第一资源集群中与第一目标外部设备进行连接;还可控制目标CPU设备响应于下发的连接指令,按照确定的第二目标拓扑链路在第二资源集群中与第二目标外部设备进行连接;需要说明的是,本说明书实施例提供的设备连接方法还可直接向第一目标外部设备的MCPU模块以及第二目标外部设备的MCPU模块下发连接指令,由各自的

MCPU模块完成目标CPU设备与第一目标外部设备相连接、目标CPU设备与第二目标外部设备相连接;本实施例对具体的连接方式则不作过多限定。

[0123] 本说明书实施例提供的设备连接方法,设备连接控制系统通过分别下发连接指令的方式,实现目标CPU设备按照第一目标拓扑链路与第一目标外部设备相连,以及目标CPU设备按照第二目标拓扑链路与第二目标外部设备相连,进而实现用户需求数量的实例的应用需求。

[0124] 为了能够实现目标CPU设备与第二资源集群中的第二目标外部设备相连接,则需要从目标CPU设备所在的资源集群中进行跨域,连接到第二资源集群中的外部设备,由此,本说明书实施例提供的设备连接方法提供了一种拓展线卡,以实现跨域连接外部设备;具体的,所述控制所述目标处理设备响应于所述连接指令,基于所述第二目标拓扑链路在所述第二资源集群中与所述第二目标外部设备进行连接,包括:

[0125] 控制所述目标处理设备响应于所述连接指令,基于所述连接指令中携带的第二目标拓扑链路确定与所述目标处理设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述目标处理设备相对应的目标交换插槽中;

[0126] 通过所述拓展线卡控制所述目标处理设备与所述第二资源集群中的第二目标外部设备进行连接。

[0127] 其中,拓展线卡可以理解为一种能够实现跨域连接,具有将PCIE连接线的协议报文转换为网络交换机能够识别的协议报文的功 能,比如,FPGA(Field Programmable Gate Array)卡,FPGA实现了跨以太网络的PCIE交换连接功能。

[0128] 交换插槽可以理解为能够实现上下行流逻辑的插槽,且可将拓展线卡安装于该交换插槽,实现跨域集群的设备连接功能。

[0129] 实际应用中,设备连接控制系统可控制目标CPU设备响应于连接指令,并基于该连接指令中携带的第二目标拓扑线路,确定该目标CPU设备对应的拓展线卡(FPGA卡),并且,该拓展线卡可安装于目标CPU设备对应的交换机中目标芯片组的目标交换插槽中,相对于目标CPU设备而言,需要跨域实现设备连接,可选取目标芯片组中的下行流的插槽作为目标交换插槽。进一步地,设备连接控制系统通过拓展线卡进而控制目标CPU设备与第二资源集群中第二目标外部设备进行连接。

[0130] 需要说明的是,设备连接控制系统也可下发连接指令到第一目标外部设备和第二目标外部设备各自的管理模块,再控制各自的管理模块实现目标CPU设备与第一目标外部设备相连接,以及目标CPU设备与第二目标外部设备相连接。

[0131] 本说明书实施例提供的设备连接方法,设备连接控制系统通过控制目标CPU设备,利用交换机中目标芯片组的交换插槽中安装的拓展线卡,实现目标CPU设备与第二目标外部设备相连接,通过交换插槽中安装的拓展线卡可实现跨以太网络的设备连接。

[0132] 进一步地,通过安装在交换机的目标交换插槽中的拓展线卡,能够实现对设备连接协议的报文转换为交换网络协议的报文,进而保证网络交换机能够识别设备连接中的设备连接协议,便于后续能够实现跨域目标CPU设备与第二目标外部设备相连接;具体的,所述通过所述拓展线卡控制所述目标处理设备与所述第二资源集群中的第二目标外部设备进行连接,包括:

[0133] 通过所述拓展线卡控制所述目标处理设备的设备连接协议的报文转换为交换网



络协议的报文；

[0134] 控制所述目标处理设备基于所述交换网络协议的报文,通过网络交换机与所述第二资源集群中的第二目标外部设备进行连接。

[0135] 具体实施时,设备连接控制系统可控制安装在目标交换插槽中的拓展线卡,对目标CPU设备与第二目标外部设备之间连接的设备连接协议的报文转换为交换网络协议的报文,并控制目标CPU设备基于交换网络协议的报文,且利用网路交换机实现目标CPU设备与第二资源集群中的第二目标外部设备进行连接;实际应用中,FPGA卡内部可将PCIE传输层的报文(TLP)封装成以太网报文,并通过FPGA上的网络接口传输。进而,可将具备这种功能的FPGA卡为PoCE线卡。

[0136] 本说明书实施例提供的设备连接方法,利用拓展线卡实现对设备连接协议的报文进行转换,转换为网络交换机能够可识别的报文格式,便于后续利用网络交换实现目标CPU设备与第二资源集群中的第二目标外部设备相连接。

[0137] 更进一步地,拓展线卡将设备连接协议的报文转换为交换网络协议的报文之后,网络交换机即可将交换网络协议的报文转发至第二资源集群中的第二目标外部设备的目标交换插槽中,进而完成目标CPU设备连接第二外部设备;具体的,所述控制所述目标处理设备基于所述交换网络协议的报文,通过网络交换机与所述第二资源集群中的第二目标外部设备进行连接,包括:

[0138] 控制所述目标处理设备通过网络交换机将接收到的所述交换网络协议的报文发送至所述第二资源集群中的第二目标外部设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述第二目标外部设备相对应的目标交换插槽中;

[0139] 控制所述目标处理设备基于所述拓展线卡将所述交换网络协议的报文进行转换,获得设备连接协议的报文;

[0140] 控制所述目标处理设备基于所述设备连接协议的报文与所述第二资源集群中的第二目标外部设备建立连接。

[0141] 实际应用中,在实现跨域设备连接的过程中,可利用到网络交换机才能实现跨域连接,因此网络交换机在接收到一个连接指令时,只能识别出交换网络协议下的报文,但在第一资源集群和第二资源集群中,设备与设备之间的连接是基于PCIE的设备连接协议,进而,即需要安装在目标CPU设备的目标交换插槽中的拓展线卡,实现对PCIE的设备连接协议的报文封装成以太网报文,能够使得网络交换机识别,同时,网络交换机识别到以太网报文后,还可将该以太网报文转发至第二目标外部设备对应的交换机中目标芯片组的目标交换插槽安装的拓展线卡,也即是说,在第二资源集群中,若想要实现与第二目标外部设备的连接,依旧需要通过PCIE总线的连接,进而第二目标外部设备对应的拓展线卡还需将以太网报文再次转换为设备连接协议报文,才能实现目标CPU设备与第二目标外部设备相连接。

[0142] 参见图6,图6示出了本说明书实施例提供的一种设备连接方法跨域设备连接处理过程的结构示意图。

[0143] 图6中具有一个机头对应的PCIE sw(芯片组)上的交换插槽中安装的FPGA卡,相应地,有3个机尾对应的PCIE sw(芯片组)上的交换插槽中安装的FPGA卡,另外,还有个网络交换机,其中,每个FPGA卡上均具有两个数据转换流,第一个可以理解为上行流(up port-MAC),一个可以理解为下行流(MAC-down port);具体实施时,与机头相连接的FPGA卡可将

