

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04Q 3/68 (2006.01)

H04L 12/56 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200680034179.6

[43] 公开日 2008年9月17日

[11] 公开号 CN 101268703A

[22] 申请日 2006.7.13

[21] 申请号 200680034179.6

[30] 优先权

[32] 2005.7.15 [33] US [31] 11/182,242

[86] 国际申请 PCT/IB2006/001936 2006.7.13

[87] 国际公布 WO2007/010351 英 2007.1.25

[85] 进入国家阶段日期 2008.3.17

[71] 申请人 艾利森电话股份有限公司

地址 瑞典斯德哥尔摩

[72] 发明人 H·-Y·张

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王岳 张志醒

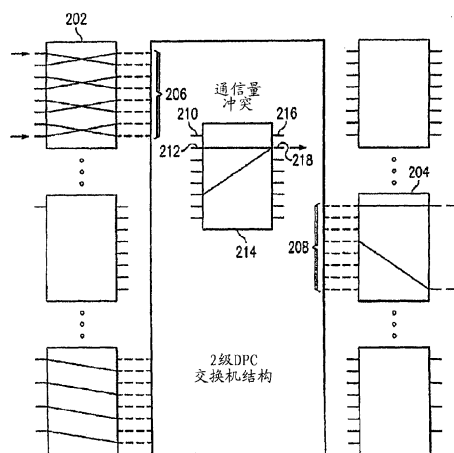
权利要求书4页 说明书24页 附图16页

[54] 发明名称

用于多级交换元件的增强的虚拟电路分配方法和系统

[57] 摘要

一种用于路由数据的系统和方法，其包括：交换节点，所述交换节点包括多个交换元件以及耦合到所述交换元件的多条链路，以便提供用于所述数据的路由通道；以及虚拟电路标识算法，其与该交换节点进行通信以便提供通过该交换节点的均匀的数据通信量分发，其中，所述虚拟电路标识算法确定通道选择公式。



1、一种分配通过互连网络的通道的方法，所述互连网络包括多个交换元件以及耦合所述交换元件的多条链路，该方法包括以下步骤：

使用所述互连网络的物理限制来得到该互连网络的体系结构的逻辑表示；

确定该互连网络的通信量模式，以便平衡通过耦合所述交换元件的所述链路的数据通信量；以及

使用该互连网络的所述逻辑表示和通信量模式来建立表示通道选择公式的虚拟电路标识符表，从而通过所述互连网络更为均匀地分发双工数据通信量。

2、权利要求1所述的方法，还包括：基于分组报头信息，确定将使用所述虚拟电路标识符表的数据通道控制器是处在执行通信量分条的高速板上还是处在低速板上。

3、权利要求2所述的方法，其中，所述交换元件被设置成行和列，其中每一列代表一级。

4、权利要求3所述的方法，其中，增大所述行和列的数目允许所述方法具有可伸缩性。

5、权利要求3所述的方法，其中，所述级数在所述互连网络中是4。

6、权利要求5所述的方法，其中，如果所述数据通道控制器不处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=i[2]^j[5]$$

$$b=i[1]^j[4]$$

$$c=i[0]^j[3]^i[6]$$

$$d=i[1]^j[4]^i[4]$$

$$e=i[2]^j[5]^i[5]$$

其中，^是异或，并且

a、b、c、d、e是通道选择值，其中a、b、c、d和e是纯二进制比特（0或1）。

7、权利要求5所述的方法，其中，如果所述数据通道控制器处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=j[2]$$

$$b=j[1]$$

$$c=j[0]$$

$$d=j[1]^i[4]$$

$$e=j[2]^i[5]$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ ；并且

a、b、c、d、e是通道选择值，其中a、b、c、d和e是纯二进制比特（0或1）。

8、权利要求3所述的方法，其中，所述级数在所述互连网络中是3。

9、权利要求8所述的方法，其中，如果所述数据通道控制器不处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=i[2]^j[5]^i[5]$$

$$b=i[1]^j[4]^i[4]$$

$$c=i[0]^j[5]^i[3]$$

其中， \wedge 是异或，并且

a、b和c是通道选择值，其中a、b和c是纯二进制比特（0或1）。

10、权利要求8所述的方法，其中，如果所述数据通道控制器处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=j[2]$$

$$b=j[1]$$

$$c=j[0]$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ ；并且

a、b和c是通道选择值，其中a、b和c是纯二进制比特（0或1）。

11、一种用于路由数据的系统，包括：

交换节点，其包括多个交换元件以及耦合到所述交换元件的多条链路，以便提供用于所述数据的路由通道；以及

虚拟电路标识算法，其与该交换节点进行通信以便提供通过该交换节点的均匀的数据通信量分发，其中，所述虚拟电路标识算法确定通道选择公式，所述通道选择公式被用来填充虚拟电路标识符表。

12、权利要求11所述的系统，还包括：用于基于分组报头信息来确定将使用所述虚拟电路标识符表的数据通道控制器是处在执行通信量分条的高速板上还是处在低速板上的装置。

13、权利要求12所述的系统，其中，所述交换元件被设置成行和列，其中每一列代表一级。

14、权利要求13所述的系统，其中，增大所述行和列的数目允许所述方法具有可伸缩性。

15、权利要求13所述的系统，其中，所述级数在所述交换节点中是4。

16、权利要求15所述的系统，其中，如果所述数据通道控制器不处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=i[2]^j[5]$$

$$b=i[1]^j[4]$$

$$c=i[0]^j[3]^i[6]$$

$$d=i[1]^j[4]^i[4]$$

$$e=i[2]^j[5]^i[5]$$

其中，^是异或，并且

a、b、c、d、e是通道选择值，其中a、b、c、d和e是纯二进制比特（0或1）。

17、权利要求15所述的系统，其中，如果所述数据通道控制器处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=j[2]$$

$$b=j[1]$$

$$c=j[0]$$

$$d=j[1]^i[4]$$

$$e=j[2]^i[5]$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ ；并且

a、b、c、d、e是通道选择值，其中a、b、c、d和e是纯二进制比特（0或1）。

18、权利要求13所述的系统，其中，所述级数在所述互连节点中是3。

19、权利要求18所述的系统，其中，如果所述数据通道控制器不处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=i[2]^j[5]^i[5]$$

$$b=i[1]^j[4]^i[4]$$

$$c=i[0]^j[5]^i[3]$$

其中，^是异或，并且

a、b和c是通道选择值，其中a、b和c是纯二进制比特（0或1）。

20、权利要求18所述的系统，其中，如果所述数据通道控制器处在执行通信量分条的高速板上，则把所述通道选择公式设置成：

$$a=j[2]$$

$$b=j[1]$$

$$c=j[0]$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ ；并且

a、b和c是通道选择值，其中a、b和c是纯二进制比特（0或1）。

用于多级交换元件的增强的虚拟电路分配方法和系统

相关申请的交叉引用:

本申请涉及 2002 年 1 月 2 日提交的标题为 “Methods and System Virtual Circuit Identification based on BIT Permutation of link numbers for multi-stage elements (用于多级元件的基于链路号的比特置换的虚拟电路标识方法和系统)” 的申请序列号 10/033,039, 其要求 2001 年 2 月 28 日提交的 Ho-Yen Chang、Tyan-shu Jou、Ritesh Ahuja 和 James da Silva 的标题为 “VCI ALLOCATING SCHEME BASED ON BIT PERMUTATION OF LINK NUMBERS FOR MULTI-STAGE ELEMENTS (用于多级元件的基于链路号的比特置换的 VCI 分配方案)” 的临时申请号 60/271,973 的权益。这两个申请都被全文结合在此以作参考。

技术领域

本发明总体上涉及一种数据网络, 更具体来说, 本发明涉及在数据网络的各节点之间分配电路通道。更进一步具体来说, 本发明涉及一种用于多级交换元件的基于链路号的比特置换的虚拟电路标识 (VCI) 分配方案。

背景

数据网络使用多种交换机、路由器和通信量链路来分发数据传输。许多现今的网络交换机和路由器采用由多级 ATM (异步传输模式) 交换元件 (在交换机结构中) 构成的互连网络来提供快速、高带宽数据传输能力。每个交换元件具有多个输入和输出端口和一个数据通道控制器 (DPC), 该数据通道控制器具有经配置的交换机查找表。当已封装的分组到达给定输入端口上时, 该交换元件通过使用分组报头信息进行表查找, 以便决定应当在其哪个输出端口上把该分组发送出去。为了增大路由器系统所需的输入/输出端口的数目, 交换机结构布局包括处在几个级中的多个交换元件, 所述交换元件通过物理链路或总线互连。在物理链路下, 所述数据分组可以通过把不同的虚拟电路标识 (VCI) 置于其已封装的 ATM 报头中来利用多条虚拟电路通道。

在开发这种系统的过程中的一个挑战在于把所述虚拟电路通道（或者所述虚拟电路标识（VCI）映射）分配给每个交换元件的查找表，以便最大化所述互连网络的利用率和吞吐量，并且满足对于总体网络的任何特定通信量模式要求。例如，在互连网络中可能发生通信量通道冲突。当来自两个不同输入源端口的输入通信量使用一个数据通道控制器（DPC）上的相同物理链路到两个不同的输出目的地端口时，就可能发生所述冲突。冲突意味着无法同时利用所述两条通信量通道（即使是对不同的输入和输出端口对）来提供全带宽，这是由于在所述交换机结构中的所述相同物理链路上的争用而造成的。

在一种已知的路由器配置中，该路由器具有 128 个充当“入口”点的输入端口以及 128 个充当该路由器的“出口”点的输出端口。从该路由器的入口侧到出口侧，数据通道的分布常常被分配成使得所述数据被不均匀地传输，并且从而使得所述数据通信量在某些虚拟通道上将更重。这常常导致所述路由器内的数据瓶颈。在此问题之外还包括了具有变化的数据传输速度的交换元件。另外，这些交换元件可能需要把数据分解到多条不同通道中以进行正确操作。

例如，在许多现今的路由器系统体系结构中，使用支持单端口高速板与多端口低速板进行通信的分配方案来路由数据。因此，从该低速板的所有端口到达的数据通信量可以通过多条并行通道没有冲突地聚集到该单端口高速板中。类似地，从高速端口到达的通信量可以被分解到给定低速板的多个端口中。此外，为了支持两个高速交换元件之间的通信，可能必须把通信量均匀地分解到多条较小通道（或“条带（stripe）”）中，并且将其发送到第一级中的多个端口。通过在所述互连网络中预先分配的多条无冲突通道，随后必须在最后级中把通信量聚集到所期望的输出端口。把所述数据划分并重组到对应于各板的所需的多条路径中可能会导致在对所述数据的路由过程中发生冲突，从而可能会导致严重的拥塞和分组丢失。

因此，需要一种设计所述系统的方法，其中，通过交换机结构中的多级交换元件在源端口与目的地端口之间均匀地传输通信量，并且在所传输的数据通道中不会发生冲突。

概要

本发明提供一种对应于根据链路号的比特表示和比特置换进行 VCI 分配的方案的方法和系统。该方案可以被扩展到用在互连网络中的任何多级 ATM 交换机结构，其具有变化的级数和变化的物理布局。

根据一个实施例，公开了一种分配通过互连网络的通道的方法，所述互连网络包括多个交换元件以及耦合所述交换元件的多条链路。该方法包括以下步骤：使用所述互连网络的物理限制在所述交换机结构中得到该互连网络的体系结构的逻辑表示。接下来，确定该互连网络的通信量模式，以便平衡通过耦合所述交换元件的所述多条链路的数据通信量。所述互连网络的所述逻辑表示和通信量模式被用来建立虚拟信道标识符，所述虚拟信道标识符确定通过所述交换元件和链路的通道，从而通过所述互连网络更为均匀地分发数据通信量。

本发明的某些方面提供一种新的“通道选择公式”，其在每个 ATM 交换机元件（DPC，即数据通道控制器）中建立所述 VCI 查找表，并且在每个通信量单元行经不同级上的不同 DPC 时为其分配一条 VCI 通道。

本发明的某些方面提供了在现有技术基础上的增强，从而使得在通过所述交换机结构的任何两个端口之间的双向通信量通道上都不会发生争用，以便在所述接口上提供全双工性能，而先前的解决方案则无法解决这一问题。另外，应当在所有可能的通道之间平衡通信量，以保证所有的 DPC 都有均等的机会被利用。

本发明的某些方面提供了对应于 3 级和 4 级交换机结构的通道选择公式。

根据一个实施例，所述方法包括以下步骤：利用虚拟电路标识符映射所述互连网络；以及使用该虚拟电路标识符来为从输入端口到输出端口通过该互连网络的数据分配通道。

根据另一个实施例，所述分配通过互连网络的通道的方法包括以下步骤：把所述互连网络映射到虚拟电路标识符表，其中所述虚拟电路标识符包括该互连网络的物理限制和通信量模式。接下来，使用所述虚拟标识符来为从所述互连网络的输入端口到输出端口通过该互连网络的数据分配通道。在一个实施例中，所述虚拟电路标识符内的所述物理限制是基于所述互连网络中的交换元件的速度。在另一个实施例中，所述物理限制包括所述互连网络的状态的数目。其他实施例包括这样的虚拟电路标识符，其中所述物理限制是基于所述互连网络的每一级中的交换

元件的数目、对应于该互连网络中的每个交换元件的端口数、该互连网络中的每个输入端口与每个输出端口的链路数、以及该互连网络中的每个输入端口与每个输出端口之间的物理链路布局。

