



(12) 发明专利申请

(10) 申请公布号 CN 113177018 A

(43) 申请公布日 2021.07.27

(21) 申请号 202110449133.7

(22) 申请日 2021.04.25

(71) 申请人 山东英信计算机技术有限公司

地址 250101 山东省济南市高新区浪潮路
1036号浪潮科技园S05号楼北3层北区

(72) 发明人 于泉泉 王焕超 付水论 叶明洋
刘闻禹 闫玉婕 韩煦

(74) 专利代理机构 北京连和连知识产权代理有限公司 11278

代理人 刘小峰 宋薇薇

(51) Int.Cl.

G06F 15/78 (2006.01)

G06F 13/42 (2006.01)

G06F 1/18 (2006.01)

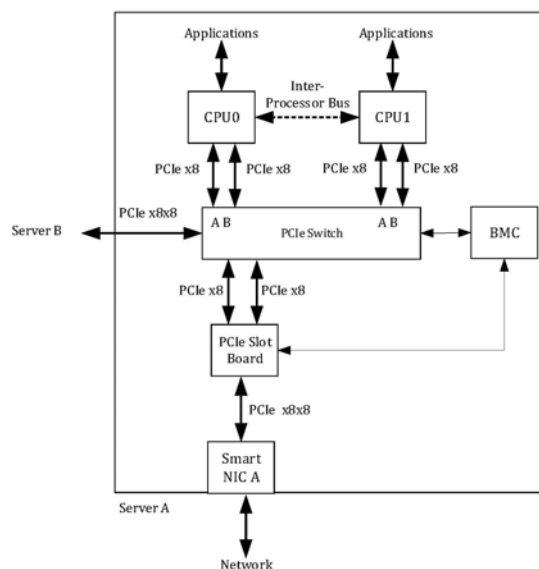
权利要求书2页 说明书8页 附图3页

(54) 发明名称

一种使用双槽CPU的服务器

(57) 摘要

本发明公开了一种使用双槽CPU的服务器，包括：多个双槽CPU，每个双槽CPU具有至少两个PCIE x8插槽；PCIE切换器，具有多个PCIE x8插槽，PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽；第一PCIE插接板，通过两个PCIE x8通道连接到PCIE切换器；第一智能网卡，通过两个PCIE x8通道连接到第一PCIE插接板，同时通过外部接口连接到外部网络；基板管理控制器，连接到PCIE切换器和第一PCIE插接板，基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板。本发明能够提升Socket Direct拓扑在实际应用中的适配能力和可用性，同时降低维护成本和设备资源浪费。



1. 一种使用双槽CPU的服务器,其特征在于,包括:
多个双槽CPU,每个所述双槽CPU具有至少两个PCIE x8插槽;
PCIE切换器,具有多个PCIE x8插槽,所述PCIE切换器通过不同的PCIE x8通道分别连接到多个所述双槽CPU的两个PCIE x8插槽;
第一PCIE插接板,通过两个PCIE x8通道连接到所述PCIE切换器;
第一智能网卡,通过两个PCIE x8通道连接到所述第一PCIE插接板,同时通过外部接口连接到外部网络;
基板管理控制器,连接到所述PCIE切换器和所述第一PCIE插接板,所述基板管理控制器配置为响应于检测到所述第一智能网卡在位而控制所述PCIE切换器将来自两个不同的所述双槽CPU的PCIE x8通道导通到所述第一PCIE插接板。
2. 根据权利要求1所述的服务器,其特征在于,多个所述双槽CPU之间通过处理器间通信总线相互连接。
3. 根据权利要求1所述的服务器,其特征在于,所述PCIE切换器还通过两个PCIE x8通道连接到另一服务器。
4. 根据权利要求1所述的服务器,其特征在于,所述基板管理控制器通过I2C总线和/或GPIO总线连接到所述PCIE切换器和所述第一PCIE插接板。
5. 根据权利要求1所述的服务器,其特征在于,所述基板管理控制器配置为若从所述第一PCIE插接板接收到已插接所述第一智能网卡的在位信号,确认所述第一智能网卡在位,其中所述第一智能网卡兼容直接套接拓扑。
6. 根据权利要求5所述的服务器,其特征在于,所述PCIE切换器中存储有固件,所述基板管理控制器配置为向所述PCIE切换器发送第一固件配置以使得所述PCIE切换器的所述固件控制所述PCIE切换器将来自两个不同的所述双槽CPU的PCIE x8通道导通到所述第一PCIE插接板。
7. 一种使用双槽CPU的服务器,其特征在于,包括:
多个双槽CPU,每个所述双槽CPU具有至少两个PCIE x8插槽;
PCIE切换器,具有多个PCIE x8插槽,所述PCIE切换器通过不同的PCIE x8通道分别连接到多个所述双槽CPU的两个PCIE x8插槽;
第一PCIE插接板,通过两个PCIE x8通道连接到所述PCIE切换器;
第二PCIE插接板,通过两个PCIE x8通道连接到所述PCIE切换器;
第一智能网卡,通过两个PCIE x8通道连接到所述第一PCIE插接板,同时通过外部接口连接到外部网络;
第二智能网卡,通过两个PCIE x8通道连接到所述第二PCIE插接板,同时通过外部接口连接到外部网络;
基板管理控制器,连接到所述PCIE切换器、所述第一PCIE插接板和所述第二PCIE插接板,所述基板管理控制器配置为响应于检测到所述第一智能网卡和所述第二智能网卡在位而控制所述PCIE切换器将来自两个不同的所述双槽CPU的PCIE x8通道导通到所述第一PCIE插接板和所述第二PCIE插接板。
8. 根据权利要求7所述的服务器,其特征在于,所述基板管理控制器配置为若从所述第二PCIE插接板接收到已插接所述第二智能网卡的在位信号,确认所述第二智能网卡在位,

