(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2009 年 8 月 27 日 (27.08.2009)

PCT

(10) 国际公布号
WO 2009/103225 A1

(51) 国际专利分类号:
H04L 12/56 (2006.01)

(21) 国际申请号: PCT/CN2009/070378

(22) 国际申请日: 2009 年 2 月 6 日 (06.02.2009)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
200810008199.7 2008 年 2 月 18 日 (18.02.2008) CN

(71) 申请人 (对除美国外的所有指定国): **华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。

(72) 发明人; 及

(75) 发明人/申请人 (仅对美国): **张浩 (ZHANG, Hao)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 **龙志平 (LONG, Zhiping)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 **田向远 (TIAN, Xiangyuan)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 **李维 (LI, Wei)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 **顾晓浩 (GU, Xiaohao)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 **邵建树 (SHAO, Jianshu)** [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: **北京三高永信知识产权代理有限责任公司 (BEIJING SAN GAO YONG XIN INTELLECTUAL PROPERTY AGENCY CO., LTD.)**; 中国北京市海淀区学院路蓟门里和景园 A-1-102, Beijing 100088 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS,

[见续页]

(54) **Title**: A PACKET FORWARDING METHOD AND EQUIPMENT

(54) **发明名称**: 一种报文转发的方法和设备

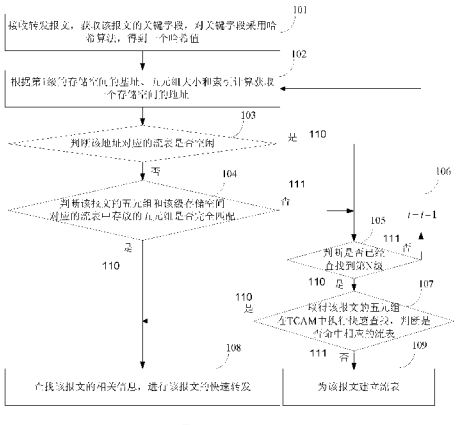

FIG. 1

101 RECEIVING A PACKET TO BE FORWARDED, ACQUIRING A KEY FIELD OF THE PACKET, APPLYING A HASH ALGORITHM TO THE KEY FIELD AND GENERATING A HASH VALUE
102 CALCULATING AND ACQUIRING AN ADDRESS OF A STORAGE SPACE ACCORDING TO A BASE ADDRESS OF THE LEVEL I STORAGE SPACE, A SIZE OF FIVE-TUPLE AND AN INDEX
103 JUDGING WHETHER THE FLOW TABLE CORRESPONDING TO THE ADDRESS IS IDLE
104 JUDGING WHETHER THE FIVE-TUPLE OF THE PACKET MATCHES THAT STORED IN THE FLOW TABLE CORRESPONDING TO THE STORAGE SPACE OF THIS LEVEL
105 JUDGING WHETHER THE LEVEL N HAS BEEN SEARCHED
107 ACQUIRING THE FIVE-TUPLE OF THIS PACKET AND EXECUTING THE FAST SEARCHING IN TCAM, JUDGING WHETHER A CORRESPONDING FLOW TABLE IS HITTED
108 SEARCHING THE RELATED INFORMATION OF THE PACKET AND EXECUTING THE FAST-FORWARDING OF THE PACKET
109 ESTABLISHING A FLOW TABLE FOR THE PACKET
110 YES
111 NO

(57) **Abstract**: A packet forwarding method and equipment belonging to the communication field is disclosed. The method comprises: receiving a packet and acquiring a key field of the packet; according to the key field, searching whether there is a flow table corresponding to the key field in a storage space of a first type, if yes, forwarding the packet according to the flow table; otherwise, according to the key field, searching whether there is a flow table corresponding to the key field in a storage space of a second type, if yes, forwarding the packet according to the flow table which is found in the storage space of the second type. The equipment comprises a receiving and acquiring module and a searching and forwarding module. By adopting a flow table mechanism that combines the multilevel hash table with the TCAM, the fast-forwarding of packet, the balance between the storage volume of packets and the forwarding speed as well as the fast convergence of flow table are achieved, the packet storage collision is reduced, the request for capability and speed of the network operator is satisfied and the cost is saved.

[见续页]

WO 2009/103225 A1

LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) **指定国** (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

(57) **摘要**:

本发明公开了一种报文转发的方法和设备，属于通信领域。方法包括：接收报文，获取所述报文的关键字段；根据所述关键字段，查找第一类存储空间中是否有与所述关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表；如果是，根据在所述第二类存储空间中查找到的流表转发所述报文。设备包括：接收及获取模块、查找及转发模块。通过采用将多级哈希表与 TCAM 相结合的流表机制，实现报文的快速转发，实现存储报文的容量和转发速度之间的平衡，减少报文存储冲突，实现流表快速收敛，满足网络运营商对容量和速度的要求，节约了成本。

说 明 书

一种报文转发的方法和设备

技术领域

本发明涉及通信领域，特别涉及一种报文转发的方法和设备。

背景技术

随着 IP(Internet Protocol,网络之间互连的协议)技术的不断发展，IP 网络从承载单一的 Internet 数据业务向承载语音、数据、视频、大客户专线、3G(3rd Generation，第三代数字通信)、NGN（Next/New Generation Network)，下/新一代网络）等运营级多业务方向转型。由于，过去 IP 网络主要承载普通 Internet 的数据服务，如收发 Email、上网页看新闻等等，因此对网络的带宽方面的要求并不高，通常 56Kpbs 的访问速度已经能够满足需求。但是随着 IP 网络的转型，需要服务于各种层出不穷的新业务，如网络游戏、P2P(Peer-to-Peer，对等技术)、视频、语音等等，于是，在转型过程中，对 IP 网络在转发性能、容量、安全、业务服务质量上等方面提出了更高的要求。

目前，在 IP 网络的骨干网的转发速率已经达到 40G 甚至 100G，汇聚层、接入层的转发速率也已经普遍达到 10G，通常是采用网络处理器以及多核处理器来实现高速转发，如何在网络处理器或多核处理器中实现对大容量的报文的快速转发，是亟需解决的问题。

当前，对于报文的快速转发主要根据报文的相应关键字段进行相应动作后将报文转发出去，其中，报文的关键字段如报文的 MAC（Media Access Control，介质访问控制）地址，IP 地址或者报文的五元组（如，由报文的源 IP 地址,目的 IP 地址,协议号,源端口号,目的端口号构成等)、七元组(如，由报文的源 IP 地址、目的 IP 地址、协议号、源端口号、目的端口号、服务类型、输入接口构成等)等等。相应地，在转发设备如防火墙、深度报文解析、会话边界控制器中，采用基于报文的关键字段组合而形成的流表(即会话表)以及由相应参数和动作组成的动作表两部分组成的流表机制（其中流表和动作表是一一对应的关系，根据实现方式的不同，可以将流表和动作表统一在一起也可以将流表和动作表进行分别存储)，从而实现报文的快速转发。但是，转发设备需要能够提供大容量的存储空间来存放大量由报文的关键字段组成的流表，同时又需要转发设备提供高速的访问和转发速度。目前转发设备中的存储装置在使用时存在不同程度的缺陷：容量大价格低的存储装置如 SRAM（Static Random-Access Memory，静态随机存取存储器)、DRAM(Dynamic Random-Access Memory，动态随机存储器)等，访问速度往往就比较低；而访问速度大的存储装置如 TCAM

（Triple Content Addressable Memory，三态内容可寻址存储器）其容量往往较小，价格也较高。

同时在进行报文存储时，以什么样的组织形式存放这些由报文的关键字段组合而成的流表也是必须考虑的问题。如果以直接表或者链表的形式直接存放大量的流表，需要大量的存储空间，难以提供高速的访问速度。目前主要的办法是采用将流表中的部分关键字段存放在直接表或者链表中，实现对存储空间和访问速度的平衡，而这样有可能由于不同的报文其部分关键字段相同，而导致存放时出现冲突；采用树的形式以最长匹配或者完全匹配的方式存放这些流表，可以解决上述冲突的问题，但这种方法比较复杂，而且性能不高。