设备连接协议的报文转换至以太网协议报文,并将该以太网协议报文转发至网络交换机,在网路交换机接收到以太网协议报文之后,将该以太网协议报文分别转发至第二资源集群中选取的第二目标外部设备对应的FPGA卡中,且第二目标外部设备对应的FPGA卡将以太网协议报文转换至设备连接协议报文,进而实现目标CPU设备与第二目标外部设备相连接。

[0144] 本说明书实施例提供的设备连接方法,通过拓展线卡实现将设备连接协议报文转换为以太网报文,以及将以太网报文再次转换为设备连接协议报文,同时利用网络交换机进行相应的转发,通过该种方式便于实现目标CPU设备与任意一个第二资源集群中的第二目标外部设备相连接。

[0145] 另外,本说明书实施例提供的设备连接方法,还包括:

[0146] 在确定所述第一目标外部设备和所述第二目标外部设备的数量之和小于所述需求数量的情况下,向所述用户发送针对所述目标处理设备的拒绝连接信息。

[0147] 实际应用中,在上述实施例中对第一目标外部设备的数量和第二目标外部设备的数量进行确定之后,若在第一资源集群以及至少一个第二资源集群中均无法找到满足用户需求数量的外部设备时,设备连接控制系统即可向用户发送拒绝连接信息,比如,拒绝连接信息为资源不足,拒绝申请的信息等,本说明书实施例对发送拒绝连接信息的内容以及方式不作任何限定。

[0148] 本说明书实施例提供的设备连接方法,在确定第一资源集群以及至少一个第二资源集群中无法找到满足用户需求数量的外部设备时,设备连接控制系统可直接向用户返回拒绝连接信息,以提高控制系统的处理效率。

[0149] 此外,本说明书实施例提供的设备连接方法还可提供另一种连接情况,即设备连接控制系统在第一资源集群中确定第一目标外部设备时,就可以实现满足用户需求的外部设备的数量,那么,设备连接控制系统可在第一资源集群中实现目标CPU设备与第一目标外部设备相连接;具体的,所述接收用户发送的针对目标处理设备的设备连接请求之后,还包括:

[0150] 在确定第一资源集群中待连接的第一目标外部设备满足所述需求数量的情况下,向所述第一目标外部设备的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备的管理模块响应于所述拓扑链路配置指令确定的目标拓扑链路;

[0151] 向所述目标处理设备与所述第一目标外部设备分别发送连接指令,其中,所述连接指令携带有目标拓扑链路。

[0152] 实际应用中,在设备连接控制系统确定第一资源集群中待连接的第一目标外部设备满足需求数量的情况下,可直接向第一目标外部设备的管理模块发送拓扑链路配置指令,同时可接收到第一目标外部设备的管理模块响应与拓扑链路配置指令确定的目标拓扑链路,最后,再向目标CPU设备和第一目标外部设备分别发送连接指令,根据连接指令中携带的目标拓扑链路,实现目标CPU设备和第一目标外部设备之间建立连接。

[0153] 需要说明的是,本实施例中的具体实现步骤均可参见上述实施例对具体的如何确定第一目标外部设备满足需求数量、如何发送拓扑链路配置指令以及确定目标拓扑链路的方式的描述,本实施例对此则不做具体赘述。

[0154] 本说明书实施例提供的设备连接方法,设备连接控制系统在确定第一资源集群中的第一目标外部设备满足需求数量时,可在第一资源集群中就实现目标CPU设备与外部设

备连接,无需跨域实现设备连接,进而设备连接的速度较快,只有在第一资源集群中的外部设备不能满足用户设备需求数量的情况下,再考虑跨域实现设备连接。

[0155] 综上,通过PCIE芯片级联来做池化,以解决GPU设备的碎片化分布越来越严重的问题以及拓展不够灵活的问题。为了从解决这个问题,让池化的规模不再受限,本说明书实施例提供的设备连接方法通过提出了拓展线卡(PCIE over Converge Ethernet,PoCE线卡)这个概念,通过以太网,传输PCIE协议的报文,让PCIE拓扑具备以太网络的拓展能力,且实际应用中,接入网络中的PoCE线卡,跨以太网实现了一个PICE switch,连通了处在不同PCIE交换机的CPU和GPU。所有的PCIE交换机中CPU/GPU都通过以太网连接到了一起,该池化互联方案以实现高性能和高扩展性这两大特点。

[0156] 本说明书实施例还提供了一种设备连接芯片,用以实现跨域设备连接的功能;具体的,该设备连接芯片,包括:至少一个交换插槽、至少一个上行端口,所述上行端口用于与第一资源集群中处理设备相连接,所述芯片通过所述交换插槽与第二资源集群中的外部设备相连接,其中,所述第一资源集群与所述第二资源集群不同。

[0157] 进一步地,参见图7,图7示出了本说明书实施例提供的一种设备连接芯片结构示意图。

[0158] 图7中包括上行端口、交换插槽,其中,对于上行端口的个数以及交换插槽的个数并不做具体限定,可根据不同的应用场景配置不同的设备连接芯片。实际应用中,该上行端口用户与CPU设备相连接,该交换插槽用于在跨域设备连接时,与跨域集群中的外部设备相连接的插槽,具体的应用方式在本实施例则不作具体限定。

[0159] 本说明书实施例提供的设备连接芯片,在配置在交换机中时,能够实现对跨域设备连接,的功能,提高对外部设备的资源利用率。

[0160] 下述结合附图8,以本说明书提供的设备连接方法在用户申请GPU实例的应用为例,对所述设备连接方法进行进一步说明。其中,图8示出了本说明书一个实施例提供的一种设备连接方法的处理流程图,具体包括以下步骤。

[0161] 需要说明的是,管控系统即可以理解为设备连接控制系统。

[0162] 步骤802:管控系统接收到客户申请GPU实例。

[0163] 步骤804:管控系统接收到申请后,盘点集群中空闲资源的状态。

[0164] 步骤806:管控系统确定单台设备交换机中的空闲资源是否能够满足客户的需求,若是,则执行步骤812,若否,则执行步骤808。

[0165] 步骤808:管控系统确定通过PCIE总线连接的双设备交换机中的空闲资源是否能够满足客户的需求,若是,则执行步骤812,若否,则执行步骤810。

[0166] 步骤810:管控系统确定交换机跨网络调度资源后,能否满足客户的需求,若是,则执行步骤812,若否,则执行步骤804。

[0167] 需要说明的是,若否,则需要管控系统向客户返回资源不足,拒绝申请的信息。

[0168] 步骤812:管控系统下发配置指令给相关的交换机中的MCPU。

[0169] 步骤814:管控系统控制MCPU配置相关的交换机中的目标芯片组和拓展线卡,改变机头的CPU设备和GPU的拓扑关系,并向客户返回开启实例应答申请。

[0170] 需要说明的是,在管控系统向客户返回开启实例应答申请时,即完成了该客户申请GPU实例的请求。

[0171] 本说明书实施例提供的设备连接方法,通过在本地的资源集群中确定空闲资源是否能够满足客户需求,在确定无法满足时,才考虑跨域进行资源调度,并通过改变拓扑关系提高资源使用率,实现资源池化。