其中所述第二智能网卡兼容直接套接拓扑;所述PCIE切换器中存储有固件,所述基板管理控制器配置为向所述PCIE切换器发送第一固件配置以使得所述PCIE切换器的所述固件控制所述PCIE切换器将来自两个不同的所述双槽CPU的PCIE x8通道导通到所述第二PCIE插接板。

9. 一种使用双槽CPU的服务器,其特征在于,包括:

多个双槽CPU,每个所述双槽CPU具有至少两个PCIE x8插槽;

PCIE切换器,具有多个PCIE x8插槽,所述PCIE切换器通过不同的PCIE x8通道分别连接到多个所述双槽CPU的两个PCIE x8插槽;

第一PCIE插接板,通过两个PCIE x8通道连接到所述PCIE切换器;

第二PCIE插接板,通过两个PCIE x8通道连接到所述PCIE切换器;

第一智能网卡,通过两个PCIE x8通道连接到所述第一PCIE插接板,同时通过外部接口连接到外部网络;

第二外部设备,通过一个PCIE x16通道连接到所述第二PCIE插接板,同时通过外部接口连接到外部网络;

基板管理控制器,连接到所述PCIE切换器、所述第一PCIE插接板和所述第二PCIE插接板,所述基板管理控制器配置为响应于检测到所述第一智能网卡和所述第二外部设备在位而控制所述PCIE切换器将来自两个不同的所述双槽CPU的PCIE x8通道导通到所述第一PCIE插接板,同时控制所述PCIE切换器将来自同一个所述双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到所述第二PCIE插接板。

10. 根据权利要求9所述的服务器,其特征在于,所述基板管理控制器配置为若从所述第二PCIE插接板接收到已插接所述第二外部设备的在位信号,确认所述第二外部设备在位,其中所述第二外部设备不兼容直接套接拓扑;所述PCIE切换器中存储有固件,所述基板管理控制器配置为向所述PCIE切换器发送第二固件配置以使得所述PCIE切换器的所述固件控制所述PCIE切换器将来自同一个所述双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到所述第二PCIE插接板。

一种使用双槽CPU的服务器

技术领域

[0001] 本发明涉及计算机领域,更具体地,特别是指一种使用双槽CPU的服务器。

背景技术

[0002] 随着数据的指数级增长,企业和云提供商要求服务器和计算资源具有更高的性能,以便对大量数据进行实时分析。支持Multi-Host(多主机)技术的智能网卡能够实现性能的提升。例如智能网卡的PCIE(快速外围组件互连拓展)接口如果是两个x8槽位,且这两个x8分别来自不同的CPU(中央处理器),这使得它能够在多个CPU处理器的内核间平衡网络通信的性能,并且减少开销。

[0003] 现有技术中,数据中心服务器常用的配置是一个Multi-Socket(多槽CPU)主板和一个高速网卡。智能网卡一般采用PCIE x16槽位,Socket Direct(直接套接拓扑)的服务器拓扑实现服务器中的每个CPU到网络的直接访问,从而提高了Dual-Socket(双槽CPU)服务器的性能。但现有技术的问题是:

[0004] (1) 方案呆板,适配不灵活。Dual-Socket有2个CPU,每个CPU都有多个PCIE x16槽位,每个x16槽位理论上又可以出2个x8槽位,可能会衍生出几十种分配组合。上行PCIE x8没法根据设计需要灵活配置,只能在设计之初固定好用哪两个x8槽位,后期无法改动。如果可以通过线缆实现不同x8槽位的选通会灵活一些,但是随之而来的是线缆种类成倍增加和走线难题。

