

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04L 12/24 (2006.01)

H04L 12/56 (2006.01)



[12] 发明专利申请公开说明书

[21] 申请号 200610034799.1

[43] 公开日 2006年11月8日

[11] 公开号 CN 1859208A

[22] 申请日 2006.3.29

[21] 申请号 200610034799.1

[71] 申请人 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 郭英辉

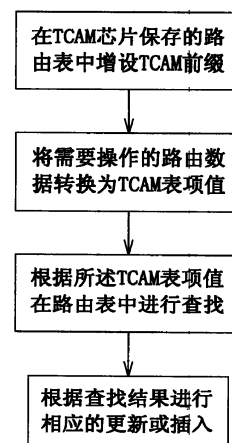
权利要求书 3 页 说明书 9 页 附图 4 页

[54] 发明名称

一种 TCAM 路由表管理方法和系统

[57] 摘要

一种涉及电通信技术的 TCAM 路由表管理方法和系统，该方法采用如下步骤：A. 在 TCAM 芯片所保存的路由表中增设 TCAM 前缀，所述的 TCAM 前缀反映相应 IPv4 地址所对应的路由前缀，B. 将需要操作的路由数据转换为 TCAM 表项值，C. 根据所述 TCAM 表项值在路由表中进行查找，根据查找结果进行相应的更新或插入，该系统包括 TCAM 芯片，所述的 TCAM 芯片中保存有路由表，其特征在于：还包括管理单元，TCAM 芯片中的路由表包括 IPv4 地址和 TCAM 前缀，TCAM 前缀反映相应 IPv4 地址所对应的路由前缀，管理单元完成有关数据转换，对 TCAM 芯片进行相关查找、更新、插入操作，本发明不需要额外建立转发信息表，提高了系统效率，具有较强的实用性和较好的性能价格比。



1. 一种 TCAM 路由表管理方法，其特征在于：它采用如下步骤：
 - A、在 TCAM 芯片所保存的路由表中增设 TCAM 前缀，所述的 TCAM 前缀反映相应 IPv4 地址所对应的路由前缀；
 - 5 B、将需要操作的路由数据转换为 TCAM 表项值；
 - C、根据所述 TCAM 表项值在路由表中进行查找，根据查找结果进行相应的更新或插入。
2. 根据权利要求 1 所述的 TCAM 路由表管理方法，其特征在于：所述的步骤 A 中，不同的 TCAM 前缀长度值对应不同的路由前缀长度值，且
10 路由前缀长度值与 TCAM 前缀长度值之间为多一对应关系。
3. 根据权利要求 2 所述的 TCAM 路由表管理方法，其特征在于：所述的 TCAM 前缀长度值（tcam_prefix_len）与路由前缀长度值（ip_prefix_len）之间采用如下方式转换：
$$\text{tcam_prefix_len} = (\text{ip_prefix_len} \gg 1) \& 0x0F$$
- 15 4. 根据权利要求 1-3 中任意一项所述的 TCAM 路由表管理方法，其特征在于：所述的步骤 B 中，所述的 TCAM 表项值包括 IPv4 地址和对应的 TCAM 前缀长度值。
5. 根据权利要求 1-3 中任意一项所述的 TCAM 路由表管理方法，其特征在于：所述的步骤 C 包括如下步骤：
 - 20 C1、根据所述 TCAM 表项值采用多匹配击中方式在路由表中进行查找，根据查找结果进行如下操作：
 - C11、若未找到匹配路由表项，则继续如下步骤 C2；
 - C12、若找到一个或一个以上匹配路由表项，则对所述 TCAM 表项值进行精确匹配，根据匹配结果进行如下操作：

