(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114445260 B
(45)授权公告日 2024.01.12

(21)申请号 202210051088.4
(22)申请日 2022.01.17
(65)同一申请的已公布的文献号
申请公布号 CN 114445260 A
(43)申请公布日 2022.05.06
(73)专利权人 苏州浪潮智能科技有限公司
地址 215168 江苏省苏州市吴中经济开发区郭巷街道官浦路1号9幢
(72)发明人 张静东 王江为 王媛丽 阚宏伟
(74)专利代理机构 北京市万慧达律师事务所 11111
专利代理师 黄玉东
(51)Int.Cl.
G06T 1/20(2006.01)
G06F 9/50(2006.01)

(56)对比文件
CN 104583933 A,2015.04.29
CN 107391432 A,2017.11.24
CN 108804376 A,2018.11.13
CN 109240832 A,2019.01.18
CN 113900793 A,2022.01.07
US 2018174268 A1,2018.06.21
Adrian M. Caulfield等.A cloud-scale acceleration architecture.2016 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO).2016,全文.

审查员 徐苏宁

权利要求书2页 说明书7页 附图2页

(54)发明名称

基于FPGA的分布式GPU通信的方法及装置

(57)摘要

本申请涉及一种基于FPGA的分布式GPU通信的方法及装置,该方法应用于通信装置,通信装置包括第一FPGA处理芯片,第一FPGA处理芯片通过接收GPU发送的请求信息,请求信息包括数据存储地址,从GPU的数据存储地址中读取数据,根据从远程资源调度中心设备接收到的配置信息向服务器发送数据,或者根据配置信息向第二FPGA处理芯片发送数据,FPGA处理芯片作为GPU与服务器之间的转接卡,降低了GPU与服务器之间的耦合度,还可以与相邻的FPGA处理芯片形成二维环形网络拓扑,极大地提高了GPU间的灵活性,减少GPU之间互相通信的时间。

CN 114445260 B

1.一种基于FPGA的分布式GPU通信的方法,其特征在于,所述方法应用于通信装置,所述通信装置包括第一FPGA处理芯片,所述第一FPGA处理芯片包括第一接口、第二接口、第一网络接口和多个第二网络接口;所述第一FPGA处理芯片的第一接口与GPU通信连接,所述第一FPGA处理芯片的第二接口与服务器通信连接;所述第一网络接口与远程资源调度中心设备通信连接;所述多个第二网络接口用于与第二FPGA处理芯片通信连接;所述方法包括:

接收所述GPU发送的请求信息,所述请求信息包括数据存储的地址;

从所述GPU的所述数据存储地址中读取数据;

根据从所述远程资源调度中心设备接收到的配置信息向所述服务器发送所述数据;或者根据所述配置信息向所述第二FPGA处理芯片发送所述数据;其中,所述配置信息包括第一指示信息和第二指示信息,所述第一指示信息和所述第二指示信息为切换PCIe RC信号,根据所述配置信息向所述服务器发送所述数据,包括:

根据所述第一指示信息,向所述服务器发送所述数据;

根据所述配置信息向所述第二FPGA处理芯片发送所述数据,包括:

根据所述第二指示信息,向所述第二FPGA处理芯片发送所述数据。

2.根据权利要求1所述的方法,其特征在于,所述第一FPGA处理芯片还包括数据处理模块,所述数据处理模块包括GPU直接数据存取单元;所述从所述GPU的所述数据存储地址中读取数据,包括:

通过所述GPU直接数据存取单元从所述GPU的所述数据存储地址中读取数据。

3.根据权利要求1或2所述的方法,其特征在于,所述第一FPGA处理芯片还包括桥梁模块;在所述根据从所述远程资源调度中心设备接收到的配置信息向所述服务器发送所述数据;或者根据所述配置信息向所述第二FPGA处理芯片发送所述数据之前,所述方法还包括:

通过所述桥梁模块接收所述远程资源调度中心设备发送的所述配置信息。

4.根据权利要求1所述的方法,其特征在于,所述数据处理模块还包括运算单元,所述配置信息还包括运算规则的信息和目标网络接口的信息;所述根据所述第二指示信息,向所述第二FPGA处理芯片发送所述数据,包括:

根据所述第二指示信息,通过所述运算单元采用所述运算规则对所述数据进行处理得到处理结果;