现有的流表实现机制中，普遍采用树的形式或者链表的方式来存放流表，在树的形式的实现方式中，根据报文的关键字段(如五元组)以最长匹配或者完全匹配等方式来建立和查找流表；在链表的实现方式中，采用多级链表的方式，每级查找表下含有多个辅助链表来建立和查找流表。这两种方式虽然都可以解决冲突问题，但仅适用于访问速度和容量要求不高的场合，但是，随着对报文的转发速度的要求的日益提高，如在目前普遍要求达到10G 的转发速度的接入层和汇聚层中，存在大量的业务报文需要大量的存储空间和高速转发，如果采用树的形式或者链表的方法存放流表就会出现树的深度越来越深，链表级数越来越大，从而导致访问速度越来越慢，难以实现报文的快速转发。

发明内容

为了实现报文的快速转发，实现流表的快速收敛，本发明实施例提供了一种报文转发的方法和设备。所述技术方案如下：

一方面，本发明实施例提供了一种报文转发的方法，所述方法包括：

接收报文，获取所述报文的关键字段；

根据所述关键字段，查找第一类存储空间中是否有与所述关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表；如果是，根据在所述第二类存储空间中查找到的流表转发所述报文。

另一方面，本发明实施例提供了一种报文转发设备，所述设备包括：

接收及获取模块，用于接收报文，获取所述报文的关键字段；

查找及转发模块，用于根据所述关键字段，查找第一类存储空间中是否有与所述接收及获取模块获取到的关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表；如果是，根

据在所述第二类存储空间中查找到的流表转发所述报文。

本发明实施例提供的技术方案的有益效果是：

通过采用多级存储空间（第一类存储空间和第二类存储空间），将多级哈希表与 TCAM 相结合的流表机制，可以实现报文的快速转发，在设备中实现存储报文的容量和转发速度之间的平衡，减少报文存储的冲突，实现流表的快速收敛，从而很好的满足网络运营商对容量和速度的要求，节约了成本，方便其开展更多的增值业务。

附图说明

图 1 是本发明实施例 2 提供的报文转发的方法的流程示意图；

图 2 是本发明实施例 2 提供的建立流表的流程示意图；

图 3 是本发明实施例 3 提供的报文转发的方法的流程示意图；

图 4 是本发明实施例 4 提供的报文转发的设备示意图；

图 5 是本发明实施例 5 提供的报文转发的设备示意图；

图 6 是本发明实施例 6 提供的报文转发的设备示意图；

图 7 是本发明实施例 7 提供的报文转发的设备示意图。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明实施方式作进一步地详细描述。

实施例 1

本发明实施例提供的方案通过采用多级哈希表与 TCAM 相结合的转发流表机制，通过使用 N+1 级存储空间，前面 N 级使用价格低廉的存储空间（如 DRAM 等）利用报文的关键字段采用哈希算法存储流表，最后一级采用 TCAM 存储流表。其中，当存储报文建立流表时，如果前面 N 级都存在报文存储的冲突，便直接利用最后一级的 TCAM 存储报文建立流表，实现快速收敛。其中，前面 N 级使用的价格低廉的存储空间称为第一类存储空间，最后一级 TCAM 称为第二类存储空间，或称为高速存储空间。方法包括：

当接收到报文后，获取报文的关键字段；然后根据关键字段，查找第一类存储空间中是否有与关键字段对应的流表，如果是，根据在第一类存储空间中查找到的流表转发报文；否则，根据关键字段，查找第二类存储空间中是否有与关键字段对应的流表；如果是，根据在第二类存储空间中查找到的流表转发报文。

其中，当第一类存储空间为多级存储空间时，需要对第一类存储空间的各级存储空间进行依次查找，根据命中的流表进行报文的转发，具体内容为：

获取第一类存储空间的当前存储空间的流表，当查找到与关键字段相对应的流表，则根据流表转发报文，否则，查找当前存储空间的下一级存储空间；当第一类存储空间的各级存储空间中没有查找到与关键字段相对应的流表时，查找第二类存储空间是否有与关键字段对应的流表，然后根据在第二类存储空间中命中的流表进行报文的转发。

进一步地，在获取第一类存储空间的当前存储空间的流表时，可以通过以下方式实现：

首先，对报文的关键字段进行哈希运算，计算得到该报文的关键字段的哈希值；通过当前存储空间的大小获取掩码；对哈希值与掩码进行相与获取一个索引；

然后，计算索引和报文的关键字段的乘积值；根据乘积值与当前的存储空间的基址之和，获取当前存储空间的地址；

最后，根据当前存储空间的地址获取当前存储空间的流表。

进一步地，当查找到与关键字段相对应的流表，根据流表转发报文时，还需要判断流表是否空闲，如果当前存储空间的流表不空闲，则需要查找当前存储空间的流表，当在当前存储空间的流表中查找到与关键字段相应的流表，则根据该查找到的流表转发报文；如果当前存储空间的流表空闲，查找当前存储空间的下一级存储空间。

综上所述，本发明实施例实现了根据报文的关键字段采用哈希算法后，在第一类存储空间逐级进行查找；如果第一类存储空间都查找失败，则可以直接在第二类存储空间中进行快速查找。当命中了流表后，就可以实现报文的快速转发。

本发明实施例提供的技术方案，当第二类存储空间中没有与关键字段对应的流表时，即此时报文无法命中第一类存储空间和第二类存储空间中的流表，说明该报文的相应的流表还没有建立，于是需要为该报文建立流表。其中，为该报文建立流表的过程如下：

首先，获取报文的关键字段的哈希值；通过第一类存储空间的当前存储空间的大小获取掩码；对哈希值与掩码相与获取索引；计算索引和关键字段的乘积值；

然后，根据乘积值和当前存储空间的基址，获取当前存储空间中的地址（该当前存储空间中的地址对应于在该存储空间中的某一流表），判断地址对应的流表是否空闲，如果是，在当前存储空间的地址为报文建立流表；否则，查找当前存储空间的下一级存储空间；当第一类存储空间的各级存储空间中的地址对应的流表均为非空闲状态时，在第二类存储空间为报文建立流表。

其中，在进行报文查流表转发的时候，还可以记录下第一类存储空间的各级存储空间中的地址对应的流表的空闲信息；于是，相应地，为报文建立流表的过程具体如下：

首先，获取报文的关键字段的哈希值；

然后，通过查找记录的第一类存储空间的各级存储空间中的地址对应的流表的空闲信息，在各级存储空间中首次出现空闲的存储空间的地址，为报文建立流表；当第一类存储空间的各级存储空间中的地址对应的流表是非空闲状态时，在第二类存储空间为报文建立流表。

通过上述方法，本发明实施例实现了为报文建立流表。在这里需要特别注意的是，当为报文进行建表时，需要判断是否有相同哈希值的报文正在被加锁，如果是，等待相同哈希值的报文解锁后，对报文执行相应的动作，如果否，则对报文加锁。

相应地，在为报文建立流表后，对被加锁的报文进行解锁。

其中，等待相同哈希值的报文解锁后，对报文执行相应的动作，具体为：

根据获取报文的关键字段的哈希值，判断报文和相同哈希值的报文是否属于同一类型，如果是，则等待相同哈希值的报文解锁后，根据相同哈希值的报文建立的流表，进行报文的快速转发；否则，报文等待相同哈希值的报文解锁后，为报文建立流表。

本发明实施例提供的报文转发的方法，通过采用了第一类存储空间和第二类存储空间相结合的转发机制，在设备中实现存储报文的容量和转发速度之间的平衡，减少报文存储的冲突，实现流表的快速收敛，从而很好的满足网络运营商对容量和速度的要求，节约了成本，方便其开展更多的增值业务。

针对本实施例提供的报文转发的方案，下面以多个实施例针对查找和建表进行详细的说明。

实施例 2

参见图 1，本发明实施例提供了一种报文转发的方法，本实施例以设置了 N+1 级别存储空间为例进行说明，其中前 N 级为由普通存储设备如 SRAM\DRAM 等提供的存储空间，第 N+1 级为由 TCAM 等存储设备提供的高速存储空间，利用本发明实施例提供的方法，报文转发时根据报文的关键字段采用哈希算法获得的键值在前 N 级存储空间逐级进行查找；如果前 N 级存储空间都查找失败，则直接在 TCAM 中进行快速查找。具体内容包括如下：