C121、若匹配路由表项中具有与所述 TCAM 表项值相同的
对应路由前缀，则对路由表更新；

C122、否则，继续如下步骤 C2；

C2、在路由表中插入所述 TCAM 表项值。

- 5 6. 根据权利要求 5 所述的 TCAM 路由表管理方法，其特征在于：所述的
步骤 C12 中，所述的精确匹配中，将匹配路由表项在 TCAM 芯片中的
存放地址区段所对应的路由前缀，与所述 TCAM 表项值对应的路由前
缀相比较。
7. 根据权利要求 5 所述的 TCAM 路由表管理方法，其特征在于：所述的
10 步骤 C2 中将所述 TCAM 表项值插入对应路由前缀的存放区段。
8. 一种 TCAM 路由表管理系统，包括 TCAM 芯片，所述的 TCAM 芯片中保
存有路由表，其特征在于：还包括管理单元，所述 TCAM 芯片中的路
由表包括 IPv4 地址和 TCAM 前缀，所述的 TCAM 前缀反映相应 IPv4
地址所对应的路由前缀；所述的管理单元完成有关数据转换，对
15 TCAM 芯片进行相关查找、更新、插入操作。
9. 根据权利要求 8 所述的 TCAM 路由表管理系统，其特征在于：所述的
管理单元包括转换模块、搜寻模块、比较模块和操作模块，其中，
所述的转换模块接收需要操作的路由数据，对所述路由数据进行
进行转换，产生相应的 TCAM 表项值，并将 TCAM 表项值和路由数据分
20 别发送至搜寻模块、比较模块或操作模块；
所述的搜寻模块根据 TCAM 表项值在路由表中查找匹配路由表
项，并将查找结果发送至比较模块；
所述的比较模块根据查找结果对 TCAM 表项值进行精确匹配，并
将匹配结果发送至操作模块；
25 操作模块根据匹配结果对路由表进行更新或插入操作。
10. 根据权利要求 9 所述的 TCAM 路由表管理系统，其特征在于：

所述的管理单元还包括 TCAM 空间分配表，所述的 TCAM 空间分配表反映 IPv4 地址所对应的路由前缀与 TCAM 芯片中的存放地址区段之间的对应关系；

5 所述的比较模块根据 TCAM 空间分配表取得查找结果中匹配路由表项所相应的路由前缀，将其与 TCAM 表项值所对应的路由前缀进行比较，进行精确匹配；

所述操作模块根据 TCAM 空间分配表将 TCAM 表项值插入对应路由前缀的存放地址区段。

一种 TCAM 路由表管理方法和系统

技术领域

- 5 本发明涉及电通信技术,尤其涉及一种 TCAM 路由表管理方法和系统。

背景技术

10 随着人们对宽带网络需求的不断扩大,路由设备的转发能力也要求越来越高,用于加速转发查表性能的路由查找引擎及硬件加速器的应用也越来越广泛,三值内容地址存储器(TCAM: Ternary Content Address Memory)芯片作为高速转发的查找引擎,在路由器上得到了广泛应用。

在 TCAM 芯片中现有的 IPv4 路由表管理方法中,由于 IPv4 路由查找要求最长匹配,根据 TCAM 最先匹配低地址的原理,如图 1 所示,掩码或路由前缀长的路由表项保存在 TCAM 的低地址处,而掩码或路由前缀短的路由表项保存在 TCAM 的高地址处,图 1 中路由前缀 Prefix 为 32 15 位的路由表项保存于 TCAM 的最低地址处,路由前缀 Prefix 为 0 位的路由表项保存于 TCAM 的最高地址处。系统中保存一个 TCAM 空间分配表用于记录各路由前缀 Prefix 所占用的 TCAM 存放地址区段,其具体存放地址区段通过表项地址索引(Entry index)进行指示。

20 在进行路由添加或者删除时,需要精确匹配到 TCAM 中的具体路由表项,由于 TCAM 芯片的路由表中只保存有 IPv4 地址,因此不能通过用 IP 地址查找 TCAM 来直接索引到对应路由,通常使用辅助的转发信息表(FIB: Forwarding Information Base)来帮助查找定位。

如图2所示，在内存中创建TCAM转发信息表FIB，该转发信息表FIB可采用如CIDR Tree等树结构，FIB的叶子内容包含了该路由表项对应应在TCAM中的地址索引，以及IP地址和掩码/路由前缀Prefix，当添加或删除路由时，首先根据IP地址和掩码/路由前缀Prefix查找转发信息表FIB，如果查到则直接根据转发信息表FIB中保存的TCAM地址索引对TCAM表项进行更新，否则在FIB中插入新的节点，在TCAM中插入新的路由表项，并将该表项对应的TCAM地址索引保存到转发信息表FIB中。