通过所述多个第二网络接口中的所述目标网络接口向所述第二FPGA处理芯片发送所述数据。

5.一种基于FPGA的分布式GPU通信的装置,其特征在于,所述装置包括第一FPGA处理芯片,所述第一FPGA处理芯片包括第一接口、第二接口、第一网络接口和多个第二网络接口;所述第一FPGA处理芯片的第一接口与GPU通信连接,所述第一FPGA处理芯片的第二接口与服务器通信连接;所述第一网络接口与远程资源调度中心设备通信连接;所述多个第二网络接口用于与第二FPGA处理芯片通信连接;所述第一FPGA处理芯片用于:

通过所述第一接口接收所述GPU发送的请求信息,所述请求信息包括数据存储的地址;

从所述GPU的所述数据存储地址中读取数据;

根据通过所述第一网络接口从所述远程资源调度中心设备接收到的配置信息向所述服务器发送所述数据;或者根据所述配置信息向所述第二FPGA处理芯片发送所述数据;其中,所述配置信息包括第一指示信息和第二指示信息,所述第一指示信息和所述第二指示

信息为切换PCIe RC信号,根据所述配置信息向所述服务器发送所述数据,包括:

根据所述第一指示信息,向所述服务器发送所述数据;

根据所述配置信息向所述第二FPGA处理芯片发送所述数据,包括:

根据所述第二指示信息,向所述第二FPGA处理芯片发送所述数据。

6.根据权利要求5所述的装置,其特征在于,所述第一FPGA处理芯片还包括数据处理模块,所述数据处理模块包括GPU直接数据存取单元;所述第一FPGA处理芯片具体用于:

通过所述GPU直接数据存取单元从所述GPU的所述数据存储地址中读取数据。

7.根据权利要求5或6所述的装置,其特征在于,所述第一FPGA处理芯片还包括桥梁模块;所述第一FPGA处理芯片具体用于:

通过所述桥梁模块接收所述远程资源调度中心设备发送的所述配置信息。

基于FPGA的分布式GPU通信的方法及装置

技术领域

[0001] 本申请涉及通信技术领域,特别是涉及一种基于FPGA的分布式GPU通信的方法及装置。

背景技术

[0002] 图形处理器(Graphics Processing Unit,GPU)是一种专用的图形处理芯片,无论是早期用于图形图像处理,还是现在广泛用于AI人工智能计算领域,都是一种重要的计算芯片。GPU作为一种高速串行计算机扩展总线标准(Peripheral Component Interconnect express,PCIe)设备插接在数据中心服务器插槽上,通过PCIe接口与主机服务器和其他GPU节点通信。

[0003] 由于GPU与服务器通过PCIe接口连接的紧耦合关系,GPU无法独立于服务器单独运行,跨节点GPU间通信只能通过网卡连接到交换机的方式进行,通信网络拓扑不够灵活,数据转发效率低、通信延时大。

发明内容

[0004] 基于此,有必要针对上述技术问题,提供一种基于FPGA的分布式GPU通信的方法及装置,FPGA处理芯片作为GPU与服务器之间的转接卡,不仅降低了GPU与服务器之间的耦合度,还可以与相邻的FPGA处理芯片形成二维环形网络拓扑,极大地提高了通信网络拓扑的灵活性,减少GPU之间互相通信的时间,提高数据转发效率。

[0005] 第一方面,提供一种基于FPGA的分布式GPU通信的方法,该方法应用于通信装置,通信装置包括第一FPGA处理芯片,第一FPGA处理芯片包括第一接口、第二接口、第一网络接口和多个第二网络接口;第一FPGA处理芯片的第一接口与GPU通信连接,第一FPGA处理芯片的第二接口与服务器通信连接;第一网络接口与远程资源调度中心设备通信连接;多个第二网络接口用于与第二FPGA处理芯片通信连接;方法包括:

[0006] 接收GPU发送的请求信息,请求信息包括数据存储的地址;

[0007] 从GPU的数据存储地址中读取数据;

[0008] 根据从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据。

[0009] 在一种可能的实现方式中,第一FPGA处理芯片还包括数据处理模块,数据处理模块包括GPU直接数据存取单元;从GPU的数据存储地址中读取数据,包括:

[0010] 通过GPU直接数据存取单元从GPU的数据存储地址中读取数据。

[0011] 在一种可能的实现方式中,第一FPGA处理芯片还包括桥梁模块;在根据从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据之前,方法还包括:

