



(12) 发明专利

(10) 授权公告号 CN 111385212 B

(45) 授权公告日 2021.08.31

(21) 申请号 201811635241.8

(22) 申请日 2018.12.29

(65) 同一申请的已公布的文献号  
申请公布号 CN 111385212 A

(43) 申请公布日 2020.07.07

(73) 专利权人 华为技术有限公司  
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 曾重 陈云 崔晓松 段小祥

(74) 专利代理机构 北京中博世达专利商标代理有限公司 11274

代理人 申健

(51) Int. Cl.

H04L 12/751 (2013.01)

H04L 12/761 (2013.01)

G06N 3/04 (2006.01)

(56) 对比文件

- US 6934293 B1, 2005.08.23
- CN 107809364 A, 2018.03.16
- CN 104734964 A, 2015.06.24
- CN 108495284 A, 2018.09.04
- CN 106878171 A, 2017.06.20
- CN 103067129 A, 2013.04.24
- US 10102449 B1, 2018.10.16
- CN 105519058 A, 2016.04.20
- CN 106453138 A, 2017.02.22
- CN 106790732 A, 2017.05.31
- CN 1612562 A, 2005.05.04
- CN 108809590 A, 2018.11.13
- CN 108880969 A, 2018.11.23
- CN 108540398 A, 2018.09.14

审查员 李华

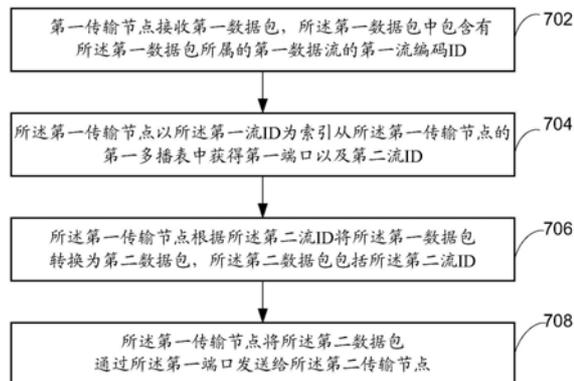
权利要求书2页 说明书11页 附图6页

(54) 发明名称

数据传输技术及神经网络系统

(57) 摘要

本申请提供了一种数据传输技术及神经网络系统。根据该数据传输技术,当第一传输节点接收到第一数据流的第一数据包后,第一传输节点可以以所述第一数据流的第一流编码ID为索引获得第一端口以及所述第一数据流的第二流ID,所述第二流ID为所述第一数据流在第二传输节点的流ID。所述第一传输节点根据所述第二流ID将所述第一数据包转换为包含所述第二流ID的第二数据包,并将所述第二数据包通过所述第一端口发送给所述第二传输节点。本申请提供的数据传输技术能够提升神经网络系统中的数据转发效率。



1. 一种数据传输方法,其特征在于,所述方法由第一传输节点执行,所述方法包括:  
接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID;

以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID;

根据所述第二流ID将所述第一数据包转换为第二数据包,所述第二数据包包括所述第二流ID;

将所述第二数据包通过所述第一端口发送给所述第二传输节点。

2. 根据权利要求1所述的方法,其特征在于:所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

3. 根据权利要求1或2所述的方法,其特征在于,所述以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID包括:

以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。

4. 根据权利要求3所述的方法,其特征在于,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

5. 根据权利要求2所述的方法,其特征在于:所述第一传输节点集合还包括第三传输节点,所述方法还包括:

接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID;

根据所述第二数据流的流ID获得传输所述第三数据包的第二端口;

通过所述第二端口将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。

6. 一种神经网络系统,其特征在于,包括第一传输节点和与所述第一传输节点连接的第二传输节点,所述第一传输节点用于:

接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID;

以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID;

根据所述第二流ID将所述第一数据包转换为第二数据包,所述第二数据包包括所述第二流ID;

将所述第二数据包通过所述第一端口发送给所述第二传输节点。

7. 根据权利要求6所述的神经网络系统,其特征在于:所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

8. 根据权利要求6或7所述的神经网络系统,其特征在于,所述以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID包括:

以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。

9. 根据权利要求8所述的神经网络系统,其特征在于,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

10. 根据权利要求6或7所述的神经网络系统,其特征在于,还包括连接所述第一传输节点的第三传输节点,所述第三传输节点和所述第一传输节点属于同一传输节点集合,所述第一传输节点还用于:

接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID;

根据所述第二数据流的流ID获得传输所述第三数据包的第二端口;

通过所述第二端口将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。

11. 一种传输节点,其特征在于,包括:

接收模块,用于接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID;

获取模块,用于以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID;

转换模块,用于根据所述第二流ID将所述第一数据包转换为第二数据包,所述第二数据包包括所述第二流ID;

发送模块,用于将所述第二数据包通过所述第一端口发送给所述第二传输节点。

12. 根据权利要求11所述的传输节点,其特征在于:第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

13. 根据权利要求11或12所述的传输节点,其特征在于,所述获取模块用于:

以所述第一流ID为索引从第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。

14. 根据权利要求13所述的传输节点,其特征在于,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

15. 根据权利要求12所述的传输节点,其特征在于:所述第一传输节点集合还包括第三传输节点;

所述接收模块还用于接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID;

所述获取模块还用于根据所述第二数据流的流ID获得传输所述第三数据包的第二端口;

所述发送模块还用于通过所述第二端口将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。

## 数据传输技术及神经网络系统

### 技术领域

[0001] 本申请涉及计算机技术领域,尤其涉及一种数据传输技术及神经网络系统。

### 背景技术

[0002] 深度学习(Deep Learning,DL)是人工智能一个重要分支,深度学习是为了模仿人脑构造的一种神经网络,可以达到比传统的浅层学习方式更好的识别效果。近年来,深度神经网络在图像识别,语音识别等领域取得了广泛的成功。由于神经网络处理的数据量很大,因此如何提高神经网络的处理速度成为神经网络的重要方向。

[0003] 基于流处理的神经网络中,一个节点处理完数据后,就将处理结果发送给其他节点进行处理。这种流水线式的处理方式能够使得多个节点并发处理同一条数据流,提高处理效率。在流处理过程中,数据需要在多个节点间传输。数据传输通常包括单播和多播两种方式。单播传输是指将数据包传递给单一目的地址。多播传输是指将数据包同时传输给多个目的地址。具体的,在以多播的形式将数据包传输给多个目的地址时,数据包在链路上只传递一次,只有在链路分叉时,数据包才会被复制分别传输给不同的节点,因此,多播传输的方式,减少量网络中重复的数据包的数量,提升了传输效率。

[0004] 现有技术中,多播传输通常以查询多播表的方式实现。具体实现时,每个传输节点配置一个多播表,所述多播表包括数据流的流编码(ID)以及传输端口。每个传输节点以流ID为索引,通过查询所述多播表获得数据包的传输端口,从而,将数据包从端口传输到下一传输节点。现有技术中通常以数据流的源地址以及所述数据流的目的地址作为数据流的全局ID,为每个传输节点配置多播表。然而这种使用全局ID指示数据流的方式,如果在每个节点都为每条流预留一个表项,则会造成多播表中的流表项的数量很大,造成存储空间的浪费。如果各节点只存储流经该节点的数据流的流表项,则会造成各节点中的多播表中的流编码不连续,查询多播表时,查询速度较慢,影响神经网络的处理速度。

### 发明内容

[0005] 本申请提供了一种数据传输技术及神经网络系统,能够提升神经网络中的数据传输速度。

[0006] 第一方面,本发明实施例提供了一种数据传输方法,根据所述数据传输方法,第一传输节点接收第一数据包后,以所述第一数据包所属的第一数据流的第一流编码ID为索引获得第一端口以及所述第一数据流的第二流ID,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID。第一传输节点根据所述第二流ID将所述第一数据包转换为第二数据包,并通过所述第一端口将所述第二数据包发送给所述第二传输节点。

