

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

H04L 12/28

H04Q 3/64

[12] 发明专利申请公开说明书

[21] 申请号 01104900.6

[43] 公开日 2001 年 9 月 12 日

[11] 公开号 CN 1312631A

[22] 申请日 2001.2.28 [21] 申请号 01104900.6

[30] 优先权

[32] 2000.3.2 [33] US [31] 09/516,859

[71] 申请人 阿尔卡塔尔互联网运行公司

地址 美国加利福尼亚州

[72] 发明人 德鲁·伯塔格纳

[74] 专利代理机构 中国国际贸易促进委员会专利商标事
务所

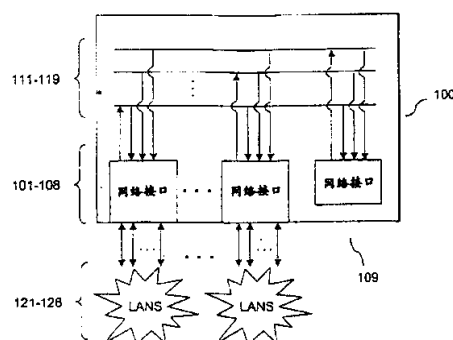
代理人 马 浩

权利要求书 2 页 说明书 12 页 附图页数 9 页

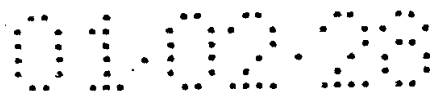
[54] 发明名称 数据通信交换机的优先权重新映射

[57] 摘要

一种用于数据通信交换机,例如一个支持信源学习桥接的 LAN 交换机的优先权处理,其中根据一个入口标签优先权和多个包含一个入口标签 VLAN 和接收端口标识的其他数值确定一个被应用于一个出口加标签分组的优先权数值。多个其他数值最初可以被用来确定一个虚拟中继线标识,虚拟中继线标识可以和入口标签优先权配合使用以确定出口标签优先权。通过把多个其他数值减少到一个较小位值并且在一个表查找中使用较小位值可以确定虚拟中继线标识。



ISSN 1008-4274



权 利 要 求 书

1. 一种在具有多个端口的数据通信交换机中重新映射分组优先权的方法，其中包括：

在一个第一端口上接收含有一个第一优先权的分组；

根据多个数值确定一个虚拟中继线数值；

根据第一优先权数值和虚拟中继线数值确定一个第二优先权数值；

在一个第二端口上发送含有第二优先权数值的分组。

2. 如权利要求 1 所述的方法，其中多个数值包含第一端口的一个标识。

3. 如权利要求 1 所述的方法，其中多个数值包含一个 VLAN 标识。

4. 如权利要求 3 所述的方法，其中 VLAN 标识被包含在所接收的分组中。

5. 如权利要求 1 所述的方法，其中第一优先权数值被包含在所接收的分组中。

6. 如权利要求 1 所述的方法，其中第二优先权数值被包含在所发送的分组中。

7. 如权利要求 1 所述的方法，其中确定虚拟中继线数值的步骤包含将多个数值缩减成一个位数较少的数值以及在一个表格查找中使用位数较少的数值。

8. 一种在具有多个端口的数据通信交换机中重新映射分组优先权的方法，其中包括：

在一个第一端口上接收含有一个第一优先权的分组；

根据第一优先权数值和多个含有第一端口标识的其它数值确定一个第二优先权数值；

在一个第二端口上发送含有第二优先权数值的分组。

9. 如权利要求 8 所述的方法，其中多个其它数值还包含一个 VLAN

标识。

10. 如权利要求 9 所述的方法，其中 VLAN 标识被包含在所接收的分组中。

11. 如权利要求 8 所述的方法，其中第一优先权数值被包含在所接收的分组中。

12. 如权利要求 8 所述的方法，其中第二优先权数值被包含在所发送的分组中。

13. 一种数据通信交换机的网络接口，其中包括：

一个访问控制器，该控制器具有一个接收含有一个第一优先权数值的分组的端口；

一个被连接到访问控制器的交换引擎，用于从访问控制器接收分组，根据该分组向一个第一单元发送多个数值，根据多个数值从第一单元接收一个虚拟中继线标识，向一个第二单元发送虚拟中继线标识和第一优先权数值，根据虚拟中继线标识和第一优先权数值从第二单元接收一个第二优先权数值，发送含有第二优先权数值的分组。

14. 如权利要求 13 所述的网络接口，其中多个数值包含一个端口标识。

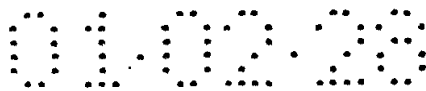
15. 如权利要求 13 所述的网络接口，其中在访问控制器上接收的分组包含一个 VLAN 标识。

16. 如权利要求 15 所述的网络接口，其中多个数值包含 VLAN 标识。

17. 一种用于数据通信交换机的网络接口，其中包括：

一个访问控制器，该控制器具有一个接收含有一个第一优先权数值和一个 VLAN 标识的分组的端口；

一个被连接到访问控制器的交换引擎，用于从访问控制器接收分组，查询多个数据库以便根据含有一个端口标识，VLAN 标识和第一优先权数值的多个数值得到一个第二优先权数值，发送含有第二优先权数值的分组。



说明书

数据通信交换机的优先权重新映射

电气电子工程师协会公布的标题为 "虚拟桥接局域网" 的标准 802.1Q 定义了一个专门支持桥接局域网中业务优先权划分的协议。该标准包含一个在创建一个符合 Std.802.1Q 的分组时在一个分组中产生一个标签的协议。该标签可以专门包含一个虚拟 LAN (VLAN) 标识和一个优先权。通常, VLAN 标识被用来确定分组可以被发送到哪里, 而优先权被用来确定相对于其他分组而言处理分组会有多快。

当一个符合 Std.802.1Q 的分组通过一个网络中的各个桥接器时, 该协议规定在所接收的分组中 "重新产生" 标签优先权 (即入口标签优先权), 并且在所发送的分组中标示重新产生的优先权 (即出口标签优先权)。根据该标准, 出口标签优先权可以不同于入口标签优先权并且根据物理端口可以在各个桥接器上独立确定出口标签优先权, 其中通过上述物理端口接收分组和入口标签优先权。然而在确定出口标签优先权时在该标准中没有规定使用其他的数值。可能期望根据不同于入口标签优先权和接收物理端口的数值来确定一个出口标签优先权。

本发明为一个数据通信交换机, 例如一个支持信源学习 (source - learned) 桥接的局域网交换机提供优先权处理, 其中根据一个来自分组的优先权和多个其他数值为一个分组中的标示确定一个优先权。多个其他数值可以包含来自分组的一个接收端口标识和一个 VLAN 标识。该分组可以是一个格式符合 Std. 802.1Q 的 "加标签" 分组。多个其他数值最初可以被用来确定一个虚拟中继线标识, 虚拟中继线标识可以和来自分组的优先权配合使用以确定分组中的标示的优先权。通过把多个其他数值减少到一个较小位值并且在一个表查找中使用较小位值可以确定虚拟中继线标识。

参照下列结合如下简述的附图所进行的详细说明可以是更好地理解本发明的这些及其他方面。

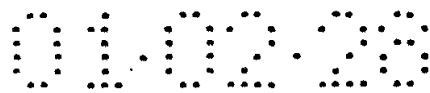


图1图解了一个局域网交换机;

图2图解了图1局域网交换机内部的一个典型网络接口;

图3图解了一个在图2的网络接口内部的接入控制器上从一个局域网接收的分组;

图4图解了一个在图2的网络接口内部的交换引擎上从图2的接入控制器接收的分组;