[0012] 通过桥梁模块接收远程资源调度中心设备发送的配置信息。

[0013] 在一种可能的实现方式中,配置信息包括第一指示信息;根据配置信息向服务器

发送数据,包括:

[0014] 根据第一指示信息,向服务器发送数据。

[0015] 在一种可能的实现方式中,配置信息包括第二指示信息;根据配置信息向第二FPGA处理芯片发送数据,包括:

[0016] 根据第二指示信息,向第二FPGA处理芯片发送数据。

[0017] 在一种可能的实现方式中,数据处理模块还包括运算单元,配置信息还包括运算规则的信息和目标网络接口的信息;根据第二指示信息,向第二FPGA处理芯片发送数据,包括:

[0018] 根据第二指示信息,通过运算单元采用运算规则对数据进行处理得到处理结果;

[0019] 通过多个第二网络接口中的目标网络接口向第二FPGA处理芯片发送数据。

[0020] 第二方面,提供了一种基于FPGA的分布式GPU通信的装置,该装置包括第一FPGA处理芯片,第一FPGA处理芯片包括第一接口、第二接口、第一网络接口和多个第一网络接口;第一FPGA处理芯片的第一接口与GPU通信连接,第一FPGA处理芯片的第二接口与服务器通信连接;第一网络接口与远程资源调度中心设备通信连接;多个第二网络接口用于与第二FPGA处理芯片通信连接;第一FPGA处理芯片用于:

[0021] 通过第一接口接收GPU发送的请求信息,请求信息包括数据存储的地址;

[0022] 从GPU的数据存储地址中读取数据;

[0023] 根据通过第一网络接口从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据设备通信连接;第二网络接口用于与第二FPGA处理芯片通信连接。

[0024] 在一种可能的实现方式中,第一FPGA处理芯片还包括数据处理模块,数据处理模块包括GPU直接数据存取单元;第一FPGA处理芯片具体用于:

[0025] 通过GPU直接数据存取单元从GPU的数据存储地址中读取数据。

[0026] 在一种可能的实现方式中,第一FPGA处理芯片还包括桥梁模块;第一FPGA处理芯片具体用于:

[0027] 通过桥梁模块接收远程资源调度中心设备发送的配置信息。

[0028] 在一种可能的实现方式中,配置信息包括第一指示信息;第一FPGA处理芯片具体用于:

[0029] 根据第一指示信息,向服务器发送数据。

[0030] 上述基于FPGA的分布式GPU通信的方法及装置,通过接收GPU发送的请求信息,请求信息包括数据存储地址,从GPU的数据存储地址中读取数据,根据从远程资源调度中心设备接收到的配置信息向服务器发送数据,或者根据配置信息向第二FPGA处理芯片发送数据,FPGA处理芯片作为GPU与服务器之间的转接卡,不仅降低了GPU与服务器之间的耦合度,减少服务器端CPU、内存、网络等资源的开销,还可以与相邻的FPGA处理芯片形成二维环形网络拓扑,极大地提高了GPU间的灵活性,减少GPU之间互相通信的时间。

附图说明

[0031] 图1为本申请一个实施例中基于FPGA的分布式GPU方法的应用环境图;

[0032] 图2为一个实施例中第一FPGA处理芯片的结构框图;

[0033] 图3为一个实施例中基于FPGA的分布式GPU方法的流程示意图。

具体实施方式

[0034] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0035] 在现有技术中,GPUDirect系列是一种GPU直接内存访问技术,可以让GPU通过PCIe芯片集总线系统访问其他子设备或主机的内存,从而减少不必要的内存拷贝、降低中央处理器(Central Processing Unit,CPU)使用开销,提高数据传输效率。其中,GPUDirect Shared Memory是GPUDirect系列的一种GPU直接共享内存的技术,能够使GPU与其他PCIe设备共享主机的内存页来实现不同PCIe设备之间数据通信的技术。

[0036] GPUDirect P2P是GPUDirect系列的一种GPU间直接共享显存的技术,是指同一PCIe根复合体(PCIe Root Complex,PCIe RC)域下两个GPU设备直接互访GPU显存,GPU间不需要将数据拷贝到主机内存里中转,相比GPUDirect Shared Memory技术,减少了将数据从GPU显存拷贝到主机内存和从主机内存拷贝到GPU显存的步骤,降低了数据通路延时,提高了数据传输效率。