[0172] 与上述方法实施例相对应,本说明书还提供了对象处理装置实施例,图9示出了本说明书一个实施例提供的一种设备连接装置的结构示意图。如图9所示,该装置应用于设备连接控制系统,包括:

[0173] 请求接收模块902,被配置为接收用户发送的针对目标处理设备的设备连接请求,其中,所述设备连接请求携带待连接外部设备的需求数量;

[0174] 设备确定模块904,被配置为在确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量的情况下,在至少一个第二资源集群中确定满足所述需求数量的第二目标外部设备,其中,所述第一目标外部设备和所述第二目标外部设备的数量之和不小于所述需求数量;

[0175] 拓扑链路确定模块906,被配置为向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令确定的对应的拓扑链路;

[0176] 指令发送模块908,被配置为向所述目标处理设备、所述第一目标外部设备和所述第二目标外部设备分别发送连接指令,其中,所述连接指令中携带有对应的拓扑链路。

[0177] 可选地,所述设备确定模块904,进一步被配置为:

[0178] 确定第一资源集群中与第一目标交换机相连接的初始外部设备数量;

[0179] 确定第一资源集群中与至少一个第二目标交换机相连接的备选外部设备数量;

[0180] 在确定所述初始外部设备数量与所述备选外部设备数量之和小于所述需求数量的情况下,确定第一资源集群中待连接的第一目标外部设备不满足所述需求数量。

[0181] 可选地,所述设备确定模块904,进一步被配置为:

[0182] 在至少一个第二资源集群中,确定与所述至少一个第二资源集群中的目标交换机相连接的待连接外部设备;

[0183] 基于所述待连接外部设备确定满足所述需求数量的外部设备为第二目标外部设备。

[0184] 可选地,所述拓扑链路确定模块906,进一步被配置为:

[0185] 基于第一目标外部设备以及第二目标外部设备确定拓扑链路配置规则;

[0186] 基于所述拓扑链路配置规则向所述第一目标外部设备和所述第二目标外部设备各自的管理模块发送拓扑链路配置指令。

[0187] 可选地,所述拓扑链路确定模块906,进一步被配置为:

[0188] 接收所述第一目标外部设备和所述第二目标外部设备各自的管理模块响应于所述拓扑链路配置指令解析的所述拓扑链路配置规则;

[0189] 接收所述管理模块基于所述拓扑链路配置规则,在预设拓扑链路图的目标芯片组中调整的处理设备与外部设备之间的更新拓扑链路图;

[0190] 接收所述管理模块基于所述更新拓扑链路图确定的所述目标处理设备与所述第一目标外部设备之间的第一目标拓扑链路,以及所述目标处理设备和所述第二目标外部设备之间的第二目标拓扑链路。

[0191] 可选地,所述装置,还包括:

[0192] 控制模块,被配置为控制所述目标处理设备响应于所述连接指令,基于所述第一目标拓扑链路在所述第一资源集群中与所述第一目标外部设备进行连接;

[0193] 控制所述目标处理设备响应于所述连接指令,基于所述第二目标拓扑链路在所述第二资源集群中与所述第二目标外部设备进行连接。

[0194] 可选地,所述控制模块,进一步被配置为:

[0195] 控制所述目标处理设备响应于所述连接指令,基于所述连接指令中携带的第二目标拓扑链路确定与所述目标处理设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述目标处理设备相对应的目标交换插槽中;

[0196] 通过所述拓展线卡控制所述目标处理设备与所述第二资源集群中的第二目标外部设备进行连接。

[0197] 可选地,所述控制模块,进一步被配置为:

[0198] 通过所述拓展线卡控制所述目标处理设备的设备连接协议的报文转换为交换网络协议的报文;

[0199] 控制所述目标处理设备基于所述交换网络协议的报文,通过网络交换机与所述第二资源集群中的第二目标外部设备进行连接。

[0200] 可选地,所述控制模块,进一步被配置为:

[0201] 控制所述目标处理设备通过网络交换机将接收到的所述交换网络协议的报文发送至所述第二资源集群中的第二目标外部设备相对应的拓展线卡,其中,所述拓展线卡安装于与所述第二目标外部设备相对应的目标交换插槽中;

[0202] 控制所述目标处理设备基于所述拓展线卡将所述交换网络协议的报文进行转换,获得设备连接协议的报文;

[0203] 控制所述目标处理设备基于所述设备连接协议的报文与所述第二资源集群中的第二目标外部设备建立连接。

[0204] 可选地,所述拓扑链路确定模块906,进一步被配置为:

[0205] 在确定第一资源集群中待连接的第一目标外部设备满足所述需求数量的情况下,向所述第一目标外部设备的管理模块发送拓扑链路配置指令,并接收所述第一目标外部设备的管理模块响应于所述拓扑链路配置指令确定的目标拓扑链路;

[0206] 可选地,所述指令发送模块908,进一步被配置为:

[0207] 向所述目标处理设备与所述第一目标外部设备分别发送连接指令,其中,所述连接指令携带有目标拓扑链路。

[0208] 本说明书实施例提供的设备连接装置,设备连接控制系统通过在控制与目标处理设备相连接外部设备的资源选择的过程中,先通过在第一资源集群中确定第一外部设备能否满足用户的需求,在该第一外部设备不满足用户的需求时,可在至少一个第二资源集群确定外部设备,作为第二目标外部设备,同时,再对第一目标外部设备和第二目标外部设备的管理模型下发对拓扑链路的配置指令,进而确定目标处理设备与第一目标外部设备和第二目标外部设备的具体链接的拓扑链路,最后再下发连接指令,便于后续目标处理设备、第一目标外部设备、第二目标外部设备响应于该连接指令,实现目标处理设备和第一目标外部设备连接,或者是目标处理设备和第二目标外部设备连接,该种对外部设备资源的选择

方式不仅能够实现在本地资源集群(第一资源集群)中选择与目标处理设备相连接的外部设备,还能够实现跨域资源集群(第二资源集群)中选择与目标处理设备相连接的外部设备,体现了连接设备的拓展性和资源选择的灵活性,进而能够进一步解决资源池化问题。

[0209] 上述为本实施例的一种设备连接装置的示意性方案。需要说明的是,该设备连接装置的技术方案与上述的设备连接方法的技术方案属于同一构思,设备连接装置的技术方案未详细描述的细节内容,均可以参见上述设备连接方法的技术方案的描述。

[0210] 图10示出了根据本说明书一个实施例提供的一种计算设备1000的结构框图。该计算设备1000的部件包括但不限于存储器1010和处理器1020。处理器1020与存储器1010通过总线1030相连接,数据库1050用于保存数据。