图5图解了图2的网络接口内部的虚拟中继线搜寻器;

图6图解了图5的虚拟中继线搜寻器内部的多路复用器排列;

图7图解了图5的虚拟中继线搜寻器内部的虚拟中继线散列RAM;

图8图解了图2的网络接口内部的优先权重新映射数据库;

图9图解了图2的网络接口内部的传送数据库;

图10图解了一个在图2的网络接口内部的交换引擎上一个底板总线接收的一个分组的局部报头;

图11图解了图2的网络接口内部的队列重新映射数据库;

图12图解了一个在图2的网络接口内部的接入控制器上从图2的交换引擎接收的分组;

图13图解了一个由图2的网络接口内部的接入控制器发送到一个局域网的分组;

图14是描述基于本发明一个最优实施例的入口优先权处理的一个流程图; 和

图15是描述基于本发明一个最优实施例的出口优先权处理的一个流程图;

在图1中, 示出了一个实施本发明的局域网交换机100。交换机100包含一个由分别被接口101 - 109驱动的分组总线111 - 119构成的矩阵。接口101 - 109包含均与一或多个局域网121 - 128, 和管理接口109关联的网络接口101 - 108。各个总线有一个根和若干个叶, 根与接口101 - 109中的一个在总线上有特权发送分组数据的接口 (即根接口) 接口, 而叶与多个从总线上接收分组数据的接口101 - 109接口 (即叶接口) 接

口。各个接口最好是一个总线111 - 119上的根接口并且最好是全部总线111 - 119上的一个叶接口，其中包含是其根接口的总线。最好以每时钟周期一个脉冲串的速率按照固定顺序在总线111 - 119上发送分组。总线111 - 119是面向广播的，因此在一个到达所有接口101 - 109的总线上传播所有数据脉冲串。除发送和接收分组数据之外，管理接口109充当交换机100的"神经中枢"，其中交换机100通过在管理总线上(未示出)发送信息协助网络接口101 - 108学习其相关的局域网121 - 128上的网络设备地址。当然，上述根-到-叶体系结构是许多针对一个根据本发明进行操作的可能体系结构中的一种。其他可能的体系结构在接口之间可以有一个单独的公共总线或"全互联"点到点连接矩阵。

在一个基本操作中，交换机100支持一个信源学习桥接功能。举例来说，在驻留于和网络接口101有关的一个局域网121之上的一个网络设备(未示出)上产生一个分组。该分组包含一个发端网络设备源地址，并且在该分组是一个单路传送分组的情况下，包含一个准备与其通信的网络设备目的地址。考虑使用诸如介质访问控制(MAC)地址的层2(数据链路)地址。由于局域网的广播性质，该分组到达接口101。如果在接口101上不能识别该分组的源地址，则该分组被送交管理控制器109以进行一个学习操作，该操作导致在接口101上"学习"该地址，即该地址被加到一个在接口101上活跃的地址列表中。其后，接口101把任何从分组总线101 - 109上接收的，将经过学习的地址作为目的地址的分组识别成定向到一个局域网121上的网络设备的分组，并且捕捉这样的分组以便转发。

由于某些例外导致不将接口101学习的地址当作目的地址的分组被接口101丢弃或"过滤"，对从分组总线101 - 109上接收的分组进行的目的地址检查通常被称作"过滤"检查。接口101 - 109单独在各个从总线111 - 119上接收的分组进行过滤检查。通常根据分组是否包含一个此前被接口按上述方式学习的目的地址来判断是否转发或过滤一个分组。但是，接口101 - 109共享这种判断结果以避免过滤其目的地址尚未被任何接口学习的分组。这样的"目的地址未知"分组被所有接口捕

捉。更具体地，在一个示例性过滤检查中，一个接口应用以下过滤规则：

- 1.如果分组有一个此前被接口学习过的目的地址，则过滤检查通过。该分组被捕捉。
- 2.如果分组有一个此前未被接口学习的目的地址，并且该目的地址此前被另一个接口学习过，该过滤检查失败。该分组被过滤。
- 3.如果分组有一个此前未被接口学习的目的地址，并且该目的地址此前未被另一个接口学习过，则过滤检查通过。该分组被捕捉。

接口101 - 109最好提出线路要求(assert claim lines)（未示出）以便彼此通知有关捕捉分组的判断。

除了支持刚描述的信源学习桥接功能之外，交换机100还支持一个本发明最初提到的优先权处理。优先权处理被考虑用于具有与其相关的标签优先权的分组，比如符合Std. 802.1Q的以太网分组。

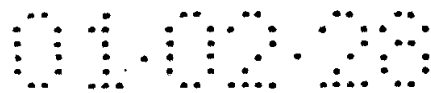
回到图2，现在参照网络接口200描述一个最优优先权处理，其中网络接口200代表网络接口101 - 108。接口200包含被连接到局域网和交换引擎211的访问控制器201。控制器201从局域网接收分组，格式化分组并且发送分组到引擎211。控制器201也从引擎211接收分组，格式化分组并且在局域网上发送分组。引擎211被连接到实现优先权处理的单元，其中包括虚拟中继线搜寻器221，优先权重新映射数据库231、内容-可寻址存储器（CAM）241、转发数据库251和队列重新映射数据库261。具体地，引擎211从控制器201接收分组，提交分组到入口优先权处理并且在接口200是其根的一个总线111 - 119上发送分组。引擎211还从总线111 - 119接收分组，提交其中选定的一个分组到出口优先权处理并且向控制器201发送其中选定的一个分组。入口优先权处理由虚拟中继线搜寻器221，优先权重新映射数据库231，CAM 241和转发数据库251协助实施，而出口优先权处理由CAM 241，转发数据库251和队列重新映射数据库261协助实施。

在图3至16中，参照具有在接口200上从一个局域网接收的格式的符合Std.802.1Q的以太网分组（此后称作“加标签”分组）更详细地描

述在交换机100中支持的优先权处理。首先参照图3,入口分组300包含一个分别跟随有一个源MAC地址(SA0-SA5),一个标签(TAG0-TAG3)和类型-长度信息(TL0和TL1)的目的MAC地址(DA0-DA5)。类型-长度信息跟随有附加信息,附加信息通常包含一个网际协议(IP)报头,而报头的前4个字节被图示成D0-D3。

在访问控制器201上,入口分组300被标识成一个加标签分组并且被放在一个入口处理就绪格式中并且被发送到交换引擎211。参照标签的一部分(TAG0和TAG1)将分组识别成一个加标签分组,正如IEEE 802.1Q标准中所定义的,如果分组是一个加标签分组,则标签值为x8100。在图4中入口就绪分组入口被图示成具有两字节宽度以反映在图解的实施例,通过两字节脉冲串把分组400从控制器201发送到引擎211。但是,可以理解在其它实施例中可以按照不同宽度进行传输。分组400包含一个物理端口标识(PORT),这个标识指明借以接收入口分组300的物理端口,并且后面跟随有控制信息(CTRL),控制信息把分组400标识成一个加标签分组。CTRL后面跟随有一部分标签(TAG2和TAG3),这部分具有如Std.802.1Q所定义的标签控制信息。更具体地,TCI包含一个3位入口标签优先权,一个一位规范格式指示符和一个12位VLAN标识。象在入口分组300那样格式化分组400的其余部分。入口分组中未加标签部分没有携带任何内容。CTRL指示在对应的入口就绪分组中的这个情况并且没有TCI被附加到这样的分组上。