[0005] (2) 维护困难。在已有的方案基础上,如果可以通过线缆调整到其他方案,需要维护人员停机、重新配置接线。

[0006] (3) 配置资源浪费。Multi-Host的智能网卡占用了x16的PCIE槽位,如果不支持Multi-Host的其它x16外插设备接入进来,也有可能导导致外插设备工作异常。如果弃用该槽位,无疑是一种大的配置浪费。当该服务器中需要支持多个智能网卡时,这种配置浪费会进一步加剧。

[0007] (4) 串行通路可靠性低。如果关键领域服务器搭配一张智能网卡,当某x8链路异常时,整个链路都没法继续使用,提高整机更换概率。

[0008] 针对现有技术中Socket Direct服务器适配死板、维护困难、资源浪费、鲁棒性低的问题,目前尚无有效的解决方案。

发明内容

[0009] 有鉴于此,本发明实施例的目的在于提出一种使用双槽CPU的服务器和使用双槽CPU的服务器,能够提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费。

[0010] 基于上述目的,本发明实施例的第一方面提供了一种使用双槽CPU的服务器,包括:

[0011] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;

[0012] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;

[0013] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;

[0014] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;

[0015] 基板管理控制器,连接到PCIE切换器和第一PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板。

[0016] 在一些实施方式中,多个双槽CPU之间通过处理器间通信总线相互连接。

[0017] 在一些实施方式中,PCIE切换器还通过两个PCIE x8通道连接到另一服务器。

[0018] 在一些实施方式中,基板管理控制器通过I2C总线和/或GPIO总线连接到PCIE切换器和第一PCIE插接板。

[0019] 在一些实施方式中,基板管理控制器配置为若从第一PCIE插接板接收到已插接第一智能网卡的在位信号,确认所述第一智能网卡在位,其中第一智能网卡兼容直接套接拓扑。

[0020] 在一些实施方式中,PCIE切换器中存储有固件,基板管理控制器配置为向PCIE切换器发送第一固件配置以使得PCIE切换器的固件控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板。

[0021] 本发明实施例的第二方面提供了一种使用双槽CPU的服务器,包括:

[0022] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;

[0023] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;

[0024] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;

[0025] 第二PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;

[0026] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;

[0027] 第二智能网卡,通过两个PCIE x8通道连接到第二PCIE插接板,同时通过外部接口连接到外部网络;

[0028] 基板管理控制器,连接到PCIE切换器、第一PCIE插接板和第二PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡和第二智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板和第二PCIE插接板。

[0029] 在一些实施方式中,基板管理控制器配置为若从第二PCIE插接板接收到已插接第二智能网卡的在位信号,确认所述第二智能网卡在位,其中第二智能网卡兼容直接套接拓扑;PCIE切换器中存储有固件,基板管理控制器配置为向PCIE切换器发送第一固件配置以使得PCIE切换器的固件控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第二PCIE插接板。

[0030] 本发明实施例的第三方面提供了一种使用双槽CPU的服务器,包括:

[0031] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;

[0032] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连

接到多个双槽CPU的两个PCIE x8插槽；

[0033] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器；

[0034] 第二PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器；

[0035] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络；

[0036] 第二外部设备,通过一个PCIE x16通道连接到第二PCIE插接板,同时通过外部接口连接到外部网络；

[0037] 基板管理控制器,连接到PCIE切换器、第一PCIE插接板和第二PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡和第二外部设备在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板,同时控制PCIE切换器将来自同一个双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到第二PCIE插接板。

[0038] 在一些实施方式中,基板管理控制器配置为若从第二PCIE插接板接收到已插接第二外部设备的在位信号,确认所述第二外部设备在位,其中第二外部设备不兼容直接套接拓扑;PCIE切换器中存储有固件,基板管理控制器配置为向PCIE切换器发送第二固件配置以使得PCIE切换器的固件控制PCIE切换器将来自同一个双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到第二PCIE插接板。

[0039] 本发明具有以下有益技术效果:本发明实施例提供的使用双槽CPU的服务器,通过使用多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;基板管理控制器,连接到PCIE切换器和第一PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板的技术方案,能够提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费。

附图说明

[0040] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0041] 图1为本发明提供的使用双槽CPU的服务器一个实施例的电路原理图；

[0042] 图2为本发明提供的使用双槽CPU的服务器另一实施例的电路原理图；

[0043] 图3为本发明提供的使用双槽CPU的服务器又一实施例的电路原理图。

具体实施方式

[0044] 为使本发明的目的、技术方案和优点更加清楚明白,以下结合具体实施例,并参照附图,对本发明实施例进一步详细说明。

[0045] 需要说明的是,本发明实施例中所有使用“第一”和“第二”的表述均是为了区分两

个相同名称非相同的实体或者非相同的参量,可见“第一”“第二”仅为了表述的方便,不应理解为对本发明实施例的限定,后续实施例对此不再一一说明。

[0046] 基于上述目的,本发明实施例的第一个方面,提出了一种提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费的使用双槽CPU的服务器的一个实施例。图1示出的是本发明提供的使用双槽CPU的服务器的电路原理图。