在现有的TCAM路由表管理中，为协助TCAM准确定位表项位置需要额外建立一个转发信息表FIB。在一些分布式转发应用场合，转发设备的控制CPU的内存通常并不大，当路由表项数量较大时，为管理TCAM而新增的转发信息表FIB所占用的存储空间与转发设备已有存储器资源的矛盾会十分突出，例如，为了支持512K IPv4路由表的管理，转发信息表FIB需要占用额外的几十MB内存空间，这就会降低系统的效率。

15 发明内容

本发明的目的在于提供一种高效率的TCAM路由表管理方法和系统，以解决现有技术中因需要额外建立转发信息表FIB，造成系统效率降低的问题。

本发明所采用的TCAM路由表管理方法采用如下步骤：

- 20 A、在TCAM芯片所保存的路由表中增设TCAM前缀，所述的TCAM前缀反映相应IPv4地址所对应的路由前缀；
- B、将需要操作的路由数据转换为TCAM表项值；
- C、根据所述TCAM表项值在路由表中查找，根据查找结果进行相应的更新或插入。

25

所述的步骤A中，不同的TCAM前缀长度值对应不同的路由前缀长

度值，且路由前缀长度值与 TCAM 前缀长度值之间为多一对应关系。

所述的 TCAM 前缀长度值 (tcam_prefix_len) 与路由前缀长度值 (ip_prefix_len) 之间采用如下方式转换：

5 tcam_prefix_len = (ip_prefix_len >> 1) & 0x0F

所述的步骤 B 中，所述的 TCAM 表项值包括 IPv4 地址和对应的 TCAM 前缀长度值。

10 所述的步骤 C 包括如下步骤：

C1、根据所述 TCAM 表项值采用多匹配击中方式在路由表中进行查找，根据查找结果进行如下操作：

C11、若未找到匹配路由表项，则继续如下步骤 C2；

C12、若找到一个或一个以上匹配路由表项，则对所述 TCAM 表项值
15 进行精确匹配，根据匹配结果进行如下操作：

C121、若匹配路由表项中具有与所述 TCAM 表项值相同的对应
 路由前缀，则对路由表更新；

C122、否则，继续如下步骤 C2；

C2、在路由表中插入所述 TCAM 表项值。

20

所述的步骤 C12 中，所述的精确匹配中，将匹配路由表项在 TCAM 芯片中的存放地址区段所对应的路由前缀，与所述 TCAM 表项值对应的路由前缀相比较。

25 所述的步骤 C2 中将所述 TCAM 表项值插入对应路由前缀的存放区段。

这种 TCAM 路由表管理系统，包括 TCAM 芯片，所述的 TCAM 芯片中保存有路由表，其特征在于：还包括管理单元，所述 TCAM 芯片中的路由表包括 IPv4 地址和 TCAM 前缀，所述的 TCAM 前缀反映相应 IPv4 地址所对应的路由前缀；所述的管理单元完成有关数据转换，对 TCAM 芯片进行相关查找、更新、插入操作。

所述的管理单元包括转换模块、搜寻模块、比较模块和操作模块，其中，

10 所述的转换模块接收需要操作的路由数据，对所述路由数据进行转换，产生相应的 TCAM 表项值，并将 TCAM 表项值和路由数据分别发送至搜寻模块、比较模块或操作模块；

所述的搜寻模块根据 TCAM 表项值在路由表中查找匹配路由表项，并将查找结果发送至比较模块；

15 所述的比较模块根据查找结果对 TCAM 表项值进行精确匹配，并将匹配结果发送至操作模块；

操作模块根据匹配结果对路由表进行更新或插入操作。

所述的管理单元还包括 TCAM 空间分配表，所述的 TCAM 空间分配表反映 IPv4 地址所对应的路由前缀与 TCAM 芯片中的存放地址区段之间的对应关系；

所述的比较模块根据 TCAM 空间分配表取得查找结果中匹配路由表项所相应的路由前缀，将其与 TCAM 表项值所对应的路由前缀进行比较，进行精确匹配；

25 所述操作模块根据 TCAM 空间分配表将 TCAM 表项值插入对应路由前缀的存放地址区段。

本发明的有益效果为：在本发明中，在 TCAM 芯片所保存的路由表中增设 TCAM 前缀，通过 TCAM 前缀反映相应 IPv4 地址所对应的路由前缀，将需要操作的路由数据转换为 TCAM 表项值，再进行查找并根据查找结果进行相应的更新或插入，本发明中的路由表就包含 IPv4 地址和相应的 TCAM 前缀，就可以通过用 IP 地址和 TCAM 前缀查找 TCAM 来直接索引到对应路由，相对于现有技术，不需要额外建立转发信息表 FIB，从而提高系统效率。