在交换引擎211上,对分组400进行入口优先权处理以便准备出在总线上传输的分组400,其中该总线的接口200是根。引擎211参照CTRL把分组400识别成一个加标签分组。一旦被识别出来,引擎211从分组400中去除PORT,CTRL和TAG2与TAG(其中包含12位VLAN标识)。PORT被重新映射成交换机100上唯一的一个8位虚拟端口标识。在这点上,可以理解两个或更多的网络接口101-108可以具有一个由相同物理端口标识表示的物理端口,并且把分组400中的PORT重新映射成一个VPI很好地消除了任何潜在的模糊性。VPI



和 12 位 VLAN 标识被成对分别提交给输入线路 511 和 512 上的虚拟中继线搜寻器 221。

在虚拟中继线搜寻器 221 上，20 位 VPI/VLAN 标识对被缩减成一个用于解析一个虚拟中继线标识(VTI)的 10 位散列键值 (key)。可以理解，这种缩减减少了实现象这里描述的解析 VTI 那样解析一个标识的一个基于索引的查找方案所需的 RAM 尺寸。回到图 5 和 6，在多路复用器阵列 510 中接收引擎 211 发送的 VPI/VLAN 对。阵列 510 是一个交错多路复用器 611-620 阵列，该阵列被用来把 VPI/VLAN 对缩减成一个散列键值，该键值包含从 VPI/VLAN 对中的位位置上得到的位，其中一个散列算法已经确定上述位位置是最能彼此有效区分不同 VPI/VLAN 对的位置。VPI/VLAN 对被分散到不同子集内以便被多路复用器 611-620 分别在输入线路 601-610 上接收。不同输入线路 601-610 发送的子集被彼此交错以便聚集中的多路复用器 611-620 可以根据提供给阵列 510 的位选择命令从 VPI/VLAN 对中的任意 10 位组合中选择出散列键值。这样，举例来说，输入线路 601 可以发送 VPI/VLAN 对的位 0 到位 12，输入线路 602 可以发送位 1 到位 13，输入线路 603 可以发送位 2 到位 14，依此类推。在多路复用器控制 520 的协助下，各个多路复用器 611-620 从其相关的一个输入线路 601-610 中选择一个单独的位，并且只发送选择的位。可选地，多路复用器控制 520 可以使一或多个多路复用器 611-620 忽略其相关的一个输入线路 601-610 上的所有位，并且可以自动选择一个 0 值。在输出线路 631-640 上发送从 VPI/VLAN 对中选择位和任意 0 值，并且共同构成散列键值。

多路复用器控制 520 控制多路复用器位选择。控制 520 包含一个存储散列屏蔽标志的存储器单元和相关逻辑。最好在输入线路 521 上对散列屏蔽标志进行编程和更新。通过一个散列算法计算散列屏蔽标志的值以便阵列 510 接收的任意 VPI/VLAN 对可以被缩减成一个散列键值，并且该键值包含从散列算法已经确定最能有效彼此区分 VPI/VLAN 对的位位置上得到的位。为了实现选择，控制 520 通过屏蔽线路 631-640 被连接到多路复用器 611-620。控制 520 根据散列屏蔽

标志确定各个多路复用器 611-620 的位选择命令，并且在屏蔽线路 631-640 上发送位选择命令。每个位选择命令完全能够识别 VPI/VLAN 对的一个单独的位，并且如果有这样的位的话，各个多路复用器 611-620 会选择把这个位包含在散列键值中。这样，多路复用器 611 在线路 631 上接收的位选择命令可以指引多路复用器 611 选择 VPI/VLAN 对位 0 至位 12 中的一个位包含在散列键值中；多路复用器 612 在线路 632 上接收的位选择命令可以指引多路复用器 612 选择 VPI/VLAN 对位 1 至位 13 中的一个位包含在散列键值中；依此类推。散列键值和一个在输入线路 531 上发送的偏移一起被发送到虚拟中继线匹配控制 530，它们共同构成一个指向虚拟中继线散列随机访问存储器(RAM)540 的指针。这里，散列 RAM 540 包含两个表格(表 1 和表 2)，根据偏移在任意指定时间选择其中的一个表格。

在虚拟中继线匹配控制 530 中还接收 20 位 VPI/VLAN 对以用作一个比较数，该比较数以一个现在要更详细描述的方式与散列 RAM 540 返回的一或多个 VPI/VLAN 对进行相关比较。参照图 7，匹配控制 530 与散列 RAM 540 接口以便使用根据散列键值和偏移构成的指针进行相关比较。指针被用作散列 RAM 540 的初始指针以便检索与索引相关的入口的内容，其中该入口的值与指针匹配。指针对于促成遍历表格内的一个入口链表有重要的作用，这种遍历连续进行直到找到一个 VPI/VLAN 对匹配或到达链表的末端。更具体地，表格 700 包含具有对应索引的入口内容。对于各个入口，入口内容包含一个 VPI/VLAN 对，一个 VTI 和一个“下一入口”键值。在图解的例子中，表格 700 包含构成 N 个链表或“桶子”的 N 个入口子集。第一个桶子包含索引 701 和 709。第二个桶子包含索引 702，705，707 和 710。第三个桶子包含索引 703 和 708。第 N 个桶子包含索引 704 和 706。如果一个指针指向索引 702 并且在匹配控制 530 中来自对应入口的 VPI/VLAN 对不与 VPI/VLAN 对匹配，则来自入口的“下一入口”键值(指明索引 705)被用作一个指向索引 705 的指针。如果从对应于索引 705 的入口返回的对不匹配，则来自入口的“下一入口”键值(指明索

引 907)被用作一个指向索引集合 707 的指针。匹配控制 530 连续遍历散列 RAM 120 直到找到一个匹配。匹配控制 530 向匹配控制 530 返回与匹配入口相关的 VTI 以便通过输出线路 531 传送到交换引擎 211。

虚拟中继线匹配控制 530 还跟踪散列算法的性能并且在性能足够差时通过输出线路 532 通知一个外部处理器(未示出)。具体地,控制 530 每当从散列 RAM 540 返回的 VPI/VLAN 对与比较数匹配失败时递增一个存储器内的一个数值。当性能降低低于一个最小性能标准时,控制 530 通过输出线路 532 向处理器发送一个失败通知,使处理器重新计算散列算法并且通过输入线路 521 更新散列屏蔽标志。根据诸如失败尝试总数,具体遍历的失败尝试最高数量,单遍历失败尝试平均数量或失败尝试数量越过标准的频率的各种考虑可以实现各种最小性能标准。最小性能标准在匹配控制 530 上是可以配置的。可以理解,当散列屏蔽标志改变时,散列 RAM 540 中的入口必须被重新写到新的索引位置上。通过在进行其它重写时使用一个用于查找的表格并且保证查找优先权高于因访问散列 RAM 540 而进行的重写,可以在最小影响性能的情况下实现入口重写。

现在参照图 8,这里更详细地描述优先权重新映射数据库 231。数据库 231 根据入口标签优先权和被解析的 VTI 为入口就绪分组 400 确定一个入口标签优先权。具体地,交换引擎 211 向数据库 231 提交分组 400 的入口标签优先权和从虚拟中继线搜寻器 221 接收的 VTI。入口标签优先权和 VTI 被用作一个指向数据库 231 中的一个对应索引,返回一个出口标签优先权的指针。

与优先权重新映射数据库 231 中进行的出口标签优先权确定相独立的是,交换引擎 211 查询 CAM 241 和传送数据库 251 对分组 400 进行优先权选择。回到图 2, CAM 241 在不同 CAM 索引上具有保留驻留在与接口 200 相关的 LAN 上的网络设备的学习地址的入口。传送数据库 251 维护通过一个公共索引被连接到 CAM 241 中的这些入口的入口。分组 400 中的源地址被提交到 CAM 241, CAM 241 返回驻留有源地址的 CAM 索引(此后称作源 CAM 索引或 SCI)。参照图 9,