[0047] 所述的使用双槽CPU的服务器如图1所示包括:

[0048] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;

[0049] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;

[0050] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;

[0051] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;

[0052] 基板管理控制器,连接到PCIE切换器和第一PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板。

[0053] 在一些实施方式中,多个双槽CPU之间通过处理器间通信总线相互连接。

[0054] 在一些实施方式中,PCIE切换器还通过两个PCIE x8通道连接到另一服务器。

[0055] 在一些实施方式中,基板管理控制器通过I2C总线和/或GPIO总线连接到PCIE切换器和第一PCIE插接板。

[0056] 在一些实施方式中,基板管理控制器检测到第一智能网卡在位包括:基板管理控制器配置为若从第一PCIE插接板接收到已插接第一智能网卡的在位信号,确认所述第一智能网卡在位,其中第一智能网卡兼容直接套接拓扑。

[0057] 在一些实施方式中,基板管理控制器控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板包括:PCIE切换器中存储有固件,基板管理控制器配置为向PCIE切换器发送第一固件配置以使得PCIE切换器的固件控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板。

[0058] 下面根据图1所示的具体实施例进一步阐述本发明的具体实施方式。

[0059] 智能网卡的核心是通过FPGA(现场可编程门阵列)协助CPU处理网络负载,编程网络接口功能。智能网卡通过FPGA本地化编程可以支持数据面和控制面功能定制,协助CPU处理网络负载。智能网卡通常包含多个端口和内部交换机,快速转发数据并基于网络数据包、应用程序套接字等智能映射到相关应用程序,同时还检测和管理网络流量。

[0060] 智能网卡也能够提升应用程序和虚拟化性能,实现软件定义网络(SDN)和网络功能虚拟化(NFV)的诸多优势,将网络虚拟化、负载均衡和其他低级功能从服务器CPU中移除,确保为应用提供最大的处理能力。与此同时,智能网卡还能够提供分布式计算资源,使得用户可以开发自己的软件或提供接入服务,从而加速特定应用程序。

[0061] 对于智能网卡而言,在传统的Dual-Socket服务器中,通常只连接到其中本地CPU的PCIE插槽。如果远程CPU需要访问网络,则必须穿过处理器间通信总线,进入本地CPU后再到达智能网卡,然后继续再访问外部网络。当数据密集型应用程序在不同的CPU上竞争访问单个网络设备时,这种拓扑可能会出现性能瓶颈。

[0062] 相反地, Dual-Socket服务器中的远程CPU不再需要通过CPU间的互联总线, 可以直接与智能网卡通过PCIE总线互联。这种拓扑下, 每个PCIE x8通道都可以直连服务器中的CPU, 每个CPU到网络的直接连接意味着互连可以绕过CPU间互联总线和另一CPU, 从而优化性能和改进延迟。正是这种Socket Direct拓扑带来了较低的延迟, 使得CPU利用率也得到了提高, 因为每个CPU只处理自己的流量, 而不处理来自另一个CPU的流量, 发挥了智能网卡的优势, 最大限度地提高数据中心的投资回报。

[0063] 这种采用Socket Direct拓扑的实现方式, 其上行两个PCIE x8分别来自Dual-Socket服务器的两个不同CPU, 而Dual-Socket服务器的每个CPU会有多个PCIE槽位。这就需要开发人员在设计之初就预留好来自不同CPU的PCIE x8线路给到同一张PCIE插接板。预留的方式可以通过固定的PCIE槽位经过主板走线到PCIE插接板, 或者选择预留某些PCIE槽位到连接器, 然后连接线缆来到PCIE插接板。智能网卡通过PCIE金手指接到PCIE插接板上, 从而实现与不同CPU的两路PCIE x8槽位的交互。PCIE插接板仅是一个结构上的转接卡, 它可以直接或者通过线缆与服务器主板相连, 将PCIE信号通过PCIE x16槽位传递给智能网卡。如果智能网卡上行支持线缆接入PCIE, 形式上它也可以被线缆代替。