本发明中，由于 IPv4 地址宽度是 32bit，而 TCAM 的宽度是 36bit，在现有技术中，这个多余的 4bit 闲置不用，本发明利用这个闲置 4bit 来设置 TCAM 前缀长度，以 TCAM 前缀长度来标识相应的路由前缀长度，因此，本发明充分利用了现有 TCAM 的硬件配置，具有两方面的明显效益：首先，相对于现有技术，本发明相当于将转发信息表 FIB 中的内容直接配置于 TCAM 硬件中，显然可以大大提高有关查询速度、效率，另一方面，本发明充分利用现有技术中的闲置配置，使得系统升级成本尽可能地保持最低水平，使本发明具有较强的实用性和较好的性能价格比。

本发明采用多匹配击中 Multi-Hit 方式在路由表中进行查找，使本发明可取得完全的匹配路由表项，再通过对应的路由前缀进行精确匹配，从而实现对 TCAM 中的具体路由表项的匹配，进一步提高了本发明的实用性和可操作性。

附图说明

图 1 为现有技术中 TCAM 的 IPv4 路由表存储示意图；

图 2 为现有技术中 TCAM 管理方法示意图；

图 3 为本发明系统结构示意图；

图 4 为本发明数据转换结构示意图；
图 5 为本发明基本控制流程示意图；
图 6 为本发明具体控制控制流程示意图；
图 7 为本发明插入实例示意图。

5

具体实施方式

下面根据附图和实施例对本发明作进一步详细说明：

根据图 3，本发明包括 TCAM 芯片 1 和管理单元 2，TCAM 芯片 1 中保存有路由表，路由表中包括 IPv4 地址和 TCAM 前缀，TCAM 前缀反映相应 IPv4 地址所对应的路由前缀。
10

如图 3 所示，管理单元 2 完成有关数据转换，对 TCAM 芯片 1 进行相关查找、更新、插入操作。管理单元 2 包括转换模块 21、搜寻模块 22、比较模块 23、操作模块 24 和 TCAM 空间分配表，TCAM 空间分配表反映 IPv4 地址所对应的路由前缀与 TCAM 芯片中的存放地址区段之间的对应关系。
15

如图 3 所示，转换模块 21 接收需要操作的路由数据，对所述路由数据进行转换，产生相应的 TCAM 表项值，并将 TCAM 表项值和路由数据分别发送至搜寻模块 22、比较模块 23 或操作模块 24。

如图 3 所示，搜寻模块 22 根据 TCAM 表项值在路由表中查找匹配路由表项，并将查找结果发送至比较模块 23；
20

如图 3 所示，比较模块 23 根据查找结果对 TCAM 表项值进行精确匹配，并将匹配结果发送至操作模块 24，具体地，比较模块 23 根据 TCAM 空间分配表取得查找结果中匹配路由表项所相应的路由前缀，将其与 TCAM 表项值所对应的路由前缀进行比较，进行精确匹配。

如图 3 所示，操作模块 24 根据匹配结果对路由表进行更新或插入操作，其中，操作模块 24 根据 TCAM 空间分配表将 TCAM 表项值插入对
25

应路由前缀的存放地址区段。

如图 3 和图 5 所示, 本发明的基本控制流程如下:

- 1) 在 TCAM 芯片 1 所保存的路由表中增设 TCAM 前缀, TCAM 前缀反映
5 相应 IPv4 地址所对应的路由前缀, TCAM 前缀与路由前缀以相应的
长度值相对应, 这样, 路由表中包括 IPv4 地址和 TCAM 前缀长度值。
- 2) 管理单元 2 接收需要操作的路由数据, 将其转换为 TCAM 表项值,
TCAM 表项值包括 IPv4 地址和对应的 TCAM 前缀长度值。
- 3) 管理单元 2 根据转化后的 TCAM 表项值在路由表中查找匹配项。
- 10 4) 管理单元 2 根据查找结果对路由表进行相应的更新或插入。

如图 3 和图 6 所示, 本发明的具体控制流程如下:

