

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200610112323.5

[51] Int. Cl.

H04L 29/02 (2006.01)

H04L 12/56 (2006.01)

H04L 29/06 (2006.01)

H04L 12/24 (2006.01)

[43] 公开日 2007 年 2 月 28 日

[11] 公开号 CN 1921477A

[22] 申请日 2006.9.1

[21] 申请号 200610112323.5

[71] 申请人 华为数字技术有限公司

地址 100085 北京市海淀区上地信息路 3 号

[72] 发明人 宋良瑜

[74] 专利代理机构 北京三高永信知识产权代理有限公司

代理人 何文彬

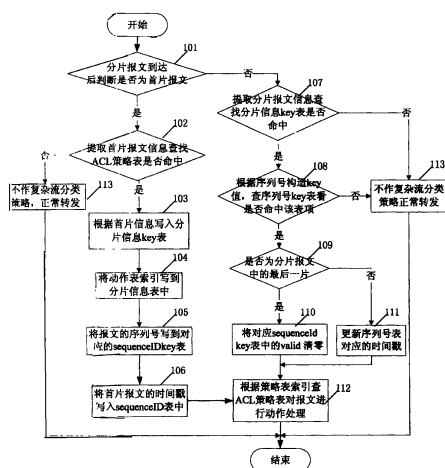
权利要求书 3 页 说明书 10 页 附图 4 页

[54] 发明名称

一种对分片报文进行复杂流分类的方法及系统

[57] 摘要

本发明提供了一种对分片报文进行复杂流分类的方法及系统，属于通信传输服务质量 QoS 技术领域。为了解决现有技术中对分片报文的后续片不能进行与首片一致的动作处理的问题，本发明提供了一种对分片报文进行复杂流分类的方法，所述方法包括对收到的分片报文进行判断，如果是首片报文，则存储首片报文信息并进行动作处理，如果是后续片报文，就查找匹配的首片报文信息，然后执行与首片报文相同的动作处理。本发明还提供了一种对分片报文进行复杂流分类的系统，所述系统包括判断模块、首片报文处理模块和后续报文处理模块。应用本发明提供的技术有效地解决了对分片报文基于传输层进行复杂流分类策略的问题，使首片与后续片报文执行相同的动作处理。



1. 一种对分片报文进行复杂流分类的方法，其特征在于，所述方法包括以下步骤：

步骤 A：收到分片报文后，判断该分片报文是否是首片报文，如果是，执行步骤 B，否则执行步骤 C；

步骤 B：提取报文中信息并存储，并根据报文携带的端口号在策略表中找到匹配的策略，执行动作操作；

步骤 C：将后续片报文的信息与存储的首片报文信息进行对比，找到对应的首片报文信息，按照对应的首片报文处理策略进行动作操作。

2. 如权利要求 1 所述的对分片报文进行复杂流分类的方法，其特征在于，所述步骤 B 具体包括：

提取首片报文的报文信息查找 ACL 策略表判断是否命中，如果命中，将首片报文的报文信息和序列号进行存储，并根据命中的策略表项进行动作处理；否则，不作复杂流分类策略处理，直接转发报文。

3. 如权利要求 2 所述的对分片报文进行复杂流分类的方法，其特征在于，所述步骤 C 具体包括：

提取后续片报文的报文信息和序列号判断是否命中存储的首片报文的报文信息和序列号，如果命中，与命中的首片报文执行相同的动作处理；否则，不作复杂流分类策略处理，直接转发报文。

4. 如权利要求 3 所述的对分片报文进行复杂流分类的方法，其特征在于，所述步骤 C 还包括判断所述后续报文是否是分片报文的最后一片报文的步骤，如果是，删除存储的序列号；否则不删除该序列号。

5. 如权利要求 4 所述的对分片报文进行复杂流分类的方法，其特征在于，

所述步骤 B 还包括对每个序列号设置一个时间戳，

相应地，步骤 C 还包括定时检测时间戳是否超过了门限值，如果超过门限值，该时间戳对应的序列号设置为无效，否则更新时间戳，继续进行检测，当一个报文所有序列号都无效

时，则将此报文的所有序列号删除。

6. 一种对分片报文进行复杂流分类的系统，其特征在于，所述系统包括以下模块：

判断模块，用于收到分片报文后，判断该分片报文是否是首片报文；

首片报文处理模块，用于当收到的报文是首片报文时，提取报文中信息并存储，然后根据报文携带的端口号在策略表中找到匹配的策略，执行动作操作；

后续片报文处理模块，用于当收到的报文是后续片报文时，将后续片报文的信息与存储的首片报文信息进行对比，找到对应的首片报文信息，按照对应的首片报文处理策略进行动作操作。