[0211] 计算设备1000还包括接入设备1040,接入设备1040使得计算设备1000能够经由一个或多个网络1060通信。这些网络的示例包括公用交换电话网(PSTN)、局域网(LAN)、广域网(WAN)、个域网(PAN)或诸如因特网的通信网络的组合。接入设备1040可以包括有线或无线的任何类型的网络接口(例如,网络接口卡(NIC))中的一个或多个,诸如IEEE802.11无线局域网(WLAN)无线接口、全球微波互联接入(Wi-MAX)接口、以太网接口、通用串行总线(USB)接口、蜂窝网络接口、蓝牙接口、近场通信(NFC)接口,等等。

[0212] 在本说明书的一个实施例中,计算设备1000的上述部件以及图10中未示出的其他部件也可以彼此相连接,例如通过总线。应当理解,图10所示的计算设备结构框图仅仅是出于示例的目的,而不是对本说明书范围的限制。本领域技术人员可以根据需要,增添或替换其他部件。

[0213] 计算设备1000可以是任何类型的静止或移动计算设备,包括移动计算机或移动计算设备(例如,平板计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等)、移动电话(例如,智能手机)、可佩戴的计算设备(例如,智能手表、智能眼镜等)或其他类型的移动设备,或者诸如台式计算机或PC的静止计算设备。计算设备1000还可以是移动式或静止式的服务器。

[0214] 其中,处理器1020用于执行如下计算机可执行指令,该计算机可执行指令被处理器执行时实现上述设备连接方法的步骤。

[0215] 上述为本实施例的一种计算设备的示意性方案。需要说明的是,该计算设备的技术方案与上述的对象处理方法的技术方案属于同一构思,计算设备的技术方案未详细描述的细节内容,均可以参见上述设备连接方法的技术方案的描述。

[0216] 本说明书一实施例还提供一种计算机可读存储介质,其存储有计算机可执行指令,该计算机可执行指令被处理器执行时实现上述设备连接方法的步骤。

[0217] 上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是,该存储介质的技术方案与上述的设备连接方法的技术方案属于同一构思,存储介质的技术方案未详细描述的细节内容,均可以参见上述设备连接方法的技术方案的描述。

[0218] 上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下,在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外,在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中,多任务处理和并行处理也是可以的或者可能是有利的。

[0219] 所述计算机指令包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是,所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减,例如在某些司法管辖区,根据立法和专利实践,计算机可读介质不包括电载波信号和电信信号。

[0220] 需要说明的是,对于前述的各方法实施例,为了简便描述,故将其都表述为一系列的动作组合,但是本领域技术人员应该知悉,本说明书实施例并不受所描述的动作顺序的限制,因为依据本说明书实施例,某些步骤可以采用其它顺序或者同时进行。其次,本领域技术人员也应该知悉,说明书中所描述的实施例均属于优选实施例,所涉及的动作和模块并不一定都是本说明书实施例所必须的。

[0221] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中沒有详述的部分,可以参见其它实施例的相关描述。

[0222] 以上公开的本说明书优选实施例只是用于帮助阐述本说明书。可选实施例并没有详尽叙述所有的细节,也不限制该发明仅为所述的具体实施方式。显然,根据本说明书实施例的内容,可作很多的修改和变化。本说明书选取并具体描述这些实施例,是为了更好地解释本说明书实施例的原理和实际应用,从而使所属技术领域技术人员能很好地理解和利用本说明书。本说明书仅受权利要求书及其全部范围和等效物的限制。