[0007] 本发明实施例所述的数据传输方法中,在第一传输节点接收到第一数据流的第一数据包后,由于第一传输节点能够以第一数据包携带的第一数据流的第一流编码获得第一数据流在第二传输节点中的第二流编码信息,并根据获得的第二流编码将第一数据包转换为第二数据包,因此,根据这种方式,虽然各传输节点只对流经该传输节点的数据流建立表

项,但各传输节点仍然能够以接收的数据包中的数据流编码为索引快速查到传输数据包的端口信息。从而,能够在减少多播表的表项数量的基础上提升各传输节点传输数据包的速度。

[0008] 在第一种可能的实现方式中,所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

[0009] 结合第一方面或第一方面的第一种可能的实现方式,在第二种可能的实现方式中,所述第一传输节点以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。

[0010] 结合第一方面的第二种可能的实现方式,在第三种可能的实现方式中,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

[0011] 在另一种可能的实现方式中,在不同的传输节点的多播表中,不同数据流的流编码可以相同。例如,第一数据流在第一传输节点的多播表中的流编码与第三数据流在第二传输节点的多播表中的流编码相同,其中,所述第一传输节点和第二传输节点属于不同的传输节点集合,所述第一数据流和第三数据流为不同的数据流。

[0012] 在又一种可能的实现方式中,所述第一传输节点集合还包括第三传输节点。所述方法还包括所述第一传输节点接收第三数据包,以第三数据包中包含的第二数据流的流ID为索引获得传输所述第三数据包的第二端口,并通过所述第二端口将所述第三数据包发送给所述第三传输节点。其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。

[0013] 第二方面,本申请提供了一种神经网络系统,包括第一传输节点和与所述第一传输节点连接的第二传输节点,所述第一传输节点用于执行上述第一方面或第一方面的任意一种可能的实现方式中所述的数据传输方法。

[0014] 在又一种可能的实现方式中,所述神经网络系统还包括第一计算节点和第二计算节点,所述第一计算节点连接所述第一传输节点,所述第二计算节点连接所述第二传输节点,所述第一计算节点,用于向所述第一传输节点发送所述第一数据包;所述第二计算节点,用于通过所述第二传输节点接收所述第二数据包。

[0015] 第三方面,本申请还提供了一种传输节点,包括处理器、接收端口和第一端口。接收端口用于接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID。所述处理器用于以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,并根据所述第二流ID将所述第一数据包转换为第二数据包,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID,所述第二数据包包括所述第二流ID。所述第一端口用于将所述第二数据包发送给所述第二传输节点。

[0016] 在一种可能的实现方式中,所述处理器802以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。具体的,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

[0017] 结合上述第三方面或第三方面的可能的实现方式,在又一种可能的实现方式中,所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所

述第一流ID和所述第二流ID不同。

[0018] 结合上述第三方面或第三方面的任意一种可能的实现方式,在又一种可能的实现方式中,在不同的传输节点的多播表中,不同数据流的流编码可以相同。例如,第一数据流在第一传输节点的多播表中的流编码与第三数据流在第二传输节点的多播表中的流编码相同,其中,所述第一传输节点和第二传输节点属于不同的传输节点集合,所述第一数据流和第三数据流为不同的数据流。

[0019] 结合上述第三方面或第三方面的任意一种可能的实现方式,在又一种可能的实现方式中,所述接收端口还用于接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID。所述处理器还用于根据所述第二数据流的流ID获得传输所述第三数据包的第二端口。所述第二端口用于将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。第三传输节点和所述第一传输节点属于同一个传输节点集合。

[0020] 第四方面,本申请提供了又一种传输节点,包括能够执行上述第一方面以及第一方面任意一种可能的实现方式中所述的数据传输方法的功能模块。

[0021] 第五方面,本申请还提供了一种计算机程序产品,包括程序代码,所述程序代码包括的指令被计算机所执行,以实现所述第一方面以及所述第一方面的任意一种可能的实现方式中所述的数据传输方法。

[0022] 第六方面,本申请还提供了一种计算机可读存储介质,所述计算机可读存储介质用于存储程序代码,所述程序代码包括的指令被计算机所执行,以实现前述第一方面以及所述第一方面的任意一种可能的实现方式中所述的方法。

## 附图说明

[0023] 为了更清楚的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。

[0024] 图1为本发明实施例提供的一种神经网络系统的结构示意图;

[0025] 图2为本发明实施例提供的一种神经网络芯片中的神经元结构示意图;

[0026] 图3为本发明实施例提供的一种神经网络芯片中的神经元集合示意图;

[0027] 图4为现有技术中的一种多播表的示意图;

[0028] 图5为本发明实施例提供的一种多播流及流编码的示意图;

[0029] 图6为本发明实施例提供的一种多播表的示意图;

[0030] 图7为本发明实施例提供的一种数据传输方法的流程图;

[0031] 图8为本发明实施例提供的一种传输节点的结构示意图;

[0032] 图9为本发明实施例提供的又一种传输节点的结构示意图。

## 具体实施方式

[0033] 为了使本技术领域的人员更好地理解本发明方案,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚描述。显然,所描述的实施例仅仅是本发明一部分的实施例,而不是全部的实施例。

[0034] 神经网络(Artificial Neural Network,ANN),简称为神经网络(Neural Network,NN)或类神经网络,在机器学习和认知科学领域,是一种模仿生物神经网络(动物的中枢神经系统,特别是大脑)的结构和功能的数学模型或计算模型,用于对函数进行估计或近似。人工神经网络可以包括卷积神经网络(Convolutional Neural Network,CNN)、深度神经网络(Deep Neural Networks,DNN)、多层感知器(Multilayer Perceptron,MLP)等神经网络。图1为本发明实施例提供的一种人工神经网络系统的结构示意图。由于卷积神经网络CNN是一种最常见的深度学习架构,也是研究最广泛的深度学习方法,因此,图1以卷积神经网络为例进行图示。如图1所示,卷积神经网络系统100可以包括主机105以及卷积神经网络电路110。卷积神经网络电路110通过主机接口与主机105连接。主机接口可以包括标准的主机接口以及网络接口(network interface)。例如,主机接口可以包括快捷外设互联标准(Peripheral Component Interconnect Express,PCIE)接口。如图1所示,卷积神经网络电路110可以通过PCIE总线106与主机105连接。因此,数据可以通过PCIE总线106输入至卷积神经网络电路110中,并通过PCIE总线106接收卷积神经网络电路110处理完成后的数据。并且,主机105也可以通过主机接口监测卷积神经网络电路110的工作状态。

[0035] 主机105可以包括处理器1052以及内存1054。需要说明的是,除了图1所示的器件外,主机105还可以包括通信接口以及作为外存的磁盘等其他器件,在此不做限制。

[0036] 处理器(Processor)1052是主机105的运算核心和控制核心(Control Unit)。处理器1052中可以包括多个处理器核(core)。处理器1052可以是一块超大规模的集成电路。在处理器1052中安装有操作系统和其他软件程序,从而处理器1052能够实现对内存1054、缓存、磁盘及外设设备(如图1中的神经网络电路)的访问。可以理解的是,在本发明实施例中,处理器1052中的Core例如可以是中央处理器(Central Processing unit,CPU),还可以是其他特定集成电路(Application Specific Integrated Circuit,ASIC)。