设备对接收到的报文进行转发时，需要从设备自身的存储空间中查找该报文对应的流表，当查找到流表后，获取到相关的动作表信息，才能完成该报文的成功转发，其中，具体内容如下：

101：接收转发报文，获取该报文的关键字段，对关键字段采用哈希算法，得到一个哈希值。

其中，本实施例中报文的关键字段以常用的五元组（源 IP 地址,目的 IP 地址,协议号,源端口号,目的端口号）为例进行说明。

102：根据第i ($1 \le i \le N$)级的存储空间的大小，获取一个掩码，将获取到的该掩码与 101 中得到的哈希值进行相与操作，从而获取一个索引；根据该第i级存储空间的基址（起始地址）、报文的五元组大小和所述索引，计算获取一个该第i级存储空间的地址。

其中，根据存储空间的大小，获取一个掩码时，掩码全 1；通常，为了避免冲突的产生，存储空间越大对应着掩码取值位数越多。在获取存储空间的地址时，通常采用计算方法如下：

存储空间的地址＝存储空间的基址＋五元组大小×索引；

103：根据获取的存储空间的地址，通过查看该地址对应流表的标志位，判断该流表是否空闲，如果是，则执行 105；否则，执行 104。

例如，如果该地址对应的流表的标志位为“0”，表示该流表为空闲状态；如果该地址对应的流表的标志位为“1”，表示该流表为非空闲状态。

104：判断该报文的五元组和该级存储空间对应的流表中存放的五元组是否完全匹配，如果是，说明该报文命中流表，执行 108；否则，说明该报文在该级存储空间进行流表查找时，出现了冲突，执行 105。

105：判断是否已经查找到第 N 级，如果是，则执行 107；否则，执行步骤 106。

106：$i = i + 1$，执行步骤 102。

107：取得该报文的五元组在 TCAM 中执行快速查找，判断是否命中相应的流表，如果是，执行 108；否则，执行 109。

108：查找该报文的相关信息，进行该报文的快速转发，结束。

本步骤中查找该报文的相关信息，主要是查找该报文对应的流表和其动作表的对应关系，通过流表和动作表结合在一起，才能完成报文的快速转发。

109：说明该报文为首包，为该报文建立流表。

本实施例将在 N+1 级存储空间中，能够命中流表的报文称为后续包，后续包取得相关信息后，读取动作表进行了相关动作后，就能够实现报文快速转发；在 N+1 级存储空间都没有命中流表的报文，称为首包，对于首包需要建立流表，同时也需要首包建立动作表，将相关动作存放在动作表中，建立并维护动作表和流表对应关系，以便以后接收到相同的报文时，该报文在命中流表查找到流表和动作表的对应关系后，根据动作表中的动作进行相应处理。

下文会针对报文为首包时如何实现建立流表进行详细的说明：

所谓流表的建立，是指设备在第一次接收到某条报文时，由于在设备的存储空间中没有记录该报文的流表，所以需要为该报文建立流表，以便以后再收到相同报文后可以命中流表，从而实现快速转发。当设备收到该报文，查找到空闲的存储空间的情况下，将报文的相应内容插入到空闲的存储空间中。如果在前 N 级的存储空间中，都没有空闲的存储空间，则将报文的相应内容直接存储在第 N+1 级的 TCAM 中，从而实现快速收敛。

通常情况，由于报文是双向的（正向报文和反向报文），因此建立流表时，相应地，需要建立双向流表，即正向流表和反向流表，下面以建立正向流表为例，其步骤如下：

201：获取报文的关键字段，对关键字段采用哈希算法，得到一个哈希值。

其中，该步骤 201 处理过程，与前述步骤 101 类似，不再赘述。

其中，本实施例中报文的关键字段以常用的五元组（源 IP 地址,目的 IP 地址,协议号,源端口号,目的端口号）为例进行说明。相应地，正向流表和反向流表的区别在于，对于正向报文而言的源 IP 地址为反向报文的目的 IP 地址、源端口号为反向报文的目的端口号，相应地，正向报文的目的 IP 地址为反向报文的源 IP 地址、目的端口号为反向报文的源端口号。

202：为了确保建立流表时不出现错误，执行加锁动作。

其中，进行加锁的目的，是为了保证为该首包报文建立流表时，只能进行原子操作（即在建立流表的时刻，此时只针对拥有相同哈希值的报文一个报文进行建表处理），防止在为该首包报文建立流表时，出现了与该目标首包报文的哈希值相同的报文（如，隶属于同一数据报文流、拥有相同的关键字段）的第二个报文，；或在为该首包报文建立流表的时间段内，出现来自其它类的数据报文流中的首包，但是该首包获取到的哈希值和目标首包报文的哈希值相同，此时就会出现报文的建立过程的冲突，为了避免出现此类型的报文冲突，所以本步骤需要进行加锁动作。即建立流表时只能进行原子操作，防止建立流表出现错误，即此时只针对该报文执行下面的操作，保证了此时只针对拥有相同哈希值的报文一个报文进行建表处理，

203：根据第$i(1 \le i \le N)$级的存储空间的大小，获取一个掩码，将获取到的该掩码与 201 中得到的哈希值进行相与操作，从而获取一个索引；根据该第i级的存储空间的基址，该报文的五元组的大小和计算得到的索引，计算获取第i级存储空间的一个地址。

204：根据获取的地址取得该地址对应的流表的内容，根据流表的标志位，判断获取的地址对应的流表是否为空闲，如果是，则执行 205；否则执行 206。

205：将报文的相应内容存储到该地址对应的位置中建立流表，并设置相应的标志位，表示此时该位置的流表已被占用，然后执行 209。

206：判断此时i是否等于 N，即判断是否已经查找到第 N 级，如果是，执行 207；否则执行 208。

207：说明在前 N 级的存储空间中，都没有空闲的存储空间来存放该报文的相应内容（即在前 N 级的存储空间都出现了建立该报文流表的冲突），则直接将报文的相应内容存放到第 N+1 级的 TCAM 中，在 TCAM 中建立流表，置相应的标志位，从而实现了快速收敛。

208：令$i=i+1$后，返回执行步骤 203。

209：当流表建立成功后，进行解锁，此时，正向流表就成功建立。

上述以建立正向流表为例进行的说，当建立反向流表时，方法类似，不再赘述。

综上所述，为本发明实施例提供的查找和建立流表的实现报文的快速转发的过程，查找时，根据报文的关键字段采用哈希算法获得键值在前 N 级逐级查找。如果前 N 级存储空间都失败，则直接在 TCAM 中快速查找；建表时，根据报文的关键字段采用哈希算法获得键值在前 N 级逐级查找空闲空间，若空闲则直接插入，若冲突则往后逐级查找；如果前 N 级哈希存储空间都冲突，则直接存储到 TCAM 中。

需要特别地注意的是，在流表建立过程中，需要保证一次只处理一个具有相同哈希值的报文。如前文所述，例如：当在进行首包的流表建立时，如果出现了隶属于同一类数据报文（即拥有相同关键字段的报文）的伪首包也需要建立流表的情况，如当在数据流中的首包 A 建立流表的过程中，由于该数据流报文对应的流表还未建立，导致该数据流中的报文 B 无法命中相应的流表，而出现了需要建立流表的需求，此时，称该报文 B 为伪首包，针对该情况，由于对首包 A 执行了加锁，所以伪首包 B 处于等待状态，等待首包 A 建立流表解锁后，伪首包 B 就可以根据首包 A 建立的流表进行转发。

还需要注意的是，当处理具有相同哈希值的不同类数据报文的首包时，如果当在进行第一类报文的首包的流表建立时，出现了另一类型的具有相同哈希值的报文的首包需要建立流表，由于该两类报文的首包计算获取到的哈希值相同，同时对该两类报文的首包进行处理会导致流表建立时出现冲突，造成建表错误。本实施例为了避免建立流表的冲突，所以采用加锁的方法，即在为第一类报文的首包建立流表时，执行加锁，当该类首包建表成功，解锁后，再执行第二类报文的首包的建表，从而确保了虽然是具有相同哈希值的报文，但是会在不同的地址对应的流表位置建立各自的流表，确保了建表的准确性。