本发明提供了许多优点，其中包括在数据通信量穿过所述互连网络时对所述数据通信量的均匀分发。

本发明的另一个优点是支持不同端口速度的交换元件，其中在对数据进行聚集或分解以便同时适应高速和低速端口时对数据通信量进行更为均匀地分发。

附图简述

通过参考下面结合附图做出的详细描述，可以理解本发明的其他方面，其中包括其优点以及具体实施例，其中：

图 1a 示出了其中可以采用本发明的某些方面的交换机结构的 4 级互连网络的物理布局；

图 1b 示出了示例性交换元件的物理布局；

图 1c 示出了对应于图 1a 中的互连网络中的交换机的虚拟电路标识 (VCI) 表的示例性体系结构；

图 1d 示出了其中可以采用本发明的某些方面的交换机结构的 3 级互连网络的物理布局；

图 2 示出了互连网络内的通道冲突；

图 3a - 3d 示出了对于图 1a 和图 1b 的互连网络的不同通信量模式要求；

图 4 是示出了根据本发明的多级 VCI 建立的示例性流程图；

图 5 是示出了根据本发明的对应于中间级的 VCI 建立的示例性流程图；

图 6 是示出了根据本发明的对应于第一级的 VCI 建立的示例性流程图；

图 7 是示出了根据本发明的最后级 VCI 建立的示例性流程图；

图 8 是示出了根据本发明对于不同通信量要求产生 VCI 分配的示例性流程图；

图 9 是示出了在互连网络中使用本发明的示例性方框图；

图 10 示出了可以在结构单元 (fabric cell) 格式中采用的示例性 4

字节 ATM 报头;

图 11 示出了对于图 1a 的示例性 4 级互连网络填充第一级和最后级表条目的过程;

图 12 示出了对于图 1a 的示例性 4 级互连网络填充第一和第二中间级表条目的过程;

图 13 示出了用于填充示例性 3 级互连网络的第一级和最后级表条目的过程; 以及

图 14 示出了用于填充示例性 3 级互连网络的中间级表条目的过程。

除非另行指出, 否则下面的详细描述中的附图标记对应于附图中的相同附图标记。

详细描述

出于本公开内容的目的使用了各种缩写, 下面列出其定义:

ATM	异步传输模式
clp	拥塞丢失优先级
DP	数据通道
DPC	数据通道控制器
dpcnum	DPC 号
dst	目的地
efci	显性前向拥塞指示
eom	消息末尾
FS	第一级
gfc	通用流控制
ilink	输入链路
IP	因特网协议
ipport	输入端口
ivci	输入 VCI 号
LS	最后级
Ms1	第一中间级
Ms2	第二中间级
Oam	操作、管理和维护

olink	输出链路
oport	输出端口
ovci	输出 VCI 号
Portnum	端口号
QoS	服务质量
VCI	虚拟电路标识
VPI	虚拟通道标识

在制造及使用下面详细讨论的本发明的各实施例之前，应当认识到，本发明提供了可以在多种特定情境中具体实现的许多适用的创造性构思。例如，虽然关于包括 4 级和 3 级的交换网络描述了本发明，但是应当理解，可以用具有不同级数和多种配置的任何多级交换网络来实践本发明。

图 1a 是可以用在诸如 IP 路由器之类的网络节点中的 4 级结构交换机或网络 10 的说明性实施例的图形表示。网络 10 代表数据通信网络中的交换元件，比如 ATM 网络或路由器。为此，互连网络 10 在功能上可以被视为“交换网络”，并且可以通篇互换使用所述术语。正如下面将详细解释的那样，该示例性网络 10 包括设置在 4 列中的 64 个 ATM 交换元件。每一列包含 16 个交换元件并且代表一级。有 4 级：A、B、C 和 D。因此，第一列 12 包括 16 个交换机元件，其表示第一级（或者在该例中是级 A）。在该说明性实施例中，列 12 中的第一个交换机元件被标记为 DPM #0。列 12 中的第二个交换机元件被标记为 DPM #1。列 12 中的第三个交换机元件被标记为 DPM #2，后面依此类推。列 12 中的最后一个交换机元件被标记为 DPM #15。第二列 14 代表具有 16 个交换机元件的第一中间级或级 B 组，所述各交换机元件也被标记为 DPM #0 到 DPM #15。类似地，第三列 16 表示具有 16 个交换机元件的第二中间级（级 C）组，并且第四列 18 表示具有 16 个交换机元件的最后级或级 D 组。在一些实施例中，所述 ATM 交换元件可以是数据通道控制器（DPC）。一个交换元件的主要组件是被称作数据通道（DP）和 DPC 的一对元件。

图 1b 是诸如交换元件 11 之类的示例性交换元件的详细图示。如图所示，每个交换元件支持 8 个输入端口 15a 到 15h 和 8 个输出端口 17a 到 17h，如在图 1b 中所示出的那样。来自输入端口的以 ATM 信元形式

封装的分组将通过输出端口离开所述交换元件。如下面所讨论的那样，分组可以通过列 12 上的各输入端口进入所述交换机结构，所述分组将通过所述交换机结构或网络 10 中示出的 64 个 DPC 当中的 4 个 DPC（每列或每级中一个 DPC）被导向列 18 的一个输出端口（目的地）。

如上所述，互连交换网络 10 由各具有 16 个交换元件的 4 级 A-D 构成，正如图 1a 中所示出的那样。在该实例中，级 A 中的具有偶数 DPC 标签的所有交换元件（例如 DPC #0、2、4、6、8、10、12、14）连接到级 B 中的 16 个交换元件当中的前半（例如列 14 中的 DPC #0、1、2、3、4、5、6、7），而级 A 中的具有奇数标签的所有交换元件（例如 DPC #1、3、5、7、9、11、13、15）则连接到级 B 中的 16 个交换元件当中的后半（例如列 14 中的 DPC #8、9、10、11、12、13、14、15）。此外，级 B 中的具有偶数 DPC 标签的所有交换元件（例如列 14 中的 DPC #0、2、4、6、8、10、12、14）连接到级 C 中的 16 个交换元件当中的前半（例如列 16 中的 DPC #0、1、2、3、4、5、6、7）。级 B 中的具有奇数标签的交换元件（例如列 14 中的 DPC #1、3、5、7、9、11、13、15）连接到级 C 中的 16 个交换元件当中的后半（例如列 16 中的 DPC #8、9、10、11、12、13、14、15）。

按照类似的方式，级 C 中的具有偶数 DPC 标签的交换元件（例如列 16 中的 DPC #0、2、4、6、8、10、12、14）连接到级 D 中的 16 个交换元件当中的前半（例如列 18 中的 DPC #0、1、2、3、4、5、6、7）。级 C 中的具有奇数标签的交换元件（例如列 16 中的 DPC #1、3、5、7、9、11、13、15）连接到级 D 中的 16 个交换元件当中的后半（例如列 18 中的 DPC #8、9、10、11、12、13、14、15）。

互连网络 10 的各目的地点的集合也被称作互连网络 10 的“出口”侧。同样地，级 A 的各输入端口形成其中实现了互连网络 10 的板的“入口”侧。级 A 和最后级 D 的 DP 和 DPC 对处在交换元件上，而级 B 和 C 对则处在互连网络 10 上。

在该实施例中，每个 DPC 包含一个其尺寸为 8192 个条目的虚拟电路标识（VCI）查找表，或者说对于 8 端口板的每个输入端口有 1024 个条目。图 1c 示出了在对应于图 1a 的互连网络 10 的 DPC 中的这种 VCI 表的体系结构。在该例中，对所述 VCI 表的 13 比特索引（0 到 8191）是 3 比特内部端口号（0 到 7）与进入的 VCI 号的 10 个最低有效位（0

到 1023) 的串联。因此, 对于图 1b 中示出的每个输入端口 15a - 15h, 至少可以支持 1024 个 VCI。注意, 对于该系统体系结构, 1024(即 8x128) 个条目可能就足以为每个 DPC 支持输入端口(其中的 8 个)与目的地链路(其中的 128 个链路)的任何组合。对于不同 QoS 等级的通信量可以使用其他 VCI 通道。

在该实施例中, 对于每个输入端口可以支持至少 1024 个 VCI。来自 DPC 的输入端口的 ATM 通信量信元(具有其 ATM 报头的输入 VCI)将被发送到新的输出端口, 其中其 ATM 报头中具有替换的 VCI 值, 这部分地是基于对应于该 DPC 的 VCI 查找表。也就是, 基于所述 ATM 信元报头的 VCI 号、所述输入端口号和所述查找表, 所述 DPC 可以决定将要使用预先分配的哪条静态通道来递送每个通信量信元。因此, 后面描述的算法涉及建立 VCI 查找表以便在任何两个输入和输出端口之间静态地分配完全网状的 VCI 通道。

图 1a 和 1b 中示出的互连设置仅仅是一个例子, 并且可以使用其他互连配置。例如, 一级的偶数编号的交换元件可以链接到后续级的后部交换元件, 或者所述偶数交换元件可以链接到后续级的仅仅奇数编号的交换元件。

图 1d 表示可以用在诸如 IP 路由器之类的网络节点中的 3 级结构交换机或网络 20 的另一个说明性实施例的图形表示。在该示例性实施例中, 网络 20 包括设置在 3 列中的 24 个 ATM 交换元件。第一列 22 交换元件表示第一级, 或者在该例中是级 A。第二列 24 表示中间级或级 B。第三列 26 表示最后级或级 C。

应当理解, 在图 1a、1b、1c 和 1d 中示出及描述的图示意图代表其中可以利用本发明的数据路由系统, 而不是限制本发明的特定物理结构。

在描述了互连网络 10 或网络 20 的配置以及所述交换元件和耦合所述交换元件的链路之后, 现在参考图 2, 该图示出了在图 1a 和 1b 的互连网络 10 内的通道冲突。VCI 分配的基本目标是确保从第一级中的任一个交换元件到最后级中的任何交换元件所存在的不同通道的数目至少可以使得所述互连网络中的给定插槽处的交换元件的每个输入端口可以到达该互连网络中的给定插槽处的交换元件的任何输出端口。例如, 如图 1a 和 1b 所示, 从级 A 中的交换元件到级 D 中的交换元件存在

至少 8 条不同通道,从而到级 A 中的交换元件中的每一个输入可以到达离开级 D 中的交换元件的每一个输出。因此,如果某一数据通信量模式是例如在交换元件 202 与 204 之间的输入端口到输出端口的置换,则可以使用通道 206 和 208 来全速传输所述数据通信量。另外,如果与此同时在特定通道上没有传输其他数据通信量,则可以并行传输所述数据通信量。

如上所述,当来自输入源端口(比如输入端口 210 和 212)的数据通信量尝试使用通过交换元件 214 到输出目的地端口(比如输出端口 216 和 218)的相同物理路径时,可能会发生冲突。冲突意味着在端口 210 和 212 处的数据通信量无法同时使用相同的物理通道并且保持所述数据路由系统的全带宽。本发明提供了一种在保持用来路由数据的系统的带宽能力的同时避免这种冲突的方法和系统。所述方法可以被称作 VCI 分配算法,而所述系统则采取包括本发明的 VCI 分配算法的数据路由系统的形式。

图 3a 到 3d 示出了示例性互连网络中的不同通信量模式要求,其中交换元件可以同时支持高速和低速端口。图 3a 用图形表示两级交换机结构或网络 300。可以看出,本发明的 VCI 分配算法支持这样的情况:其中高速交换元件 304 上的单个输出端口可以提供所需的速度以支持低速多端口交换元件 302。多端口交换元件 302 典型地包括 2 个、4 个、6 个或 8 个端口,不过其他的配置也是可能的。因此,来自第一级中的该低速交换元件 302 的多个端口的通信量可以被没有冲突地并行聚集到单个高速交换元件 304 中。

现在转向图 3b,该图示出了一种支持高速交换元件之间的通信的情况。在该情况下,通信量在两个高速接口板 306 与 308(各利用一个端口)之间流动。在该例中,所述通信量被均匀地分解成较小条带 307,并且被发送到第一级中的多个端口。通过在所述结构中预先分配的多条无冲突通道,随后可以在最后级中把所述通信量聚集到所期望的输出端口 309。换句话说,通过使用可以在所述互连网络中被预先分配的多条无冲突通道,随后可以在最后级中把数据通信量聚集到所期望的输出端口。因此,来自高速交换元件 306 的通信量可以被分解,以便输入到交换元件 308 中。

如所示,图 3c 示出了这样一种情况:高速交换元件可以把来自单

个输入端口的输入通信量分解到低速出口交换机元件的不同输出端口。

图 3d 示出了利用唯一通道在两个低速板之间进行的正常通信量。只要输入端口与输出端口的组合对是不同的，那么交换机元件上的不同输入端口就能够并行地在全速下与另一个交换元件的不同输出端口进行通信。

本发明提供了用于在诸如互连网络 10 之类的互连网络上均匀地分发数据通信量的方法和系统，其中在从输入点传输到输出点时利用了该互连网络的全带宽而不遭受数据通道冲突。可以使用不同速度的交换元件，并且仍然能够避免现有技术系统在所述互连网络内的某些点处所经历的冲突和拥塞的问题。在这方面，本发明使用所述互连网络的物理限制和所述互连网络的逻辑表示来建立虚拟信道标识符，所述标识符确定通过所述交换元件以及连接所述交换元件的链路的通道。这些 VCI 被用来在整个互连网络内均匀地分发数据通信量。

另外，如图 3a 到 3d 所示，所述互连网络可以由不同速度的交换元件构成，其中可以包括低速交换元件和高速交换元件。从低速交换元件到高速交换元件，可以把数据通信量从多个输入聚集到单个输出中。从高速交换元件到低速交换元件，可以把数据通信量从单个输入分解到多个输出中。从高速交换元件到另一个高速交换元件，可以把数据通信量分条。从低速交换元件到另一个低速交换元件，可以在不涉及任何聚集、分解或分条的情况下传输数据通信量。本发明的网络和系统支持所有这些通信量模式，以便在所述互连网络中平衡通信量。

如上所述，每个交换元件具有输入侧和输出侧，并且每一侧可以包含多个端口。所述互连网络被连接成使得第一级中的交换元件的输出端口连接到第二级中的输入端口。另外，例如可以通过使用 7 个比特来表示耦合所述交换元件的链路号。所述 7 个比特的前 4 个比特可以标识与特定交换元件相关联的编号，而剩余的 3 个比特则可以标识对应于该交换元件的端口号。因此，例如可以利用 7 个比特来指定输入源链路号（包括输入交换元件和该交换元件的端口），并且也可以利用例如 7 个比特来指定输出目的地链路号（包括输出交换元件和该交换元件的端口）。对于每一个输入/输出链路对，在特定级处的 VCI 映射条目可以被如下表示：

输入： 交换元件# 输入端口 - 输入 VCI

输出：交换元件# 输出端口 - 输出 VCI

所述互连网络的体系结构可以包括关于所述虚拟电路标识符分配算法的一些指南。回顾参考图 1a 和 1b 的讨论，每一级的偶数交换元件连接到下一级中的交换元件当中的前半，并且每一级的奇数交换元件连接到下一级中的交换元件当中的后半。因此，在该例中，可能需要一些初步指南以在所述互连网络上均匀地分发数据通信量。