[0037] 内存1054是主机105的主存。内存1054通过双倍速率(double data rate,DDR)总线和处理器1052相连。内存1054通常用来存放操作系统中各种正在运行的软件、输入和输出数据以及与外存交换的信息等。为了提高处理器1052的访问速度,内存1054需要具备访问速度快的优点。在传统的计算机系统架构中,通常采用动态随机存取存储器(Dynamic Random Access Memory,DRAM)作为内存1054。处理器1052能够通过内存控制器(图1中未示出)高速访问内存1054,对内存1054中的任意一个存储单元进行读操作和写操作。

[0038] 卷积神经网络电路110是由多个神经网络芯片(chip)组成的芯片阵列。例如,如图1所示,卷积神经网络电路110包括多个进行数据处理的神经网络芯片(chip)115和多个路由器120。为了描述方便,本发明实施例将申请中的神经网络芯片115简称为芯片115。所述多个芯片115通过路由器120相互连接。例如,一个芯片115可以与一个或多个路由器120连接。多个路由器120可以组成一种或多种网络拓扑。芯片115之间可以通过所述多种网络拓扑进行数据传输。例如,所述多个路由器120可以组成第一网络1106以及第二网络1108,其中,第一网络1106为环形网络,第二网络1108为二维网状(2D mesh)网络。从而,从输入端口1102输入的数据能够所述多个路由器120组成的网络发送给相应的chip115,任意一个芯片115处理后的数据也可以通过所述多个路由器120组成的网络发送给其他芯片115或从输出端口1104发送出去。

[0039] 进一步的,图1也示出了芯片115结构示意图。如图1所示,chip 115可以包括多个

神经元125以及多个路由器122。图1以神经元为瓦片(tile)为例进行描述。在图1所示的数据处理芯片115的架构中,一个tile 125可以与一个或多个路由器122相连。芯片115中的所述多个路由器122可以组成一种或多种网络拓扑。Tile 125之间可以通过所述多种网络拓扑进行数据传输。例如,所述多个路由器122可以组成第一网络1156以及第二网络1158,其中,第一网络1156为环形网络,第二网络1158为二维网状(2D mesh)网络。从而,从输入端口1152输入芯片115的数据能够根据所述多个路由器122组成的网络发送给相应的tile 125,任意一个tile 125处理后的数据也可以通过所述多个路由器122组成的网络发送给其他tile 125或从输出端口1154发送出去。

[0040] 需要说明的是,卷积神经网络电路110中的多个路由器120组成的一种或多种网络拓扑与数据处理芯片115中的多个路由器122组成的网络拓扑可以相同也可以不相同,只要芯片115之间或tile 125之间能够通过网络拓扑进行数据传输,且芯片115或tile 125能够通过所述网络拓扑接收数据或输出数据即可。在本发明实施例中并不对多个路由器120和122组成的网络的数量和种类进行限制。并且,本发明实施例中,路由器120和路由器122可以相同也可以不同。为了描述清楚,在图1中将连接芯片的路由器120和连接tile的路由器122在标识上进行了区分。需要说明的是,在本发明实施例中,神经网络电路110中的多个芯片115的网络结构与芯片115内部tile 125的网络结构类似。芯片115传输数据的方式与tile 125传输数据的方式也类似。为了描述方便,下述实施例主要以数据芯片115的结构为例进行详细描述,卷积神经网络电路110的结构可以参考数据芯片115的结构。为了描述方便,在本发明实施例中,也可以将卷积神经网络系统中的芯片115或tile 125称为计算节点(computing node)。

[0041] 图2为本发明实施例提供的一种神经网络芯片中的神经元结构示意图。如图2所示,芯片115中包括多个路由器122,每个路由器可以连接一个神经元,图2以神经元为tile 125为例进行示意。需要说明的是,实际应用中,一个路由器122还可以连接多个tile 125。如图2所示,每个tile 125可以包括输入输出接口(TxRx) 1252、交换装置(TSW) 1254以及多个处理器件(processing element, PE) 1256。所述TxRx 1252用于接收从router122输入tile125的数据,或者输出tile125的计算结果。换一种表达方式,TxRx 1252用于tile 125和router122之间的数据传输。交换机(TSW) 1254连接TxRx 1252,所述TSW 1254用于所述TxRx 1252以及多个PE 1256之间的数据传输。每个PE 1256中可以包括多个引擎(engine) 1258,所述多个engine 1258用于对输入PE 1256中的数据进行神经网络计算。例如,可以对输入tile 125的数据与tile125中预设的卷积核进行乘加运算。Engine 1258的计算结果可以通过TSW 1254以及TxRx 1252发送给其他tile 125。

[0042] 为了描述清楚,下面简要介绍一下卷积神经网络的流处理。本领域技术人员可以知道,卷积神经网络可以包括多个卷积层。在图1所示的卷积神经网络系统中,为了实现对数据的流处理,可以将所述多个神经元125分为多个节点集合。例如,如图3所示,可以将每个芯片115中的多个tile 125划分为多个节点集合:第一节点集合202、第二节点集合204、第三节点集合206、第四节点集合208以及第五节点集合210。其中每个节点集合中包括至少一个计算节点(例如神经元125)。同一节点集合的神经元用于处理同一卷积层的数据,不同卷积层的数据分别由不同节点集合的神经元进行处理。一个计算节点(或神经元)处理后的处理结果将传输给其他节点集合中的计算节点进行处理,这种流水线式的处理方式可以不

用缓存中间处理结果,并能够使得多个计算节点并发处理同一条数据流,提高处理效率。此外,本领域技术人员可以知道,tile125在执行神经网络计算(例如,卷积计算)时,可以将输入tile 125的数据与对应卷积层的权值(又可被称为卷积核(kernel))进行卷积计算。在本发明实施例中,同一节点集合的tile 125可以共享权值,不同节点集合中的tile125的权值可以不相同。在本发明实施例中,各tile125中的权值可以通过事先配置完成。

[0043] 如图1和图2所示,要实现上述对数据的流水线处理,例如,一个chip 115中的第一tile的处理结果要传输到另一个芯片115中的第二tile进行处理,则数据经过的传输节点可以包括四层。第一层传输节点为连接chip 115的router 120,第二层传输节点为连接tile 125的router 122,第三层传输节点为连接tile和router 122的输入输出接口TxRx 1252,第四层传输节点为处理数据的PE。具体的,每个router120和router 122均可以分别包括东(east)、西(west)、南(south)、北(north)和本地(local)5个端口。TxRx 1252可以用于将数据包从TxRx 1252发送给一个或多个PE 1258,并将PE 1258处理后的数据发送出去。以一个tile125包括6个PE为例,每个TxRx 1252至少包括6个传输端口,用于分别将数据传输给6个PE。由于每个tile可以接收数据,也可以输出处理结果,因此,tile内的PE可以作为接收端PE也可以作为发送端PE。以一个tile内包括4个引擎(engine)为例,对于接收端PE,数据包会被发送到PE内的一个或多个engine。因此,每个接收端PE至少包括4个端口,分别用于将数据传输给不同的engine。对于发送端PE,发送端PE处理后的数据包会被发送到同一tile内其他的PE或者从TxRx 1252发送出去,因此发送端PE需要包括7个端口,用于分别将数据发送给tile内的其他PE和TxRx 1252。需要说明的是,本发明实施例描述的传输节点是指用于传输数据包的节点,包括图1和2中所示的router120、router122、TxRx1252以及PE 1256。在本发明实施例中,一层传输节点也可以被称为一个传输节点集合。

[0044] 如上所述,当数据包以多播的方式在神经网络的节点间传输时,需要经过多层传输节点。多播流传输时,通常是通过查询多播表获得各个流的数据包的传输端口。现有技术中的多播表的形式可以如图4所示。多播表400可以包括流编码(又可被称为流编码ID)402和端口404,其中,流编码402为数据包所属的数据流的编码,流编码402可以作为查询多播表的地址索引。端口404用于指示数据包的传输端口。例如,图4中,对于流ID为a的数据流,可以通过本传输节点的east口和local口传输出去。由于不同的端口对应不同的接收节点,因此从不同端口传输出去的数据包就能到达不同的接收节点。在本发明实施例中,流经每个传输节点的数据流的流ID和转发数据包的端口可以预先配置。