SCI 被用作一个指向数据库 251 中的链接入口的指针以便检索包含分组 400 的一个优先权选择指示符的传送数据，该数据被返回到引擎 211。正如下面所要详细描述，优先权选择指示符确定分组 400 在交换机 100 中是否会得到与其标签优先权相符的服务质量。所示的数据库 251 包含具有具体索引的优先权选择指示符和局部优先权队列标识。在当前讨论中局部优先权队列标识不涉及入口优先权处理，但正如下面所讨论的，最好在基于出口优先权的处理中实现。应当注意，如果仍未学习分组 400 中的源地址，则没有从 CAM 241 返回有效的 SCI，并且可以代替使用从虚拟中继线搜寻器 221 返回的 VTI 来索引另一个数据库(未示出)以便对分组 400 进行优先权选择。

交换引擎 211 把一个局部分组报头追加到入口就绪分组 400(较小的 PORT, CTRL 和 TAG2 与 TAG3, 在前面被去掉了)以便把分组转到出口就绪格式。参照图 10, 其中示出了出口就绪分组的局部分组报头 1000。报头 1000 包含 SCI, 或者在 CAM 241 没有返回有效的 SCI 的情况下包含 VTI 和一个足以把 VTI 标识成从接口 200 发出的接口标识。这里可以理解两个或更多的网络接口 101-108 可以具有一个由相同 VTI 表示的 VPI/VLAN 对, 并且附加位较好地消除了任何潜在的模糊性。报头 1000 还包含一个表明是否返回一个有效的 SCI 的“无效 SCI”指示符。具体地, 如果“无效 SCI”指示符置位, 则出口就绪分组会被管理接口 109 捕捉并且经过“信源学习”, 使得源地址被加到 CAM 241 中。报头 1000 还包含从优先权重新映射数据库 231 返回的出口标签优先权, 从传送数据库 251(或者在 CAM 241 未返回一个有效 SCI 的情况下从 VTI 查找)返回的优先权选择指示符, 以及报头长度信息, 一个目的地址和分组控制信息。通过接口 200 是其根的总线发送出口就绪分组, 从而完成了基于入口优先权的处理。

至此, 已经描述了一个入口优先权处理, 其中根据一个入口标签优先权和多个其它数值确定一个分组的一个出口标签优先权并且将该优先权应用于分组; 确定一个分组的一个优先权选择指示符并且应用于该分组, 该指示符确定是否会在交换机中根据其标签优先权为分组

划分优先权。还会有其它和进一步的越过这里描述的常规优先权处理方案的优点；不管怎样，至此描述的入口优先权处理被认为代表了超越现有技术的显著进步。

现在参照代表性的网络接口 200 描述对出口就绪分组的出口优先权处理。为了避免不必要地增加复杂度，假定接口 200 是与驻留有出口就绪分组要到达的网络设备的 LAN 相关的网络接口 101-108 中的一个接口，并且假定授权发出分组的网络设备与分组要到达的网络设备通信。因而不会更详细地描述验证这些假定的出口处理方面，例如使用报头 1000 的 SCI 分量进行的授权检查。交换引擎 211 从分组中去除报头 1000，查询 CAM 241 和传送数据库 251 以确定是否识别报头 1000 中的目的地址，并且为其确定一个局部优先权队列标识(LPQID)。如上所述，CAM 241 在不同 CAM 索引上具有保留驻留在与接口 200 相关的 LAN 上的网络设备的学习地址的入口。由于在所考虑的例子中识别了目的地址，CAM 241 返回目的地址驻留其上的 CAM 索引(此后称作目的 CAM 索引或 DCI)。回到图 9，DCI 被用作一个指向传送数据库 251 中检索包含分组的 LPQID 的传送数据的链接入口的指针，其中 LPQID 被返回到引擎 211。由于交换机 100 是在没有参考任何标签优先权的情况下确定出 LPQID，所以 LPQID 是“局部的”。因而 LPQID 提供了一种划分出口就绪分组优先权的能力，尤其是在相对于其它分组计划从接口 200 释放该分组的时候，LPQID 不依赖其标签优先权。根据标签优先权确定是否划分优先权，或者参照优先权选择指示符以不依赖标签优先权的方式进行确定，现在更详细地加以描述。

交换引擎 211 从报头 1000 查询优先权选择指示符以确定是否应当对出口就绪分组进行基于标签的优先权划分。如果优先权选择指示符指示不应当对分组进行基于标签的优先权划分，则根据一个基于优先权的调度算法在 LPQID 指定的优先权队列中对分组(重新格式化)进行排队以便释放到访问控制器 201。但如果优先权选择指示符指示应当对分组进行基于标签的优先权划分，则引擎 211 接入队列重新映射数据库 261 以确定一个标签优先权队列标识(TPQID)。现在参照图 11 更

详细地图解数据库 261。数据库 261 根据出口标签优先权和 LPQID 确定分组的 TPQID。具体地，引擎 211 向队列重新映射数据库 261 提交来自报头 1000 的出口标签优先权和从传送数据库 251 得到的 LPQID。出口标签优先权和 LPQID 被用作一个指向数据库 261 中一个返回 TPQID 的对应索引的指针。TPQID 被返回到引擎 211 并且根据一个基于优先权的调度算法在 TPQID 指定的优先权队列中对分组(重新格式化)进行排队以便释放到访问控制器 201。

参照图 12，其中示出了引擎 211 向访问控制器 201 发送分组的格式。分组 1200 包含一个指定物理端口的物理端口标识(PORT)，其中通过该物理端口发送出口分组，该标识后面跟随有控制信息(CTRL)，控制信息把分组 1200 标识成一个加标签分组。CTRL 后面跟随有一部分具有包含出口标签优先权的 TCI 的标签(TAG2 和 TAG3)。分组 1200 还具有目的地址(DA0-DA5)，源地址(SA0-SA5)，类型长度信息(TL0-TL1)和其它信息。在访问控制器 201 上，根据 CTRL 把分组 1200 标识成一个加标签分组。一旦被标识成一个加标签分组，则控制器 201 从分组 1200 中去除 PORT，CTRL，TAG2 和 TAG3，并且在 IEEE 802.1Q 标准指示的位置上提供完整的标签(TAG0-TAG3)以产生出口分组 1300。控制器 201 通过 PORT 指示的物理端口把分组 1300 发送到一个 LAN 121。

在图 14 中参照流程图描述了入口优先权处理。在交换机 100 的一个物理端口上接收一个加标签分组(1410)，并且根据一个与物理端口相关的标识确定一个虚拟端口标识(1420)。根据虚拟端口标识和入口标签的 VLAN 标识分量确定一个虚拟中继线标识(1430)。虚拟中继线标识和入口标签的优先权分量被用来确定一个出口标签优先权(1440)，并且单独使用入口分组的一个源地址进行优先权选择(1450)。优先权选择指示符和出口标签优先权被应用于分组的一个局部报头(1460)并且发送分组(1470)。

在图 15 中参照流程图描述了出口优先权处理。接收一个经过入口处理的分组(1510)并且根据分组的一个目的地址确定一个局部优先权

队列标识(1520)。检查分组的优先权选择指示符以确定是否根据其加标签优先权对分组进行优先权划分(1530)。如果答案为否，则根据局部优先权队列标识把分组提供到一个优先权队列(1550)。如果答案为是，则根据局部优先权队列标识和出口标签优先权确定一个标签优先权队列标识(1540)并且根据标签优先权队列标识把分组提供到一个优先权队列(1550)。在两种情况下，一个出口标签被提供到分组(1560)，其中还包含出口标签优先权，并且从交换机 100 发送分组(1570)。