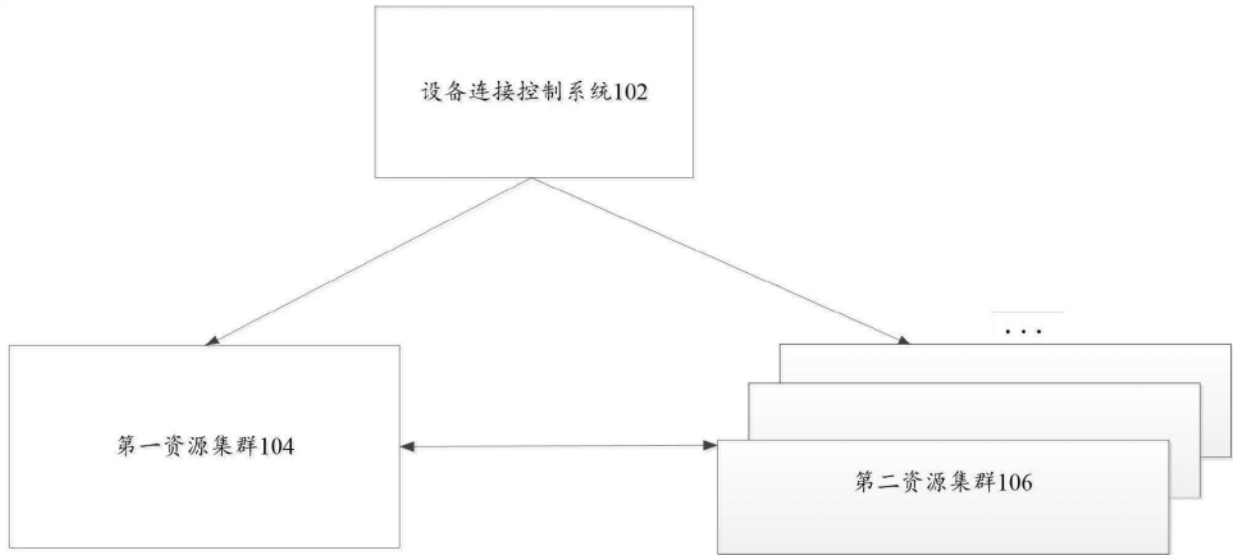


图1

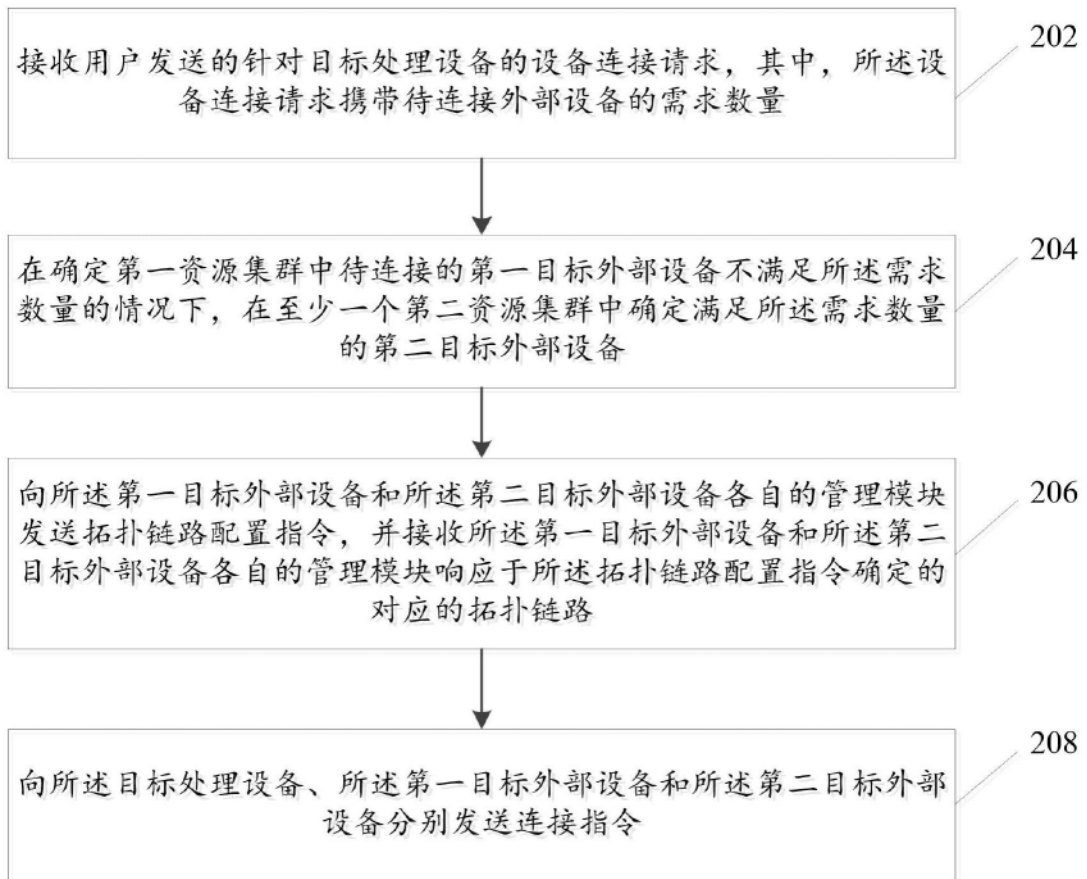


图2



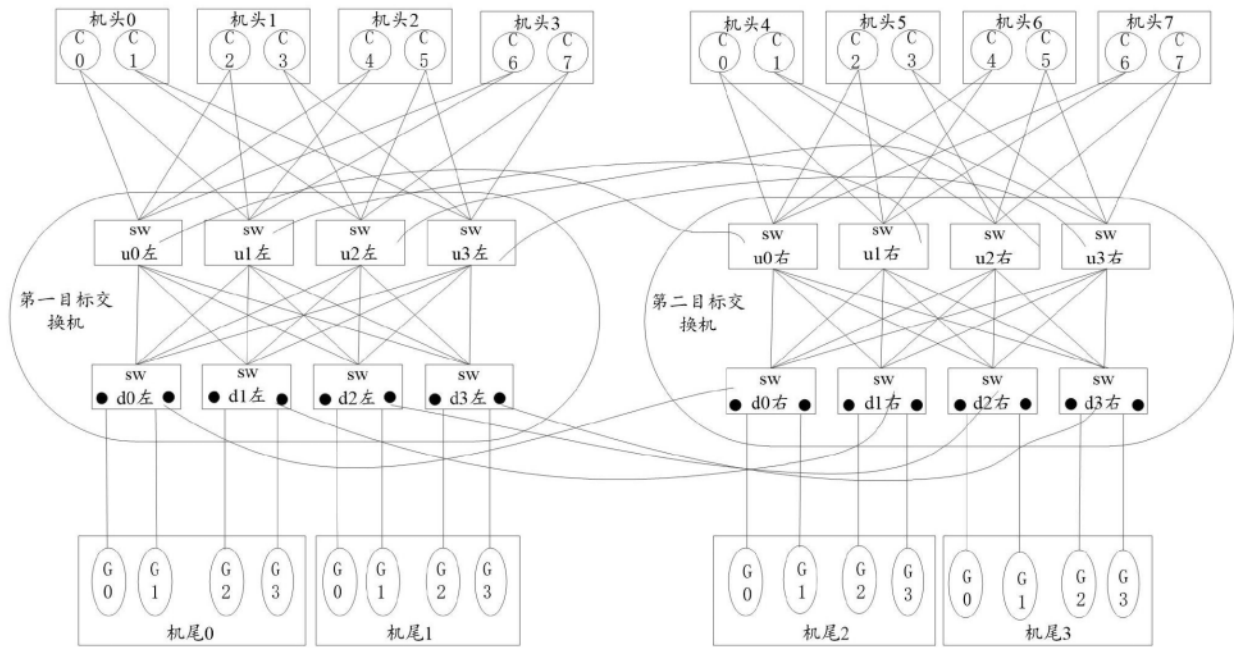


