



(12) 发明专利

(10) 授权公告号 CN 102907057 B

(45) 授权公告日 2015. 12. 02

(21) 申请号 201080066907. 8

栏第 34-41 行, 图 1、图 3-4、图 20-21.

(22) 申请日 2010. 05. 20

US 20020097672 A1, 2002. 07. 25, 说明书第 23 段至 95 段, 图 1-7.

(85) PCT国际申请进入国家阶段日
2012. 11. 20

US 20030163555 A1, 2003. 08. 28, 全文.

(86) PCT国际申请的申请数据
PCT/US2010/035661 2010. 05. 20

US 7330467 B2, 2008. 02. 12, 说明书第 2 栏第 14 行至 26 行, 第 51 行至第 8 栏第 61 页, 图 1-3.

(87) PCT国际申请的公布数据
W02011/146066 EN 2011. 11. 24

审查员 薛乐梅

(73) 专利权人 惠普发展公司, 有限合伙企业
地址 美国德克萨斯州

(72) 发明人 N. 宾克特 M. 麦克拉伦 M. 谭

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 段俊峰 卢江

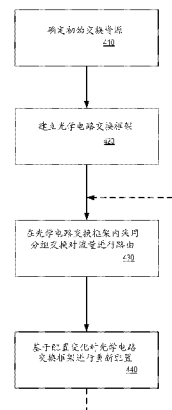
(51) Int. Cl.
H04L 12/931(2013. 01)

(56) 对比文件
US 7317873 B2, 2008. 01. 08, 说明书第 7 栏第 25-67 行, 第 20 栏第 20-67 行, 第 21 栏第 1-67 行, 第 22 栏第 1-52 行, 第 23 栏第 24-62 行, 第 24

权利要求书1页 说明书4页 附图4页

(54) 发明名称
网络装置中的交换

(57) 摘要
本文中描述的网络装置包括执行电路交换的光学电路交换机。所述网络装置还具有多个可移除线卡, 每一线卡包括分组交换机。交换管理器基于所述可移除线卡的配置对所述光学电路交换机自动重新配置, 从而在活动线卡之间保持所保证的分组交换带宽。



1. 一种网络装置,包括:
用以执行电路交换的光学电路交换机;
多个可移除线卡,每一线卡具有分组交换机;以及
交换管理器,其基于所述可移除线卡的配置对所述光学电路交换机进行自适应重新配置,从而在活动线卡之间保持所保证的分组交换带宽。
2. 根据权利要求 1 所述的网络装置,其中,所述光学电路交换机分布于活动线卡当中。
3. 根据权利要求 1 所述的网络装置,其中,每一线卡能够代表其他线卡对分组进行路由。
4. 根据权利要求 1 所述的网络装置,所述光学电路交换机还处理以下至少之一:
网络上的链路冗余;
网络上的链路故障。
5. 根据权利要求 1 所述的网络装置,其中,每一线卡还通过所述光学电路交换机连接至无源光学背板。
6. 根据权利要求 1 所述的网络装置,其中,所述网络装置是以太网交换机。
7. 一种交换方法,包括:
在网络上建立光学电路交换框架;
通过在所述光学电路交换框架的范围内执行分组交换的可移除线卡对所述网络上的流量进行路由;
基于所述可移除线卡的配置的变化而对所述光学电路交换框架自动重新配置,所述重新配置用于在活动线卡之间保持所保证的分组交换带宽。
8. 根据权利要求 7 所述的方法,还包括:
使所述光学电路交换框架分布在活动线卡当中。
9. 根据权利要求 7 所述的方法,其中,所述流量的路由还包括:
活动线卡代表另一活动线卡对分组进行路由。
10. 根据权利要求 7 所述的方法,还包括:
所述光学电路交换框架检测链路状态事件;以及
自动解决所述链路状态事件。
11. 根据权利要求 10 所述的方法,其中,链路状态事件包括以下中的一者或多者:
链路故障;
分组交换机故障。