前面已经描述了一种优先权处理，其中在一个交换机上可以为一个加标签分组提供(i)一个独立于标签优先权的局部优先权划分或(ii)一个依赖标签优先权的加标签优先权划分。并且，已经描述了一种基于优先权的处理，其中不管是否在交换机上提供出口标签优先权以供优先权划分，均可以保留一个出口标签优先权以便通过一个出口加标签分组发送。这些特性被认为是代表了超越现有技术的显著进步。

本领域的技术人员可以理解，在不偏离本发明的宗旨或必要特征的前提下可以通过其它特定形式实现本发明。因而本描述在总体上被认为是说明性的和非限制性的。本发明的范围由所附权利要求书规定，在本发明描述的含义和范围内的所有修改均被包括其中。

说明书附图

图1

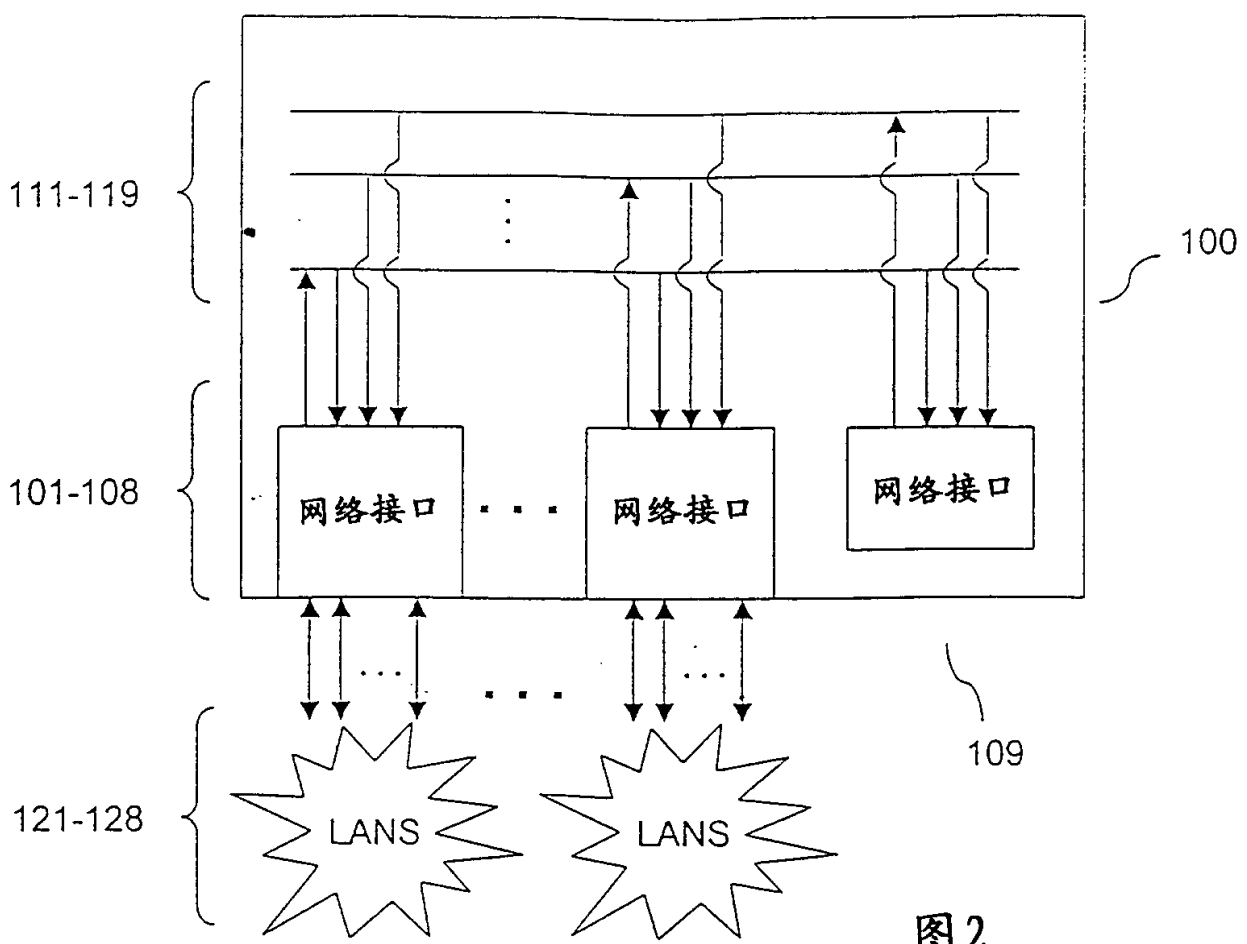


图2

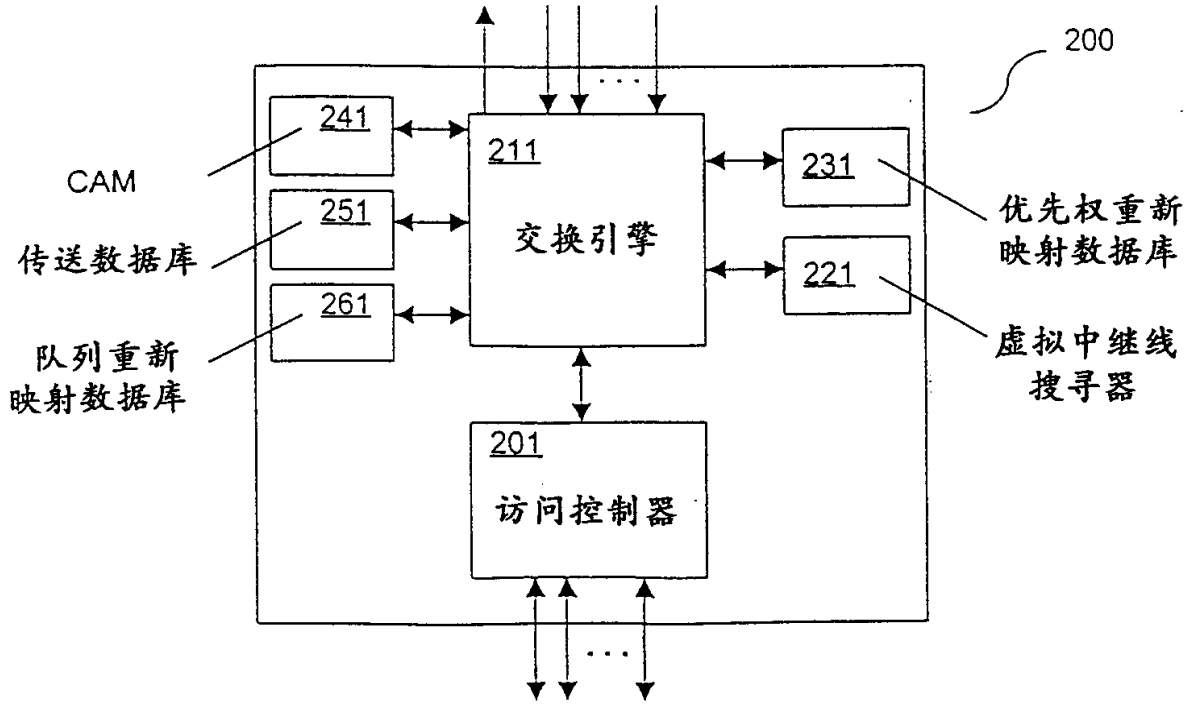


图 3

[7:0]	
0	DA0
1	DA1
2	DA2
3	DA3
4	DA4
5	DA5
6	SA0
7	SA1
8	SA2
9	SA3
10	SA4
11	SA5
12	TAG0
13	TAG1
14	TAG2
15	TAG3
16	TL0
17	TL1
18	D0
19	D1
20	D2
21	D3