[0045] 经研究发现,在实现流处理的神经网络中,多播流的数量比较多,但大多数多播流的传输路径较短,为局部多播。因此,在神经网络中,如果每个传输节点都为每条流预留一个表项,则会造成多播表中的流表项的数量很大,造成存储空间的浪费。如果各传输节点只存储流经该传输节点的数据流的流表项,则会造成各传输节点中的多播表中的流编码不连续,查询多播表时,查询速度较慢,影响神经网络的处理速度。为了在减少各个传输节点中的流表项的数量的基础上,提升神经网络的处理速度,本发明实施例提供了一种数据传输方法,用于在图1-3所示的神经网络中传输数据包。为了实现数据包在神经网络中的多播传输,本发明实施例提供了一种多播表。下面先介绍一下本发明实施例提供的一种多播表及其配置方法。

[0046] 在本发明实施例中,各个传输节点的多播表可以预先配置。为了减少多播表中的

流表项,每个传输节点的多播表中只记录流经该节点的数据流的信息。然而,为了提高多播表的查询速度,一个节点的多播表中流经该节点的数据流的流编码(ID)会保持连续,使得与流ID对应的存储空间能够组成连续的一段存储空间。根据这种方式,在查询多播表时,能够以流ID为地址索引,快速的从对应的地址空间中获得该流ID的输出端口信息。为了实现这个目的,在本发明实施例中,在配置各传输节点的多播表时,对流入该传输节点的数据流分配编码,其中,不同的数据流分配不同的编码,且为不同的数据流分配的多个编码保持连续。并且,在分配编码的过程中,流经同一层的不同传输节点的同一条数据流采用相同的流ID,不同的数据流在同一层传输节点中的流ID不同。例如,如图5所示,流入Router\_1的流包括流2、3、4、5、8,在配置Router\_1的多播表时,如510所示,可以依次为流2、3、4、5、8分配编码a、b、c、d、e。流入Router\_2的流包括数据流2、6和7,则在配置Router\_2的多播表时,如512所示,可以依次为流经第二router的流2、6、7分配编码a、b和c。由于Router\_1和Router\_2均是同一层router,例如均是连接tile的router,则同一条流在Router\_1和Router\_1的多播表中的编码相同,例如流2在Router\_1和Router\_1的多播表中的编码均为a。

[0047] 在本发明实施例中,按照局部编码的原则,每个传输节点都只对流经该传输节点的数据流进行编码。如图5所示,通过Router\_1流入tile中的TxRx\_1的流包括流3、4和5,则按照局部编码的原则,在对传输节点TxRx\_1配置多播表时,如图5中的514所示,可以分别为流3、4和5分配编码a、b和c。类似的,按照局部编码原则,对流入PE\_1的流的编码可以如516所示,为流入PE\_1的流3和流1分别分配编码a和b。对流入PE\_2的流的编码可以如518所示,为流入PE\_2的流5和流4分别分配编码a和b。并且,为了减少编码的数量,在编码过程中,不同数据流在不同传输节点中的编码可以相同。例如,流3在PE\_1的编码为a,流5在PE\_2的编码也可以为a。

[0048] 由于对于一个PE来说,可以作为接收端,也可以作为发送端,因此,在本发明实施例中,对每个PE可以分别独立配置两个多播表,一个多播表为PE作为接收端时的多播表(也可以被称为发送端多播表),另一个多播表为PE作为发送端时的多播表(也可以被称为接收端多播表)。例如,如图5所示,由于PE\_2既作为接收端PE又作为发送端PE,因此,对PE\_2配置可以配置接收端多播表以及发送端多播表。接收端多播表的流编码可以参见518,发送端多播表的流编码可以参见520。如520所示,在配置PE\_2的发送端多播表时,对PE\_2发送的数据流8和1分别配置流编码a和b。在本发明实施例中,在配置PE的多播表时,可以使同一条数据流在发送端PE的多播表和接收端PE的多播表中的编码相同。例如,如图5所示,流1从PE\_2流入PE\_1,因此,在发送端PE\_2的多播表以及作为接收端的PE\_2的多播表中,流1的编码均为b。

[0049] 在本发明实施例中,在获得各传输节点编码过程中,可以采用图着色的方式并结合上述编码规则获得流经各传输节点的数据流的编码。在获得每一个传输节点的流编码后,由于各传输节点的多播表中的数据流均是独立编码,虽然对一个传输节点来说,流入该传输节点的不同流的编码能够保持连续。但是,不同的流在不同的节点中可能采用相同的编码,例如,在Router\_1中,流2的编码为a,在TxRx\_1中,流3的编码为a。并且,同一条流在不同节点中的编码也可能不同,例如,流3在Router\_1中的编码为b,而流3在TxRx\_1中的编码为a。因此,在本发明实施例中,为了提高多播表的查询速度,在设置各传输节点的多播表时,可以在本节点的多播表中记录同一条流在该节点的下一个节点中的编码,从而获得本

节点的多播表。例如,如图5所示,流3在Router\_1中的编码为b,而流3在TxRx\_1中的编码为a,则在Router\_1的多播表,可以在流3的表项中,记录流3在TxRx\_1中的编码为a(如图5中流3对应的表项所示)。根据这种方式,可以得到本发明实施例提供的一种多播表(如图6所示)。其中,第一流编码602用于指示流在第一传输节点的多播表中的流编码,第二流编码604用于指示同一条数据流在第二传输节点的多播表中的流编码,端口606用于指示第一传输节点发送数据包的端口信息。其中,第二传输节点为第一传输节点的下一个传输节点。从上述描述可以知道,在本发明实施例提供的多播表中,不仅包括一个数据流在本节点的流编码,还包括该数据流在下一个传输节点的流编码信息。

[0050] 下面将结合图6所示的本发明实施例提供的多播表,对本发明实施例提供的数据传输方法进行详细的描述。图7为本发明实施例提供的数据传输方法流程图。图7所示的方法可以由前述图1-2中所示的任意一个传输节点执行。如图7所示,该方法可以包括下述步骤。

[0051] 在步骤702中,第一传输节点接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID。具体的,第一传输节点可以为上述神经网络中所示的任意一个传输节点。例如,第一传输节点可以为连接chip115的router 120、连接tile 125的router 122、TxRx 1252或PE 1258。所述第一数据包为待传输的数据包。为了描述方便,本发明实施例以多播流中的一个数据包为例进行描述。

[0052] 在步骤704中,所述第一传输节点以所述第一流ID为索引从所述第一传输节点的第一多播表中获得第一端口以及第二流ID。其中,所述第二流ID为所述第一数据流在第二传输节点的第二多播表中的流ID。在本发明实施例中,第一传输节点在接收到所述第一数据包后,会通过解析所述第一数据包,获得所述第一数据包所属的数据流的流ID。实际应用中,每个数据包的包头部分会包括所述数据包所属的数据流的流ID。本发明实施例的流ID可以是按照图5所示方法预先对数据流分配获得。流ID可以以字母表示,也可以以数字标识,在此不对流ID的表示方式进行限制。