[0064] 还有一种带Adapter(适配器)的Dual-Socket服务器的Socket Direct系统方案。其采用Socket Direct拓扑的实现方式, 上行两个PCIE x8分别来自Dual-Socket服务器的两个不同CPU。其中一个CPU直接与PCIE插接板互联, 另一个CPU通过直连CPU适配卡(Socket Direct Adapter)与PCIE插接板的另一路PCIE x8接口连接。智能网卡通过PCIE金手指接到PCIE插接板上, 从而实现与不同CPU的两路PCIE x8槽位的交互。由于采用直连CPU适配卡可以占用服务器上一个通用的PCIE槽位, 大大提高了该拓扑的兼容性。直连CPU适配卡拥有一个NIC(智能网卡), 上行通过PCIE x8槽位与CPU连接, 下行通过连接器、线缆与智能网卡连接, 可以作为CPU与智能网卡之间的信号桥接。它可以给Multi-Socket服务器带来的一个关键好处是消除了通过多CPU的内部总线的网络通信量, 从而显著减少了CPU开销和延迟。

[0065] 然而, 上述两方案存在以下问题:

[0066] (1) 方案呆板, 适配不灵活。Dual-Socket有2个CPU, 每个CPU都有多个PCIE x16槽位, 每个x16槽位理论上又可以出2个x8槽位, 可能会衍生出几十种分配组合。上行PCIE x8没法根据设计需要灵活配置, 只能在设计之初固定好用哪两个x8槽位, 后期无法改动。如果可以通过线缆实现不同x8槽位的选通会灵活一些, 但是随之而来的是线缆种类成倍增加和走线难题。

[0067] (2) 维护困难。在已有的方案基础上, 如果可以通过线缆调整到其他方案, 需要维护人员停机、重新配置接线。

[0068] (3) 配置资源浪费。Multi-Host的智能网卡占用了x16的PCIE槽位, 如果不支持Multi-Host的其它x16外插设备接入进来, 也有可能导导致外插设备工作异常。如果弃用该槽位, 无疑是一种大的配置浪费。当该服务器中需要支持多个智能网卡时, 这种配置浪费会进一步加剧。

[0069] (4) 串行通路可靠性低。如果关键领域服务器搭配一张智能网卡, 当某x8链路异常时, 整个链路都没法继续使用, 提高整机更换概率。

[0070] 相对于此, 本发明在Socket Direct拓扑的基础上, 进一步提出一种自适应多主智能网卡的服务器拓扑。它可兼容传统和Socket Direct两种拓扑, 采用BMC识别智能网卡在

位和类型后,运用PCIE Switch(PCIE切换器)灵活切换上下行PCIE通道,从而自适应适配智能网卡的Multi-Host技术,具有方案灵活、维护方便、易于扩展等优点,是性能和投资成本之间的良好选择。

[0071] 参见图1,本发明实施例采用Socket Direct拓扑的实现方式,上行多个PCIE x8分别来自Dual-Socket服务器的两个CPU(例如每个CPU出2个PCIE x8)。两个PCIE x8接入到PCIE Switch(可以搭配在主板上或者是PCIE外插卡)。PCIE Switch可以根据定义的固件配置,将下行的PCIE灵活配置。例如2个PCIE x8可以来自不同的两个CPU,也可以来自相同的两个CPU。组成PCIE x8x8给到PCIE插接板。如果是Multi-Host智能网卡通过PCIE金手指接到PCIE插接板上,从而实现与不同CPU的两路PCIE x8槽位的交互;如果不支持Multi-Host的外插设备通过PCIE金手指接到另一组PCIE槽位上,也可以实现PCIE x16的交互。根据PCIE Switch上行打开的PCIE通道数量,可以兼容多个Multi-Host智能网卡设备,灵活满足各种服务器拓扑配置所需。该拓扑中,BMC先通过I2C或者GPIO连接到PCIE插接板上,通过识别接入设备的在位和类型来识别是否是满足Multi-Host智能网卡;如果是,BMC通过I2C总线告知PCIE Switch,PCIE Switch配置寄存器识别到下行哪两个x8槽位已经接入了智能网卡,就会加载对应的固件配置,从而自动使能两个不同CPU的2个x8 PCIE连接到该下行链路上。如果需要,PCIE Switch还可以将两个PCIE x8槽位接到另一台服务器的PCIE Switch上,实现更大的跨域的Fabric互联,接入更多CPU或者智能网卡设备。

[0072] 在工作流程上,BMC先通过I2C或者GPIO连接到PCIE插接板上,通过识别接入设备的在位和类型来识别是否是满足Multi-Host智能网卡BMC读取网卡类型是否为PCIE x8x8;(PCIE协议定义:X16槽位上的B82管脚的电平用于指示当前智能网卡是需要工作在PCIE x8x8,支持Multi-Host;A32/A33管脚用于提供PCIE lane15-8的时钟信号;A50管脚用于提供PCIE lane15-8的复位信号)。如果是,BMC通过I2C总线告知PCIE Switch,PCIE Switch配置寄存器识别到下行哪两个x8槽位已经接入了智能网卡,就会加载对应的固件配置,从而自动使能两个不同CPU的2个x8 PCIE连接到该下行链路上。如果不支持Multi-Host,BMC告知PCIE Switch为非Direct Socket,分配同一个CPU的两个x8槽位给到下行2个x8 PCIE链路上。由此,智能网卡完成PCIE链路初始化。