本发明实施例采用普通存储空间中多级哈希表存储和 TCAM 直接存储相结合的办法，减少了存储流表的冲突概率。本发明实施例采用的 N＋1 级存储空间模式，前 N 级存储空间还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用 TCAM 实现了快速收敛，从而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

实施例 3

参见图 3，本发明实施例提供了一种报文转发的方法，本实施例仍以设置了 N+1 级别存储空间为例进行说明，其中前 N 级为普通存储设备如 SRAM\DRAM 等提供的存储空间，第 N+1 级为 TCAM 等存储设备提供的高速存储空间，利用本发明实施例提供的方法，报文转发时根据报文的关键字段采用哈希算法获得的键值在前 N 级存储空间逐级进行查找；如果前 N 级存储空间都查找失败，则直接在 TCAM 中进行快速查找，并在查找的过程中通过记录下的存储空间的地址对应的空闲流表的相关信息，当报文为首包时，可以直接通过记录的信息，将报文存储到对应的存储空间的空闲流表中。具体内容包括如下：

301：接收转发报文，获取该报文的关键字段，对关键字段采用哈希算法，得到一个哈希值。

其中，本实施例中报文的关键字段以常用的五元组（源 IP 地址,目的 IP 地址,协议号,源端口号,目的端口号）为例进行说明。

302：根据第 i ($1 \le i \le N$) 级的存储空间的大小，获取一个掩码，将获取到的该掩码与 301 中得到的哈希值进行相与操作，从而获取一个索引；根据该第 i 级存储空间的基址（起始地址）、报文的五元组大小和所述索引，计算获取一个该第 i 级存储空间的地址。

其中，根据存储空间的大小，获取一个掩码时，掩码全 1；通常，为了避免冲突的产生，存储空间越大对应着掩码取值位数越多。在获取存储空间的地址时，通常采用计算方法如下：

存储空间的地址＝存储空间的基址＋五元组大小×索引；

303：根据获取的存储空间的地址，通过查看该地址对应流表的标志位，判断该流表是否空闲，如果是，则执行 305；否则，执行 304。

304：判断该报文的五元组和该级存储空间对应的流表中存放的五元组是否完全匹配，如果是，说明该报文命中流表，执行 309；否则，说明该报文在该级存储空间进行流表查找时，出现了冲突，执行 306。

305：记录下该空闲流表的相关信息后，执行 306。

306：判断是否已经是查找第 N 级，如果是，则执行 307；否则，执行步骤 308。

307：取得该报文的五元组在 TCAM 中执行快速查找，判断是否命中相应的流表，如果是，执行 309；否则，执行 310。

308：$i = i + 1$，执行步骤 302。

309：查找该报文的相关信息，进行该报文的快速转发，结束。

本步骤中查找该报文的相关信息，主要是查找该报文对应的流表和其动作表的对应关

系，通过流表和动作表结合在一起，才能完成报文的快速转发。

310：说明该报文为首包，需要为该报文建立流表；记录 TCAM 中下该空闲流表的相关信息，然后执行 311。

311：执行加锁动作。

312：根据记录的空闲流表的相关信息，将该报文存入相应的存储空间中。

例如，在查找过程中，记录下的空闲流表的相关信息具体为存储空间的标识、存储空间的地址的标识以及空闲状态等信息。参见表 1，提供了一种记录相关信息的示意表。

表 1

存储空间的标识	存储空间的地址的标识	空闲状态
第一级	A	否
第二级	B	是
……	……	……

如表 1 所示，此时将该报文存入第一次出现空闲的存储空间中（即在第二级存储空间的地址为 B 的位置建立流表），当前 N 级的存储空间都处于非空闲状态，则直接在第 N+1 级的 TCAM 存储空间为该报文建立流表，从而实现快速收敛。

313：执行解锁动作。

本发明实施例采用普通存储空间中多级哈希表存储和 TCAM 直接存储相结合的办法，减少了存储流表的冲突概率；同时，收到报文后，通过在查找流表的过程中记录下空闲流表的相关信息的方式，当该报文为首包时，利用记录的空闲流表的相关信息为该报文建立流表，节约了建立流表的时间。本发明实施例采用的 N＋1 级存储空间模式，前 N 级存储空间还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用 TCAM 实现了快速收敛，从而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

实施例 4

参见图 4，本发明实施例提供了一种报文转发的设备，设备包括：

接收及获取模块 401，用于接收报文，获取报文的关键字段；

查找及转发模块 402，用于根据关键字段，查找第一类存储空间中是否有与接收及获取模块 401 获取到的关键字段对应的流表，如果是，根据流表转发报文；否则，根据关键字段，查找第二类存储空间是否有与关键字段对应的流表；如果是，根据在第二类存储空间中查找到的流表转发报文。

本发明实施例提供的报文转发的设备，采用多级存储空间，减少了存储流表的冲突概

率。多级存储空间的最后一级为第二类存储空间（即高速存储空间），各级还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用高速存储空间实现了快速收敛，从而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

实施例 5

参见图 5，本发明实施例提供了一种报文转发的设备，设备包括：

接收及获取模块 501，用于接收报文，获取报文的关键字段；

查找及转发模块 502，用于根据关键字段，查找第一类存储空间中是否有与接收及获取模块 501 获取到的关键字段对应的流表，如果是，根据流表转发报文；否则，根据关键字段，查找第二类存储空间是否有与关键字段对应的流表；如果是，根据在第二类存储空间中查找到的流表转发报文。

其中，该查找及转发模块 502 具体包括：

获取单元 5021，用于获取第一类存储空间的当前存储空间的流表；

查找及转发单元 5022，用于根据获取单元 5021 获取到的第一类存储空间的当前存储空间的流表，当查找到与关键字段相对应的流表，则根据查找到的与关键字段相对应的流表转发报文，否则，查找当前存储空间的下一级存储空间；当第一类存储空间的各级存储空间中没有查找到与关键字段相对应的流表时，查找第二类存储空间。

其中，获取单元 5021 具体包括：

计算子单元，用于计算关键字段的哈希值；通过当前存储空间的大小获取掩码；对哈希值与掩码相与获取索引；计算索引和关键字段的乘积值；根据乘积值与存储空间的基址之和，获取当前存储空间的地址；

获取子单元，用于根据计算子单元获取的当前存储空间的地址获取查找当前存储空间的流表。

本发明实施例提供的报文转发的设备，采用多级存储空间，减少了存储流表的冲突概率。多级存储空间的最后一级为第二类存储空间（即高速存储空间），各级还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用高速存储空间实现了快速收敛，从而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

实施例 6

参见图 6，本发明实施例提供了一种报文转发的设备，设备包括：

接收及获取模块 601，用于接收报文，获取报文的关键字段；

查找及转发模块602，用于根据关键字段，查找第一类存储空间中是否有与接收及获取模块 601 获取到的关键字段对应的流表，如果是，根据流表转发报文；否则，根据关键字段，查找第二类存储空间是否有与关键字段对应的流表；如果是，根据在第二类存储空间中查找到的流表转发报文。

建表模块603，用于当查找及转发模块602查找第二类存储空间没有与关键字段对应的流表时，为报文建立流表，其中，该建表模块603具体包括：

获取单元，用于获取报文的关键字段的哈希值；

处理单元，用于根据获取单元获取的哈希值，通过第一类存储空间的当前存储空间的大小获取掩码；对哈希值与掩码相与获取索引；计算索引和关键字段的乘积值；根据乘积值和当前存储空间的基址，获取当前存储空间中的地址，判断地址对应的流表是否空闲，如果是，在当前存储空间的地址为报文建立流表；否则，查找当前存储空间的下一级存储空间；当第一类存储空间的各级存储空间中的地址对应的流表是非空闲状态，在第二类存储空间为报文建立流表。