1. 在 TCAM 芯片 1 所保存的路由表中增设 TCAM 前缀, TCAM 前缀反映
相应 IPv4 地址所对应的路由前缀, TCAM 前缀与路由前缀以相应的
15 长度值相对应, 路由表中包括 IPv4 地址和 TCAM 前缀长度值, 其中,
IPv4 地址为 32bit, TCAM 前缀长度值为 4bit。
2. 管理单元 2 中的转换模块 21 接收需要操作的路由数据, 将其转换
为 TCAM 表项值, 在本发明中, 由于 IP 路由前缀长度为 0 到 32,
需要 6bit 才能表示, 因此, 如图 4 所示, 需要将其中的 IP 路由前
20 缀长度值 `ip_prefix_len` 从 6bit 转换为 4bit TCAM 前缀长度值
`tcam_prefix_len`, 所采用的转换方式如下:

$$\text{tcam_prefix_len} = (\text{ip_prefix_len} \gg 1) \& 0x0F$$

- 在本发明中, TCAM 前缀长度值 `tcam_prefix_len` 与路由前缀
长度值 `ip_prefix_len` 的对应关系如下表 1 所示, 由该表可知, 不
25 同的 TCAM 前缀长度值对应不同的路由前缀长度值, 且路由前缀长
度值与 TCAM 前缀长度值之间为多一对应关系。

TCAM前缀 长度值	路由前缀 长度值	TCAM前缀 长度值	路由前缀 长度值
0	0, 16, 32	8	8, 24
1	1, 17	9	9, 25
2	2, 18	10	10, 26
3	3, 19	11	11, 27
4	4, 20	12	12, 28
5	5, 21	13	13, 29
6	6, 22	14	14, 30
7	7, 23	15	15, 31

表 1

5 经过上述转换，转换模块 21 将路由数据转换为 TCAM 表项值，
如图 4 所示，TCAM 表项值包括 IPv4 地址和对应的 TCAM 前缀长度
值，同时，转换模块 21 将 TCAM 表项值分别发送至搜寻模块 22 和
操作模块 24，将路由数据发送至比较模块 23。

3. 搜寻模块 22 根据 TCAM 表项值采用多匹配击中 Multi-Hit 方式在
TCAM 芯片 1 中的路由表中进行查找，即查找匹配路由表项，在本
10 发明中，由于路由前缀长度值与 TCAM 前缀长度值之间为多一对应
关系，则可能在路由表中查找到一个或多个匹配路由表项，所以采
用这种多匹配击中 Multi-Hit 方式进行查找，搜寻模块 22 将查找
结果发送至比较模块 23。

4. 比较模块 23 根据查找结果进行如下操作：
15 41、若查找结果为一个或一个以上匹配路由表项，则对所述 TCAM
表项值进行精确匹配，比较模块 23 对查找结果通过 TCAM 空
间分配表进行解析，将查找结果中的匹配路由表项在 TCAM 芯

片 1 中的存放地址区段所对应的路由前缀长度值, 与 TCAM 表项值对应的路由前缀长度值相比较, TCAM 表项值对应的路由前缀长度值可由路由数据中获得, 根据比较结果进行如下操作:

5 411、若存在相同的路由前缀长度值, 说明查找结果中具有精确匹配项, 比较模块 23 将该匹配结果信息发送至操作模块 24, 操作模块 24 对 TCAM 芯片 1 中的路由表进行更新。

 412、否则, 说明查找结果中没有精确匹配项, 比较模块 23
10 将该匹配结果信息发送至操作模块 24, 继续如下步骤 5。

 42、否则, 查找结果为没有匹配路由表项, 比较模块 23 将该信息发送至操作模块 24, 继续如下步骤 5。

5. 操作模块 24 将 TCAM 表项值插入路由表中, 在本发明中的插入中,
15 操作模块 24 根据 TCAM 空间分配表将 TCAM 表项值插入对应路由前缀的存放地址区段, 例如, 若插入两条路由, 路由数据分别是 10.0.0.0/8 和 10.0.0.0/24, 对应的 TCAM 表项值均为 10.0.0.0/8, 则如图 7 所示, 两条路由分别被插入对应路由前缀的存放地址区段, 实际上, 通过 TCAM 空间分配表的存放地址区段确立了被插入
20 TCAM 表项值在逻辑上的路由前缀长度值。