参考图 1a，所述指南可以包括如下内容：

对于交换元件的输入侧，与级 A 中的交换元件相关联的交换元件号的奇/偶比特等于与级 B 中的交换元件相关联的交换元件号的前/后比特。另外，与级 A 中的交换元件 11 相关联的交换元件号中的剩余比特等于与级 B 中的交换元件 118 相关联的输入端口号的各比特。此外，级 A 中的交换元件 11 的输出端口的各比特等于与级 B 中的交换元件 118 相关联的交换元件号的剩余比特。最后，对应于级 A 中的交换元件 11 的输出 VCI 号等于级 B 中的交换元件 118 的输入 VCI 号。

现在转向图 4，该图中示出了一个示例性流程图，该流程图示出了根据本发明的多级 VCI 建立的方法。在该例中，操作开始于建立所述互连网络虚拟电路标识符（步骤 402）。随后，识别出在其中建立了虚拟电路标识符的该互连网络中的各中间级（步骤 404）。随后建立交换元件虚拟电路标识符（步骤 406）。识别出在其中建立了虚拟电路标识符的该互连网络中的第一级（步骤 408），并且识别出在其中建立了虚拟电路标识符的该互连网络中的最后级（步骤 410）。

图 5 是示出了根据本发明的对应于中间级的虚拟电路标识符建立方法的示例性流程图。特别地，图 5 的方法是对图 4 中的步骤 404 的进一步说明。在该例中，操作开始于识别所述互连网络中的中间级交换元件（步骤 502）。识别出耦合到该中间级交换元件中的端口的输入链路（步骤 504）。随后，识别出耦合到该中间级交换元件中的该端口的输出链路（步骤 506）。为该中间级交换元件的输出选择通道（步骤 508）。建立对应于该中间级交换元件的输入端口和输出端口的 VCI 表条目（步骤 510）。

接下来，分配与所述中间级交换元件相关联的输入端口的 VCI 输入号（步骤 512）。分配与该中间级交换元件相关联的输出端口的 VCI 号（步骤 514）。随后确定是否有另一个端口与需要在其中分配 VCI 号的

该中间级交换元件相关联（步骤 516）。如果另一个端口与需要在其中分配 VCI 号的该中间级交换元件相关联（步骤 516：是），则操作返回到步骤 514，在该步骤中识别出到该交换元件中的端口的输入链路。

如果没有另一个端口与需要在其中分配 VCI 号的中间级交换元件相关联（步骤 516：否），则随后确定在所述互连网络中是否还有尚未识别出的另一个中间级（步骤 518）。如果在该互连网络中没有另一个尚未识别出的中间级（步骤 518：否），则操作终止。如果在该互连网络中有另一个尚未识别出的中间级（步骤 518：否），则操作返回到步骤 502，在该步骤中识别出该互连网络的中间级。

图 6 是示出根据本发明的对应于第一级的虚拟电路标识符建立方法的示例性流程图。具体来说，图 6 是图 4 的步骤 406 的进一步说明。在该例中，操作开始于识别出所述互连网络中的第一级交换元件（步骤 602）。识别出对应于第一级的该交换元件上的输入端口（步骤 604）。随后识别出耦合到对应于第一级的该交换元件上的输入端口的输出链路（步骤 606）。随后确定该第一级交换元件是否为高速交换元件（步骤 608）。如果该第一级交换元件是高速交换元件（步骤 608：是），则基于输出端口参数选择通道（步骤 610）。如果该第一级交换元件不是高速交换元件（步骤 608：否），则基于输入端口参数选择通道（步骤 612）。

不管所述通道是基于所述输入端口参数还是所述输出端口参数选择的，所述操作随后都建立对应于该第一级交换元件的输入端口和输出端口的 VCI 表条目（步骤 614）。为该 VCI 表中的该输入端口分配 VCI 输入号（步骤 616）。为该 VCI 表中的该输出端口分配 VCI 输出号（步骤 618）。随后确定在该第一级交换元件上是否还有另一个尚未识别出的端口（步骤 620）。如果在该第一级交换元件上没有另一个尚未识别出的端口（步骤 620：否），则确定在第一级中是否还有尚未识别出的任何交换元件（步骤 622）。如果在第一级中没有任何尚未识别出的交换元件（步骤 622：否），则操作终止。如果在第一级中还有尚未识别出的另一个交换元件（步骤 622：否），则操作返回到步骤 602，在该步骤中对于第一级识别出输入交换元件。返回到步骤 620，如果在第一级上还有尚未识别出的另一个端口（步骤 620：是），则操作返回到步骤 604，在该步骤中识别出该第一级交换元件上的输入端口。

图7是示出根据本发明的某些方面的用于最后级虚拟电路标识符建立的方法的示例性流程图。图7是图4中的步骤408的进一步说明。在该例中，操作开始于识别出所述互连网络中的最后级交换元件（步骤702）。随后识别出该最后级交换元件上的输出端口（步骤704）。识别出耦合到该最后级交换元件上的输出端口的输入链路（步骤706）。随后确定该输出交换元件是否为高速交换元件（步骤708）。如果该输出交换元件是高速交换元件（步骤708：是），则基于输出端口参数选择通道（步骤710）。如果该输出交换元件不是高速交换元件（步骤708：否），则基于输入端口参数选择通道（步骤712）。

不管所述通道是基于输入端口参数还是输出端口参数选择的，所述操作随后都建立对应于该最后级交换元件的VCI表条目（步骤714）。为所述输入端口分配VCI输入端口号（步骤716）。为所述输出端口分配VCI输出端口号（步骤718）。随后确定在该最后级交换元件上是否还有另一个尚未识别出的端口（步骤720）。如果在该最后级交换元件上没有另一个尚未识别出的端口（步骤720：否），则确定在最后级上是否还有尚未识别出的任何交换元件（步骤722）。如果在最后级上没有任何尚未识别出的交换元件（步骤722：否），则操作终止。如果在最后级上还有尚未识别出的更多交换元件（步骤722：是），则操作返回到步骤702，在该步骤中对于最后级识别输出交换元件。返回到步骤720，如果在最后级交换元件上还有尚未识别出的另一个端口（步骤720：是），则操作返回到步骤704，在该步骤中识别出该最后级交换元件上的输出端口。

图8是示出根据本发明的对应于支持不同通信量模式要求的VCI分配方案的方法的示例性流程图。在该例中，操作开始于分析所述互连网络的物理限制和布局分析（步骤802）。对诸如互连网络10之类的互连网络的布局分析可以包括确定交换元件的数目以及耦合到所述交换元件的链路的数目。在一个实施例中，所述物理限制被用来得到该互连网络的体系结构的逻辑表示。随后实施通信量模式要求分析（步骤804）。该通信量模式分析被用来平衡通过耦合所述交换元件的链路的数据通信量。

随后转换比特表示和置换函数，其可以对于所述互连网络中的给定源和目的地链路数目动态地计算每一个交换机元件的特定VCI表条目

(步骤 806)。所述比特表示和置换函数可以是该互连网络的体系结构的逻辑表示。确定所述不同通信量要求是否被验证(步骤 808)。如果所述不同通信量要求未被验证(步骤 808: 否), 则操作返回到步骤 806, 在该步骤中转换比特表示和置换函数。如果所述不同通信量要求被验证(步骤 808: 是), 则在该互连网络中建立交换机元件 VCI(步骤 810), 并且随后操作终止。该互连网络的逻辑表示和通信量模式被用来建立虚拟信道标识符, 所述虚拟信道标识符确定把所述数据路由通过其中的通道。所述数据被路由通过所述交换元件和所述链路, 从而通过该互连网络更均匀地分发所述数据通信量。

在图 8 中示出的操作的附加优点在于允许交换元件的热插拔, 并且满足了不同通信量模式的许多特征。交换元件的热插拔允许在不对与部件相连的板断电并且不干扰数据通信量流的情况下用另一个交换元件来替换特定交换元件。通过利用本发明, 允许对交换元件进行热插拔而不会导致通信量模式中的任何中断。另外, 在把数据传输到目的地端口的整个过程中可以保留源链路号。保留源链路号是符合期望的, 这是因为目的地链路能够以容易且高效的方式确定通信量的源。

图 9 是示出在交换网络 900 内使用本发明的示例性方框图。交换网络 900 可以包括输入端 902、虚拟电路标识算法 904、第 1 级 906、第 2 级 908、第 n 级 910 以及输出端 916。在最一般的意义下, 交换网络 900 是用于路由数据的系统。交换网络 900 包括每一级 906-910 内的多个交换元件以及耦合所述交换元件和各级的多条链路, 以便提供用于路由数据的通道。虚拟电路标识算法 904 与交换网络 900 通信, 并且提供指令以用于建立通过交换网络 900 的均匀的数据通信量分发。交换网络 900 包括用于确定网络 900 的物理限制的逻辑 912。另外, 交换网络 900 还包括用于利用交换网络 900 的通信量模式平衡通过所述多条链路的数据通信量的逻辑 914。

因此, 在本发明所提供的方法和系统中, 在所述互连网络中的各级之间均匀地传输源链路和目的地链路之间的通信量, 从而在数据传输之间不会发生冲突。通过本发明的过程, 不会浪费所述互连网络中的可用通路。数据通信量被沿着所述互连网络均匀地传输, 从而一条数据通道不会干扰另一条数据通道上的数据传输。此外, 可以使用各种速度的交换元件, 并且可以根据所述交换元件的速度要求来聚集或分解数据通信

量。本发明在所述互连网络中没有冲突和拥塞的情况下提供了系统且有序的数据通信量分配。

本发明的某些方面提供了增强，从而使得在通过所述交换机结构的任何两个端口之间的双向通信量通道上都不会发生争用，以便在所述接口上提供全双工性能。这适用于低速到低速、低速到高速以及高速到高速之间的全部 8 条通道。高速端口的第一级 VCI 分配需要独立于出口插槽号，以便实现高速板上的“条带上负载平衡”实现方式。换句话说，如同先前的解决方案一样，在决定对应于 HS 通道的 <a>、和<c>的值时，不能使用比特 $j_3 - j_6$ 。

因此，可以在全双工性能下使用所述 VCI 通道。另外，在所述结构上，应当在所有可能的通道之间平衡通信量，以便保证所有 DPC 都具有均等的机会被利用。本发明还实现了对应于所述 3 级交换机结构的比特表示和通道选择公式。

对应于 4 级交换机结构的比特表示可以被概括如下：

级 [0] :: 第一级

输入 = $(i[6], i[5], i[4], i[3]) : (i[2], i[1], i[0]) - (j[6], j[5], j[4], j[3], j[2], j[1], j[0])$

输出 = $(i[6], i[5], i[4], i[3]) : (a, b, c) - (j[6], j[5], j[4], j[3], j[2], j[1], j[0])$

级 [1] :: 中间级1

输入 = $(i[3], a, b, c) : (i[6], i[5], i[4]) - (j[6], j[5], j[4], j[3], j[2], j[1], j[0])$

输出 = $(i[3], a, b, c) : (d, e, j[6]) - (i[6], j[5], j[4], j[3], j[2], j[1], j[0])$

级 [2] :: 中间级2

输入 = $(c, d, e, j[6]) : (i[3], a, b) - (i[6], j[5], j[4], j[3], j[2], j[1], j[0])$

输出 = $(c, d, e, j[6]) : (j[5], j[4], j[3]) - (i[6], i[5], i[4], i[3], j[2], j[1], j[0])$

级 [3] :: 最后级

输入 = $(j[6], j[5], j[4], j[3]) : (c, d, e) - (i[6], i[5], i[4], i[3], j[2], j[1], j[0])$

输出 = $(j[6], j[5], j[4], j[3]) : (j[2], j[1], j[0]) - (i[6], i[5], i[4], i[3], i[2], i[1], i[0])$

其中，不同级中的交换元件的每个条目具有如下表示的输入和输出部分：

	(4 比特)	(3 比特)	(7 比特)
输入:	Dpc # (插槽号)	输入端口号 (<i>iport</i>)	输入 VCI 号 (<i>ivci</i>)
输出:	Dpc # (插槽号)	输出端口号 (<i>oport</i>)	输出 VCI 号 (<i>ovci</i>)

在该实施例中，除了最后级中的输出 VCI 之外，VCI 号的最后 3 个比特是 $j[2,1,0]$ ，最后级中的该输出 VCI 使用 $i[2,1,0]$ 。因此，在给出输出插槽（对应于目的地端口）的情况下，对应于该输出插槽的通信量的 VCI 号将不同于到其他输出插槽的通信量。另外，最后级中的输出 VCI 号是 i ，即源链路号，其允许最后级确定所述通信量的源链路是哪一个。

由于(a, b, c, d, e)值决定一条物理通道，因此问题仍然是如何分配(a, b, c, d, e)值以及如何选择每一级中的输出 VCI 号。

如上面参考图 1a 所讨论的那样，4 级交换机结构（第一级、中间级 1、中间级 2、最后级）包含 128 条链路（编号为 0~127）。存在 64 个 DPC（即 ATM 交换机）。将通过所述交换机结构中的 64 个 DPC 当中的 4 个 DPC 把来自一条输入链路（源）的以 ATM 信元形式封装的分组指定给一条输出链路（目的地）。

在将其发送到所述结构之前，可以把每个分组封装到结构 ATM 信元格式中（利用适当的填充和分段）。结构 ATM 信元类似于 ATM 信元（例如其可以具有 4 字节报头和 48 字节净荷）。在图 10 中示出了一个示例性的 4 字节报头。如图 10 中所示，前 4 个比特表示所述 *gfc*，接着 5 个比特保存所述 *vpi*，后面的 11 个比特保存所述 *vci*。后面的 4 个比特分别保存所述 *oam*、*efci*、*eom* 和 *clp* 比特。

对于该说明性实例，所述 VCI 字段将被用来基于所述 VCI 值在所述路由器结构内部转发所述通信量。该 VCI 字段可以被每个结构级用来决定接下来把所述单元发送到哪里（例如哪个 DPC 输出端口）。虽然该说明性实例中的每条物理链路可以载送 16 比特 (2^{16}) VCI 号，但是为简单起见，这里的讨论将使用 7 比特 VCI 号。