[0037] GPU直接远程直接内存访问(GPUDirect Remote Direct Memory Access,GPUDirect RDMA)技术是利用RDMA相关技术、物理网卡和传输网络实现GPU间直接交互显存数据,该技术解决了传统网络数据传输过程中处理延时大、CPU占用率高等问题,实现了不同节点间直接互访GPU显存的功能。

[0038] GPUDirect Shared Memory技术和GPUDirect P2P技术都是基于GPU作为主机下的PCIe设备开发实现,与其他设备通信依赖CPU、主机内存和PCIe交换系统等参与,设备与服务器CPU、内存等通过PCIe紧耦合,通信仅限于单节点内的GPU。采用GPUDirect Shared Memory技术进行单节点内的GPU间交换显存数据时,需要经过每个CPU及CPU1内存等模块,CPU额外开销大、数据传输处理延时大。采用GPUDirect P2P技术进行GPU间显存数据交换时,仅限于同一PCIe RC域下的GPU间通过PCIe芯片集进行显存数据直接交互,如果跨两个CPU的PCIe RC域还需要CPU及CPU内存参与数据传输,导致GPU间显存数据交互延时和CPU开销依然很大。

[0039] GPUDirect RDMA技术虽然利用RDMA技术实现了跨节点间的GPU通信问题,但需要在同一PCIe域下的高性能网卡以及本地服务器CPU等参与帮助GPU完成跨节点的数据传输,GPU与服务器仍是通过PCIe连接的紧耦合关系,GPU无法独立于服务器单独运行,跨节点GPU间通信只能通过网卡连接到交换机的方式进行,通信网络拓扑不够灵活,数据包转发效率低、通信延时大。

[0040] 为了解决现有技术问题,本申请实施例提供了一种基于FPGA的分布式GPU通信的方法及装置。下面首先对本申请实施例所提供的基于FPGA的分布式GPU通信的方法进行介绍,该方法应用于图1所示的应用环境,如图1所示,通信装置100包括多个FPGA处理芯片,每个FPGA处理芯片包含多个网络接口,其中第一网络接口230通过交换机与远程资源调度中心设备通信连接,第二网络接口240与周围其他FPGA处理芯片组成2D环形通信拓扑,第二网络接口240的数量可以根据用户需求进行调整,不以设置4个为限。FPGA处理芯片通过第二

网络接口240可以向相邻的FPGA处理芯片发送数据,或者接收相邻的FPGA处理芯片发送的数据。第一网络接口和第二网络接口为100G网络光口,使GPU之间以及GPU与远程资源调度中心设备进行高效通信。

[0041] 将其中任意一个FPGA处理芯片定义为第一FPGA处理芯片,与第一FPGA处理芯片相连接的FPGA处理芯片定义为第二FPGA处理芯片,第一FPGA处理芯片和第二FPGA处理芯片具有相同的结构。

[0042] 如图2所示,第一FPGA处理芯片200包括第一接口210、第二接口220、第一网络接口230和多个第二网络接口240;第一FPGA处理芯片的第一接口210与GPU通信连接,第一FPGA处理芯片200的第二接口220与服务器通信连接;第一网络接口230与远程资源调度中心设备通信连接;多个第二网络接口240用于与第二FPGA处理芯片通信连接。

[0043] 第一FPGA处理芯片200还包括指令配置模块270、路由模块280、第一收发模块290、第二收发模块2100,其中,第一收发模块290为RoCE收发模块,与第二网络接口240和路由模块280通信连接,通过RoCE协议接收来自相邻FPGA处理芯片的数据包,解析该数据包,并对第一FPGA处理芯片的数据进行组包,向相邻FPGA处理芯片发送数据包。

[0044] 第二收发模块2100为RoCEv2收发模块,与第一网络接口230、路由模块280和指令配置模块270通信连接,通过RoCEv2协议接收远程资源调度中心设备通信向第一网络接口230发送的配置信息,并解析配置信息,将解析结果分发至对应的模块,比如说将运算法则的信息发送至指令配置模块270,完成GPU资源的注册、初始化等任务。同时,还可以将与第一FPGA处理芯片连接的GPU数据进行组包,通过第一网络接口230连接到交换机与其他FPGA处理芯片连接的GPU进行通信。

[0045] 路由模块280在第一FPGA处理芯片上电后,通过第二网络接口240与相邻的FPGA处理芯片进行通信,获取相邻的FPGA处理芯片的第二网络接口240的MAC地址信息并保存至内存中,以便通过RoCE协议进行通信。