 对于本发明, 在转发过程中进行 TCAM 查找时, 将查找到的路由表项中的掩码 Mask 设置为忽略 TCAM 前缀长度值 `tcam_prefix_len` 这 4bit, 只用 IPv4 地址进行匹配, 虽然在 TCAM 芯片 1 中增加了 TCAM 前
25 缀长度值, 但同样能实现最长匹配查找, 这样, 所涉及的后续操作就与现有技术完全一致, 不影响转发, 此处不再赘述。

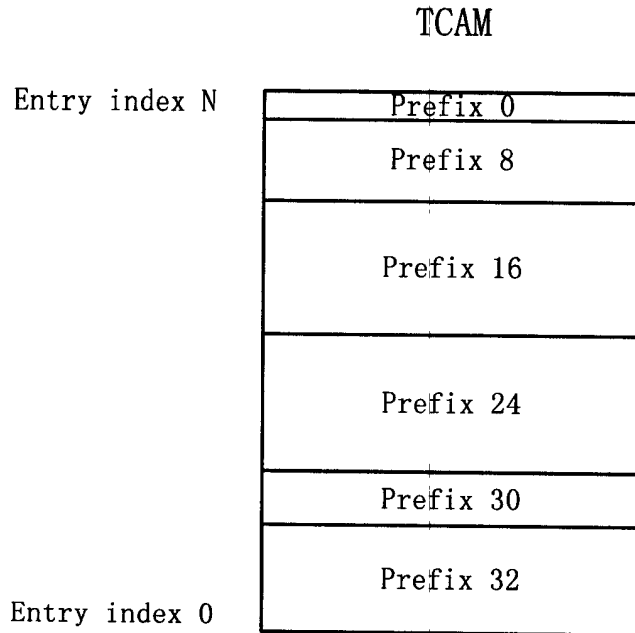


图1

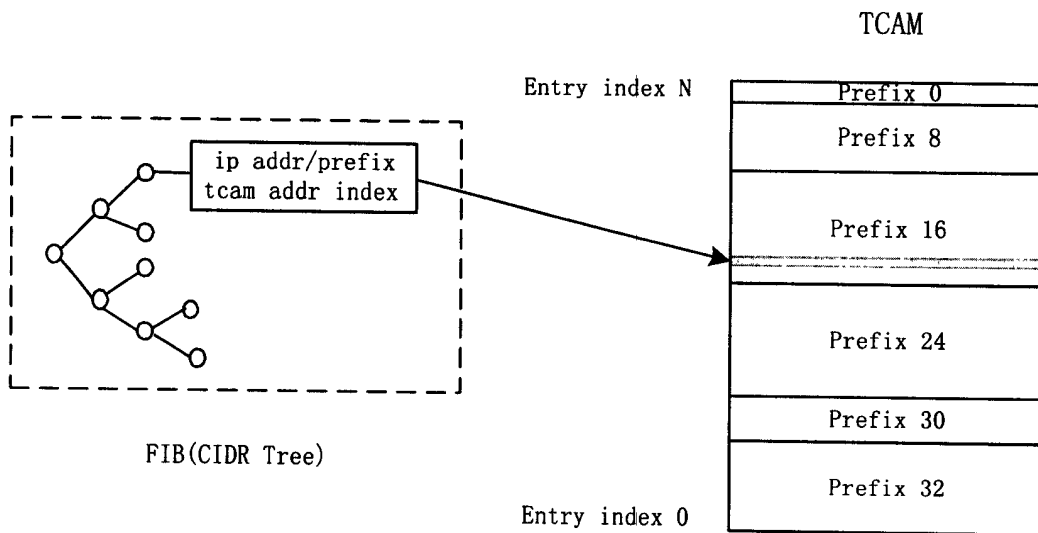


图2

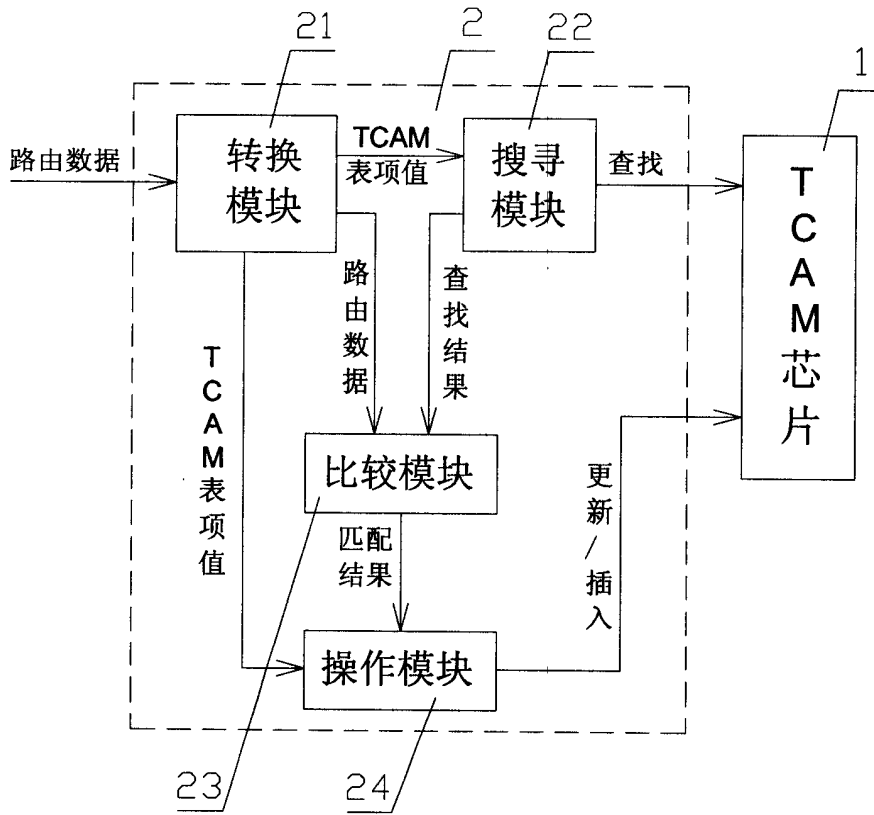


图3