网络装置中的交换

背景技术

[0001] 各种类型的网络交换机(例如,大型以太网交换机)包括几个能够用线卡填充的插槽。可以按照各种组合填充这些插槽。可以在运行时间插入或者移除线卡,并且任何线卡可能在任何时间发生故障。考虑到线卡故障、添加和 / 或移除,网络操作员 / 管理员通常负责对网络可操作性进行维护。

附图说明

[0002] 下述说明包括对附图的讨论,附图具有作为本发明的实施例的实现的例子给出的图解。

[0003] 图 1 是示出了根据各种实施例的网络装置的方框图。

[0004] 图 2 是示出了根据各种实施例的网络装置的方框图。

[0005] 图 3 是示出了根据各种实施例的网络装置的方框图。

[0006] 图 4 是根据各种实施例的系统中的操作的流程图。

具体实施方式

[0007] 具有多个可用来由线卡填充的插槽的网络装置可以利用若干集中式分组交换机芯片(例如,位于结构卡(fabric card)上或者安装在网络装置的背板上)的使用。可以将交换机芯片上的每一端口硬接线至网络装置上的插槽之一。照此,交换机芯片上的端口专用于相应的插槽,以及交换机芯片能够根据特定的插槽是否填充了运转 / 活动的线卡而开启和关闭这些端口。本文中采用的线卡是指印刷电路板上的一个或多个电子电路,其来自订户的电信线路对接到电信网络的其他部分。

[0008] 一般而言,交换机是指在 OSI (开放系统互连)模型的数据链路层(2层)处理和路由数据的网桥。经常将额外地在网络层(OSI 模型的 3 层)处理数据的交换机称为 3 层交换机、路由器或多层交换机。本文中采用的交换机可以指 2 层桥或 3 层路由器。本文中采用的网络交换机更一般地是指连接网络段的联网装置。

[0009] 上述的集中式分组交换机芯片可以采用分组交换(例如,与电路交换形成对照)以促进与网络装置一起使用分组。这样的交换机芯片往往是价格高昂,并且在所有的插槽都填充有线卡时被最大化利用。因而,部分填充的交换机的购买者为可能从不使用的功能进行支付。

[0010] 本文中描述的实施例将分组交换功能直接结合到网络装置内的线卡上。然后,在可动态重新配置的光学电路交换框架的上下文内实施分组交换。

[0011] 图 1 是示出了根据各种实施例的网络装置的方框图。如图所示,网络装置 100 包括连接至多个线卡 120-150 的集中式电路交换机 110。图 1 所示的线卡的数量只是各种实施例的一个例子;在不同的实施例中可以采用更多的线卡或者更少的线卡。还应当指出,图示的线卡的数量未必等于网络装置 100 上的线卡插槽的数量。换言之,四个图示的线卡可以仅占用网络装置 100 中的可用插槽的子集。一般而言,网络装置 100 可以是以太网交换

机,或者其他适当的用于对分组流量进行路由的装置。

[0012] 线卡 120-150 每个分别具有分组交换模块 122-152。可以将分组交换模块 122-152 实现为一个或多个硬件模块、特殊用途硬件(例如,专用硬件、专用集成电路(ASIC)、嵌入式控制器、硬接线电路等)、软件模块或者这些的某种组合。顾名思义,分组交换模块 122-152 执行分组交换。本文中采用的分组交换是指通过网络输送可变比特率数据流(即,分组序列)。在各种实施例中,每一分组交换模块能够代表网络装置 100 上的其他分组交换模块对分组进行路由。例如,分组交换模块 122-152 中的每个可以采用诸如 Valiant (*L. G. Valiant. A scheme for fast parallel communication. SIAM Journal on Computing, 11(2):350 - 361, 1982*) 提出的路由方案之类的通用路由方案对分组进行路由。

[0013] 将线卡 120-150 中的每个连接至电路交换机 110。如图所示,每一线卡包括多条通往电路交换机 110 的链路(例如,双向链路)。尽管来自每一线卡的链路的具体数量可以发生变化,但是各种实施例至少包括用于提供一定水平冗余的多条链路。