[0046] 图3示出了本申请一个实施例提供基于FPGA的分布式GPU通信的方法的流程示意图。如图3所示,该方法可以包括以下步骤:

[0047] S310,接收GPU发送的请求信息,请求信息包括数据存储地址。

[0048] 在进行数据通信之前,远程资源调度中心设备对通信装置进行初始化设置,通过服务器向GPU发送待处理的数据,GPU对接收到的数据进行处理,保存处理后的数据,向通信装置发送请求信息,以使通信装置将GPU保存的数据进行传输。其中,请求信息包括数据存储地址,以便通信装置准确地获取GPU需要传送的数据。

[0049] 通信装置通过第一FPGA处理芯片200的第一接口210接收GPU发送的请求信息,其中,第一接口210为PCIe接口,以Root Point模式与GPU的PCIe接口采用Gen5x16标准金手指连接。

[0050] S320,从GPU的数据存储地址中读取数据。

[0051] 第一FPGA处理芯片根据接收到的数据存储地址,采用Direct Memory Access直接数据存取(Direct Memory Access,DMA)的方式从GPU显存中的数据存储地址读取数据,使得数据的传输时延小,提高数据的读取效率。

[0052] S330,根据从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据。

[0053] 配置信息包括切换信息,其决定了数据的传输路径,当切换信息为由GPU向服务器通信时,第一FPGA处理芯片通过第二接口220向服务器发送数据,由于GPU与服务器之间引入了基于FPGA处理芯片的通信装置,降低了GPU和服务器的耦合度,便于GPU独立池化管理。其中,第二接口220为Endpoint模式的PCIe接口,与GPU的PCIe接口采用Gen5x16标准连接,以匹配GPU的通信接口模式。

[0054] 当切换信息为由GPU向其他FPGA处理芯片通信时,第一FPGA处理芯片向第二FPGA处理芯片发送数据,GPU通过FPGA处理芯片的多个网络接口与其他FPGA处理芯片的GPU通信网络拓扑更加灵活。通过多个网络接口与多个FPGA处理芯片进行通信实现多个维度同时计算传输数据,减少数据计算传输经过PCIe总线的次数,降低数据更新所需时间,从而减小数据通信时间开销。

[0055] 在本申请实施例中,通过接收GPU发送的请求信息,请求信息包括数据存储地址,从GPU的数据存储地址中读取数据,根据从远程资源调度中心设备接收到的配置信息向服务器发送数据,或者根据配置信息向第二FPGA处理芯片发送数据,FPGA处理芯片作为GPU与服务器之间的转接卡,不仅降低了GPU与服务器之间的耦合度,减少服务器端CPU、内存、网络等资源的开销,还可以与相邻的FPGA处理芯片形成二维环形网络拓扑,极大地提高了GPU间的灵活性,减少GPU之间互相通信的时间。

[0056] 在一些实施例中,第一FPGA处理芯片还包括数据处理模块250,数据处理模块包括GPU直接数据存取单元251;从GPU的数据存储地址中读取数据,包括:

[0057] 通过GPU直接数据存取单元251从GPU显存中的数据存储地址读取数据,第一FPGA处理芯片200通过PCIe和GPU直接数据存取单元251直接访问GPU内部显存,读取待传输数据,有效地降低了GPU间的通信延时。

[0058] 在一些实施例中,第一FPGA处理芯片200还包括桥梁模块260;在根据从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据之前,方法还包括:

[0059] 通过桥梁模块260接收远程资源调度中心设备发送的配置信息。

[0060] 通过桥梁模块260与指令配置模块270、第一接口210、第二接口220和数据处理模块连接,远程资源调度中心设备通过交换机网络将配置信息发送至第一网络接口230,第一网络接口230对配置信息进行解析得到切换PCIe RC信号,并切换PCIe RC信号发送至指令配置模块270,指令配置模块270对切换PCIe RC信号进行判断,并将判断结果发送至桥梁模块260,桥梁模块260根据判断结果可以切换GPU的PCIe总线的连接关系。

[0061] 在通信装置初始化阶段,远程资源调度中心设备通过交换机网络向第一FPGA处理芯片发送配置信息,控制切换GPU的PCIe总线的连接关系,确定数据传输路径,灵活选择GPU中数据的传输对象,解除GPU对服务器主机的依赖性。

[0062] 在一些实施例中,配置信息包括第一指示信息;根据配置信息向服务器发送数据,包括:

[0063] 根据第一指示信息,向服务器发送数据。

[0064] 指令配置模块270接收配置信息中的切换PCIe RC信号后并对其进行解析,判断切换PCIe RC信号对应的值是1还是0。其中,第一指示信息为0,当配置信息的解析结果为0时,GPU与服务器进行数据交互,桥梁模块260直接将数据发送至服务器。

[0065] 所述配置信息包括第二指示信息;根据所述配置信息向所述第二FPGA处理芯片发送所述数据,包括:

[0066] 根据所述第二指示信息,向所述第二FPGA处理芯片发送所述数据。

[0067] 第二指示信息为1,当配置信息的解析结果为1时,GPU与第二FPGA处理芯片进行数据交互,桥梁模块260先将数据发送至数据处理模块250,数据处理模块250对其进行处理,将处理后的数据发送至第一收发模块290,最后通过第二网络接口240送至第二FPGA处理芯片。

[0068] 在一些实施例中,数据处理模块250还包括运算单元252,配置信息还包括运算规则的信息和目标网络接口的信息;根据第二指示信息,向第二FPGA处理芯片发送数据,包括:

[0069] 根据第二指示信息,通过运算模块采用运算规则对数据进行处理得到处理结果;

[0070] 通过多个第二网络接口中的目标网络接口向第二FPGA处理芯片发送数据。

[0071] 在GPU与第二FPGA处理芯片进行数据交互的过程中,若数据处理模块250的接收单元未收到其他GPU对应的FPGA处理芯片发送的待处理数据,指令配置模块270根据配置信息控制运算单元252不进行任何计算,直接将GPU直接数据存取单元251从GPU读取的数据发送至发送单元253,发送单元253将数据发送至第一收发模块290,第一收发模块290从路由模块280中获取配置信息中目标网络接口的AMC地址信息,通过目标网络接口向其他FPGA处理芯片发送该数据。

[0072] 若数据处理模块的接收单元收到其他GPU对应的FPGA处理芯片发送的待处理数据,指令配置模块270根据运算规则的信息控制运算单元252进行相应的计算,GPU直接数据存取单元251从GPU获取到数据后,将数据发送至运算单元252,运算单元252根据预先配置的运算规则将GPU读取到的数据和通过接收单元254接收的FPGA处理芯片发送的数据进行混合计算,将计算结果发送至发送单元253,发送单元253将数据发送至第一收发模块290,第一收发模块290从路由模块280中获取配置信息中目标网络接口的AMC地址信息,通过目标网络接口向其他FPGA处理芯片发送该数据,FPGA处理芯片通过GPU直接数据存取单元251写入到对应的GPU显存中,完成GPU的数据的更新迭代。

[0073] 在一个实施例中,提供了一种基于FPGA的分布式GPU通信的装置,该装置包括第一FPGA处理芯片,第一FPGA处理芯片包括第一接口、第二接口、第一网络接口和多个第一网络接口;第一FPGA处理芯片的第一接口与GPU通信连接,第一FPGA处理芯片的第二接口与服务器通信连接;第一网络接口与远程资源调度中心设备通信连接;多个第二网络接口用于与第二FPGA处理芯片通信连接;第一FPGA处理芯片用于:

[0074] 通过第一接口接收GPU发送的请求信息,请求信息包括数据存储的地址;

[0075] 从GPU的数据存储地址中读取数据;

[0076] 根据通过第一网络接口从远程资源调度中心设备接收到的配置信息向服务器发送数据;或者根据配置信息向第二FPGA处理芯片发送数据设备通信连接;第二网络接口用于与第二FPGA处理芯片通信连接。

[0077] 在本申请实施例中,第一FPGA处理芯片作为GPU与服务器之间的转接卡,不仅降低了GPU与服务器之间的耦合度,还可以与相邻的FPGA处理芯片形成二维环形网络拓扑,极大地提高了通信网络拓扑的灵活性,减少GPU之间互相通信的时间,提高数据转发效率。

[0078] 在一个实施例中,第一FPGA处理芯片还包括数据处理模块,数据处理模块包括GPU直接数据存取单元;第一FPGA处理芯片具体用于:

[0079] 通过GPU直接数据存取单元从GPU的数据存储地址中读取数据。

[0080] 在一个实施例中,第一FPGA处理芯片还包括桥梁模块;第一FPGA处理芯片具体用于:

[0081] 通过桥梁模块接收远程资源调度中心设备发送的配置信息。

[0082] 在一个实施例中,配置信息包括第一指示信息;第一FPGA处理芯片具体用于:

[0083] 根据第一指示信息,向服务器发送数据。

[0084] 在一些实施例中,配置信息包括第二指示信息;第一FPGA处理芯片具体用于:

[0085] 根据第二指示信息,向第二FPGA处理芯片发送数据。

[0086] 在一些实施例中,数据处理模块还包括运算单元,配置信息还包括运算规则的信息和目标网络接口的信息;第一FPGA处理芯片具体用于:

[0087] 根据第二指示信息,通过运算单元采用运算规则对数据进行处理得到处理结果;

[0088] 通过多个第二网络接口中的目标网络接口向第二FPGA处理芯片发送数据。

[0089] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0090] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

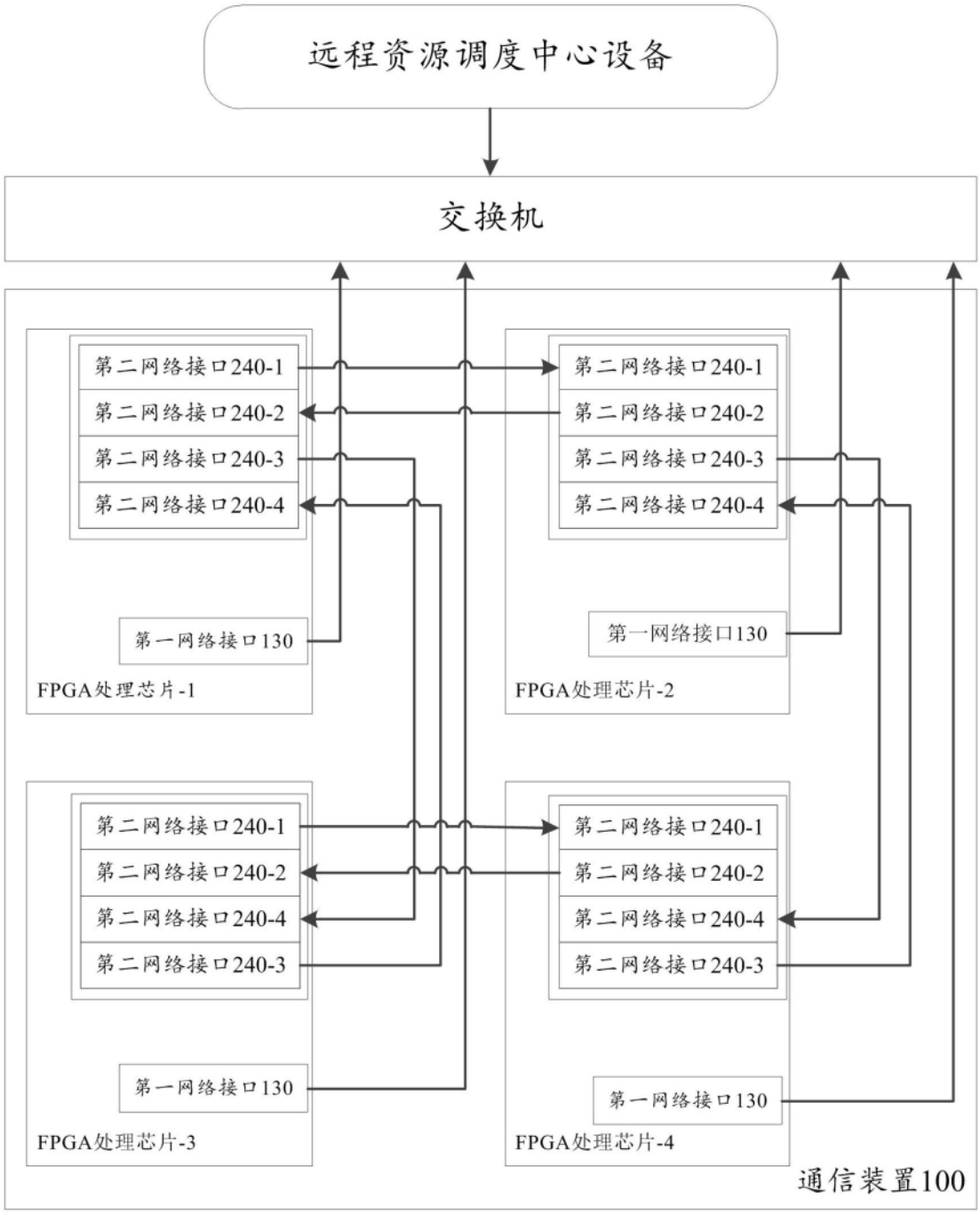
远程资源调度中心设备
交换机
第二网络接口240-1
第二网络接口240-2
第二网络接口240-3
第二网络接口240-4
第一网络接口130
FPGA处理芯片-1
第二网络接口240-1
第二网络接口240-2
第二网络接口240-3
第二网络接口240-4
第一网络接口130
FPGA处理芯片-2
第二网络接口240-1
第二网络接口240-2
第二网络接口240-4
第二网络接口240-3
第一网络接口130
FPGA处理芯片-3
第二网络接口240-1
第二网络接口240-2
第二网络接口240-4
第二网络接口240-3
第一网络接口130
FPGA处理芯片-4
通信装置100