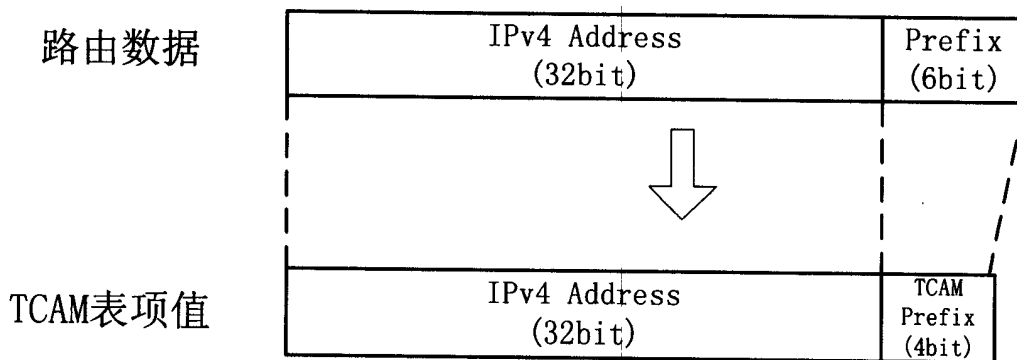


图4

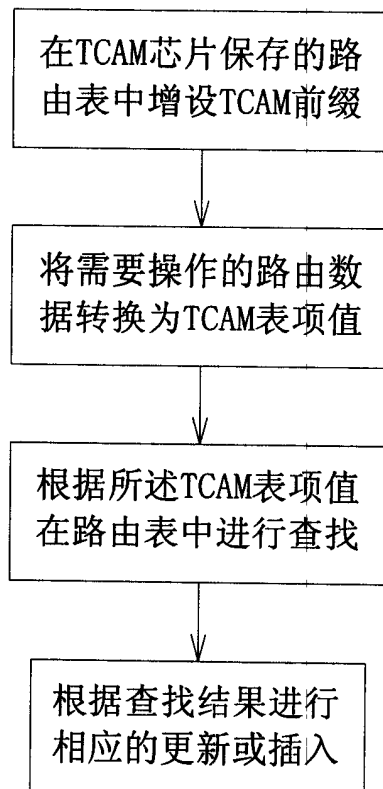


图5

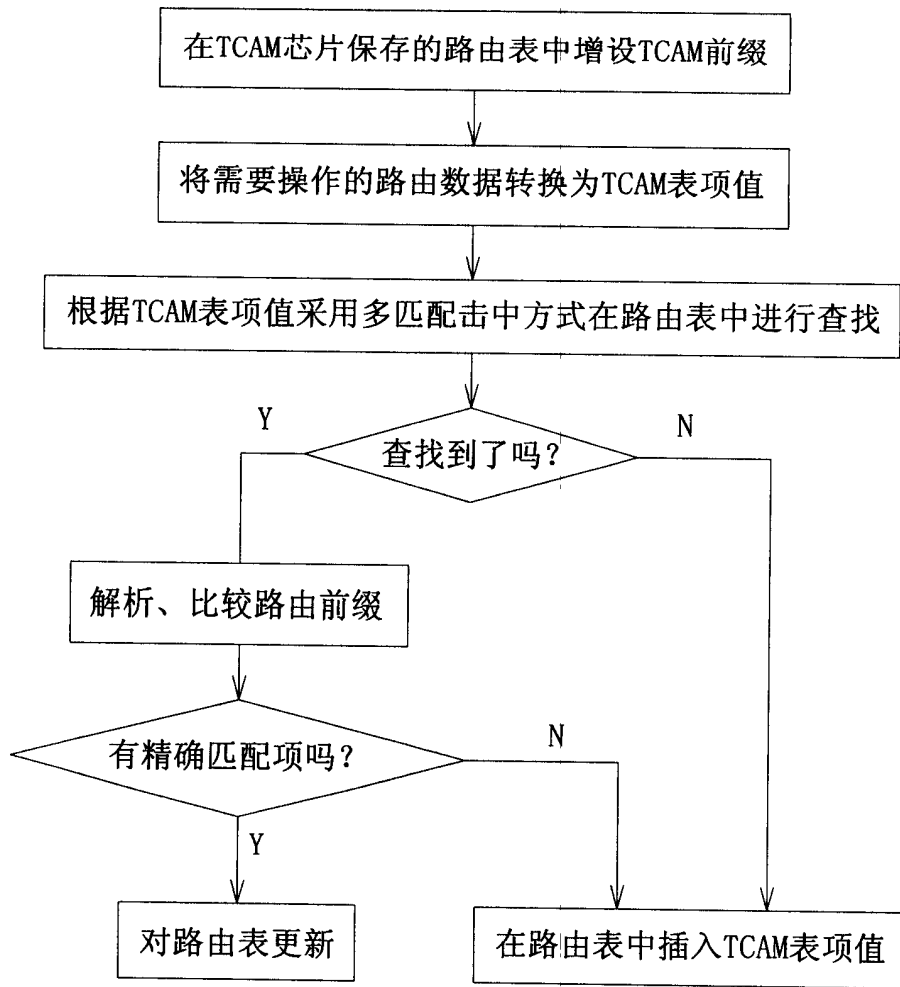


图6

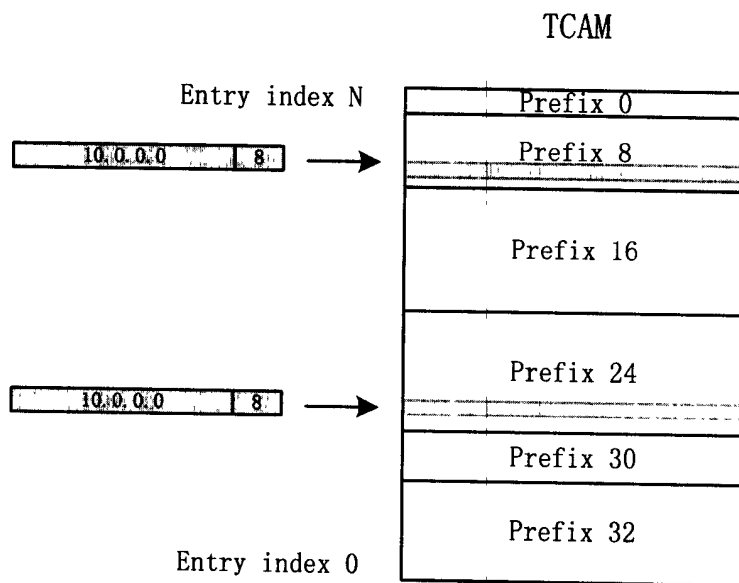


图7