[0073] 从上述实施例可以看出,本发明实施例提供的使用双槽CPU的服务器,使用多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;基板管理控制器,连接到PCIE切换器和第一PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板的技术方案,能够提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费。

[0074] 基于上述目的,本发明实施例的第二个方面,提出了一种提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费的使用双槽CPU的服务器的一个实施例。

[0075] 所述的使用双槽CPU的服务器如图2所示包括:

- [0076] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;
- [0077] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;
- [0078] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;
- [0079] 第二PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;
- [0080] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;
- [0081] 第二智能网卡,通过两个PCIE x8通道连接到第二PCIE插接板,同时通过外部接口连接到外部网络;
- [0082] 基板管理控制器,连接到PCIE切换器、第一PCIE插接板和第二PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡和第二智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板和第二PCIE插接板。
- [0083] 在一些实施方式中,基板管理控制器检测到第二智能网卡在位包括:基板管理控制器配置为若从第二PCIE插接板接收到已插接第二智能网卡的在位信号,确认所述第二智能网卡在位,其中第二智能网卡兼容直接套接拓扑;基板管理控制器控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第二PCIE插接板包括:PCIE切换器中存储有固件,基板管理控制器配置为向PCIE切换器发送第一固件配置以使得PCIE切换器的固件控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第二PCIE插接板。
- [0084] 从上述实施例可以看出,本发明实施例提供的使用双槽CPU的服务器,使用多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;基板管理控制器,连接到PCIE切换器和第一PCIE插接板,基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板的技术方案,能够提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费。
- [0085] 基于上述目的,本发明实施例的第三个方面,提出了一种提升Socket Direct拓扑在实际应用中的适配能力和可用性,同时降低维护成本和设备资源浪费的使用双槽CPU的服务器的一个实施例。
- [0086] 所述的使用双槽CPU的服务器如图3所示包括:
- [0087] 多个双槽CPU,每个双槽CPU具有至少两个PCIE x8插槽;
- [0088] PCIE切换器,具有多个PCIE x8插槽,PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽;
- [0089] 第一PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;
- [0090] 第二PCIE插接板,通过两个PCIE x8通道连接到PCIE切换器;
- [0091] 第一智能网卡,通过两个PCIE x8通道连接到第一PCIE插接板,同时通过外部接口连接到外部网络;
- [0092] 第二外部设备,通过一个PCIE x16通道连接到第二PCIE插接板,同时通过外部接

口连接到外部网络;

[0093] 基板管理控制器, 连接到PCIE切换器、第一PCIE插接板和第二PCIE插接板, 基板管理控制器配置为响应于检测到第一智能网卡和第二外部设备在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板, 同时控制PCIE切换器将来自同一个双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到第二PCIE插接板。

[0094] 在一些实施方式中, 基板管理控制器检测到第二外部设备在位包括: 基板管理控制器配置为若从第二PCIE插接板接收到已插接第二外部设备的在位信号, 确认所述第二外部设备在位, 其中第二外部设备不兼容直接套接拓扑; 基板管理控制器控制PCIE切换器将来自同一个双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到第二PCIE插接板包括: PCIE切换器中存储有固件, 基板管理控制器配置为向PCIE切换器发送第二固件配置以使得PCIE切换器的固件控制PCIE切换器将来自同一个双槽CPU的两个PCIE x8通道合并作为一个PCIE x16通道导通到第二PCIE插接板。

[0095] 从上述实施例可以看出, 本发明实施例提供的使用双槽CPU的服务器, 使用多个双槽CPU, 每个双槽CPU具有至少两个PCIE x8插槽; PCIE切换器, 具有多个PCIE x8插槽, PCIE切换器通过不同的PCIE x8通道分别连接到多个双槽CPU的两个PCIE x8插槽; 第一PCIE插接板, 通过两个PCIE x8通道连接到PCIE切换器; 第一智能网卡, 通过两个PCIE x8通道连接到第一PCIE插接板, 同时通过外部接口连接到外部网络; 基板管理控制器, 连接到PCIE切换器和第一PCIE插接板, 基板管理控制器配置为响应于检测到第一智能网卡在位而控制PCIE切换器将来自两个不同的双槽CPU的PCIE x8通道导通到第一PCIE插接板的技术方案, 能够提升Socket Direct拓扑在实际应用中的适配能力和可用性, 同时降低维护成本和设备资源浪费。