[0053] 如前所述,在本发明实施例中,每个传输节点配置有多播表,用于指示该传输节点传输不同数据流的数据包的端口信息。换一种表达方式,多播表用于指示不同数据流的传输路径。以第一传输节点的多播表如图6所示为例。所述第一传输节点的多播表用于记录流入所述第一传输节点的数据流的转发信息。所述第一传输节点的多播表可以包括第一流编码602、第二流编码604以及端口606等信息。其中,所述第一流编码602用于指示数据包所属的数据流(例如第一数据流)在第一传输节点的编码。第二流编码604用于指示所述第一数据流在第一传输节点的下一个传输节点(例如第二传输节点)的多播表中的流编码。端口606用于指示第一传输节点输出所述第一数据流的数据包的端口信息。如图6所示,以第一数据流的流编码ID为“b”为例,第一流编码602用于指示第一数据流的编码为“b”,第二流编码604用于指示第一数据流在第二传输节点的流编码为“a”,端口606指示第一传输节点输出所述第一数据流的数据包的端口为East端口和West端口。

[0054] 在实际应用中,多播表中各表项的信息可以分别被存储在该传输节点的多个地址空间中,每一个地址空间与流ID一一对应,所述多个地址空间构成一段连续的地址空间。当传输节点在查找多播表时,可以以流ID为地址索引,从对应的地址空间中获取与所述流ID对应的表项信息。例如,第一传输节点在查询多播表时,以第一数据流的编码为“b”为例,第

一传输节点以第一流ID“b”为地址索引,能够从对应的地址空间中查找到第二流编码“a”以及端口“East,West”。

[0055] 在步骤706中,所述第一传输节点根据所述第二流ID将所述第一数据包转换为第二数据包,所述第二数据包包括所述第二流ID。具体的,在步骤704中,在第一传输节点获得第一数据流的第二流ID后,所述第一传输节点使用第二流ID替换所述第一数据包中的第一流ID,从而将所述第一数据包转换为第二数据包。例如,若第一数据包的流ID为“b”,第一传输节点的多播表如图6所示,则第一传输节点可以将第一数据包中的第一流ID“b”替换为第二流ID“a”,从而得到所述第二数据包。实际应用中,第二流ID可以被携带在第二数据包的包头部分。

[0056] 在步骤708中,所述第一传输节点将所述第二数据包通过所述第一端口发送给所述第二传输节点。具体的,在第一传输节点将所述第一数据包转换为第二数据包后,所述第一传输节点可以根据从所述第一传输节点的多播表中获得的转发所述第一数据包的端口信息将所述第二数据包转发出去。例如,在将第一数据包中的第一流ID“b”替换为第二流ID“a”而获得第二数据包后,所述第一传输节点可以将所述第二数据包分别从端口East和West发送出去。

[0057] 根据这种方式,第二传输节点(例如,如图5中的TxRx\_1 506)在接收到所述第二数据包后,能够通过解析所述第二数据包获得所述第二数据包所属的数据流的第一流ID为“a”,从而能够以流ID“a”为地址索引,从第二传输节点的多播表中快速查找到转发所述数据包对应的端口信息。可以理解的是,第二传输节点接收到所述第二数据包后的处理方式,可以参考上述如图7描述的第一传输节点处理第一数据包的处理方式。在此不再赘述。

[0058] 可以理解的是,上述实施例只是以一个传输节点将数据传输给其他层的传输节点为例进行描述。根据上述实施例提供的数据传输方法,对于一个传输节点来说,由于流入该传输节点的数据流的编码采用的是独立的局部编码的方式,只对流经该传输节点的数据流建立表项,且不同的流在不同的传输节点中可以采用相同的编码,从而能减少多播表中的表项的数量。并且,由于该传输节点中各数据流的流编码是连续的,从而,在查找数据包的转发端口时,能够根据接收的数据包中携带的流ID为地址索引,从存储多播表的一段连续地址空间中快速查找到转发所述数据包的端口信息。

[0059] 进一步的,由于第一传输节点能够从本传输节点的多播表中获得该数据包所属的数据流在下一个传输节点(例如,第二传输节点)中的流编码信息,并根据获得的流编码信息对接收的数据包进行转换,将转换后的第二数据包从对应的端口信息中发送出去。因此,第二传输节点在接收到第二数据包后,通过解析所述第二数据包,能够获得该数据包所属的数据流在第二传输节点中的流ID信息。由于第二传输节点中的流ID也是连续编码的,因此,第二传输节点也能够以第二流ID为索引,从第二数据包的多播表中快速获得第二传输节点转发所述第二数据包的端口信息。根据这种方式,在神经网络中,可以使每个传输节点的多播表的表项数较少,且各传输节点能够快速查找多播流的转发端口信息,提升各个传输节点转发数据包的速度,提升神经网络对海量数据的处理性能。

[0060] 需要说明的是,在另一种情形下,由于同一节点集合(如图3所示的第一节点集合202)的不同计算节点间也可能进行数据包的传输,因此,如图5所示,数据包还可以在同一层的传输节点之间进行传输。由于本发明实施例提供的多播表中,同一个流在发送端传输

节点和接收端传输节点的流编码相同,因此,在这种情况下,在发送端节点的多播表中并不能查询出第二流ID,因此,发送端节点在转发数据包前并不需要对其接收的数据包进行转换。以第一传输节点接收的第三数据包为如图5中的流2的数据包为例,第三数据包从第一传输节点(例如,Router\_1)传输到第二传输节点(例如,Router\_2),由于流2在第一传输节点的多播表中的流编码为a,在第二传输节点的多播表中的流编码也为a,因此,在如图6所示的第一传输节点的多播表中,第一流编码为“a”的流并没有对应的第二流编码。换一种表达方式,第一传输节点并不需要对编码为a的数据流进行流编码转换。第一传输节点在查询多播表后,只会获得第三数据包的转发端口信息为“East,Local”,从而,第一传输节点可以直接将接收的第三数据包从East端口和Local端口传输出去。

[0061] 图8本发明实施例提供的一种传输节点的结构示意图。图8所示的传输节点800可以是神经网络中的任意一个传输节点,例如,可以是如上述图1-图3中所示的任意一个传输节点。如图8所示,传输节点800可以包括处理器802以及多个端口,所述多个端口可以包括接收端口804、第一端口806以及第二端口808。处理器802可以是一块集成电路。其中,所述接收端口804用于接收第一数据包,所述第一数据包中包含有所述第一数据包所属的第一数据流的第一流编码ID。所述处理器802用于以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,并根据所述第二流ID将所述第一数据包转换为第二数据包,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID,所述第二数据包包括所述第二流ID。所述第一端口用于将所述第二数据包发送给所述第二传输节点。

[0062] 具体的,所述处理器802以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。实际应用中,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

[0063] 在一种可能的实现方式中,在不同的传输节点的多播表中,不同数据流的流编码可以相同。例如,第一数据流在第一传输节点的多播表中的流编码与第三数据流在第二传输节点的多播表中的流编码相同,其中,所述第一传输节点和第二传输节点属于不同的传输节点集合,所述第一数据流和第三数据流为不同的数据流。

[0064] 在又一种情形下,所述接收端口804还用于接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID。所述处理器802还用于根据所述第二数据流的流ID获得传输所述第三数据包的第二端口。所述第二端口808用于将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。第三传输节点和所述第一传输节点属于同一个传输节点集合。

[0065] 图9为本发明实施例提供的又一种传输节点的结构示意图,图9所示的传输节点可以为前述神经网络中的任意一个传输节点,例如,可以是如上述图1-图3中所示的任意一个传输节点。如图9所示,传输节点900可以包括接收模块902、获取模块904、转换模块906以及发送模块908。

[0066] 所述接收模块902用于接收第一数据包,所述第一数据包中包含有所述第一数据

包所属的第一数据流的第一流编码ID。所述获取模块904用于以所述第一流ID为索引获得第一端口以及所述第一数据流的第二流ID,其中,所述第二流ID为所述第一数据流在第二传输节点的流ID。所述转换模块906用于根据所述第二流ID将所述第一数据包转换为第二数据包,所述第二数据包包括所述第二流ID。所述发送模块908用于将所述第二数据包通过所述第一端口发送给所述第二传输节点。实际应用中,所述第一传输节点属于第一传输节点集合,所述第二传输节点属于第二传输节点集合,所述第一流ID和所述第二流ID不同。