[0014] 在各种实施例中,电路交换机 110 是光学电路交换机。一般而言,电路交换机是指在网络节点之间建立电路或通道,好像所述节点利用电气电路被物理连接一样的交换机。因此,本文中采用的光学电路交换机是指采用光学信号和链路进行通信的电路交换机。下面是可能如何构建光学电路交换机的例子。可以在本文中描述的各种实施例中采用其他适当的(一个或多个)光学电路交换机和 / 或光学交换框架。

[0015] 可以在背板内形成波导(例如,空心金属、聚合物芯等),以创建光学交叉杆(crossbar)框架。例如,第一组并行的交叉杆可以是横贯背板的长度的总线线路,而第二组并行的交叉杆可以是与所述总线线路交叉并连接至计算元件的分接线路。

[0016] 可以将光学元件或者光学元件的组合放置到每一交叉处,从而将光学信号有选择地从所述总线线路引导到分接线路中,反之亦然。这些光学元件可以包括棱镜、反射镜、光阀以及其他光学元件。所述光学元件可以是动态的或者可以是无源的。所述交叉处的光学元件的至少其中之一可以改变状态,从而使互连从贯通(through)状态切换至交叉(crossed)状态。

[0017] 将计算元件连接至所述光学交叉杆框架。例如,可以将主计算装置或网络连接至总线线路,以及可以将若干其他计算装置连接至分接线路。可以将总线线路连接至较大的计算网络或路由器,以及可以将分接线路连接至若干刀片计算机元件。可以将所述刀片计算机元件中的每个连接至多条分接线路。计算装置可以采用更多或者更少的分接线路进行双向通信。此外,所述计算装置可以采用波分复用在给定的线路组上进行双向通信。

[0018] 如上所述,可以在沿所述总线线路的各个位置中连接总线端口。在一些实施例中,使总线端口位于总线线路的一端处可以是有利的。在其他实施例中,总线端口可以附接至总线线路的中央,并通过所述总线线路沿两个方向分配光学信号。一般而言,总线端口在总线线路上的位置可以由若干因素确定,包括:空间限制、连接限制、光损耗预算或者其他相关标准。

[0019] 可以在所述光学交叉杆框架内对光学元件进行动态切换,从而将希望的分接线路连接至总线线路。例如,对光学元件的动态切换可以包括分开切换若干单独的元件或者采用单个致动器移动元件块。作为例子而非限制,这一切换可以包括将实心潜望镜棱镜移到总线线路的路径内,将反射镜倾斜到总线线路的路径内,或者打开光阀从而允许光学信号

从总线线路通过并进入分接线路内。

[0020] 在分接线路和总线线路之间引导光学信号。例如,可以采用一系列组合器元件将来自多条分接线路的光学信号连接至单条总线线路。附加地或可替换地,可以将单条分接线路同时连接至两条或更多总线线路。

[0021] 在一些实施例中,光学电路交换机 110 是单个集中式电路交换机。在其他实施例中,光学电路交换机 110 由几个通过光学链路连接的较小的电路交换机构成。在这样的实施例中,电路交换机 100 可以分布于例如活动的线卡当中(例如,参考图 3)。换言之,除了每一线卡具有分组交换模块之外,每一线卡(或者至少某些线卡)可以具有电路交换模块。可以对分布式电路交换模块加以控制,从而使其作为单个电路交换机协调地工作,如在下文中更详细说明的。

[0022] 交换管理器 160 基于网络装置 100 上的线卡的配置自适应地对光学电路交换机 110 进行重新配置。换言之,交换管理器对网络装置 100 上的线卡插槽进行监视,并确定哪些插槽填充有活动的线卡。基于活动线卡配置,交换管理器 160 对电路交换机 110 进行自适应配置,从而相应地缩放分组交换带宽。可以将交换管理器 160 实现为一个或多个硬件模块、特殊用途硬件(例如,专用硬件、专用集成电路(ASIC)、嵌入式控制器、硬接线电路等)、软件模块或者这些的某种组合。