图1

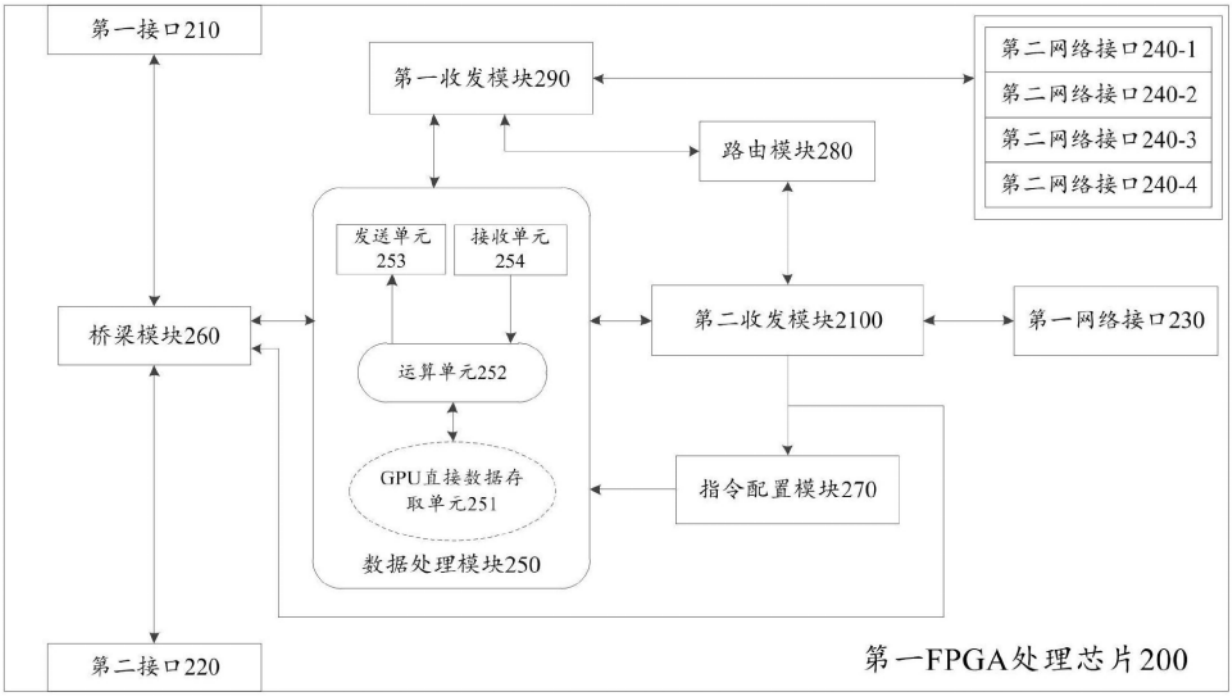
第一接口210
第一收发模块290
第二网络接口240-1
第二网络接口240-2
第二网络接口240-3
第二网络接口240-4
路由模块280
发送单元253
接收单元254
桥梁模块260
第二收发模块2100
第一网络接口230
运算单元252
GPU直接数据存取单元251
指令配置模块270
数据处理模块250
第二接口220
第一FPGA处理芯片200

图2


S310
接收GPU发送的请求信息，请求信息包括数据存储地址
S320
从GPU的数据存储地址中读取数据
S330
根据从远程资源调度中心设备接收到的配置信息向服务器发送数据；
或者根据配置信息向第二FPGA处理芯片发送数据

图3