300

图 4

	[15:8]	[7:0]
0	PORT	CTRL
1	TAG2	TAG3
2	DA0	DA1
3	DA2	DA3
4	DA4	DA5
5	SA0	SA1
6	SA2	SA3
7	SA4	SA5
8	TL0	TL1
9	D0	D1
10	D2	D3

400

图 5

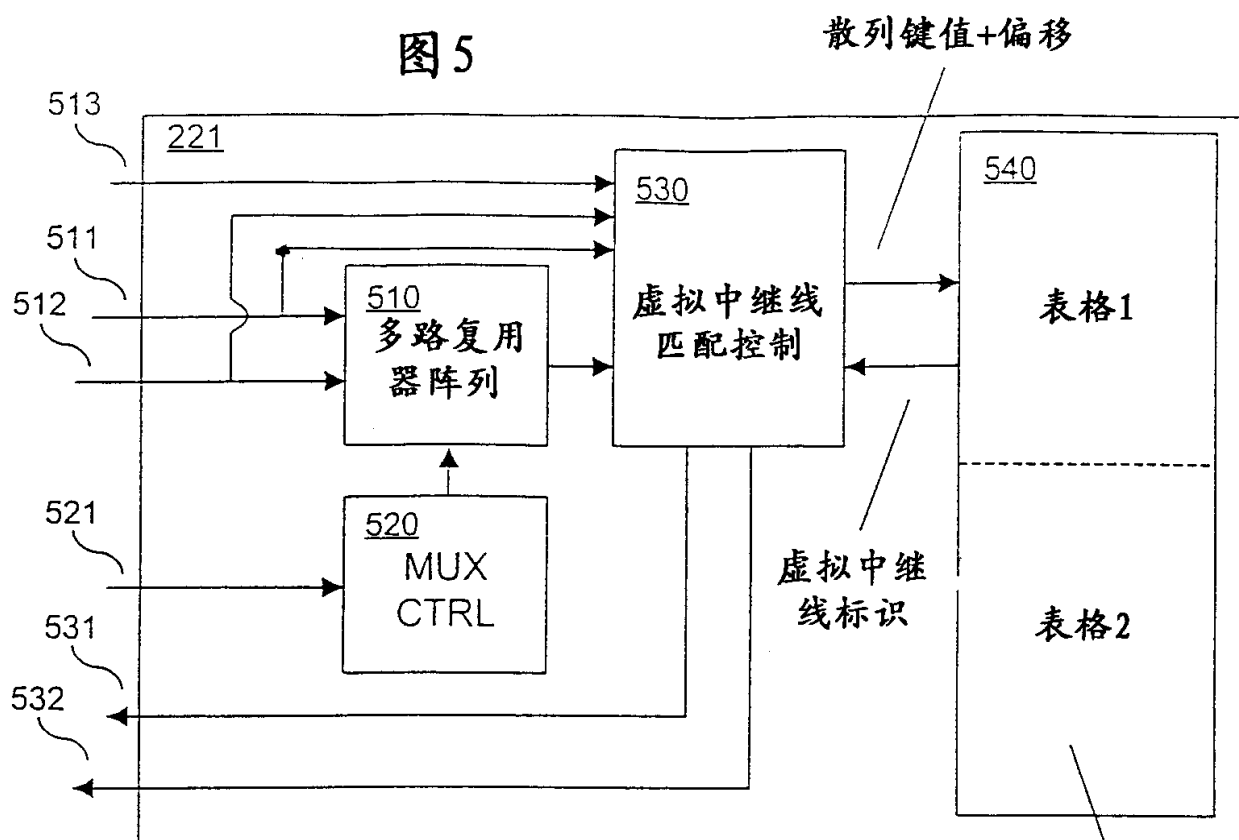


图 6

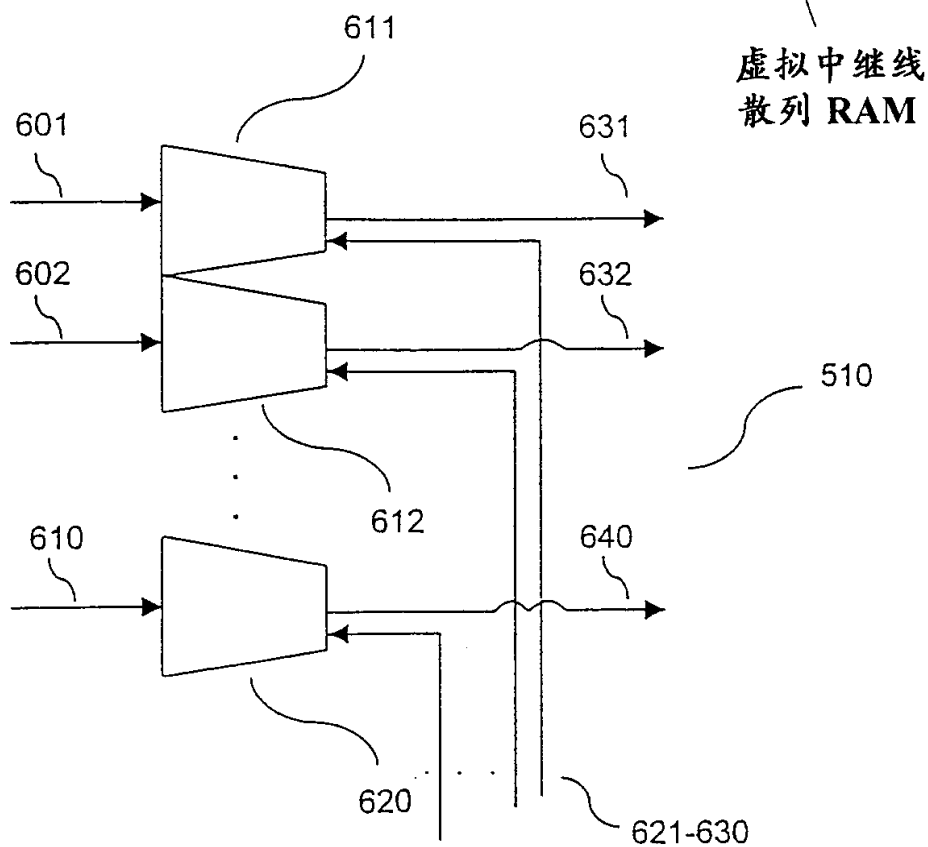
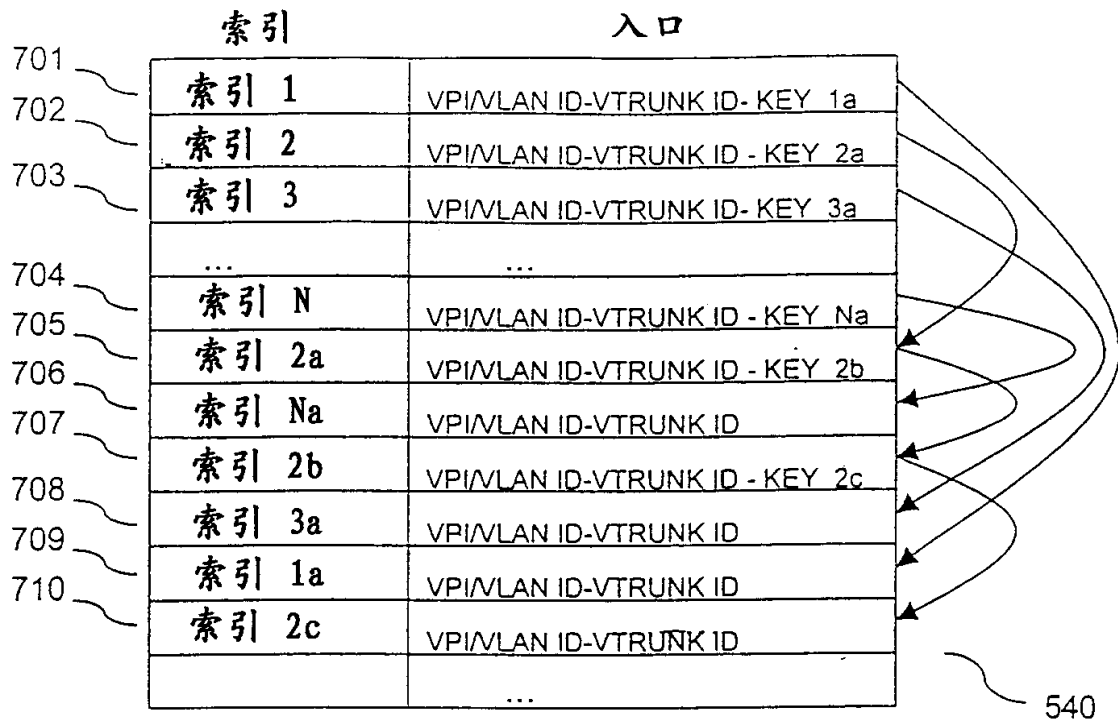