[0023] 在交换管理器 160 的指示下,光学电路交换机 110 (与例如分组交换模块 122-152 形成对照)处理网络装置 100 的链路冗余和 / 或链路故障。例如,如果链路发生了故障,交换管理器 160 对光学电路交换机 110 重新配置,以补偿链路故障,并在其余的活动链路上保持最大带宽。尽管交换管理器 160 指示、计算和 / 或确定光学电路交换机 110 的重新配置,但是最终对此予以执行的是光学电路交换机 110。

[0024] 图 2 是示出了根据各实施例的系统的方框图。与图 1 类似,网络装置 200 包括各种可移除线卡 220-250、光学背板 270、光学电路交换机 210 和交换管理器 260。与上述实施例一样,电路交换机 210 可以是例如通过光学链路分布在所述线卡当中的光学电路交换机。

[0025] 如图所示,将线卡 220-250 通过光学背板 270 连接至光学电路交换机 210。在各种实施例中,光学背板 270 可以是无源光学背板,其不提供有源总线驱动电路,但是不被认为是网络上的单个故障点。

[0026] 光学电路交换机 210 可被配置(例如,通过交换管理器 260)为将一个线卡以全带宽速率连接至另一线卡(有可能经由中间交换机)。就常规集中式分组交换而言,集中式分组交换机可能是过度供应的,从而保持全带宽连接。通过建立电路交换机 210 的光学电路交换框架,能够以正常的供应保证分组交换模块 222-252 的全带宽连接。

[0027] 此外,经常将集中式分组交换机硬接线至每一线卡。相反,在本文中描述的各种实施例中,如果移除了线卡 220-250 中的一个,那么交换管理器 260 将对电路交换机 210 的电路交换框架重新配置,从而在其余的活动线卡之间保持所保证的分组交换带宽。电路交换机 210 还通过交换管理器 260 处理链路故障和链路冗余。在各种实施例中,将交换管理器 260 实现为管理处理器。在其他实施例中,可以将交换管理器 260 实现为一个或多个硬件模块、特殊用途硬件(例如,专用硬件、专用集成电路(ASIC)、嵌入式控制器、硬接线电路等)、软件模块或者这些的某种组合。

[0028] 图3是示出了根据各种实施例的网络装置的方框图。具体而言,网络装置300示出了一种分布式光学电路交换机。所述分布式光学电路交换机至少跨越线卡320-350分布,所述线卡320-350全部光学连接至光学背板370。在某些实施例中,所述分布式电路交换机可以额外包括独立的电路交换机380和/或电路交换机381。换言之,线卡上的电路交换模块的各种组合和独立电路交换机可以被组合以形成总的光学电路交换机。通过交换管理器360对光学电路交换功能进行控制和管理。如果一个或多个线卡发生了故障,特定链路发生了故障或者发生了其他事件,那么交换管理器360通过协调对各个光学电路交换模块323-353、电路交换机380-381和/或光学背板370的变化而对总的光学电路交换机进行重新配置。与其他实施例一样,可以将交换管理器360实现为一个或多个硬件模块、特殊用途硬件(例如,专用硬件、专用集成电路(ASIC)、嵌入式控制器、硬接线电路等)、软件模块或者这些的某种组合。

[0029] 图4是根据各种实施例的系统中的操作的流程图。图4包括根据某些实施例的特定操作和执行顺序。然而,在不同的实施例中,也可以根据本文中描述的教导采用其他操作,省略一个或多个所描绘的操作,并且/或者按照其他执行顺序进行。

[0030] 确定410初始交换资源。例如,网络装置上的交换管理器可以确定哪些插槽具有活动线卡,哪些线卡具有光学电路交换模块以及网络装置上的任何其他交换资源(例如,分组交换模块)。基于可用资源,建立410用于所述装置的光学电路交换框架。具体而言,考虑所述网络装置中安装的活动(以及可移除)线卡在所述网络装置上建立光学电路交换框架。将所述光学电路交换框架配置为使每一活动线卡上包含的分组交换功能(例如,分组交换模块)最大化。