图3

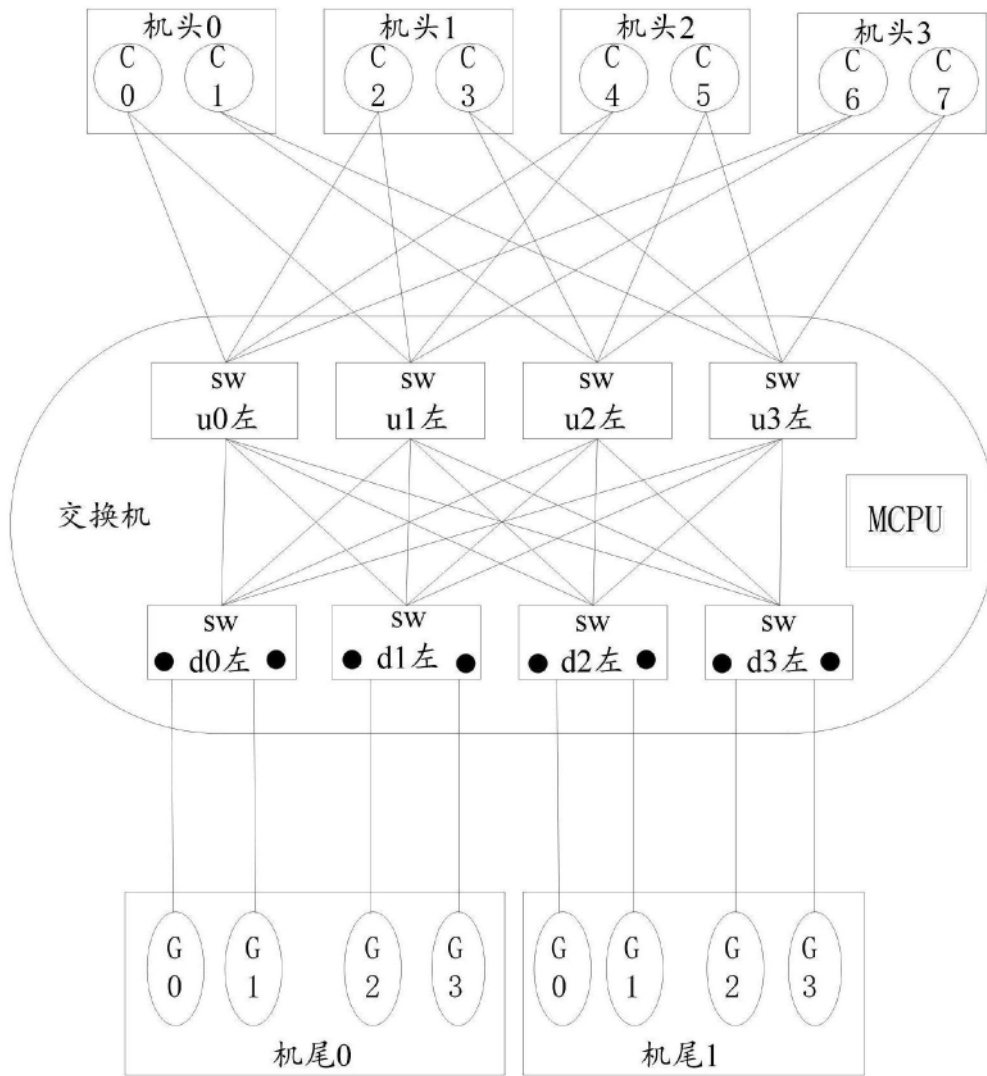


图4

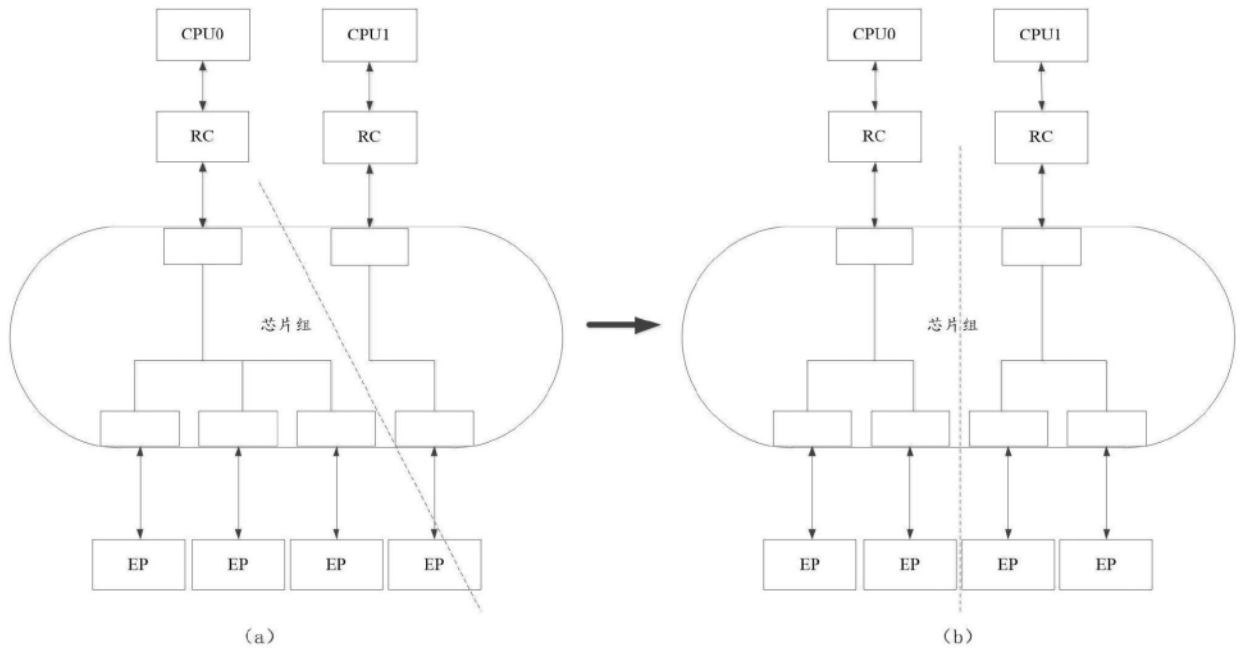


图5