图7



入口标签优先权 - 虚拟中继线标识

图8

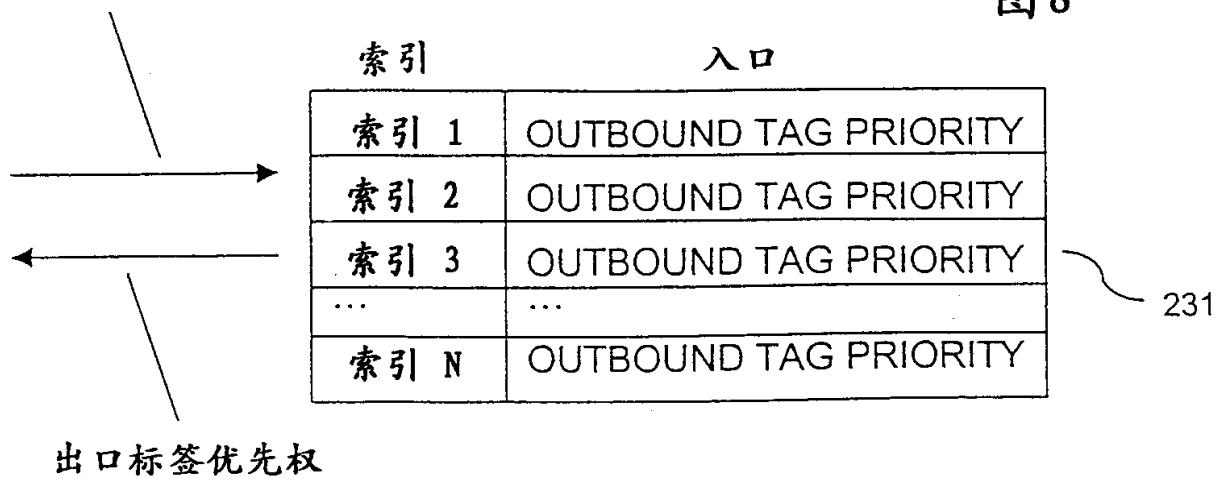


图9

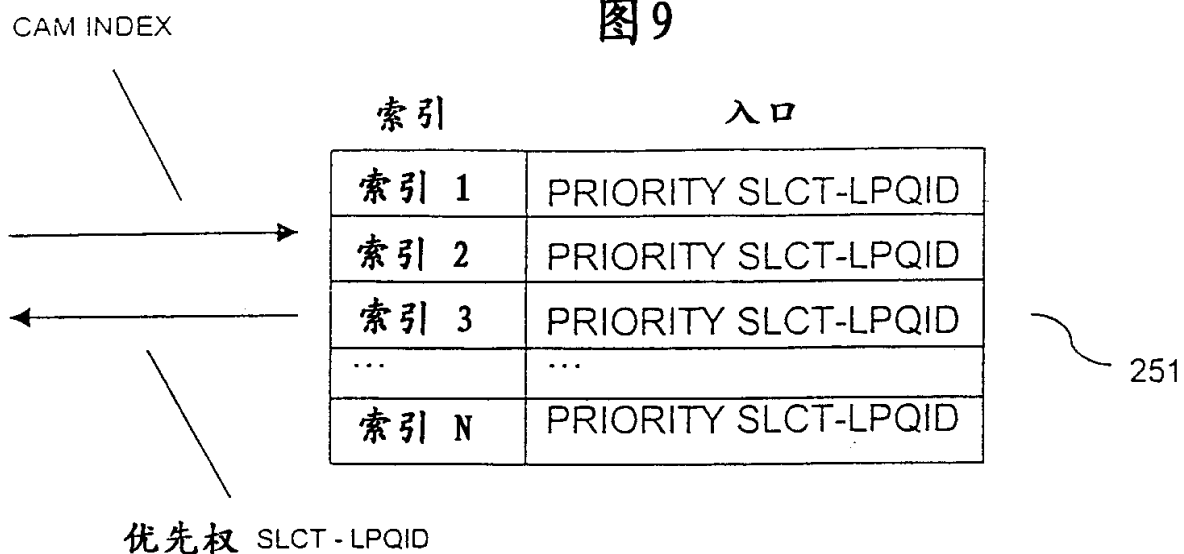


图10

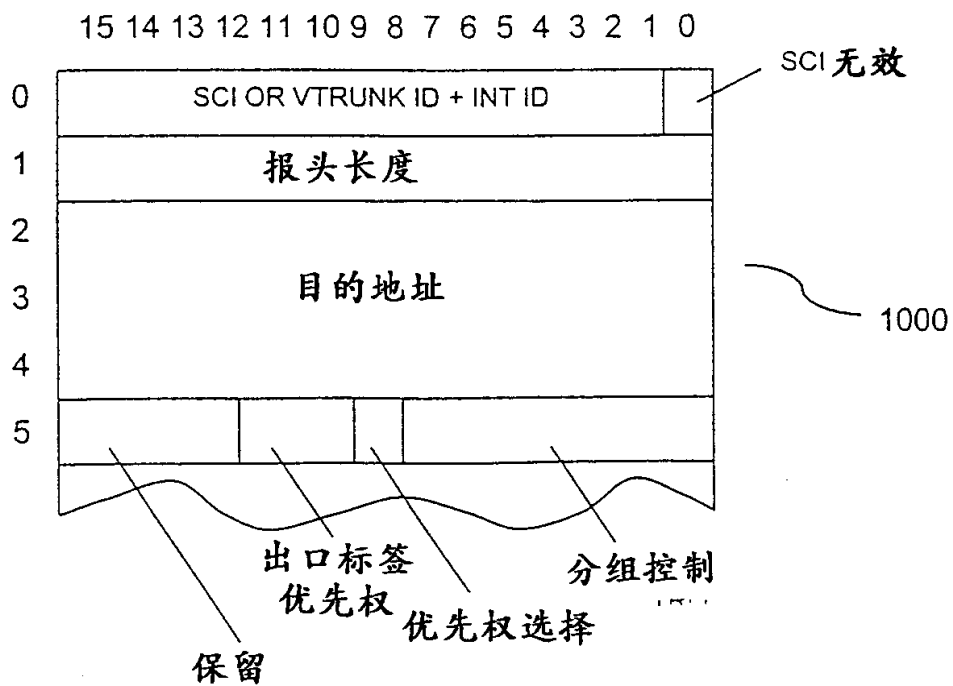




图 12

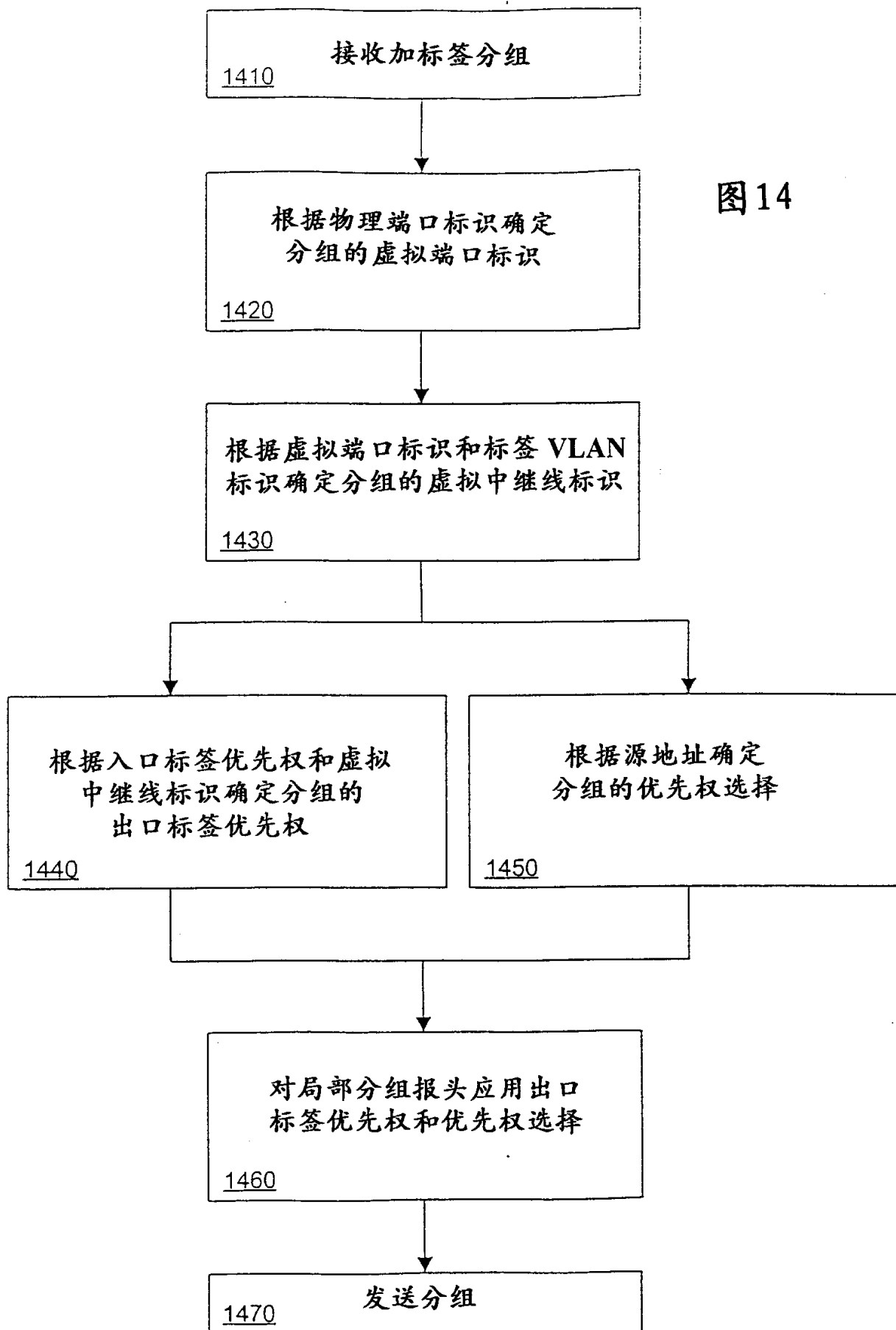
	[15:8]	[7:0]
0	PORT	CTRL
1	TAG2	TAG3
2	DA0	DA1
3	DA2	DA3
4	DA4	DA5