[0067] 在一种情形下,所述获取模块904具体用于以所述第一流ID为索引从所述第一传输节点的第一多播表中获得所述第一端口以及所述第一数据流的第二流ID,所述第一多播表用于指示所述第一传输节点传输所述第一数据流的数据包的端口信息。其中,在所述第一多播表中,不同数据流的编码不同,且不同数据流的编码为连续编码。

[0068] 在又一种情形下,所述接收模块902还用于接收第三数据包,所述第三数据包包括所述第三数据包所属第二数据流的流ID。所述获取模块904用于根据所述第二数据流的流ID获得传输所述第三数据包的第二端口。所述发送模块908还用于通过所述第二端口将所述第三数据包发送给所述第三传输节点,其中,所述第二数据流在所述第一传输节点的流ID与所述第二数据流在所述第三传输节点的流ID相同。

[0069] 需要说明的是,图8和图9所示的传输节点均可以用于人工神经网络中,并可以实现图7所示的数据传输方法。相关描述可以参见前述图1-7实施例的详细描述。

[0070] 本发明实施例还提供一种数据传输方法的计算机程序产品,包括存储了程序代码的计算机可读存储介质,所述程序代码包括的指令用于执行前述任意一个方法实施例所述的方法流程。本领域普通技术人员可以理解,前述的存储介质包括:U盘、移动硬盘、磁碟、光盘、随机存储器(Random-Access Memory,RAM)、固态硬盘(Solid State Disk,SSD)或者非易失性存储器(non-volatile memory)等各种可以存储程序代码的非短暂性的(non-transitory)机器可读介质。

[0071] 需要说明的是,本申请所提供的实施例仅仅是示意性的。所属领域的技术人员可以清楚的了解到,为了描述的方便和简洁,在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述。在本发明实施例、权利要求以及附图中揭示的特征可以独立存在也可以组合存在。在本发明实施例中以硬件形式描述的特征可以通过软件来执行,反之亦然。在此不做限定。