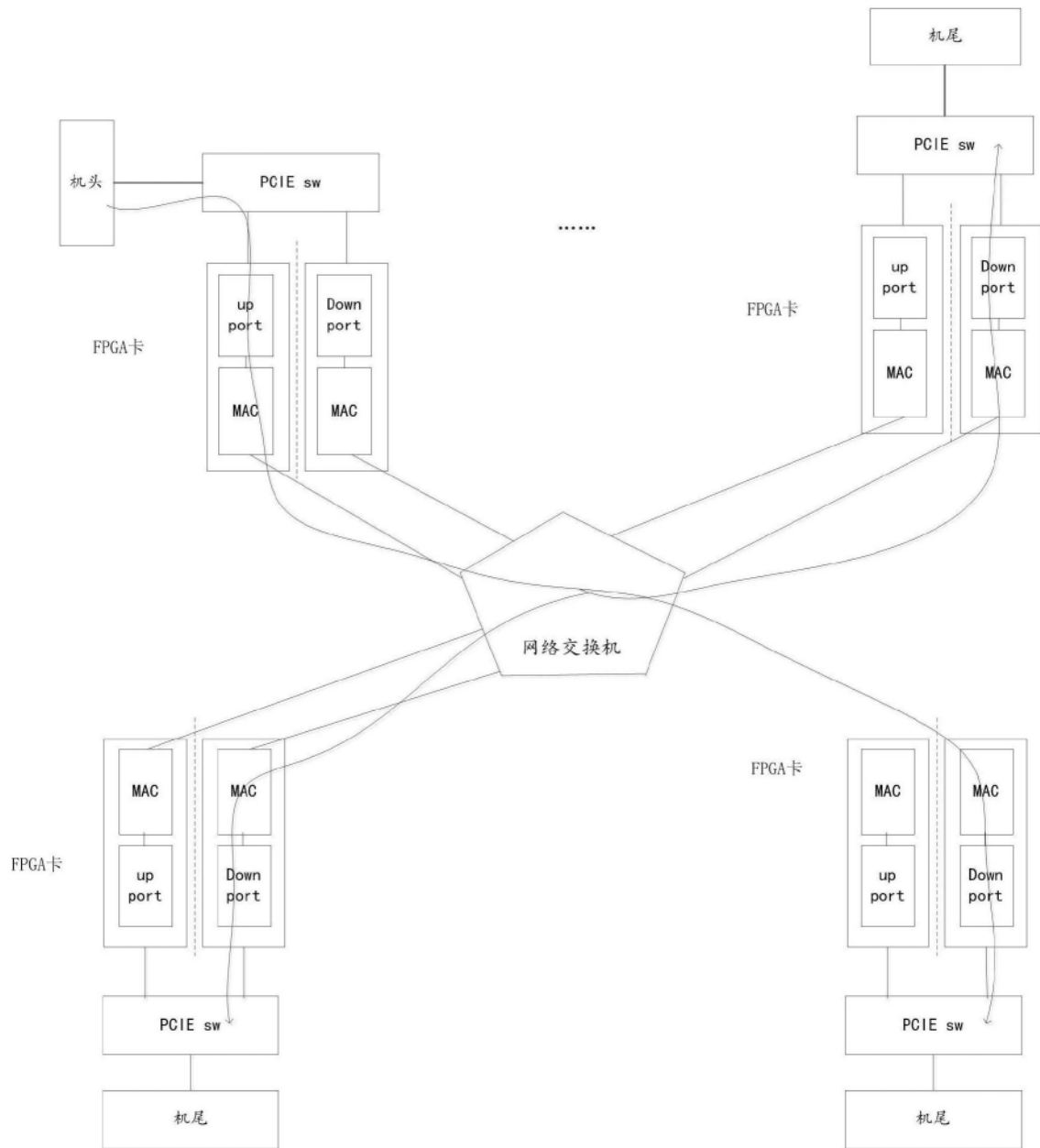


图6

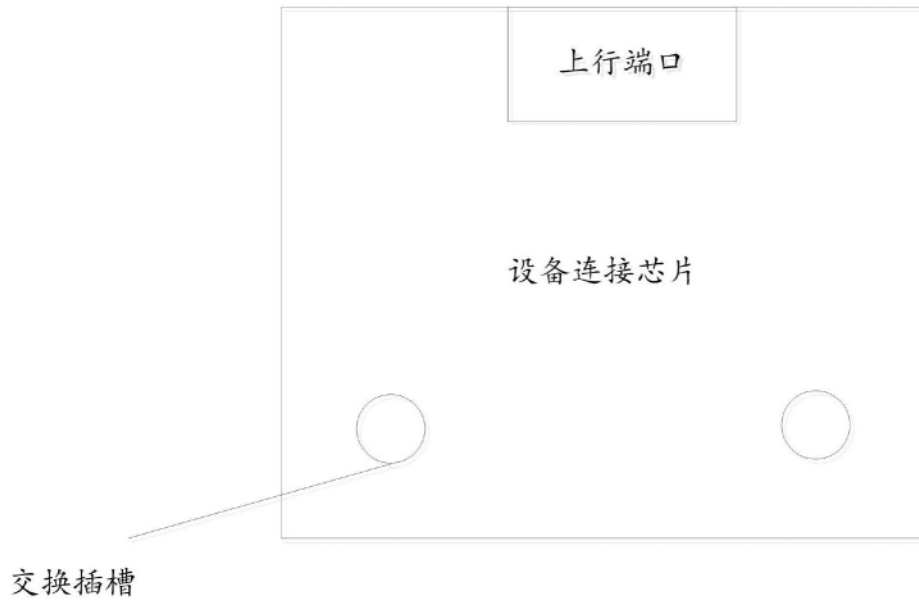


图7

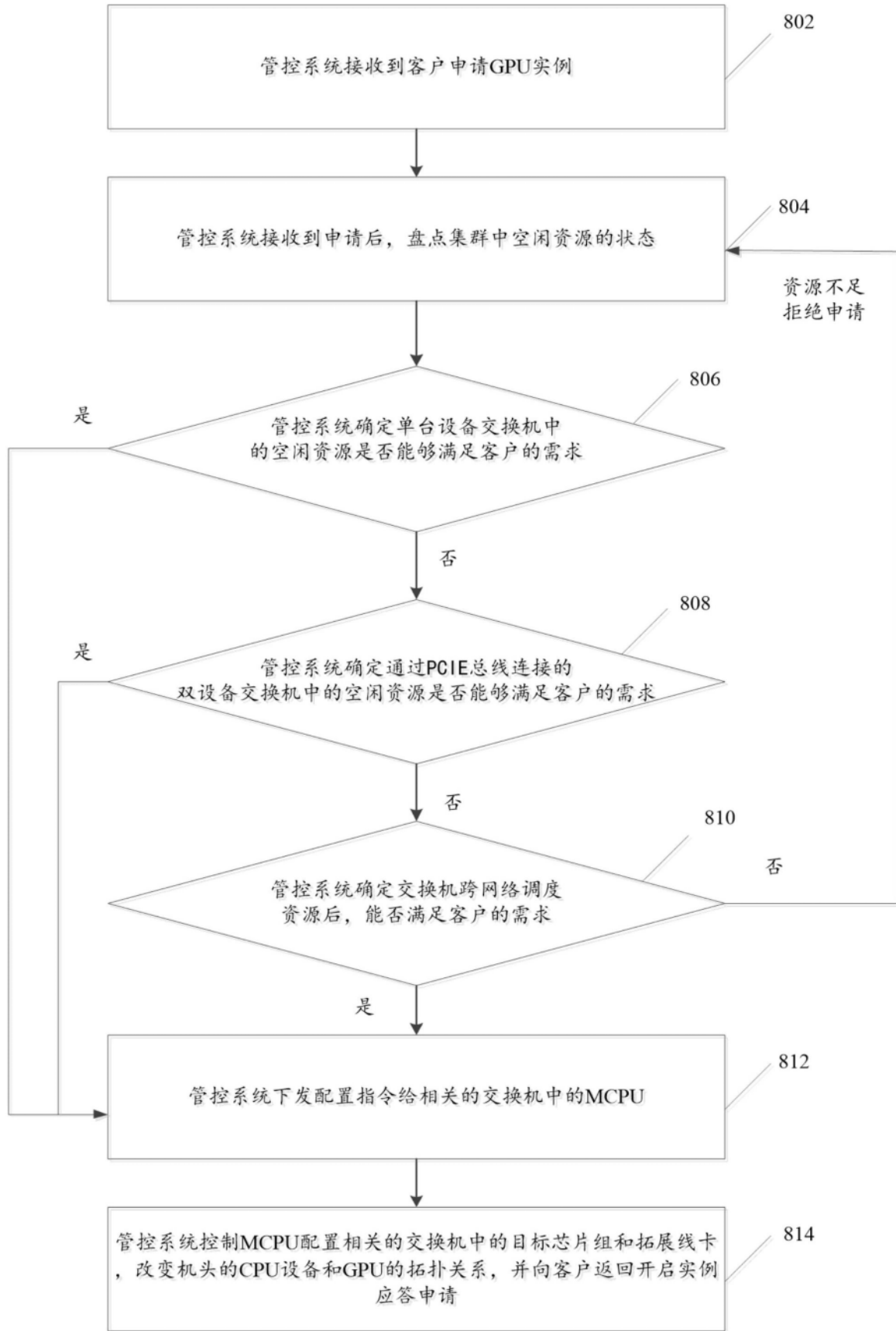