5	SA0	SA1
6	SA2	SA3
7	SA4	SA5
8	TL0	TL1
9	D0	D1
10	D2	D3

1200

图 13

	[7:0]
0	DA0
1	DA1
2	DA2
3	DA3
4	DA4
5	DA5
6	SA0
7	SA1
8	SA2
9	SA3
10	SA4
11	SA5
12	TAG0
13	TAG1
14	TAG2
15	TAG3
16	TL0
17	TL1
18	D0
19	D1
20	D2
21	D3

1300



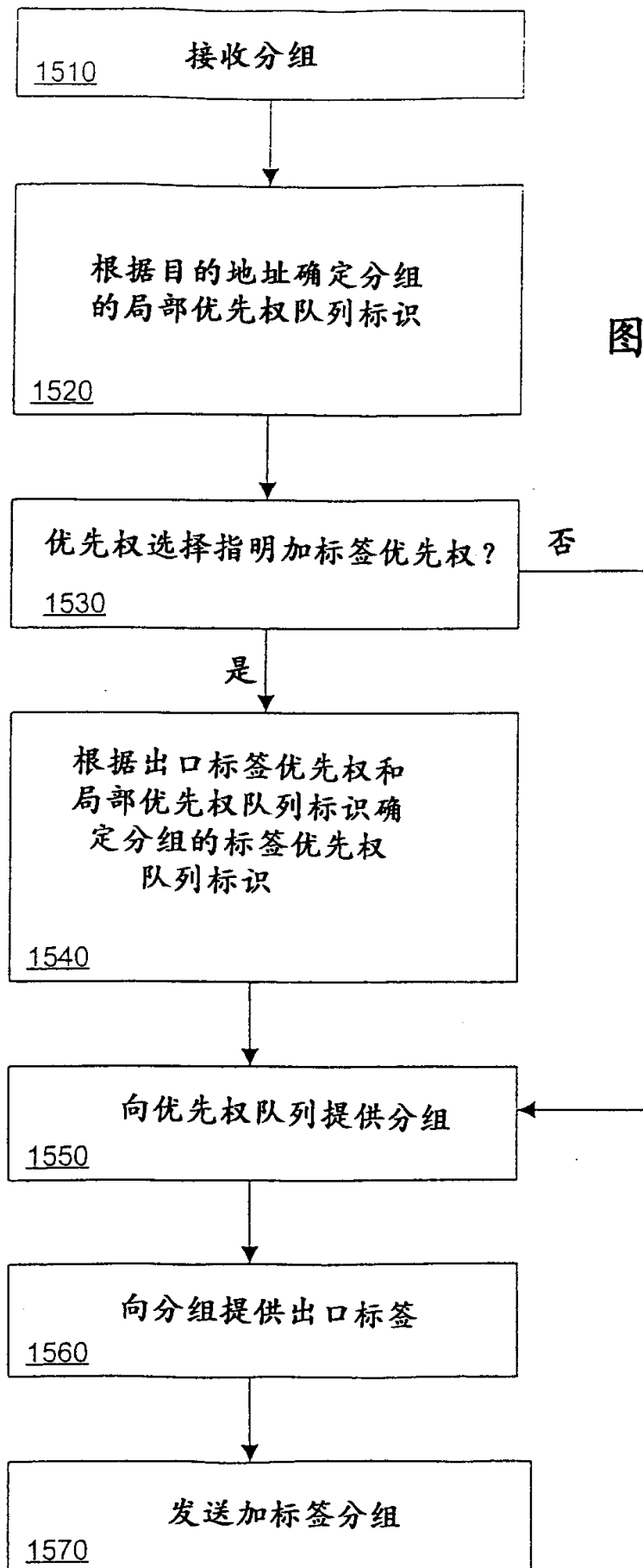


图15