进一步地，建表模块603还包括：

判断单元，用于当获取单元获取到报文的关键字段的哈希值后，判断是否有相同哈希值的报文正在被加锁，

第一处理单元，用于当判断单元判断的结果为是，报文等待相同哈希值的报文解锁后，执行相应的动作；

加锁单元，用于当判断单元判断的结果为否，对报文加锁；

相应地，建表模块603还包括：

解锁单元，用于当为报文建立流表后，对被加锁的报文进行解锁。

第一处理单元具体包括：

处理子单元，用于根据获取单元获取的报文的关键字段的哈希值；判断报文和相同的哈希值的报文是否属于同一类型，如果是，则报文等待相同哈希值的报文解锁后，根据相同哈希值的报文建立的流表，进行报文的快速转发；否则，报文等待相同哈希值的报文解锁后，为报文建立流表。

本发明实施例提供的报文转发的设备，采用多级存储空间，减少了存储流表的冲突概率。多级存储空间的最后一级为第二类存储空间（即高速存储空间），各级还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用高速存储空间实现了快速收敛，从而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

实施例 7

参见图 7，本发明实施例提供了一种报文转发的设备，设备包括：

接收及获取模块 701，用于接收报文，获取报文的关键字段；

查找及转发模块 702，用于根据关键字段，查找第一类存储空间中是否有与接收及获取模块 701 获取到的关键字段对应的流表，如果是，根据流表转发报文；否则，根据关键字段，查找第二类存储空间是否有与关键字段对应的流表；如果是，根据在第二类存储空间中查找到的流表转发报文。

记录模块 703，用于当查找及转发模块 702 进行查找和转发时，记录第一类存储空间的各级存储空间中的地址对应的流表的空闲信息；

建表模块 704，用于当查找及转发模块 702 查找第二类存储空间没有与关键字段对应的流表时，为报文建立流表，其中，该建表模块 704 具体包括：

获取单元，用于获取报文的关键字段的哈希值；

处理单元，用于根据获取单元获取的哈希值，根据记录模块 703 记录的空闲信息，在各级存储空间中首次出现空闲的存储空间的地址，为报文建立流表；当第一类存储空间的各级存储空间中的地址对应的流表是非空闲状态，在第二类存储空间为报文建立流表。

进一步地，建表模块 704 还包括：

判断单元，用于当获取单元获取到报文的关键字段的哈希值后，判断是否有相同哈希值的报文正在被加锁，

第一处理单元，用于当判断单元判断的结果为是，报文等待相同哈希值的报文解锁后，执行相应的动作；

加锁单元，用于当判断单元判断的结果为否，对报文加锁；

相应地，建表模块 704 还包括：

解锁单元，用于当为报文建立流表后，对被加锁的报文进行解锁。

第一处理单元具体包括：

处理子单元，用于根据获取单元获取的报文的关键字段的哈希值；判断报文和相同的哈希值的报文是否属于同一类型，如果是，则报文等待相同哈希值的报文解锁后，根据相同哈希值的报文建立的流表，进行报文的快速转发；否则，报文等待相同哈希值的报文解锁后，为报文建立流表。

本发明实施例提供的报文转发的设备，采用多级存储空间，减少了存储流表的冲突概率。多级存储空间的最后一级为第二类存储空间（即高速存储空间），各级还可以根据流表规格需要及性能需求进行动态配置级数，最后一级采用高速存储空间实现了快速收敛，从

而实现了成本规格及性能的平衡，有利于网络运营商开展各种丰富类型的业务。

本发明实施例中的部分步骤，可以利用软件实现，相应的软件程序可以存储在可读取的存储介质中，如光盘或硬盘等。

以上所述仅为本发明的具体实施例，并不用以限制本发明，对于本技术领域的普通技术人员来说，凡在不脱离本发明原理的前提下，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权 利 要 求 书

1. 一种报文转发的方法，其特征在于，所述方法包括：

接收报文，获取所述报文的关键字段；

根据所述关键字段，查找第一类存储空间中是否有与所述关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表；如果是，根据在所述第二类存储空间中查找到的流表转发所述报文。

2．如权利要求 1 所述的报文转发的方法，其特征在于，所述第一类存储空间为多级存储空间，相应地，所述根据所述关键字段，查找第一类存储空间中是否有与所述关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表，包括：

获取所述第一类存储空间的当前存储空间的流表，当查找到与所述关键字段相对应的流表，则根据所述流表转发所述报文，否则，查找所述当前存储空间的下一级存储空间；当所述第一类存储空间的各级存储空间中均没有查找到与所述关键字段相对应的流表时，查找所述第二类存储空间是否有与所述关键字段对应的流表。

3．如权利要求 2 所述的报文转发的方法，其特征在于，所述获取所述第一类存储空间的当前存储空间的流表，具体包括：

计算所述关键字段的哈希值；

通过当前存储空间的大小获取掩码；

对所述哈希值与所述掩码相与获取索引；

计算所述索引和所述关键字段的乘积值；

根据所述乘积值与所述存储空间的基址之和，获取所述当前存储空间的地址；

根据所述当前存储空间的地址，获取所述当前存储空间的流表。

4．如权利要求 2 所述的报文转发的方法，其特征在于，所述当查找到与所述关键字段相对应的流表，则根据所述流表转发所述报文，否则，查找所述当前存储空间的下一级存储空间包括：

判断所述流表是否空闲，如果所述流表不空闲，则查找所述流表，当查找到与所述关键字段相应的流表，则根据所述流表转发所述报文；如果所述流表空闲，查找所述当前存储空间的下一级存储空间。

5．如权利要求 2 所述的报文转发的方法，其特征在于，所述方法还包括：

当查找所述第二类存储空间没有与所述关键字段对应的流表时，为所述报文建立流表。

6．如权利要求 5 所述的报文转发的方法，其特征在于,所述为所述报文建立流表，具体包括：

获取所述报文的关键字段的哈希值；通过第一类存储空间的当前存储空间的大小获取掩码；对所述哈希值与所述掩码相与获取索引；计算所述索引和所述关键字段的乘积值；根据所述乘积值和所述当前存储空间的基址，获取所述当前存储空间中的地址，判断所述地址对应的流表是否空闲，如果是，在所述当前存储空间的地址为所述报文建立流表；否则，查找所述当前存储空间的下一级存储空间；当所述第一类存储空间的各级存储空间中的地址对应的流表都是非空闲状态，在所述第二类存储空间为所述报文建立流表。

7．如权利要求 5 所述的报文转发的方法，其特征在于，所述方法还包括：记录所述第一类存储空间的各级存储空间中的地址对应的流表的空闲信息；

相应地，所述为所述报文建立流表，具体包括：

获取所述报文的关键字段的哈希值，通过查找记录的所述第一类存储空间的各级存储空间中的地址对应的流表的空闲信息，在各级存储空间中首次出现空闲的存储空间的地址，为所述报文建立流表；当所述第一类存储空间的各级存储空间中的地址对应的流表都是非空闲状态，在所述第二类存储空间为所述报文建立流表。

8．如权利要求 6 或 7 所述的报文转发的方法，其特征在于，所述获取所述报文的关键字段的哈希值之后，还包括：

判断是否有相同哈希值的报文正在被加锁，如果是，等待所述相同哈希值的报文解锁后，对所述报文执行相应的动作，如果否，则对所述报文加锁。

9．如权利要求 8 所述的报文转发的方法，其特征在于，所述方法还包括：

当为所述报文建立流表后，对所述报文解锁。

10．如权利要求 8 所述的报文转发的方法，其特征在于，所述等待所述相同哈希值的报文解锁后，对所述报文执行相应的动作，具体为：

根据获取所述报文的关键字段的哈希值，判断所述报文和所述相同哈希值的报文是否属于同一类型，如果是，则所述报文等待所述相同哈希值的报文解锁后，根据所述相同哈希值的报文建立的流表，进行报文的快速转发；否则，所述报文等待所述相同哈希值的报文解锁后，为所述报文建立流表。

11．如权利要求 4 所述的报文转发的方法，其特征在于，所述判断所述流表是否空闲具体为：

通过查看所述当前存储空间的地址对应的标识位判断所述流表是否空闲。

12．如权利要求 1 或 2 所述的报文转发的方法，其特征在于，所述第二类存储空间包

括三态内容可寻址存储器。

13．一种报文转发设备，其特征在于，所述设备包括：

接收及获取模块，用于接收报文，获取所述报文的关键字段；

查找及转发模块，用于根据所述关键字段，查找第一类存储空间中是否有与所述接收及获取模块获取到的关键字段对应的流表，如果是，根据所述流表转发所述报文；否则，根据所述关键字段，查找第二类存储空间是否有与所述关键字段对应的流表；如果是，根据在所述第二类存储空间中查找到的流表转发所述报文。