当分组信元到达每一级的 DPC 时，该 DPC 将基于该分组的输入端口号和输入 VCI 号执行表查找（在其预先填充的交换机表上），以便决定用哪个输出端口将该分组信元导向下一级 DPC，以及对于下一级 DPC

把哪个新的输出 VCI 放在该分组信元的报头上。注意，在该实施例中，可以从链路号（0~127）导出 DPC 号（0~15，即插槽号）和特定 DPC 上的端口号（0~7）。对于所述分组在第一级 DPC 中从输入链路 i 穿越到输出链路 j ，该分组可以在其 ATM 报头中携带输出链路 j （目的地）以作为其输入 vci 号，从而可以确定其目的地。在最后级 DPC 中，该穿越分组携带输入链路 i （源）作为其输出 VCI 号，从而可以确定其源。

下面的讨论说明了用来填充（在所有级处的）每个 DPC 中的交换机的算法的各实施例，从而可以通过所述 DPC 经由表查找转发通信量来满足某些通信量要求。

现在转向图 11，其中给出了用于填充可用在如图 1a 中所示的示例性 4 级交换机结构中的 VCI 表的第一级和最后级 DPC 建立。（下面将参考图 12 讨论中间级 DPC 建立。）对于该示例性 4 级交换机结构，所述讨论将使用 7 比特链路号（0~127）来表示输入（src）和输出（dst）链路。换句话说，所述 DPC 通道建立是用于从链路 $i[6,5,4,3,2,1,0]$ 到输出链路 $j[6,5,4,3,2,1,0]$ 的通信量。在这种情况下，存在输入与输出链路的 128×128 种可能的组合。

因此，图 11 示出了对于第一级和最后级 DPC 计算 VCI 条目并且填充所述表条目（即每个 DPC 中 128×128 个条目）的示例性过程。在该实施例中，使用一系列嵌套循环来填充各表条目。步骤 1102 表示第一循环（或循环 i ）的开始，其执行填充对应于第一级和最后级中的每个 DPC 的表的过程（即利用插槽号 0~15）。步骤 1104 表示一个嵌套循环的开始，其对每个输入端口号（0~7）执行一个过程。类似地，步骤 1106 表示一个嵌套循环的开始，其对每个输出链路号（0~127）执行一个过程。

在步骤 1108 中，所述过程确定所讨论的 DPC 是否处在执行通信量分条的高速板上。如果不是的话，则所述过程继续到步骤 1110。另一方面，如果所讨论的 DPC 处在执行通信量分条的高速板上，则所述过程继续到步骤 1112。步骤 1110 和步骤 1112 计算所述通道选择值(a, b, c, d, e)，其中 a、b、c、d 和 e 是纯（only）二进制比特（0 或 1）。所述通道选择值(a, b, c, d, e)还决定在从第一级输入端到最后级输出端的所述虚拟电路通道中使用哪些 DPC。

步骤 1110 计算对应于普通板的通道选择公式。在该计算中，通过

所述输入端口号 ($i[2], i[1], i[0]$) 与最后级输出 DPC 插槽号的 3 个比特 ($j[5], j[4], j[3]$) 以及输入 DPC 插槽号的一个比特 ($j[6]$) 的异或的值来选择第一级 DCP 输出端口号 (即 a, b, c)。此外, 通过反转的第一级 DPC 输入端口号 ($i[0], i[1], i[2]$) 与部分输入 DPC 插槽号比特 ($i[6], i[4], i[5]$) 以及最后级输出 DPC 插槽号比特 ($j[3], j[4], j[5]$) 的异或来选择 (c, d, e), 其决定中间级 DPC 输出端口 (和哪个 DPC) 以及最后级 DPC 输入端口号。于是, 所述通道选择公式将是如下:

$$\begin{aligned} a &= i[2] \wedge j[5], \\ b &= i[1] \wedge j[4], \\ c &= i[0] \wedge j[3] \wedge i[6], \\ d &= i[1] \wedge j[4] \wedge i[4], \\ e &= i[2] \wedge j[5] \wedge i[5] \end{aligned}$$

其中, \wedge 是异或。

另一方面, 步骤 1112 计算对应于利用了分条的高速板的通道选择公式。在这种情况下, 该高速板对于给定 DPC 仅仅具有一个输入端口 $i=0$ (即 $i[2]=0, i[1]=0, i[0]=0$)。因此, 可以使用最后级输出 DPC 端口号来决定 (a, b, c); 并且 (c, d, e) 是反转的最后级 DPC 输出端口号 ($j[0], j[1], j[2]$) 与部分输入 DPC 插槽号 ($i[4], i[5]$) 的异或。所述通道选择公式将是如下:

$$\begin{aligned} a &= j[2], \\ b &= j[1], \\ c &= j[0], \\ d &= j[1] \wedge i[4], \\ e &= j[2] \wedge i[5] \end{aligned}$$

其中, $i[2,1,0]=\langle 0,0,0 \rangle$ 。

一旦在步骤 1110 或 1112 中计算出所述通道选择公式, 随后就可以在步骤 1114 中确定并填充关于给定的第一级 DPC 号的表条目值。如图 11 中所示, 对应于给定的第一级 DPC 号的输入端口值可以如下确定:

输入= $(i[6], i[5], i[4], i[3]):(i[2], i[1], i[0])-(j[6], j[5], j[4], j[3], j[2], j[1], j[0])$
类似地, 对应于给定的第一级 DPC 号的输出端口值可以被计算为:

输出=(i[6],i[5],i[4],i[3]):(a, b, c)-(j[6],j[5],j[4],j[3],j[2],j[1],j[0])

换句话说:

DPC#=(i[6],i[5],i[4],i[3]);

输入端口=(i[2],i[1],i[0]); 输入 VCI=(j[6],j[5],j[4],j[3],j[2],j[1],j[0]);

输出端口=(a, b, c); 输出 VCI=(j[6],j[5],j[4],j[3],j[2],j[1],j[0])

在步骤 1116 中, 随后可以确定并填充关于给定的最后级 DPC 号的表条目值。如图 11 中所示, 对应于给定的最后级 DPC 号的输入端口可以被确定为:

输入=(j[6],j[5],j[4],j[3]):(c, d, e)-(i[6],i[5],i[4],i[3],j[2],j[1],j[0])

类似地, 对应于给定的最后级 DPC 号的输出端口可以被计算为:

输出=(j[6],j[5],j[4],j[3]):(j[2],j[1],j[0])-(i[6],i[5],i[4],i[3],i[2],i[1],i[0])

换句话说:

DPC#=(j[6],j[5],j[4],j[3]);

输入端口=(c, d, e); 输入 VCI=(i[6],i[5],i[4],i[3],j[2],j[1],j[0]);

输出端口=(j[2],j[1],j[0]); 输出 VCI=(i[6],i[5],i[4],i[3],i[2],i[1],i[0])

步骤 1118 表示开始于步骤 1106 的循环结束。类似地, 步骤 1120 表示开始于步骤 1102 的循环结束。

图 12 示出了对于第一和第二中间级 DPC 计算 VCI 条目并且填充所述表条目 (即每个 DPC 中 128x128 个条目) 的示例性过程。在该实施例中, 使用一系列嵌套循环来填充各表条目。步骤 1202 表示第一循环 (或循环 i) 的开始, 其执行填充对应于第一级和最后级中的每个 DPC 的表的过程 (即利用插槽号 0~15)。步骤 1204 表示一个嵌套循环的开始, 其对每个输入端口号 (0~7) 执行一个过程。类似地, 步骤 1206 表示一个嵌套循环的开始, 其对每个输出链路号 (0~127) 执行一个过程。

在步骤 1208 中, 所述过程确定所讨论的 DPC 是否处在执行通信量分条的高速板上。如果不是的话, 则所述过程继续到步骤 1210。另一方面, 如果所讨论的 DPC 处在执行通信量分条的高速板上, 则所述过程继续到步骤 1212。步骤 1210 和步骤 1212 计算所述通道选择值(a, b, c, d, e), 其中 a、b、c、d 和 e 是纯二进制比特 (0 或 1)。

在步骤 1210 中, 如前面参考图 11 所示的过程的步骤 1110 所描述的那样利用对应于普通板的通道选择公式来计算所述通道选择公式。因

此，所述通道选择公式将是如下：

$$\begin{aligned} a &= i[2] \wedge j[5], \\ b &= i[1] \wedge j[4], \\ c &= i[0] \wedge j[3] \wedge i[6], \\ d &= i[1] \wedge j[4] \wedge i[4], \\ e &= i[2] \wedge j[5] \wedge i[5] \end{aligned}$$

其中， \wedge 是异或。

类似地，另一方面，步骤 1212 按照前面参考步骤 1110 所描述的那样计算对应于利用了分条的高速板的通道选择公式。因此，所述通道选择公式将是如下：

$$\begin{aligned} a &= j[2], \\ b &= j[1], \\ c &= j[0], \\ d &= j[1] \wedge i[4], \\ e &= j[2] \wedge i[5] \end{aligned}$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ 。

一旦在步骤 1210 或 1212 中计算出所述通道选择公式，随后就可以在步骤 1214 中确定并填充关于给定的第一中间级 DPC 号的表条目值。如图 12 中所示，对应于给定的第一中间级 DPC 号的输入端口可以如下确定：

$$\text{输入}=(i[3], a, b, c):(i[6],i[5],i[4])-(j[6],j[5],j[4],j[3],j[2],j[1],j[0])$$

类似地，对应于给定的第一中间级 DPC 号的输出端口可以被计算为：

$$\text{输出}=(i[3], a, b, c):(d, e, j[6])-(i[6],j[5],j[4],j[3],j[2],j[1],j[0])$$

换句话说：

$$\text{DPC}\#=(i[3], a, b, c);$$

$$\text{输入端口}=(i[6],i[5],i[4]); \text{输入 VCI}=(j[6],j[5],j[4],j[3],j[2],j[1],j[0]);$$

$$\text{输出端口}=(d, e, j[6]); \text{输出 VCI}=(i[6],j[5],j[4],j[3],j[2],j[1],j[0])$$

在步骤 1216 中，随后可以确定并填充关于给定的第二中间级 DPC 号的表条目值。如图 12 中所示，对应于给定的第二中间级 DPC 号的输

入端口可以被确定为:

输入=(c, d, e, j[6]):(i[3], a, b)-(i[6],j[5],j[4],j[3],j[2],j[1],j[0])

类似地, 对应于给定的最后级 DPC 号的输出端口可以被计算为:

输出=(c, d, e, j[6]):(j[5],j[4],j[3])-(i[6],i[5],i[4],i[3],j[2],j[1],j[0])

换句话说:

DPC#=(c, d, e, j[6]);

输入端口=(i[3], a, b); 输入 VCI=(i[6],j[5],j[4],j[3],j[2],j[1],j[0]);

输出端口=(j[5],j[4],j[3]); 输出 VCI=(i[6],i[5],i[4],i[3],j[2],j[1],j[0])

步骤 1218 表示开始于步骤 1206 的循环结束。类似地, 步骤 1220 表示开始于步骤 1202 的循环结束。

对于如图 1d 所示的 3 级交换机结构, 由于第一级上的每个 DPC 连接到第二级上的所有 DPC, 因此对于所述比特表示的限制较少。在该实施例中, 第一级 DPC 的输出端口号被用来确定第二级的 DPC 号, 并且仅仅(a, b, c)被用来选择物理通道。所述比特表示于是可以是如下:

级 [0]::第一级

输入=(i[5],i[4],i[3]):(i[2],i[1],i[0])-(j[5],j[4],j[3],j[2],j[1],j[0])

输出=(i[5],i[4],i[3]):(a, b, c)-(j[5],j[4],j[3],j[2],j[1],j[0])

级 [1]::中间级 1

输入=(a, b, c):(i[5],i[4],i[3])-(j[5],j[4],j[3],j[2],j[1],j[0])

输出=(a, b, c):(j[5],j[4],j[3])-(i[5],i[4],i[3],j[2],j[1],j[0])

级 [2]::最后级

输入=(j[5],j[4],j[3]):(a, b, c)-(i[5],i[4],i[3],j[2],j[1],j[0])

输出=(j[5],j[4],j[3]):(j[2],j[1],j[0])-(i[5],i[4],i[3],j[2],j[1],j[0])

现在转向图 13, 其中给出了对应于 3 级交换机结构的第一级和最后级 DPC 建立。(下面将参考图 14 讨论中间级 DPC 建立)。对于在图 13 中使用的该 3 级交换机结构, 所述讨论将使用 6 比特链路号来表示输入(src)和输出(dst)链路或端口。换句话说, 所述 DPC 通道建立是用于从 i[5,4,3,2,1,0]发送到输出链路 j[5,4,3,2,1,0]的通信量。在这种情况下, 存在输入与输出端口的 64x64 种可能的组合。

因此, 图 13 示出了对于第一级和最后级 DPC 计算 VCI 条目并且填充所述表条目(即每个 DPC 中 64x64 个条目)的示例性过程。在该实

施例中，使用一系列嵌套循环来填充各表条目。步骤 1302 表示第一循环（或循环 i ）的开始，其执行填充对应于第一级和最后级中的每个 DPC 的表的过程（即利用插槽号 0~15）。步骤 1304 表示一个嵌套循环的开始，其对每个输入端口号（0~7）执行一个过程。类似地，步骤 1306 表示一个嵌套循环的开始，其对每个输出链路号（0~64）执行一个过程。

在步骤 1308 中，所述过程确定所讨论的 DPC 是否处在执行通信量分条的高速板上。如果不是的话，则所述过程继续到步骤 1310。另一方面，如果所讨论的 DPC 处在执行通信量分条的高速板上，则所述过程继续到步骤 1312。步骤 1310 和步骤 1312 计算所述通道选择值(a, b, c)。

步骤 1310 计算对应于普通板的通道选择公式。由于任何第一级 DPC 可以物理地到达任何中间级 DPC，因此可能必须通过使用所述输入端口号 ($i[2]$, $i[1]$, $i[0]$) 与所有输入和输出插槽号 ($i[5, 4, 3]$ 和 $j[5, 4, 3]$) 的异或来提高所述置换分发。也就是，所述通道选择公式将是如下：

$$\begin{aligned} a &= i[2] \wedge j[5] \wedge i[5], \\ b &= i[1] \wedge j[4] \wedge i[4], \\ c &= i[0] \wedge j[3] \wedge i[3] \end{aligned}$$