[0031] 相应地,经由可移除线卡上的分组交换通过所述网络装置对流量进行路由430。在所述光学电路交换框架的范围内执行分组交换。换言之,所述光学电路交换框架(至少逻辑上)定义了网络装置的网络拓扑结构。在给定了该拓扑结构的情况下,线卡上的分组交换机进行操作,从而考虑所述拓扑结构通过分组交换对分组进行路由。

[0032] 在检测到链路状态事件(例如,链路故障、线卡插入、线卡移除等)时,自动解决所述事件。例如,如果线卡的配置发生了变化,那么对光学电路交换框架自动重新配置440(例如,通过交换管理器),从而在活动线卡当中保持链路冗余以及所保证的分组交换带宽。所保证的分组交换带宽可以是全带宽保证或者低于全带宽的保证。

[0033] 本文中描述的各种部件、模块等可以是用于执行相关功能、操作等的机构。

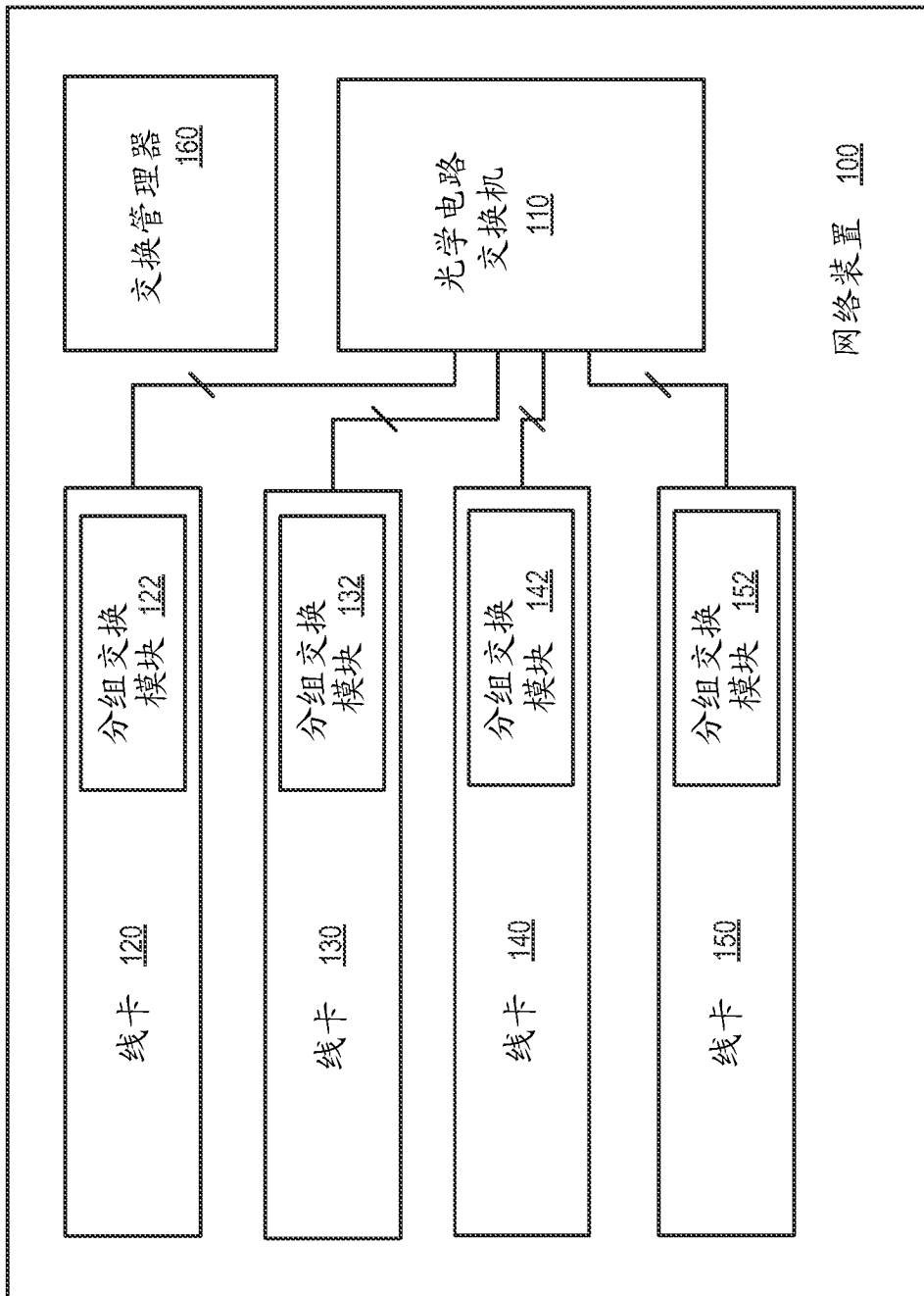


图 1