14．如权利要求 13 所述的报文转发设备，其特征在于，所述查找及转发模块具体包括：

获取单元，用于获取所述第一类存储空间的当前存储空间的流表；

查找及转发单元，用于根据所述获取单元获取到的流表，当查找到与所述关键字段相对应的流表，则根据所述流表转发所述报文，否则，查找所述当前存储空间的下一级存储空间；当所述第一类存储空间的各级存储空间中没有查找到与所述关键字段相对应的流表时，查找所述第二类存储空间。

15．如权利要求 14 所述的报文转发设备，其特征在于，所述获取单元具体包括：

计算子单元，用于计算所述关键字段的哈希值；通过当前存储空间的大小获取掩码；对所述哈希值与所述掩码相与获取索引；计算所述索引和所述关键字段的乘积值；根据所述乘积值与所述存储空间的基址之和，获取所述当前存储空间的地址；

获取子单元，用于根据所述计算子单元获取的当前存储空间的地址，获取查找所述当前存储空间的流表。

16．如权利要求 14 所述的报文转发设备，其特征在于，当所述查找及转发模块查找所述第二类存储空间没有与所述关键字段对应的流表时，所述设备还包括建表模块；

所述建表模块具体包括：

获取单元，用于获取所述报文的关键字段的哈希值；

处理单元，用于根据所述获取单元获取的哈希值，通过第一类存储空间的当前存储空间的大小获取掩码；对所述哈希值与所述掩码相与获取索引；计算所述索引和所述关键字段的乘积值；根据所述乘积值和所述当前存储空间的基址，获取所述当前存储空间中的地址，判断所述地址对应的流表是否空闲，如果是，在所述当前存储空间的地址为所述报文建立流表；否则，查找所述当前存储空间的下一级存储空间；当所述第一类存储空间的各级存储空间中的地址对应的流表都是非空闲状态，在所述第二类存储空间为所述报文建立流表。

17．如权利要求 14 所述的报文转发设备，其特征在于，所述设备还包括：

记录模块，用于当所述查找及转发模块进行查找和转发时，记录所述第一类存储空间的各级存储空间中的地址对应的流表的空闲信息；

相应地，所述建表模块，包括：

获取单元，用于获取所述报文的关键字段的哈希值；

处理单元，用于根据所述获取单元获取的哈希值，根据所述记录模块记录的空闲信息，在各级存储空间中首次出现空闲的存储空间的地址，为所述报文建立流表；当所述第一类存储空间的各级存储空间中的地址对应的流表都是非空闲状态，在所述第二类存储空间为所述报文建立流表。

18．如权利要求 16 或 17 所述的报文转发设备，其特征在于，所述建表模块还包括：

判断单元，用于当所述获取单元获取到所述报文的关键字段的哈希值后，判断是否有相同哈希值的报文正在被加锁；

第一处理单元，用于当所述判断单元判断的结果为是，所述报文等待所述相同哈希值的报文解锁后，执行相应的动作；

加锁单元，用于当所述判断单元判断的结果为否，对所述报文加锁；

解锁单元，用于当为所述报文建立流表后，对所述被加锁的报文进行解锁。

19．如权利要求 18 所述的报文转发设备，其特征在于，所述第一处理单元具体包括：

处理子单元，用于根据所述获取单元获取的所述报文的关键字段的哈希值；判断所述报文和所述相同的哈希值的报文是否属于同一类型，如果是，则等待所述相同哈希值的报文解锁后，根据所述相同哈希值的报文建立的流表，对所述报文进行快速转发；否则，所述报文等待所述相同哈希值的报文解锁后，为所述报文建立流表。

说 明 书 附 图

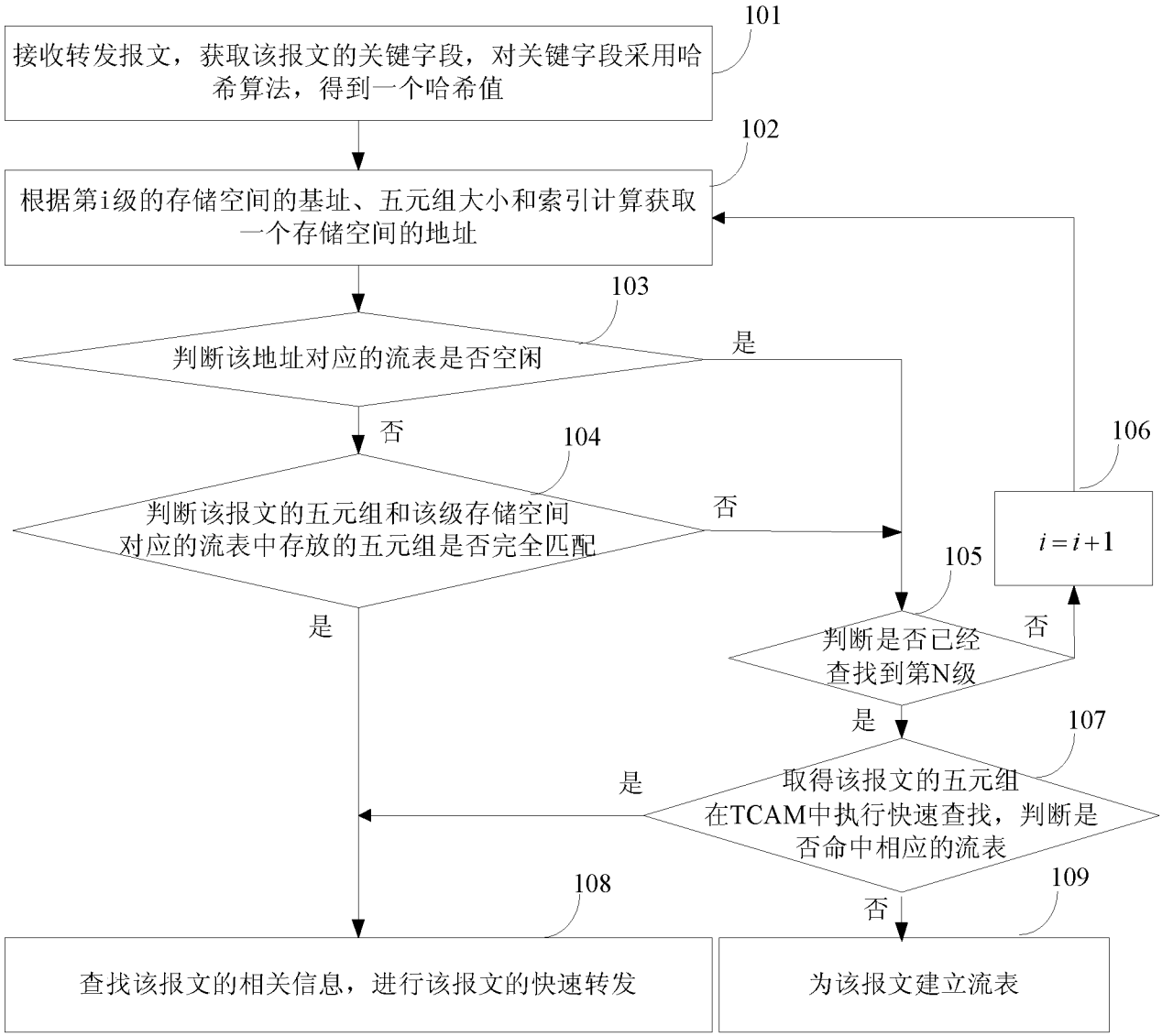
101
接收转发报文，获取该报文的关键字段，对关键字段采用哈希算法，得到一个哈希值
102
根据第i级的存储空间的基址、五元组大小和索引计算获取一个存储空间的地址
103
判断该地址对应的流表是否空闲
是
否
104
判断该报文的五元组和该级存储空间对应的流表中存放的五元组是否完全匹配
否
是
106
$i=i+1$
105
判断是否已经查找到第N级
否
是
107
取得该报文的五元组在TCAM中执行快速查找，判断是否命中相应的流表
是
否
108
查找该报文的相关信息，进行该报文的快速转发
109
为该报文建立流表