其中， \wedge 是异或。

另一方面，步骤 1312 计算对应于利用了分条的高速板的通道选择公式。在这种情况下，该高速板对于给定 DPC 仅仅具有一个输入端口 $i=0$ （即 $i[2]=0$, $i[1]=0$, $i[0]=0$ ）。因此，可以使用最后级输出端口号来决定(a, b, c)，其是第一级输出端口（因此还确定中间级 DPC）。所述通道选择公式将是如下：

$$\begin{aligned} a &= j[2], \\ b &= j[1], \\ c &= j[0], \end{aligned}$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ 。

一旦在步骤 1310 或 1312 中计算出所述通道选择公式，随后就可以在步骤 1314 中确定并填充关于给定的第一级 DPC 号的表条目值。如图 13 中所示，对应于给定的第一级 DPC 号的输入端口可以如下确定：

输入=(i[5],i[4],i[3]):(i[2],i[1],i[0])-(j[5],j[4],j[3],j[2],j[1],j[0])

类似地，对应于给定的第一级 DPC 号的输出端口可以被计算为：

输出=(i[5],i[4],i[3]):(a, b, c)-(j[5],j[4],j[3],j[2],j[1],j[0])

换句话说：

DPC#=(i[5],i[4],i[3]);

输入端口=(i[2],i[1],i[0]); 输入 VCI=(j[5],j[4],j[3],j[2],j[1],j[0]);

输出端口=(a, b, c); 输出 VCI=(j[5],j[4],j[3],j[2],j[1],j[0])

在步骤 1316 中，随后可以确定并填充关于给定的最后级 DPC 号的表条目值。如图 13 中所示，对应于给定的最后级 DPC 号的输入端口可以被确定为：

输入=(j[5],j[4],j[3]):(a, b, c)-(i[5],i[4],i[3],j[2],j[1],j[0])

类似地，对应于给定的最后级 DPC 号的输出端口可以被计算为：

输出=(j[5],j[4],j[3]):(j[2],j[1],j[0])-(i[5],i[4],i[3],i[2],i[1],i[0])

换句话说：

DPC#=(j[5],j[4],j[3]);

输入端口=(a, b, c); 输入 VCI=(i[5],i[4],i[3],j[2],j[1],j[0]);

输出端口=(j[2],j[1],j[0]); 输出 VCI=(i[5],i[4],i[3],i[2],i[1],i[0])

步骤 1318 表示开始于步骤 1306 的循环结束。类似地，步骤 1320 表示开始于步骤 1302 的循环结束。

图 14 示出了对于所述示例性 3 级结构交换机的中间级 DPC 计算 VCI 条目并且填充所述表条目（即每个 DPC 中 64x64 个条目）的示例性过程。在该实施例中，使用一系列嵌套循环来填充各表条目。步骤 1402 表示第一循环（或循环 i）的开始，其执行对于该中间级中的每个 DPC 的过程（即利用插槽号 0~15）。步骤 1404 表示一个嵌套循环的开始，其对每个输入端口号（0~7）执行一个过程。类似地，步骤 1406 表示一个嵌套循环的开始，其对每个输出链路号（0~63）执行一个过程。

在步骤 1408 中，所述过程确定所讨论的 DPC 是否处在执行通信量分条的高速板上。如果不是的话，则所述过程继续到步骤 1410。另一方面，如果所讨论的 DPC 处在执行通信量分条的高速板上，则所述过程继续到步骤 1412。步骤 1410 和步骤 1412 计算所述通道选择值(a, b, c)。

步骤 1410 通过利用与步骤 1310 中相同的通道选择公式来计算普通板的中间级 DPC 处的表条目，从而这三个级元件中的 DPC 建立可以匹

配并且形成良好的 VCI 通道。所述通道选择公式将是如下：

$$\begin{aligned} a &= i[2] \wedge j[5] \wedge i[5], \\ b &= i[1] \wedge j[4] \wedge i[4], \\ c &= i[0] \wedge j[5] \wedge i[3] \end{aligned}$$

其中， \wedge 是异或。

类似地，步骤 1412 对于利用了分条的高速板使用与步骤 1312 中相同的通道选择公式来计算所述表条目。所述通道选择公式将是如下：

$$\begin{aligned} a &= j[2], \\ b &= j[1], \\ c &= j[0], \end{aligned}$$

其中， $i[2,1,0]=\langle 0,0,0 \rangle$ 。

一旦在步骤 1410 或 1412 中计算出所述通道选择公式，随后就可以在步骤 1415 中确定并填充关于给定的中间级 DPC 号的表条目值。如图 14 中所示，对应于给定的中间级 DPC 号的输入端口可以如下确定：

$$\text{输入}=(a, b, c):(i[5],i[4],i[3])-(j[5],j[4],j[3],j[2],j[1],j[0])$$

类似地，所述输出端口可以被计算为：

$$\text{输出}=(a, b, c):(j[5],j[4],j[3])-(i[5],i[4],i[3],j[2],j[1],j[0])$$

换句话说：

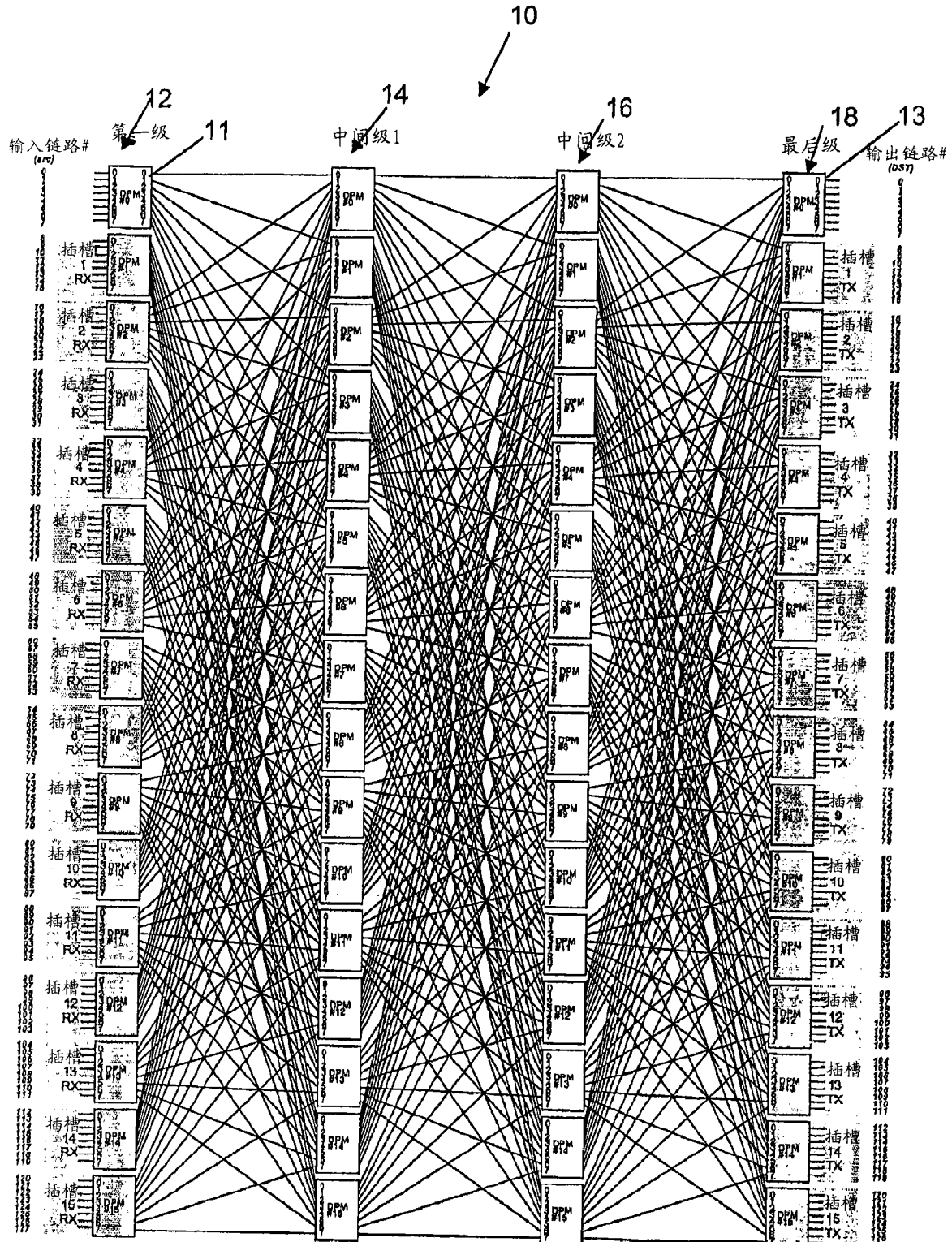
$$\text{DPC}\#=(a, b, c);$$

$$\text{输入端口}=(i[5],i[4],i[3]); \text{输入 VCI}=(j[5],j[4],j[3],j[2],j[1],j[0]);$$

$$\text{输出端口}=(j[5],j[4],j[3]); \text{输出 VCI}=(i[5],i[4],i[3],j[2],j[1],j[0])$$

步骤 1416 表示开始于步骤 1406 的循环结束。类似地，步骤 1418 表示开始于步骤 1402 的循环结束。

上面示出及描述的实施例仅仅是示例性的。虽然在前面的描述中已经与本发明的结构和功能的细节一起阐述了本发明的许多特性和优点，但是本公开内容仅仅是说明性的，并且在所附权利要求书中使用的术语的一般广义含义的最大范围内，可以在本发明的原理内做出改变。



结构交换机存储器配置

图 1a

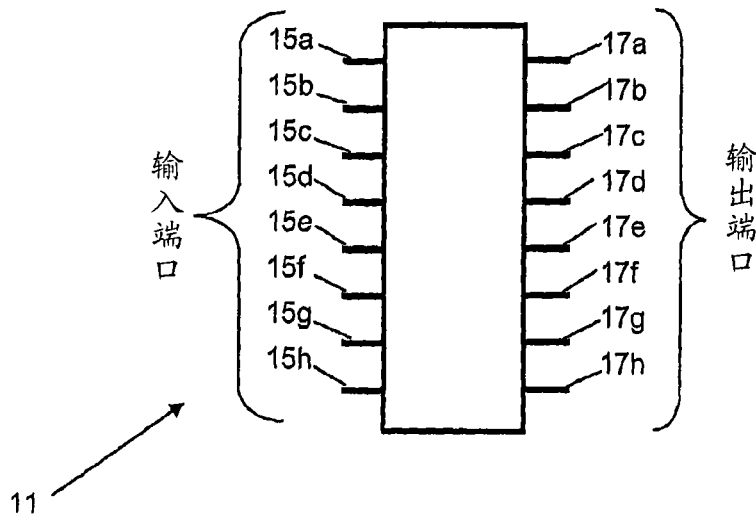
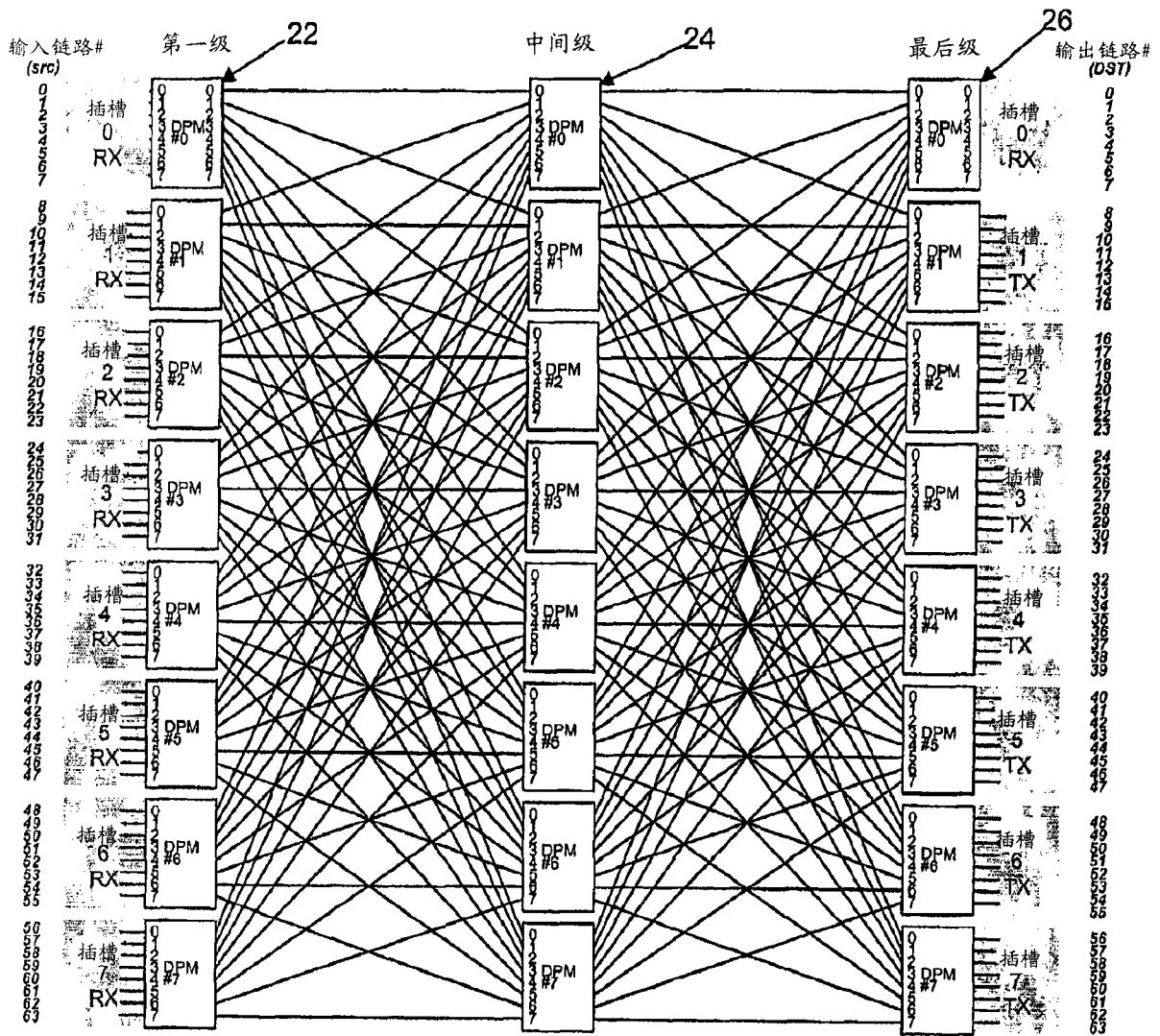