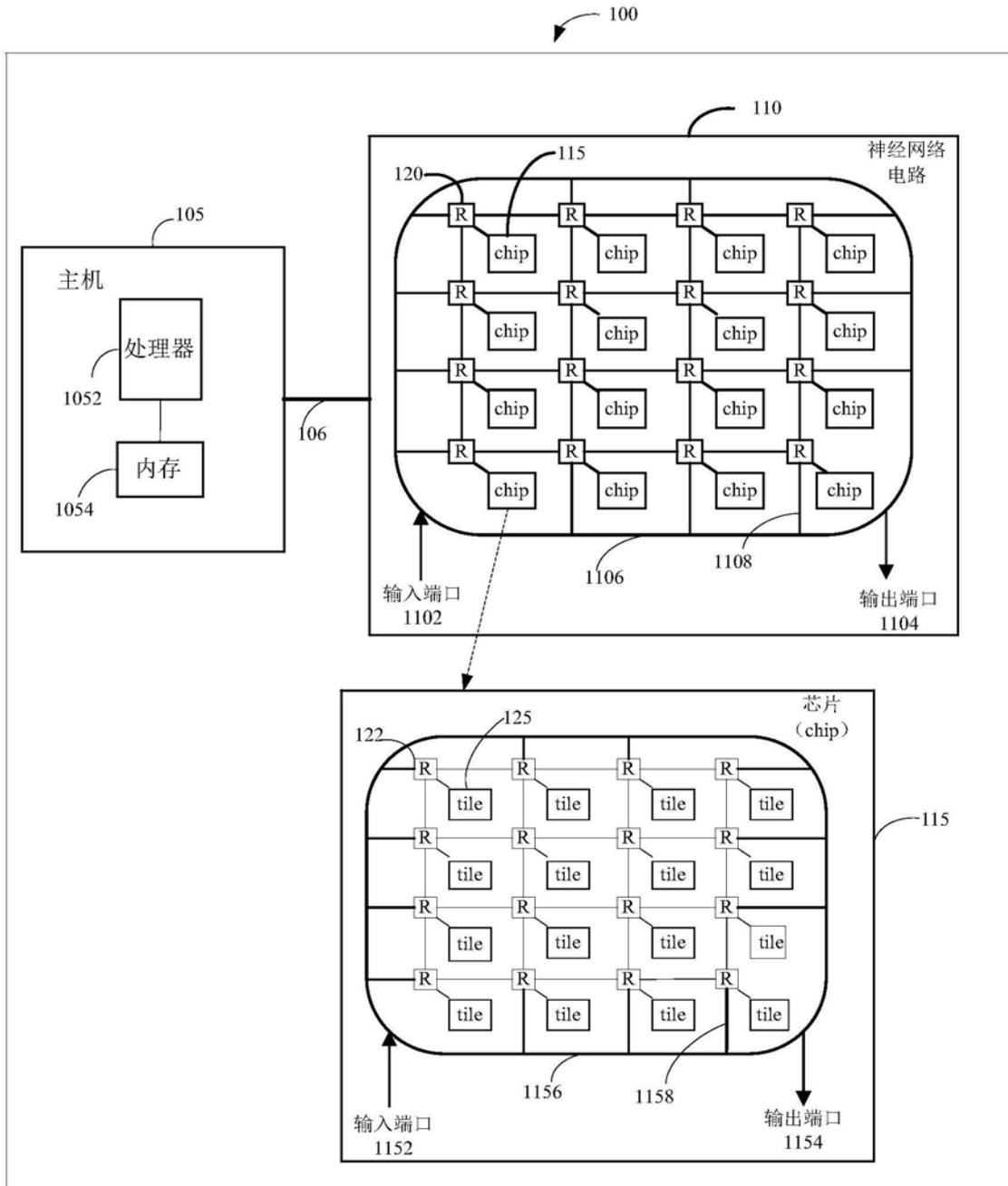


图1

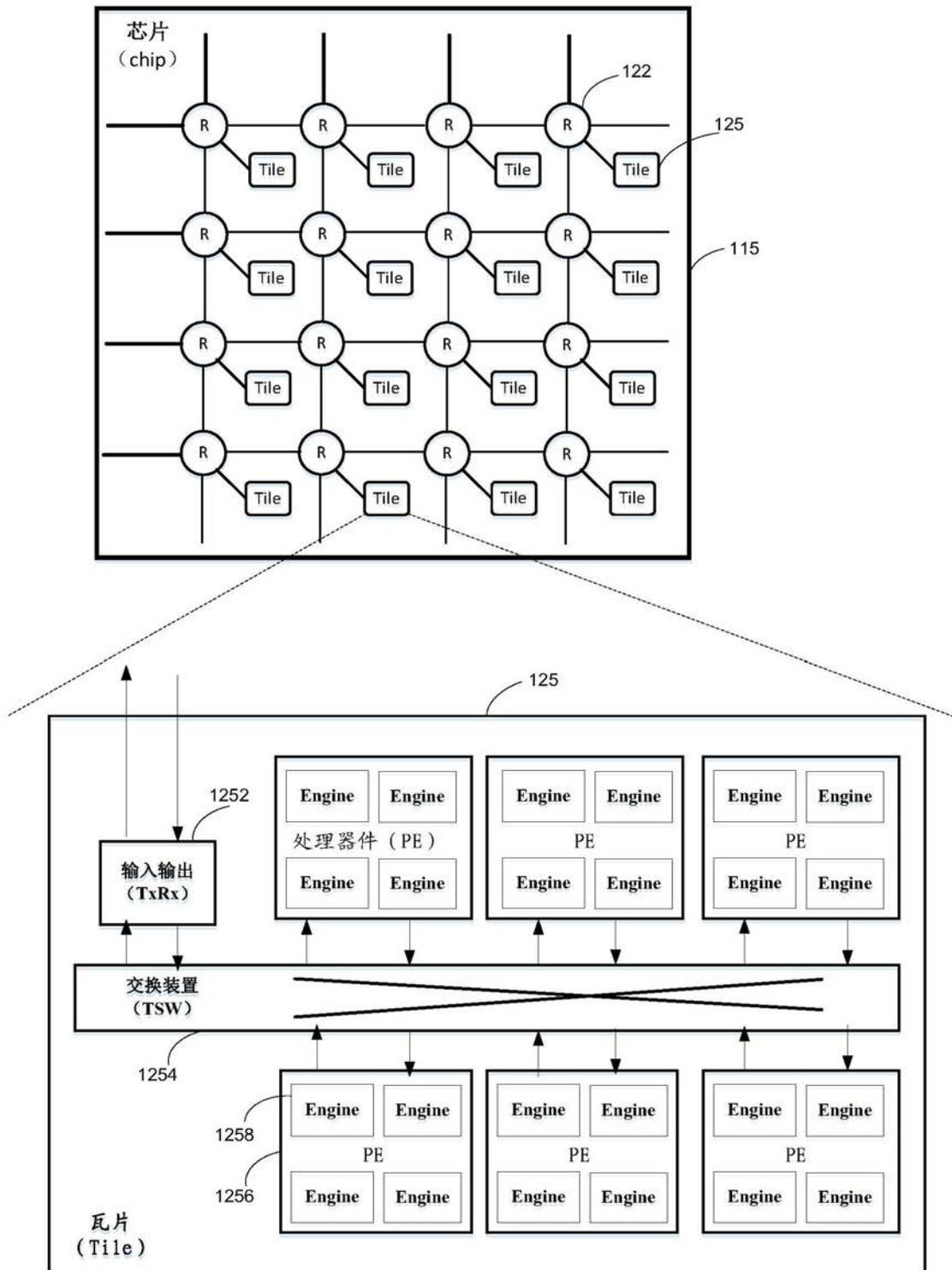


图2

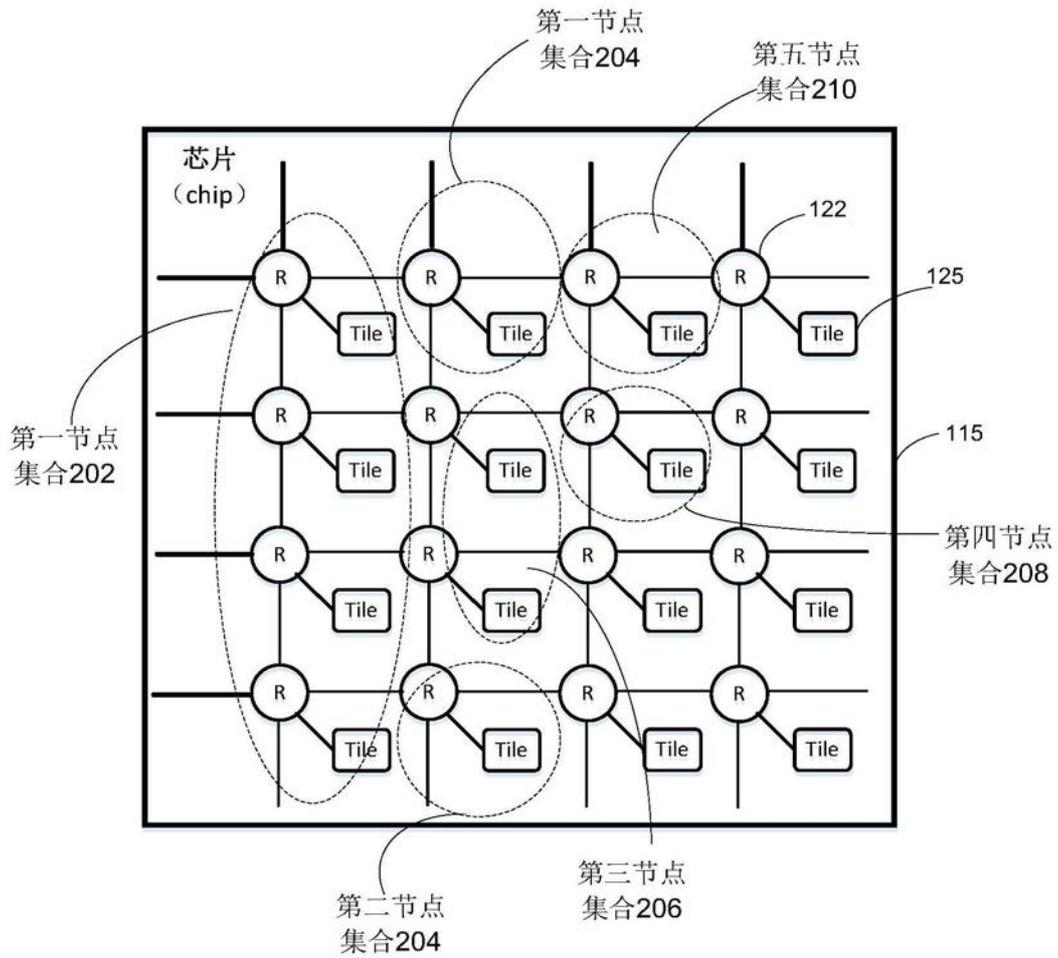


图3

400

流编码	端口
a	East, Local
b	East, West
c	West, North
d	East, South, North
...	...