图8

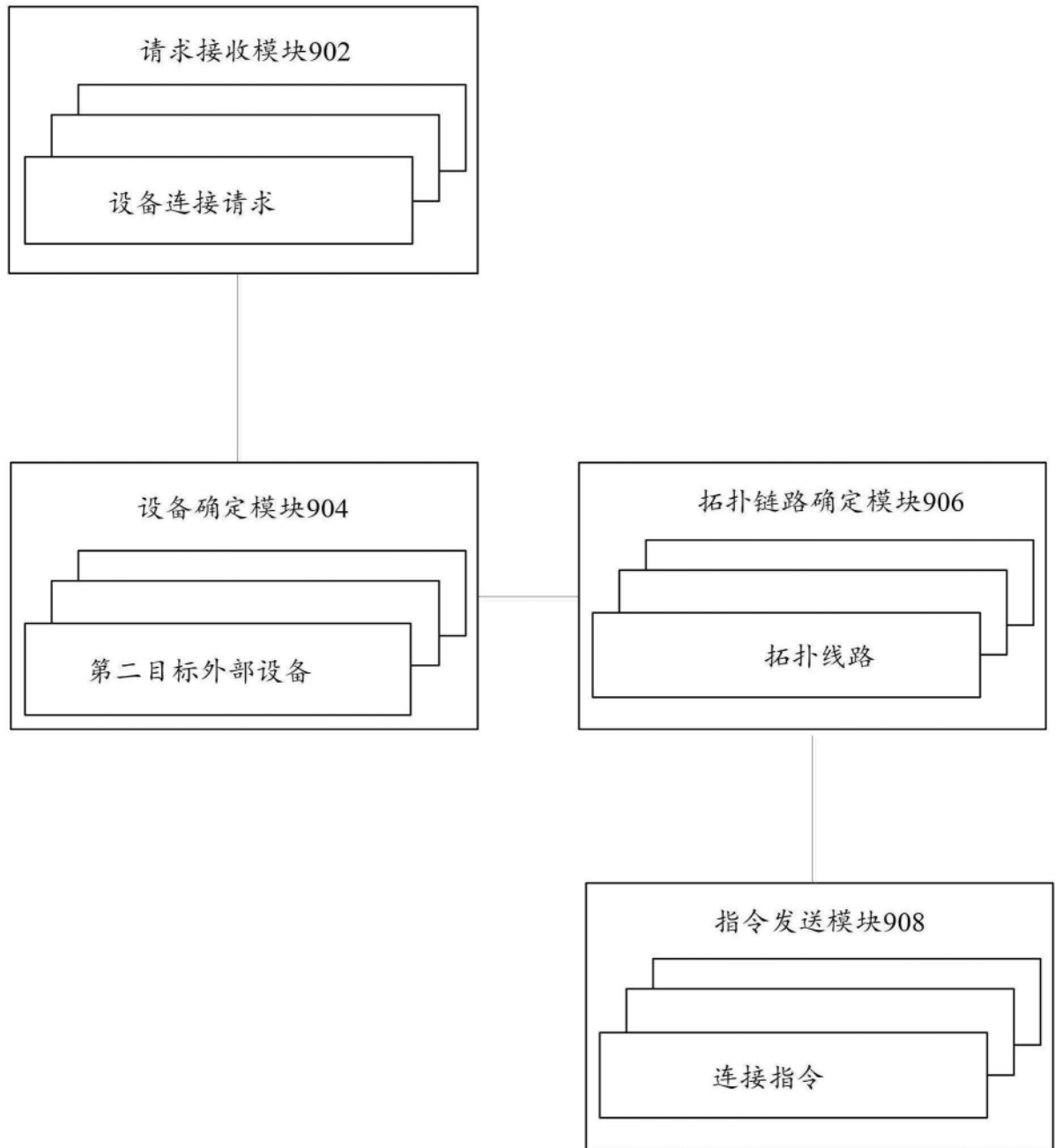


图9

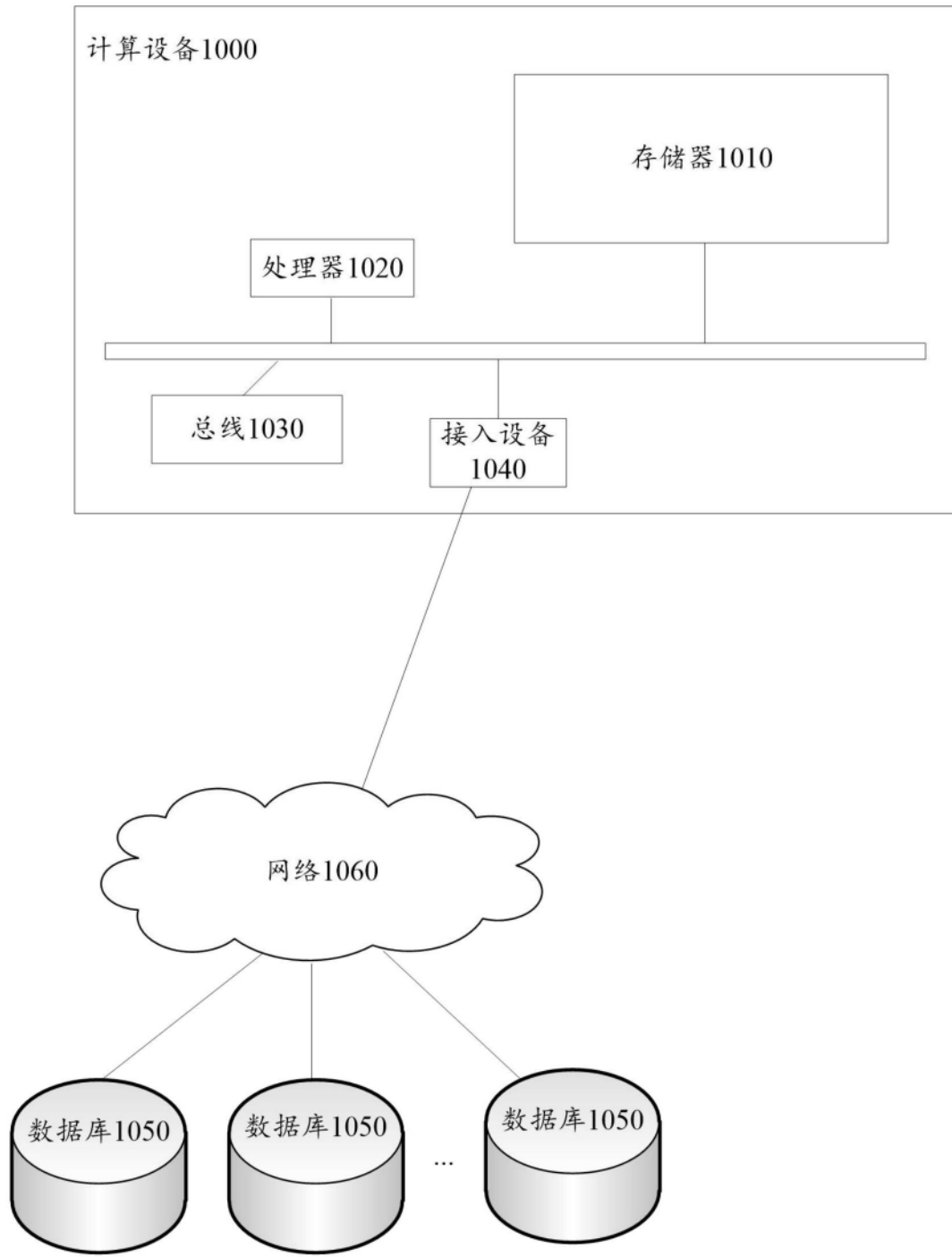


图10