图 1b

输入端口 (<i>lport</i>)	输入VCI (<i>lvcl</i>)	输出端口 (<i>oport</i>)	输出VCI (<i>ovcl</i>)
0	0	0	0
0	1	0	1
0	2	0	2
⋮	⋮	⋮	⋮
0	127	7	127
1	0	1	0
1	1	1	1
1	2	1	2
⋮	⋮	⋮	⋮
1	127	6	127
2	0	2	0
⋮	⋮	⋮	⋮
2	127	5	127
⋮	⋮	⋮	⋮
7	127	0	127
⋮	⋮	⋮	⋮

图 1c



结构交换机存储器配置



20

图 1d

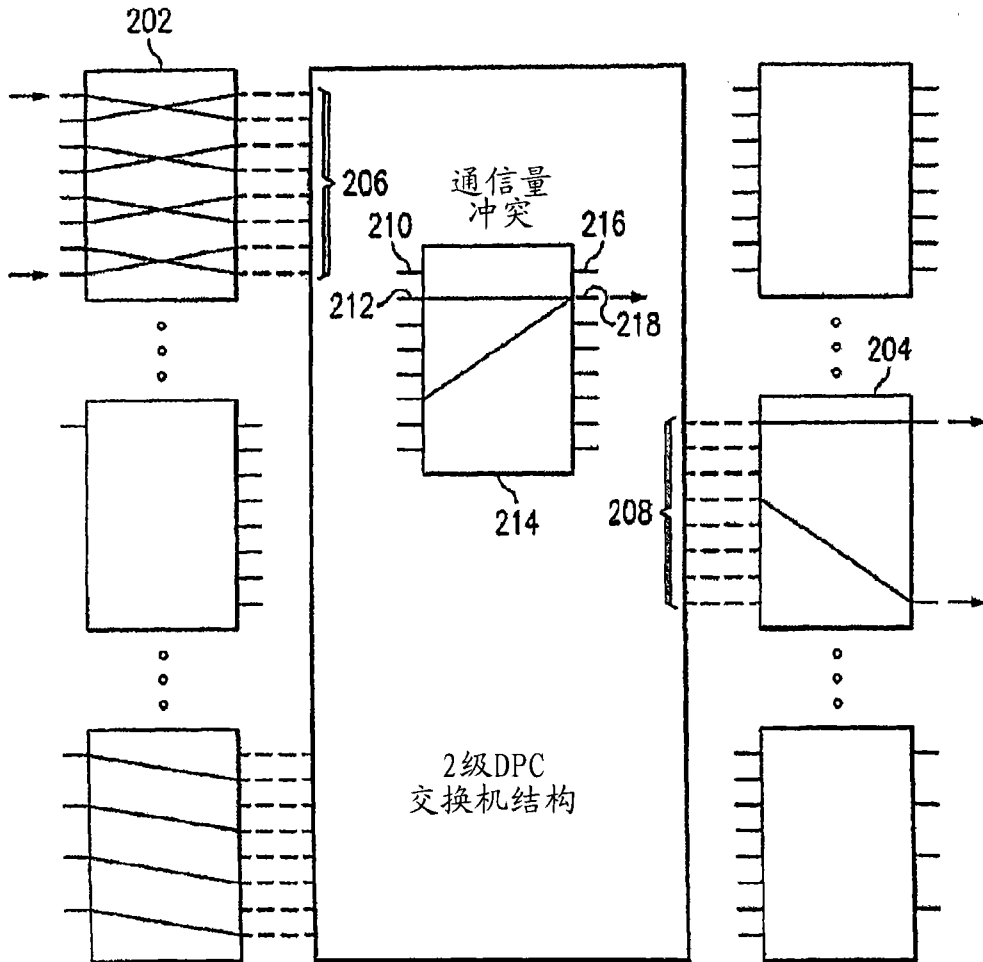


图 2

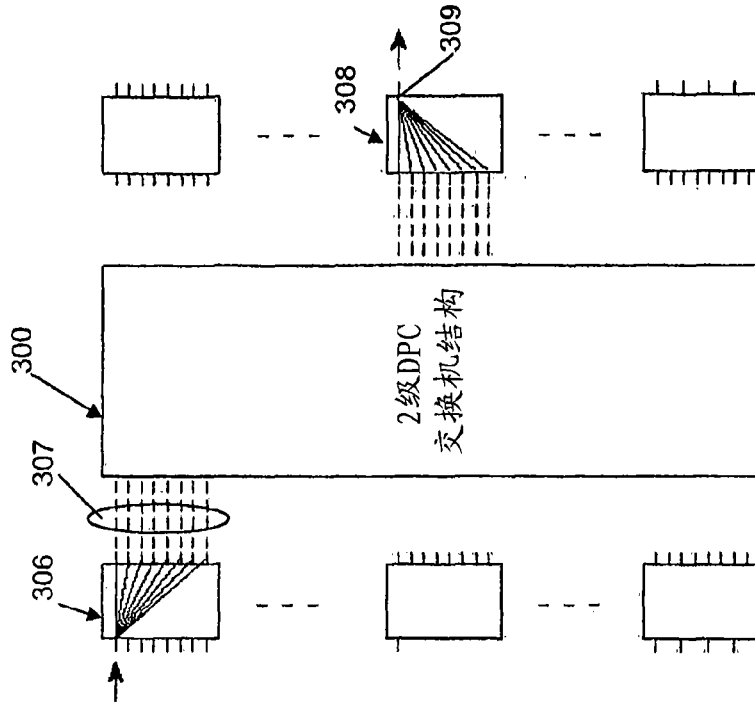