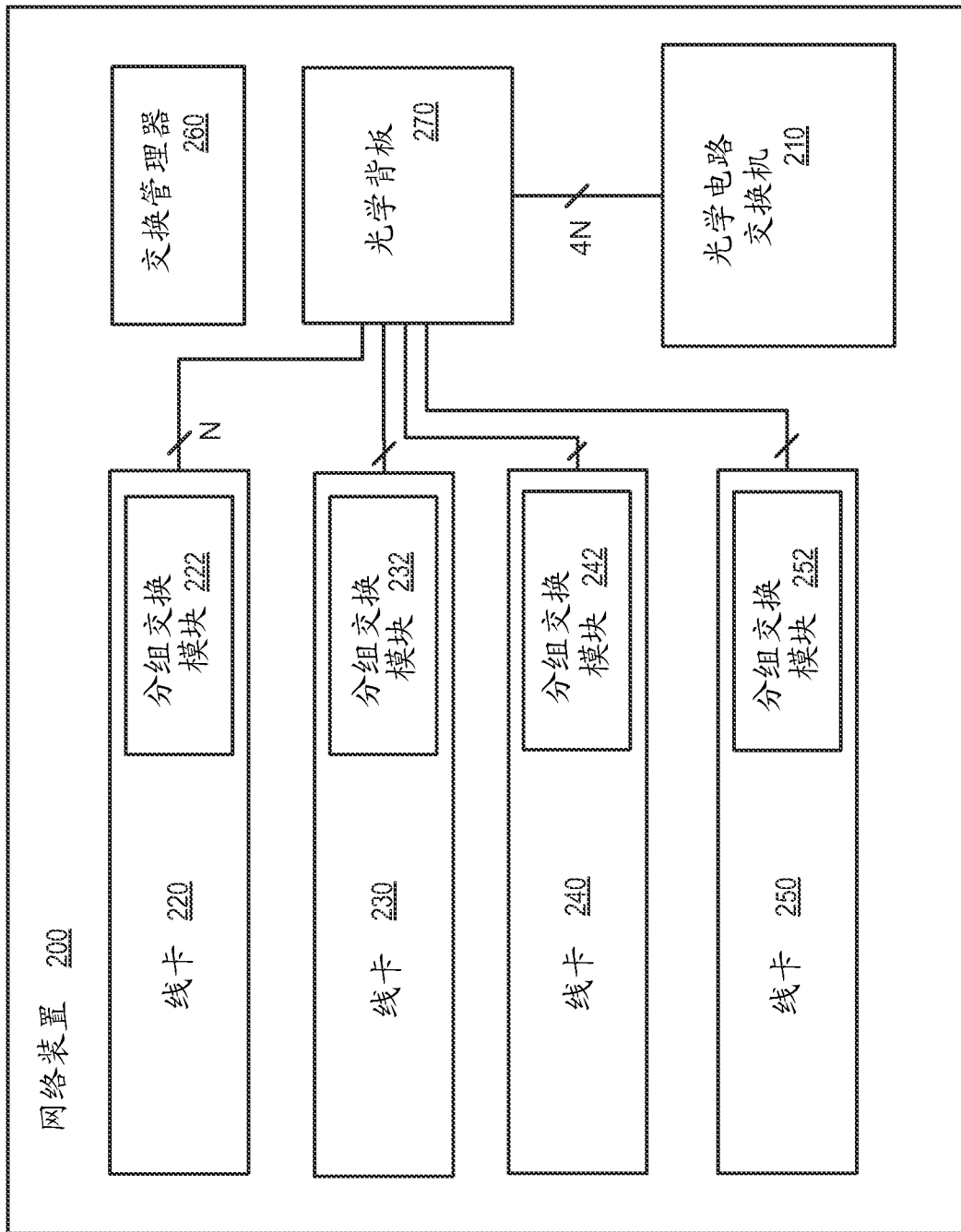


图 2

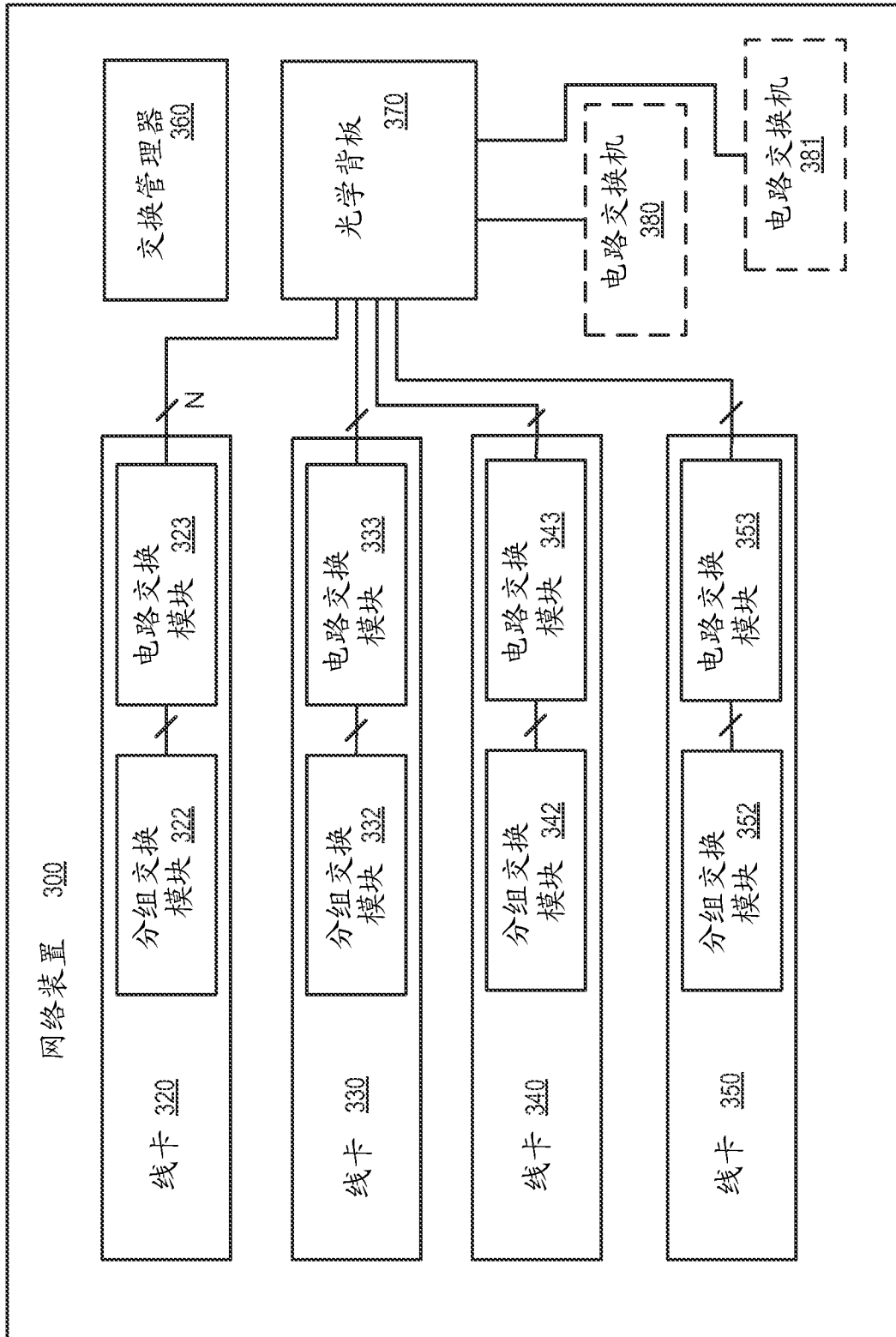


图 3

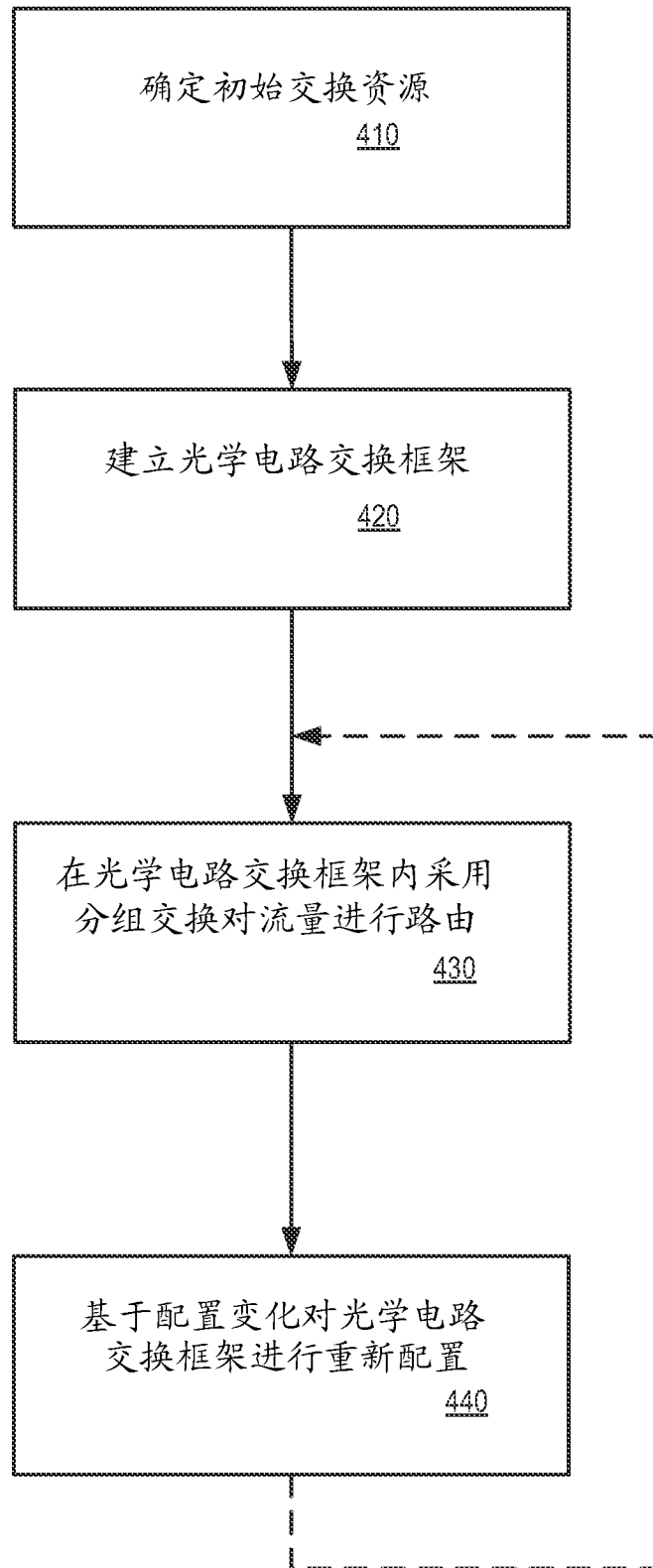


图 4