[0096] 所属领域的普通技术人员应当理解: 以上任何实施例的讨论仅为示例性的, 并非旨在暗示本发明实施例公开的范围 (包括权利要求) 被限于这些例子; 在本发明实施例的思路下, 以上实施例或者不同实施例中的技术特征之间也可以进行组合, 并存在如上所述的本发明实施例的不同方面的许多其它变化, 为了简明它们没有在细节中提供。因此, 凡在本发明实施例的精神和原则之内, 所做的任何省略、修改、等同替换、改进等, 均应包含在本发明实施例的保护范围之内。

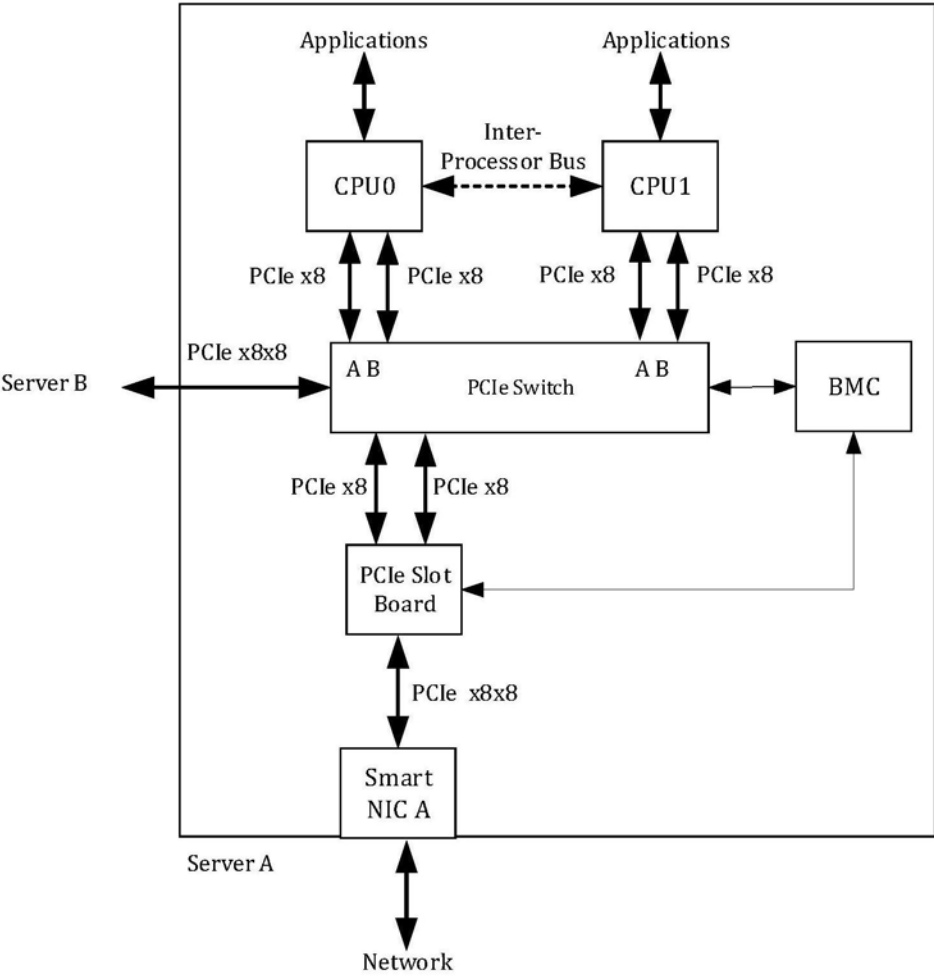


图1

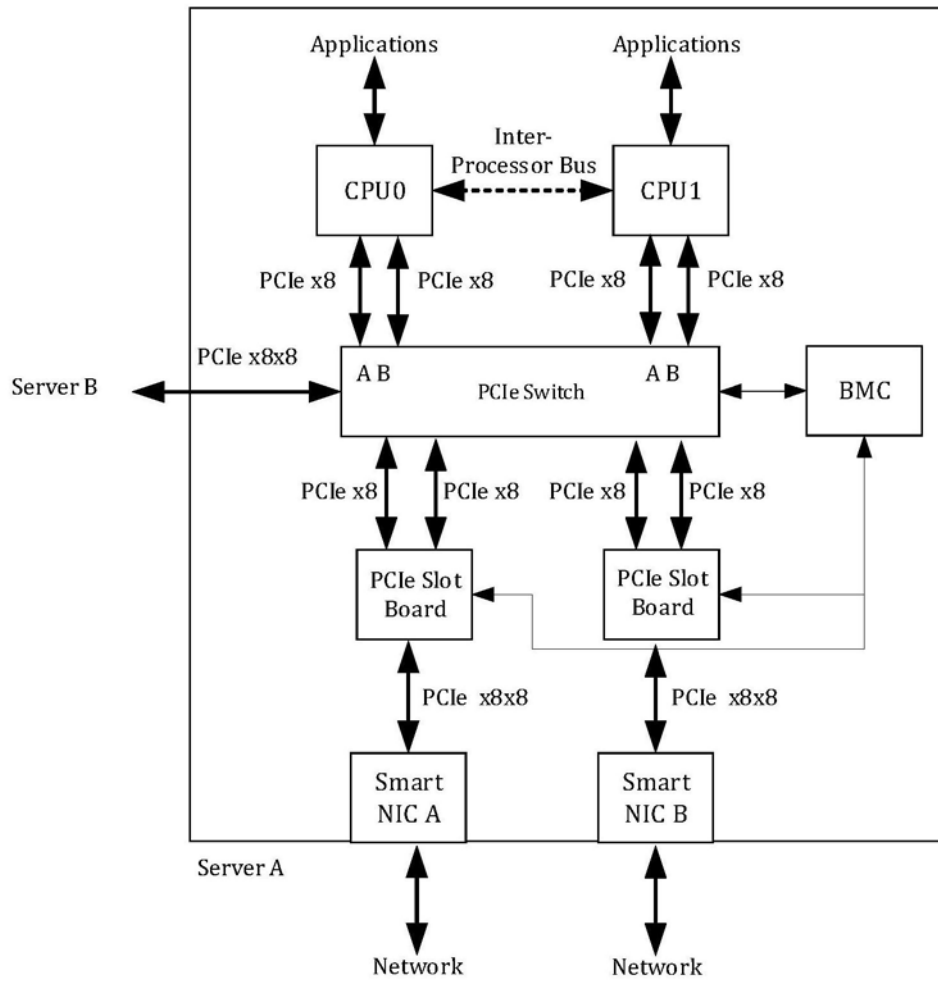


图2

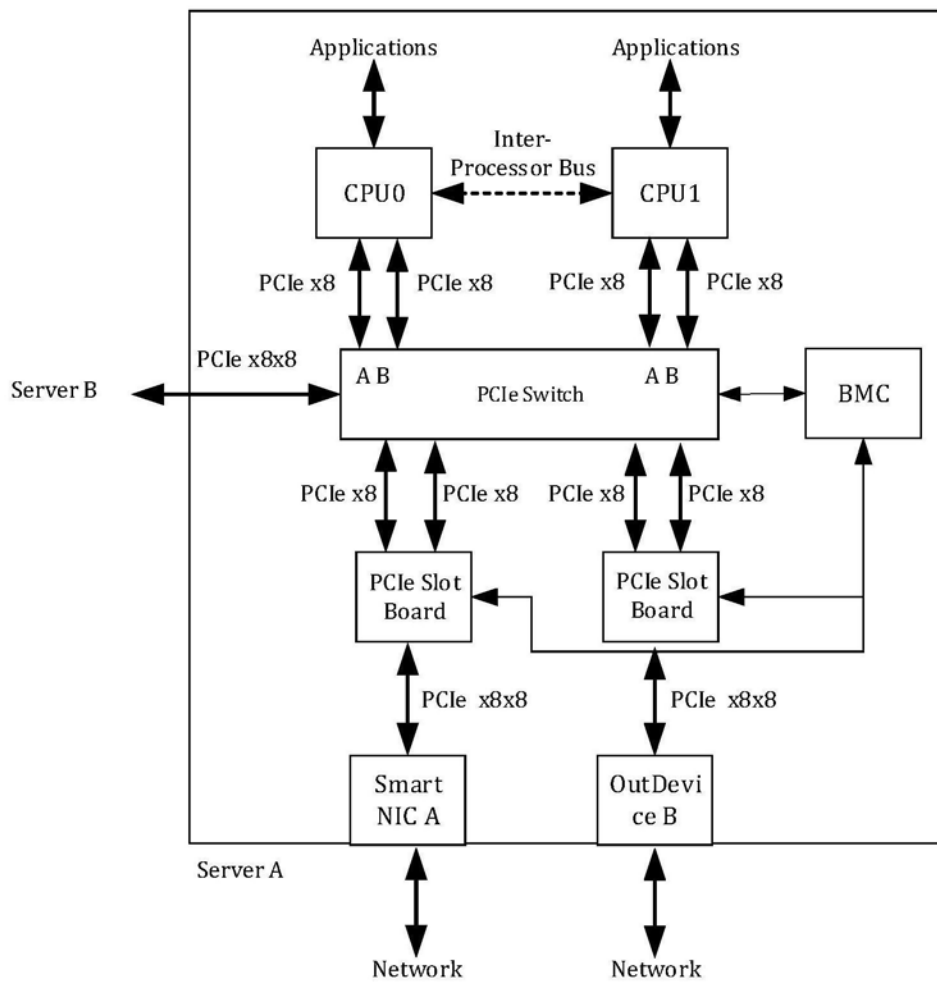


图3