图 3b

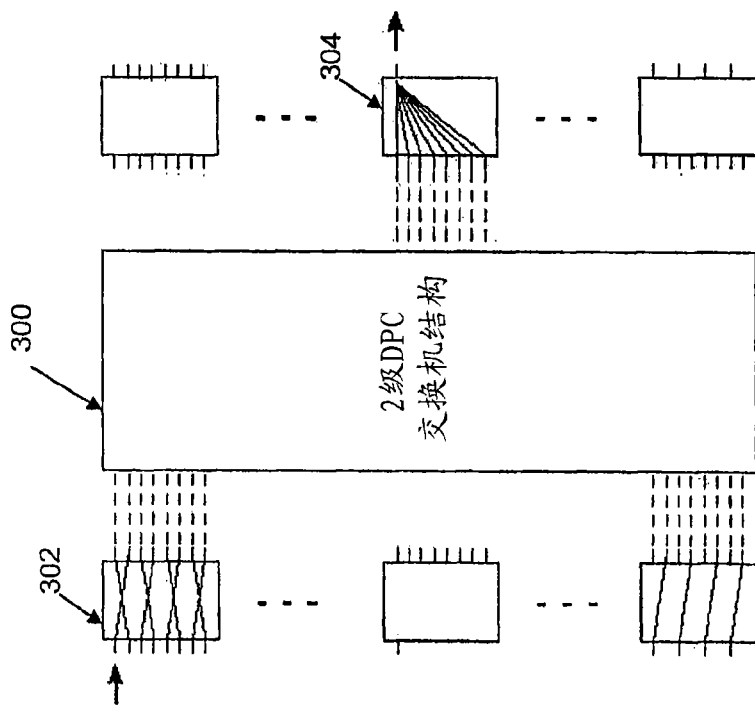


图 3a

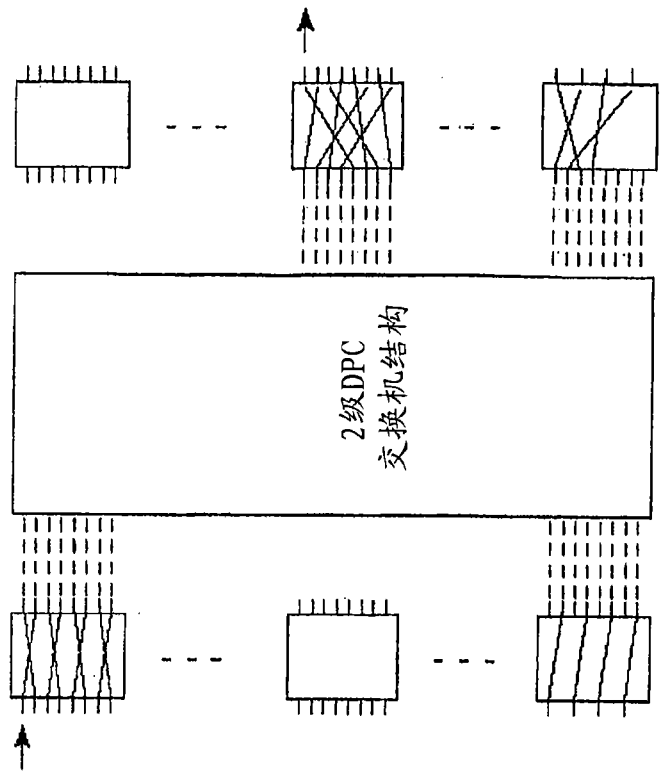


图 3d

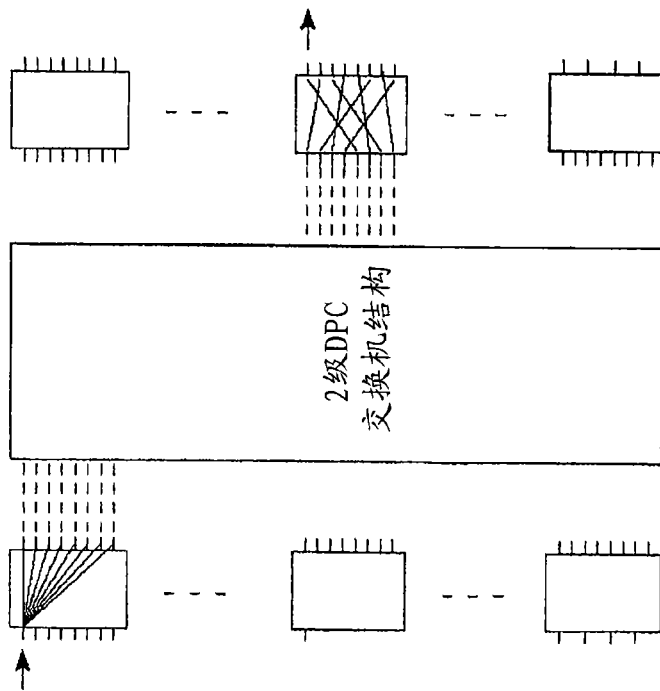


图 3c

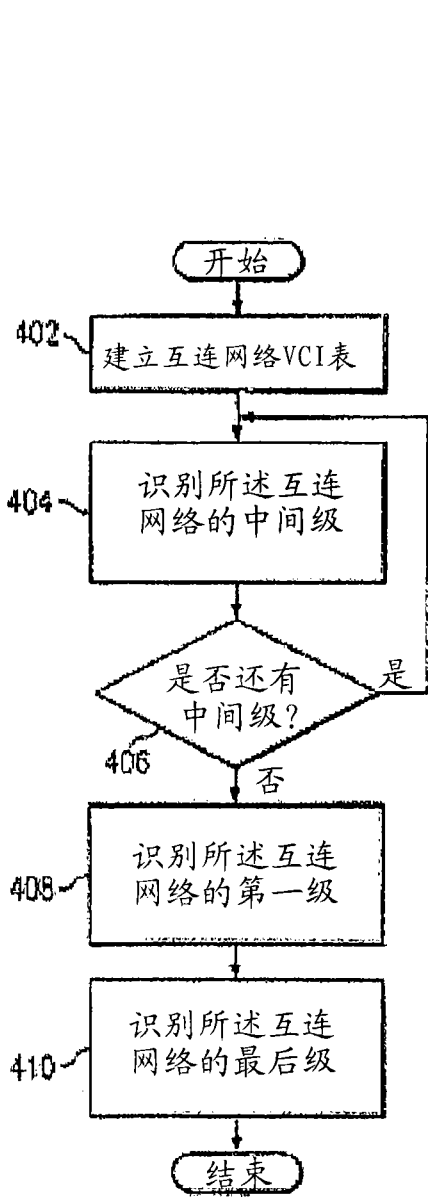


图 4

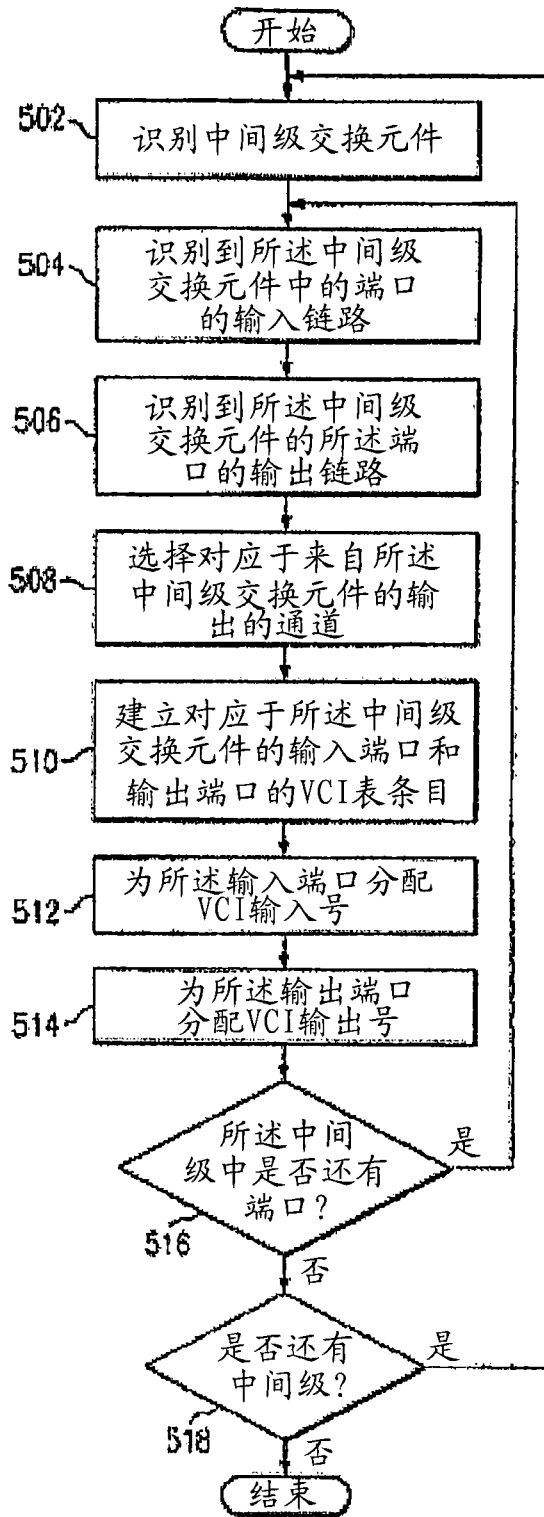


图 5

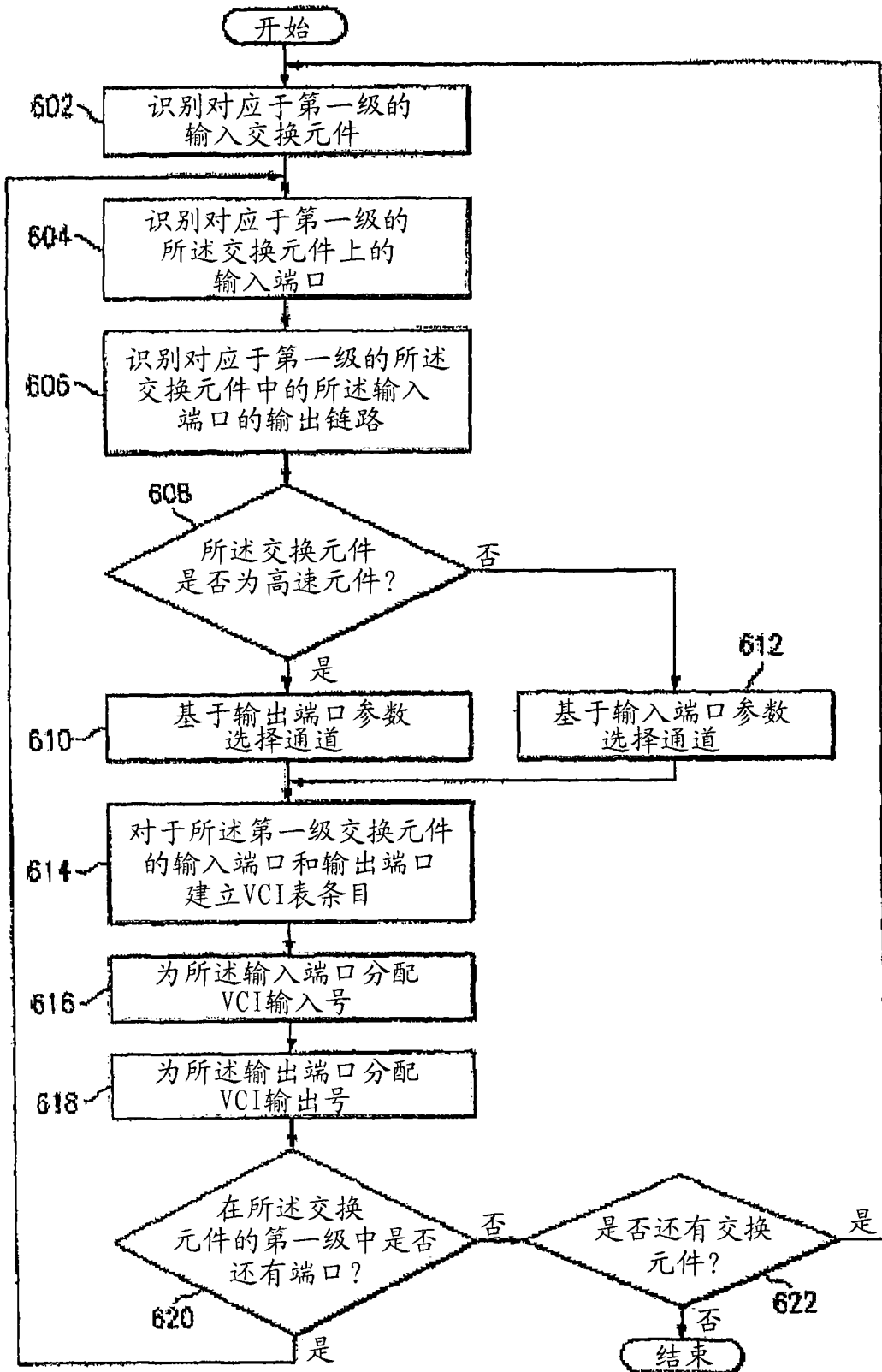


图 6

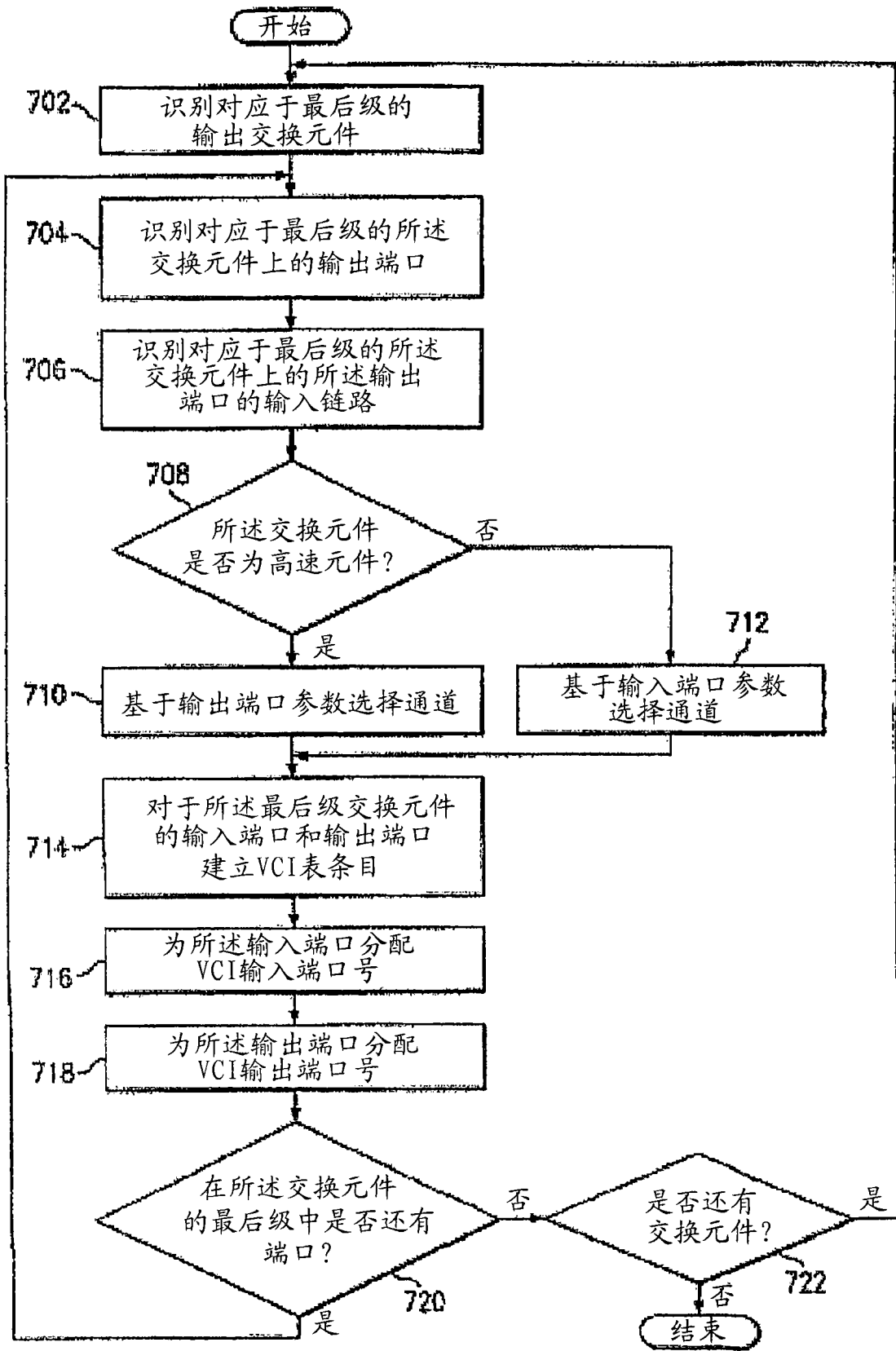