图4

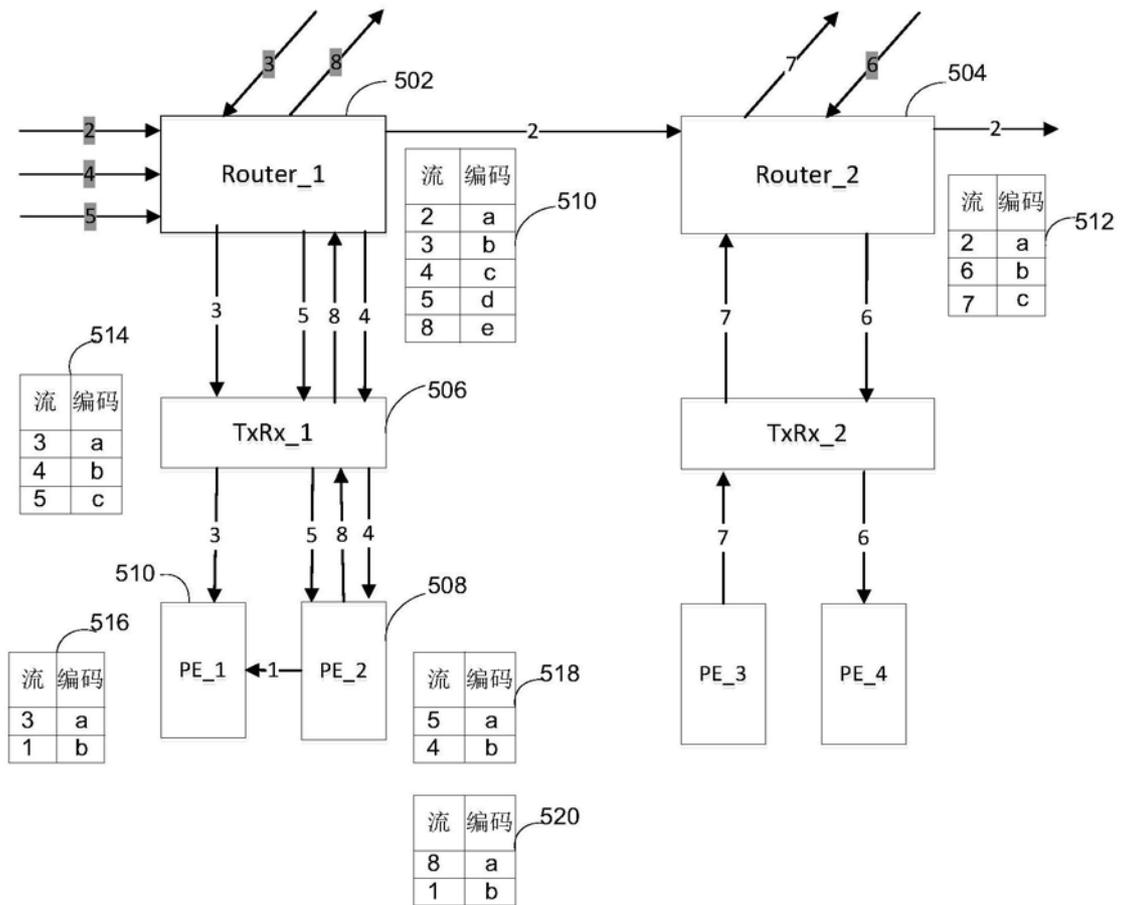


图5

600

	602 第一流编码	604 第二流编码	606 端口
流2 →	a		East, Local
流3 →	b	a	East, West
流4 →	c	b	West, North
流5 →	d	c	East, South, North
	...	...	...

图6

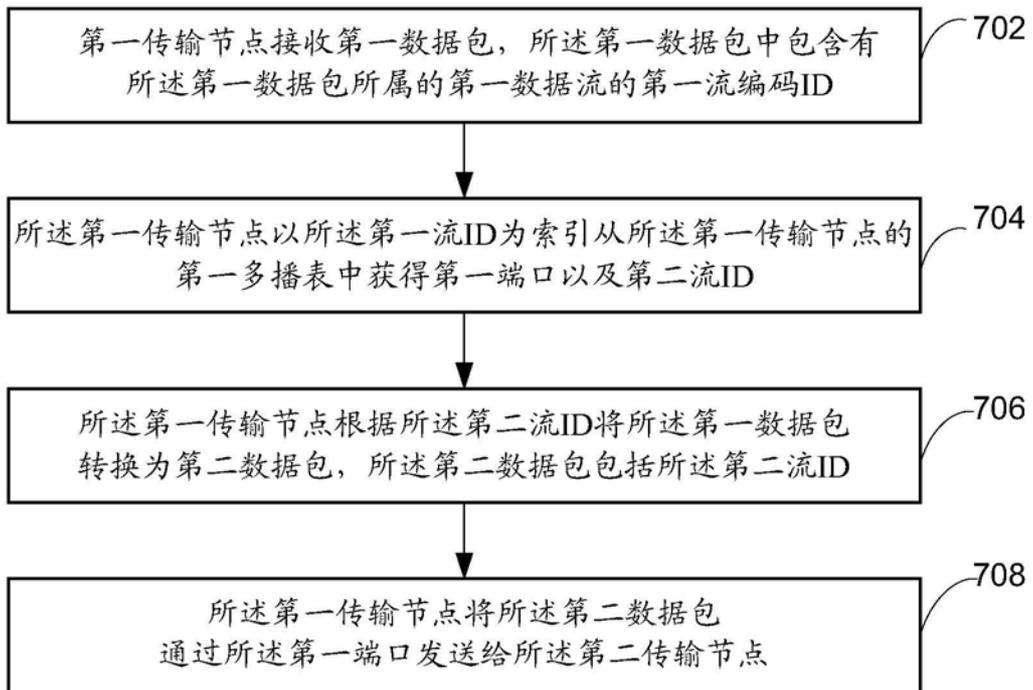


图7

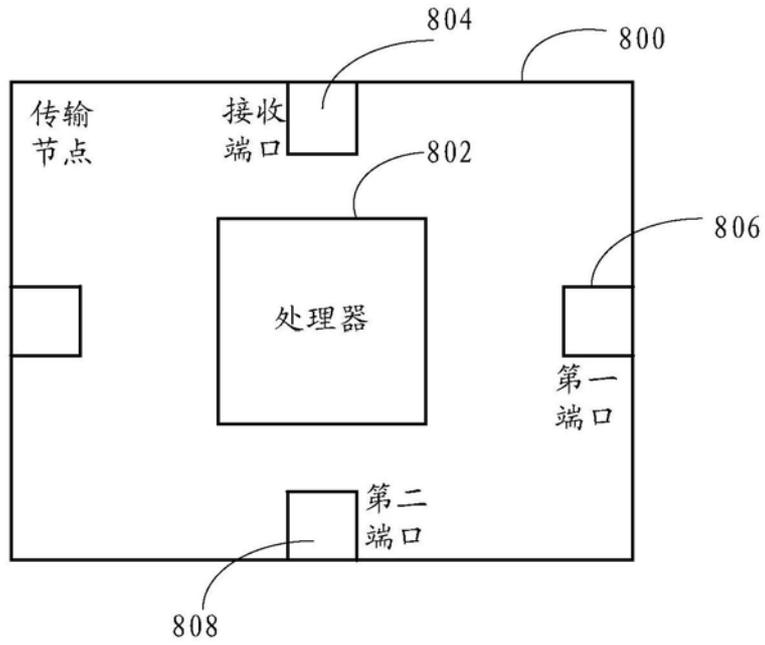


图8

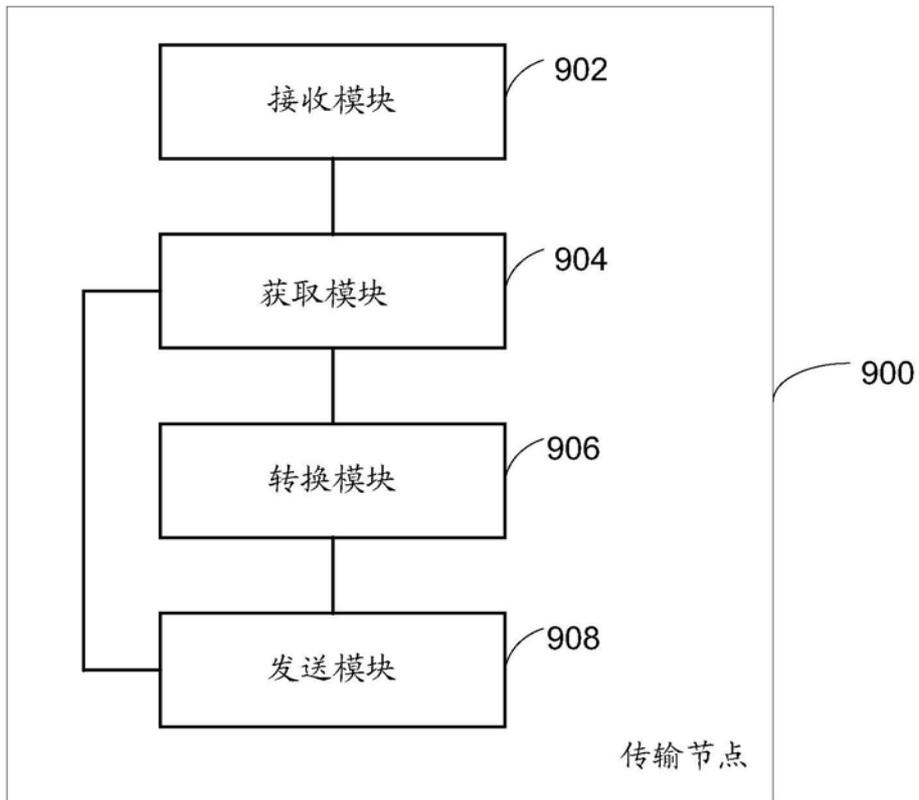


图9