图 1

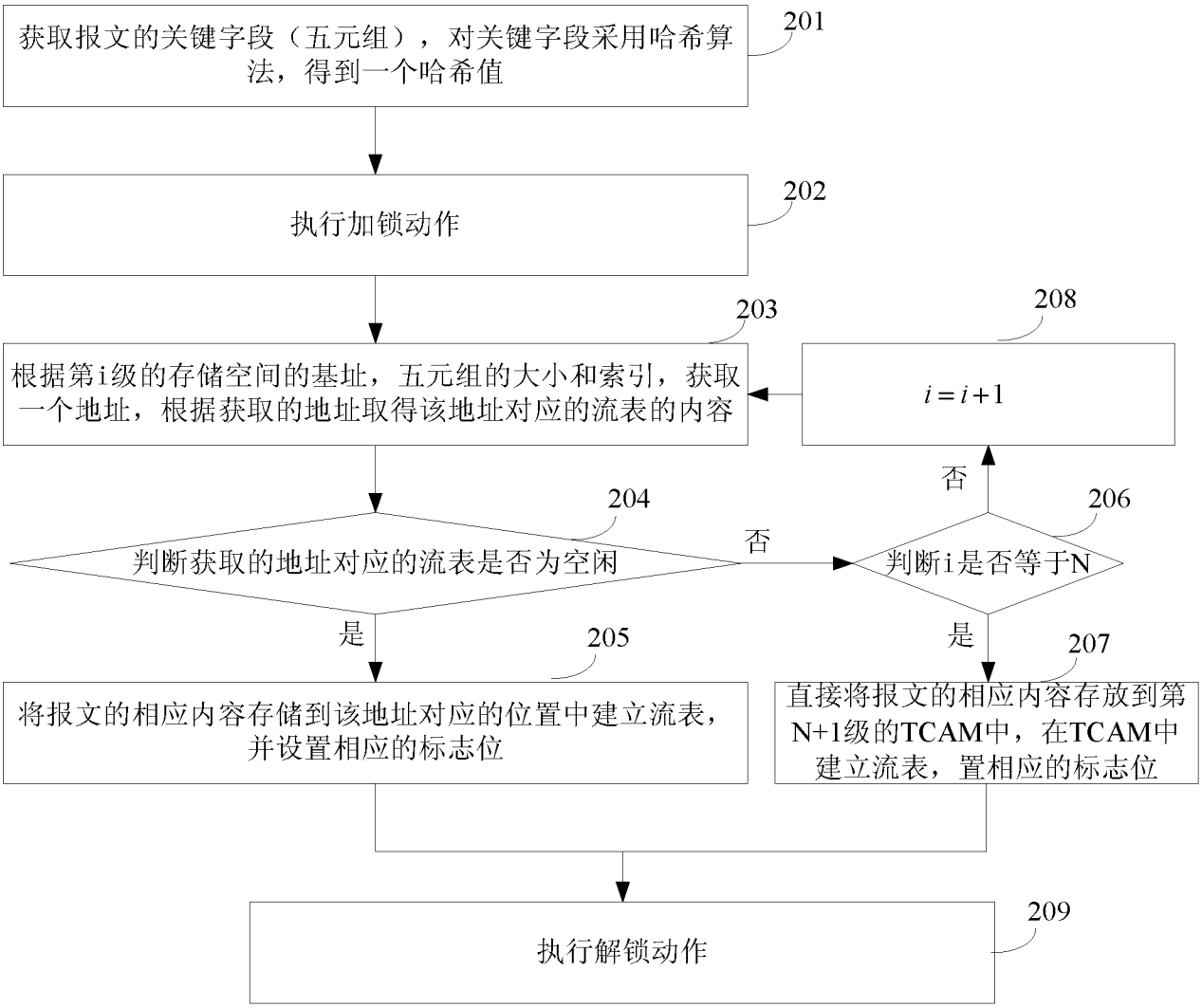
获取报文的关键字段（五元组），对关键字段采用哈希算法，得到一个哈希值
201
执行加锁动作
202
根据第i级的存储空间的基址，五元组的大小和索引，获取一个地址，根据获取的地址取得该地址对应的流表的内容
203
判断获取的地址对应的流表是否为空闲
204
否
是
将报文的相应内容存储到该地址对应的位置中建立流表，并设置相应的标志位
205
判断i是否等于N
206
否
是
i = i + 1
208
直接将报文的相应内容存放到第N+1级的TCAM中，在TCAM中建立流表，置相应的标志位
207
执行解锁动作
209

图 2

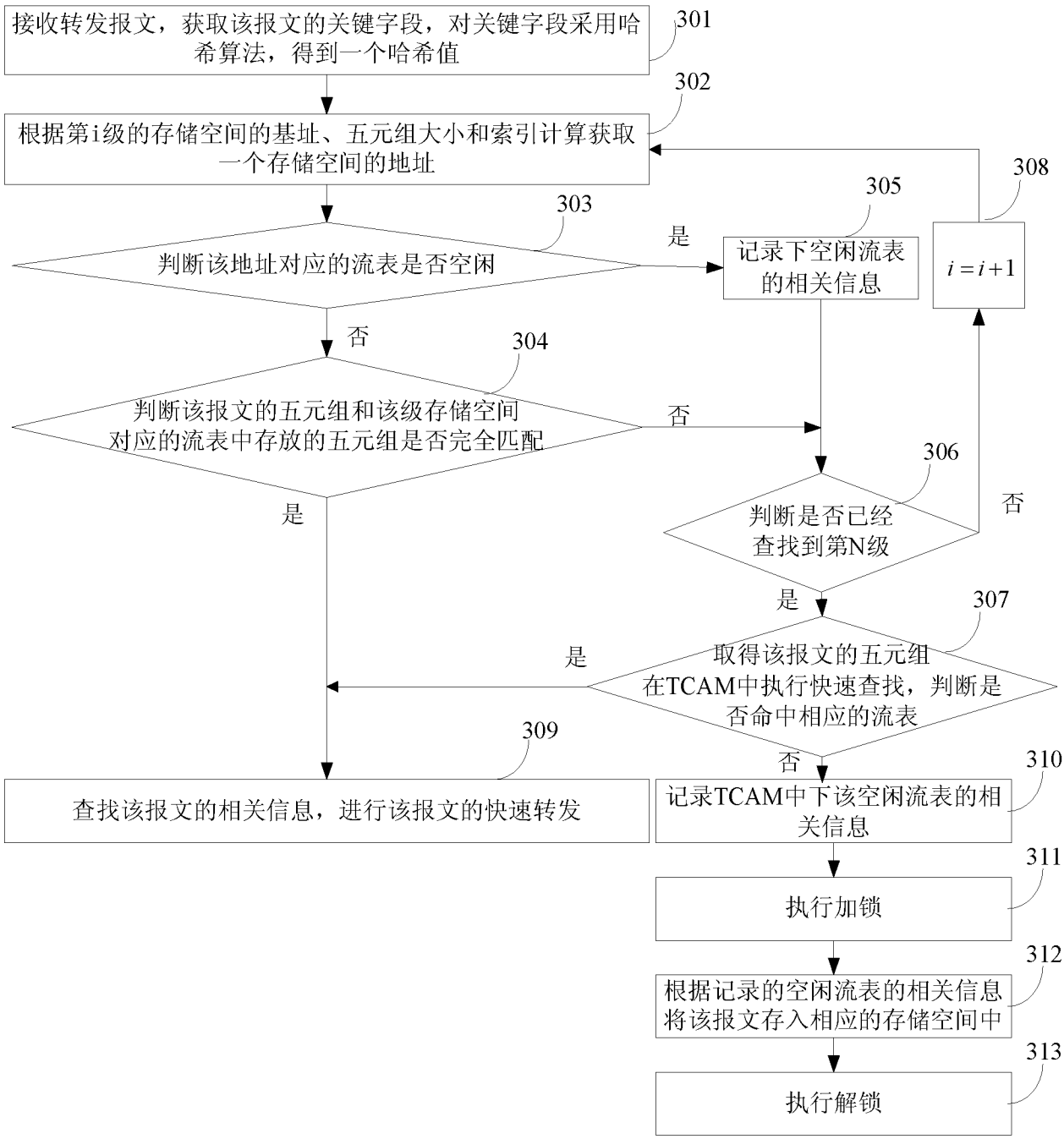
接收转发报文，获取该报文的关键字段，对关键字段采用哈希算法，得到一个哈希值
301
根据第i级的存储空间的基址、五元组大小和索引计算获取一个存储空间的地址
302
判断该地址对应的流表是否空闲
303
是
记录下空闲流表的相关信息
305
i = i + 1
308
否
判断该报文的五元组和该级存储空间对应的流表中存放的五元组是否完全匹配
304
否
是
判断是否已经查找到第N级
306
否
是
取得该报文的五元组在TCAM中执行快速查找，判断是否命中相应的流表
307
是
否
查找该报文的相关信息，进行该报文的快速转发
309
记录TCAM中下该空闲流表的相关信息
310
执行加锁
311
根据记录的空闲流表的相关信息将该报文存入相应的存储空间中
312
执行解锁
313