图 7

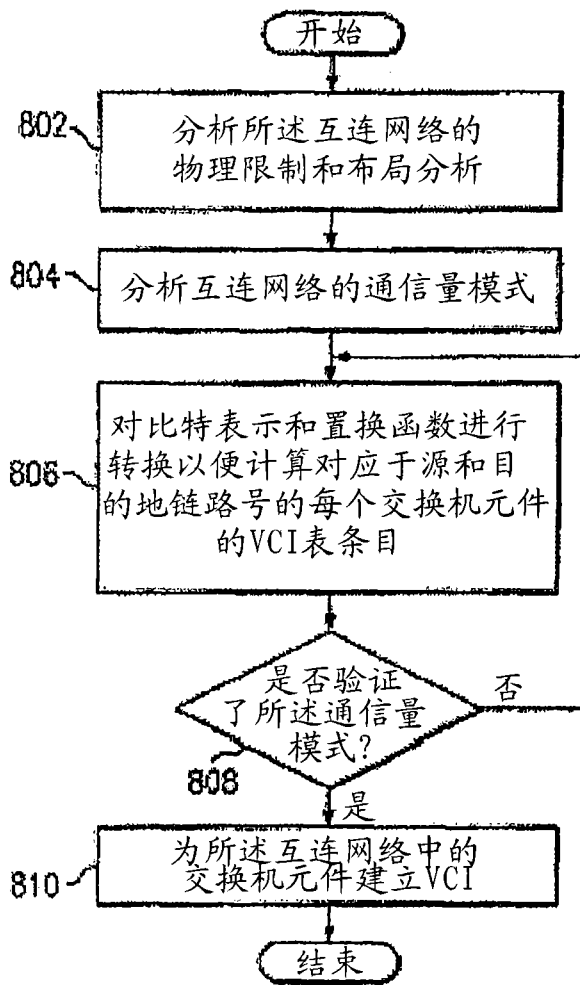


图 8

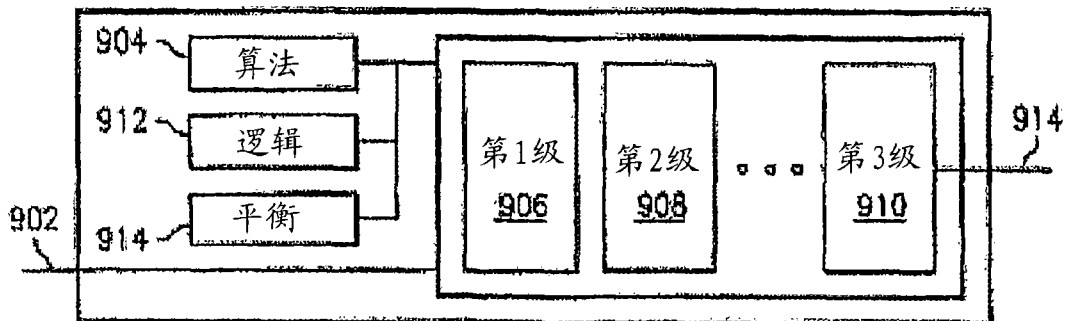


图 9

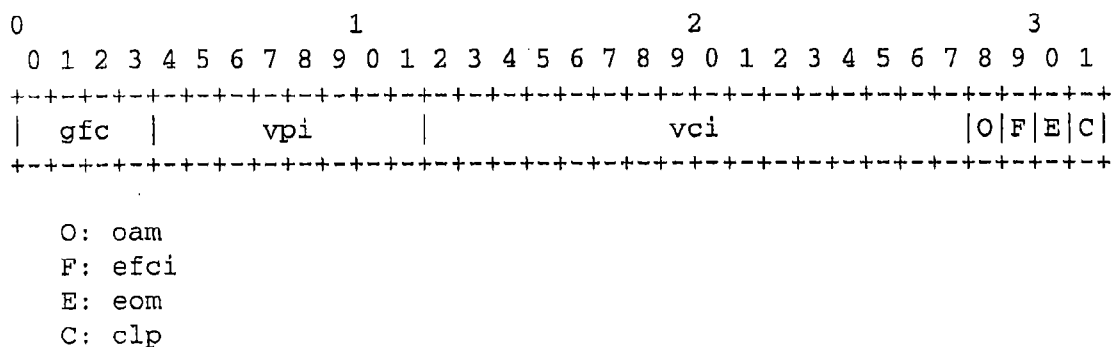


图 10

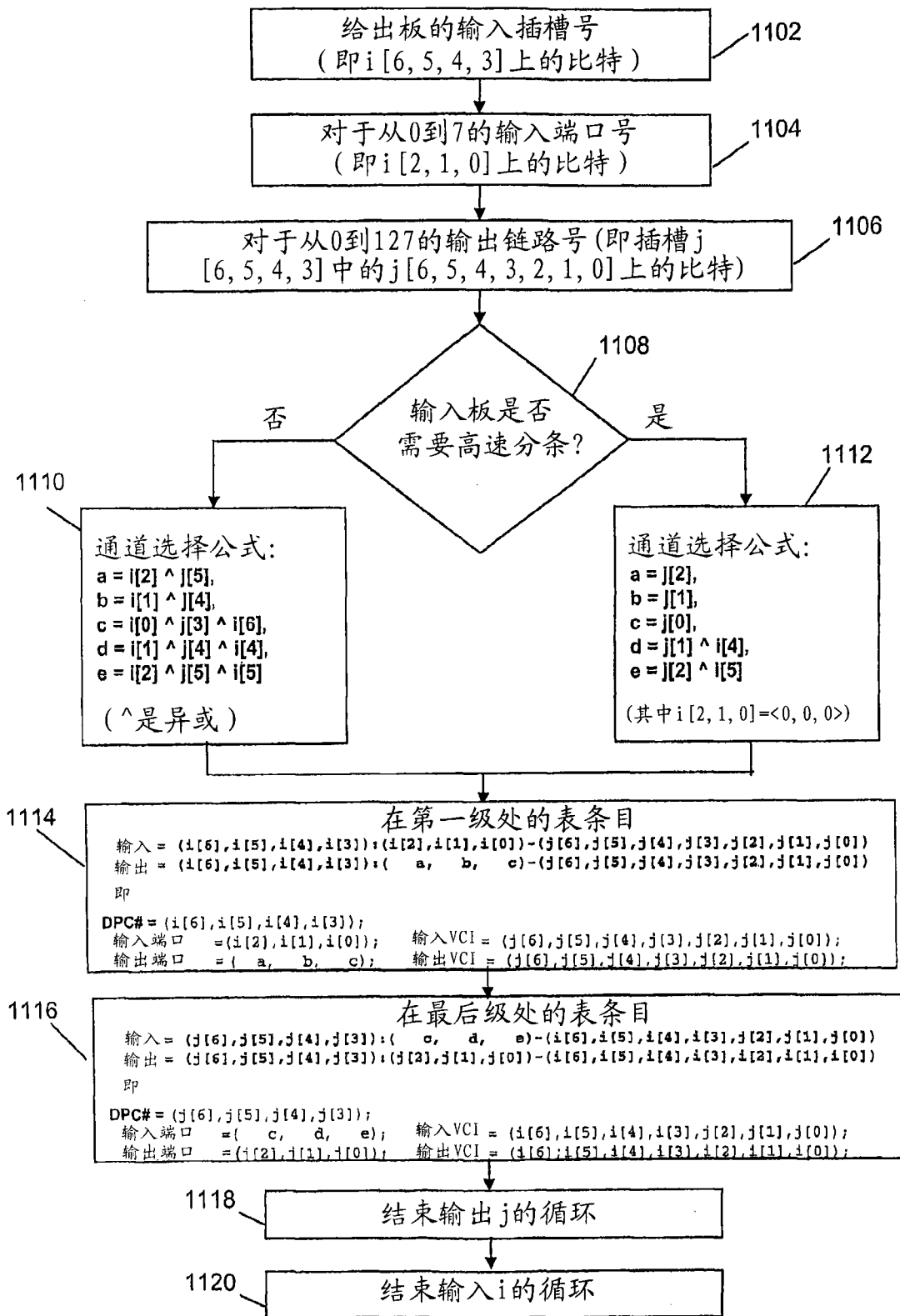


图 11

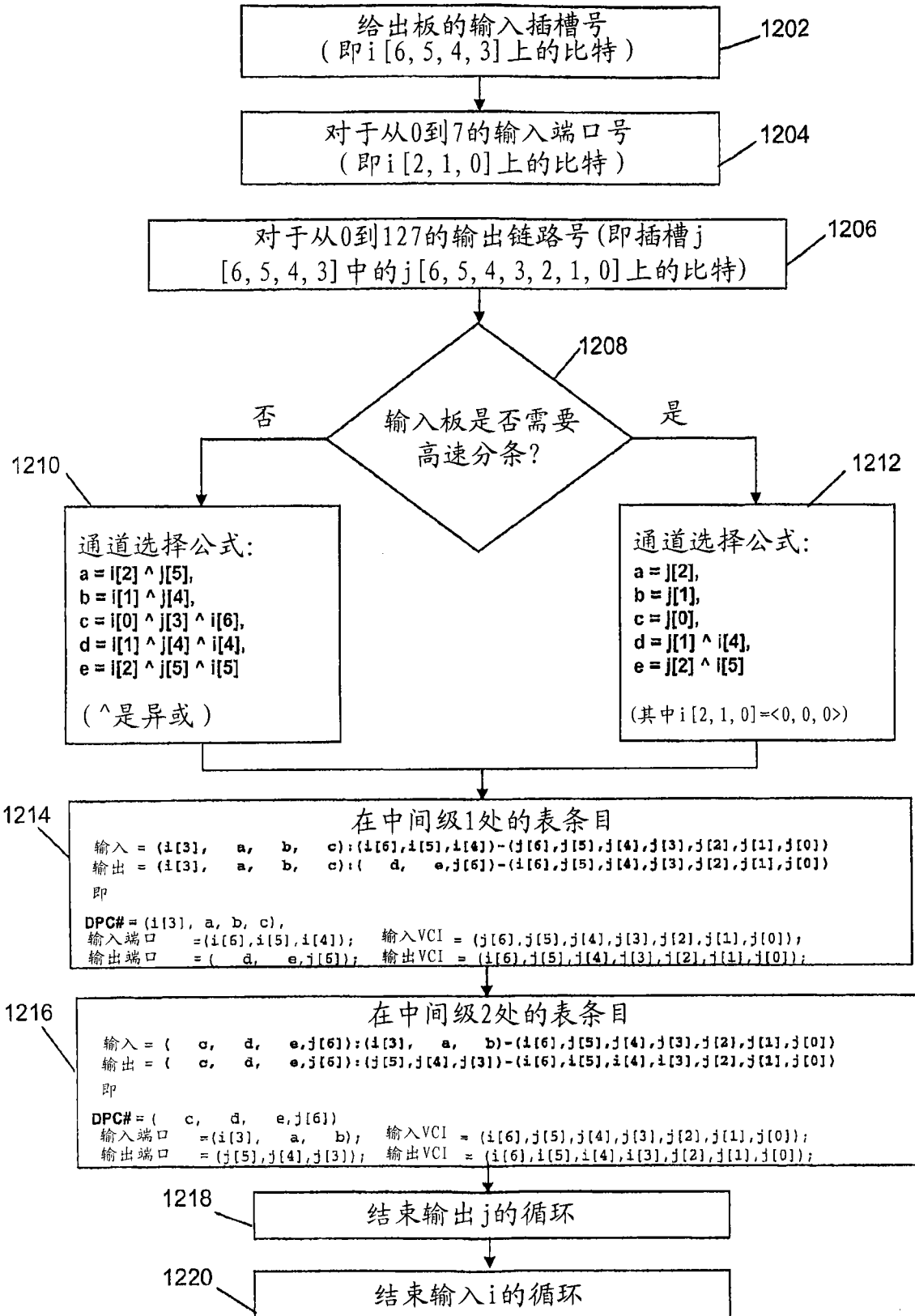


图 12

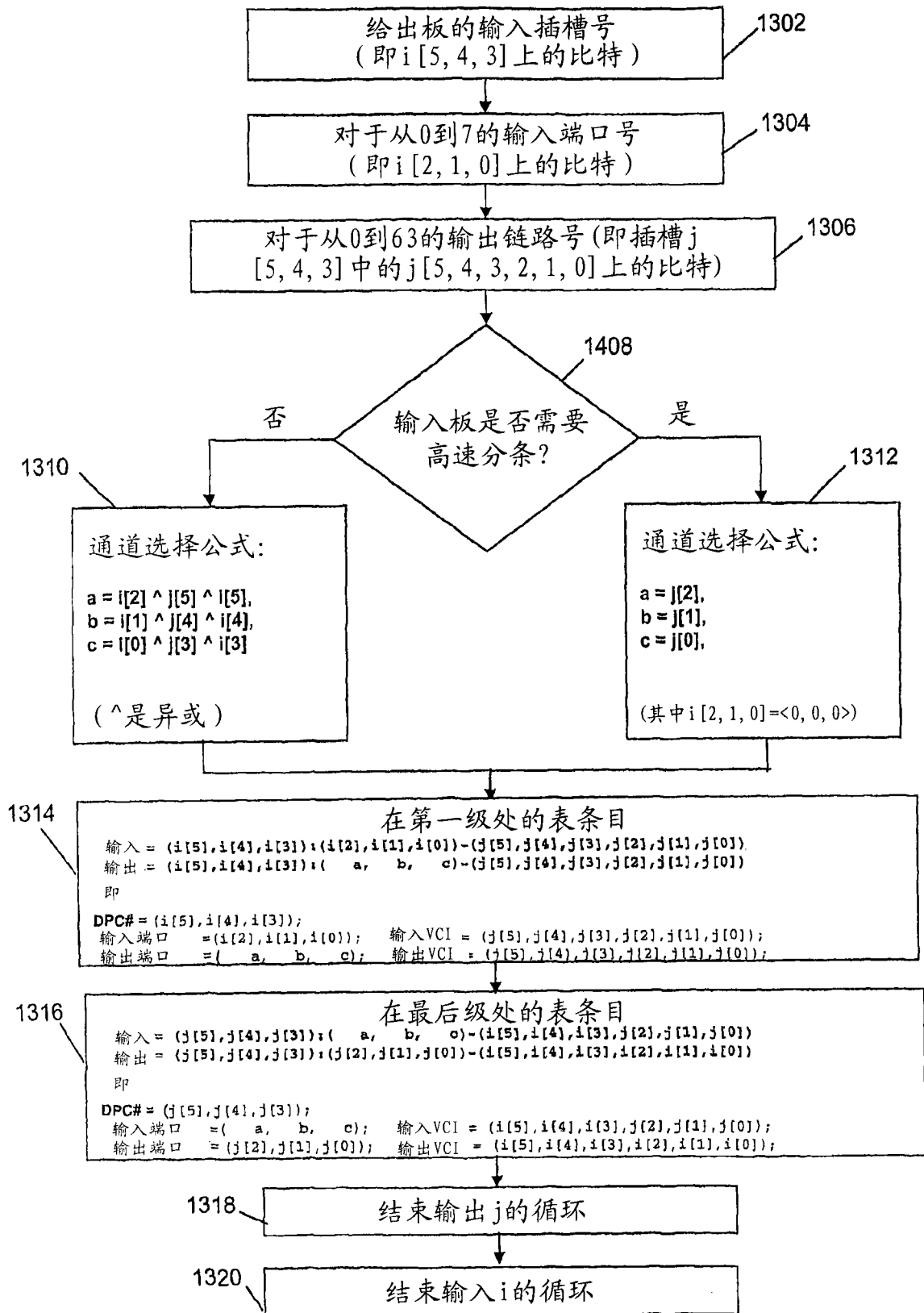


图 13

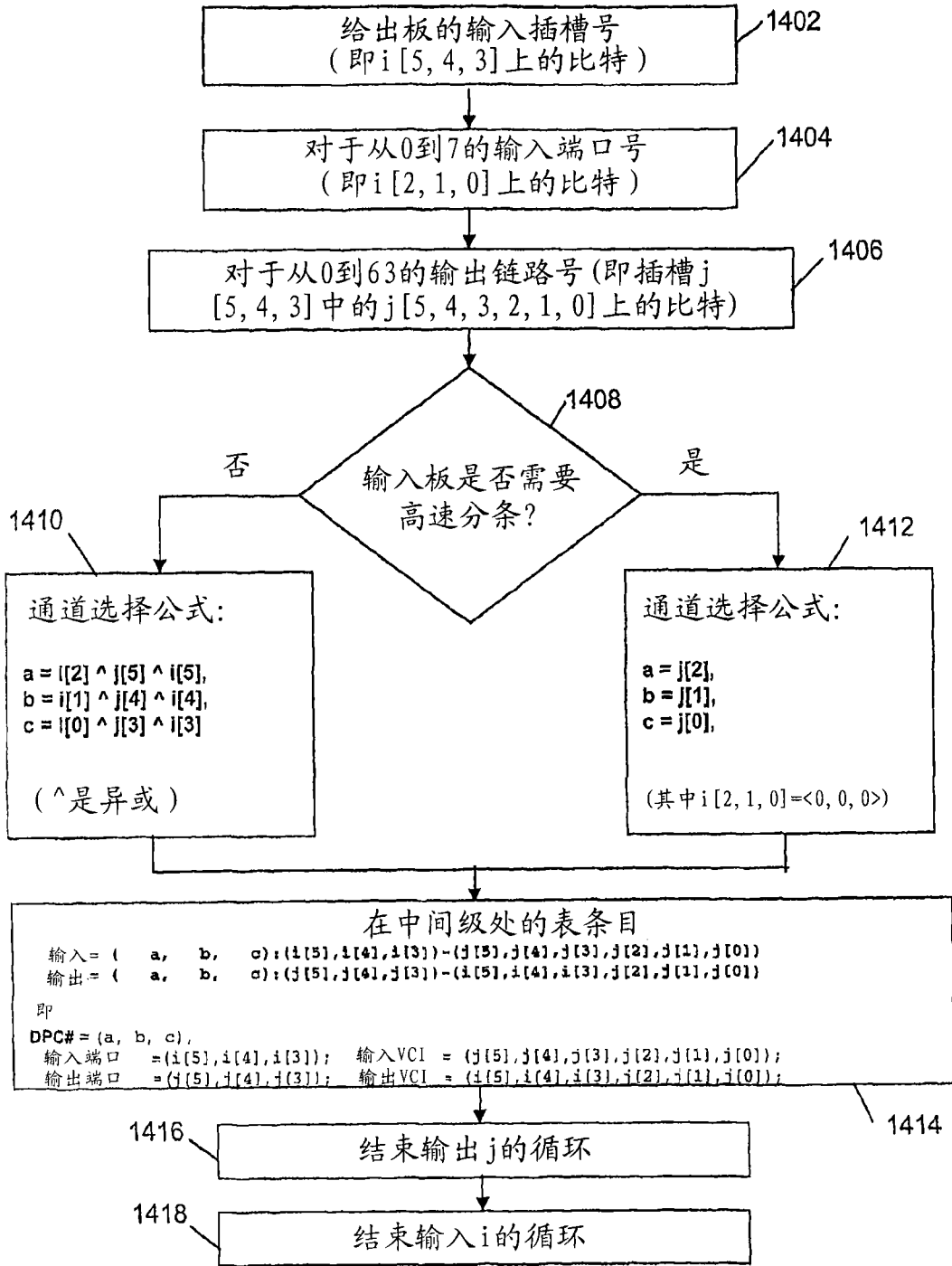


图 14