图 3

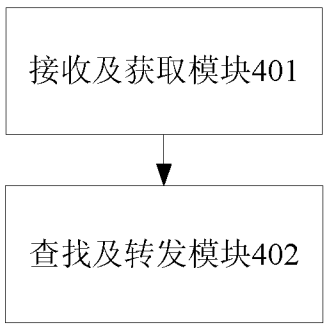
接收及获取模块401
查找及转发模块402

图 4

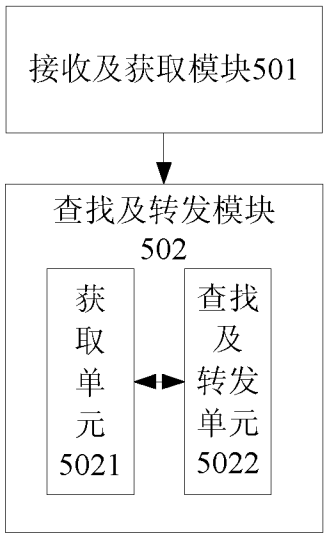
接收及获取模块501
查找及转发模块502
获取单元5021
查找及转发单元5022

图 5

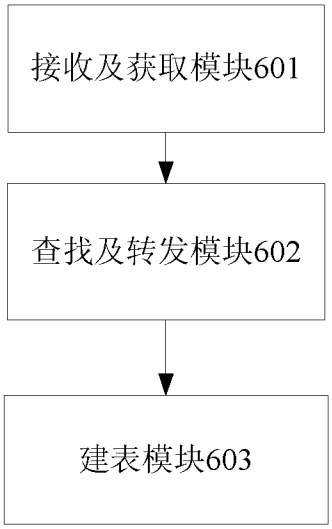
接收及获取模块601
查找及转发模块602
建表模块603

图 6

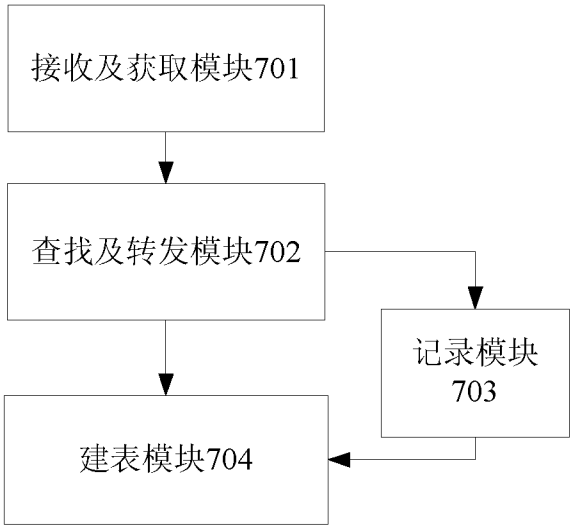
接收及获取模块701
查找及转发模块702
记录模块703
建表模块704

图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/070378

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/56 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L12/56 EC: H04L12/56C, H04L12/56S4 ICO: T04L212/002L, T04L12/56Q, T04L12/56W23

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC, PAJ： hash，TCAM，storage, memory, buffer, cache, packet, forward+, rout+, multi+, hierarchical, list, table, search+, lookup, look w up, first, second+, primary

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN101247337A ((HUAW-N) HUAWEI TECHNOLOGIES CO LTD) 20 Aug. 2008(20.08.2008) see the whole document	1-19
X	US2003214948A1 (KOREA ELECTRONICS TELECOMM [KR]) 20 Nov. 2003(20.11.2003) see the abstract, the claims	1,13
Y		2,4,5,11,12,14
Y	CN101094179A (ZHONGXING COMM CO LTD [CN]) 26 Dec. 2007(26.12.2007) see the claims	2,4,5,11,12,14
A	US2004255045A1 (UNIV EWHA IND COLLABORATION [KR]) 16 Dec. 2004(16.12.2004) see the whole document	1-19
A	CN1406073A (ZHONGXING COMM CO LTD [CN]) 26 Mar. 2003(26.03.2003) see the whole document	1-19

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 May 2009(05.05.2009)	**14 May 2009 (14.05.2009)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer PENG Rui Telephone No. (86-10)62411220

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2009/070378

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN101247337A	20.08.2008	None	
US2003214948A1	20.11.2003	KR20030089747A	28.11.2003
		KR100429904B	03.05.2004
		FR2839835A1	21.11.2003
		FR2839835B1	12.11.2004
		US7274700B2	25.09.2007
CN101094179A	26.12.2007	None	
US2004255045A1	16.12.2004	KR20040101742A	03.12.2004
		US7418505B2	26.08.2008
		KR20050043035A	11.05.2005
		KR100504387B	27.07.2005
CN1406073A	26.03.2003	CN1330190C	01.08.2007

国际检索报告

国际申请号
PCT/CN2009/070378

A. 主题的分类

H04L12/56 (2006.01) i

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L12/56 EC: H04L12/56C, H04L12/56S4 ICO: T04L212/002L, T04L12/56Q, T04L12/56W23

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

CNPAT, CNKI, WPI, EPODOC, PAJ：报文，转发，路由，多级，存储，缓存，表，搜索，查找，关键，hash，TCAM，storage, memory, buffer, cache, packet, forward+, rout+, multi+, hierarchical, list, table, search+, lookup, look w up, first, second+, primary

C. 相关文件

类　型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN101247337A (华为技术有限公司)　20.8 月 2008(20.08.2008) 参见全文	1－19
X	US2003214948A1 (KOREA ELECTRONICS TELECOMM [KR]) 20.11 月 2003(20.11.2003) 参见摘要，权利要求书	1，13
Y		2，4，5，11，12，14
Y	CN101094179A (中兴通讯股份有限公司) 26.12月 2007(26.12.2007) 参见权利要求书	2，4，5，11，12，14
A	US2004255045A1 (UNIV EWHA IND COLLABORATION [KR]) 16.12 月 2004(16.12.2004) 参见全文	1－19
A	CN1406073A (中兴通讯股份有限公司) 26.3 月 2003(26.03.2003) 参见全文	1－19

☐ 其余文件在 C 栏的续页中列出。　　☒ 见同族专利附件。

* 引用文件的具体类型：

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
05.5 月 2009(05.05.2009)	**14.5 月 2009 (14.05.2009)**
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号：　(86-10)62019451	受权官员 **彭锐** 电话号码：(86-10) **62411220**

国际检索报告

关于同族专利的信息

国际申请号
PCT/CN2009/070378

检索报告中引用的专利文件	公布日期	同族专利	公布日期
CN101247337A	20.08.2008	无	
US2003214948A1	20.11.2003	KR20030089747A	28.11.2003
		KR100429904B	03.05.2004
		FR2839835A1	21.11.2003
		FR2839835B1	12.11.2004
		US7274700B2	25.09.2007
CN101094179A	26.12.2007	无	
US2004255045A1	16.12.2004	KR20040101742A	03.12.2004
		US7418505B2	26.08.2008
		KR20050043035A	11.05.2005
		KR100504387B	27.07.2005
CN1406073A	26.03.2003	CN1330190C	01.08.2007