7. 如权利要求 6 所述的对分片报文进行复杂流分类的系统，其特征在于，所述首片报文处理模块具体包括：

首片报文处理单元，用于提取首片报文的报文信息查找 ACL 策略表判断是否命中，如果命中，将首片报文的报文信息和序列号进行存储，并根据命中的策略表项进行动作处理；否则，不作复杂流分类策略处理，直接转发报文。

8. 如权利要求 7 所述的对分片报文进行复杂流分类的系统，其特征在于，所述后续片报文处理模块具体包括：

后续片报文处理单元，用于提取后续片报文的报文信息和序列号判断是否命中存储的首片报文的报文信息和序列号，如果命中，与命中的首片报文执行相同的动作处理；否则，不作复杂流分类策略处理，直接转发报文。

9. 如权利要求 8 所述的对分片报文进行复杂流分类的系统，其特征在于，所述后续片报文处理单元具体还包括：

后续片报文处理子单元，用于判断所述后续报文是否是分片报文的最后一块报文，如果是，删除存储的序列号；否则不删除该序列号。

10. 如权利要求 7 所述的对分片报文进行复杂流分类的系统，其特征在于，所述首片报文处理单元具体还包括：

设置时间戳子单元，用于对每个序列号设置一个时间戳；

相应地，后续片报文处理单元还包括：

定时检测子单元，用于定时检测时间戳是否超过了门限值，如果超过门限值，该时间戳对应的序列号设置为无效，否则更新时间戳，继续进行检测，当一个报文所有序列号都无效时，则将此报文的所有序列号删除。

一种对分片报文进行复杂流分类的方法及系统

技术领域

本发明涉及通信传输服务质量 QoS (Quality of Service) 技术领域, 特别涉及一种对分片报文进行复杂流分类的方法及系统。

背景技术

随着计算机网络的高速发展, 除了传统的 WWW、E-Mail、FTP (File Transfer Protocol, 文件传输协议) 应用外, 用户已经在 Internet 上拓展新业务, 比如远程教学、远程医疗、可视电话、电视会议、视频点播等。企业用户也希望通过 VPN (Virtual Private Network, 虚拟专用网络) 技术, 将分布在各地的分支机构连接起来, 开展一些事务性应用。这些新业务有一个共同特点, 即对带宽、延迟、延迟抖动等传输性能有着特殊的需求。

新业务的不断涌现对 IP 网络的服务能力提出了更高的要求, 用户已不再满足于能够简单地将报文送达目的地, 还希望在投递过程中得到更好的服务, 诸如支持为用户提供专用带宽、减少报文的丢失率、管理和避免网络拥塞、调控网络的流量、设置报文的优先级, 也就是根据不同的数据流提供不同的服务。所有这些, 都要求网络应当具备更为完善的服务能力。

服务质量 QoS 是各种存在服务供需关系的场合中普遍存在的概念, 它评估服务方满足客户服务需求的能力。目前 QoS 的模型目前一般采用以下两种方式, 集成业务 (Inter-Serv) 模型和差分业务 (Diff-Serv) 模型。而 Diff-Serv 模型主要采用简单流分类 (Behavior traffic classification) 和复杂流分类 ACL (Access Control List, 访问控制列表), 其中复杂流分类具体就是通过配置报文头中的五元组 (DSCP (Differentiated Services Code Point, 差分服务) 代码点值、协议类型、IP 地址、传输层端口号、分片报文的类型) 等参数对不同的业务提供差别服务, 实现基于复杂流分类的流量策略。

在网络端到端报文传输中, 如从网络 N1 端到网络的 N2 端报文传输中, 很多时候是采用分片报文传输, 首片报文头在该报文中有一个协议域。

协议域的定义参见 RFC790, 如表 1。

Version	IHL	ToS	Total Length	
Indentification			Flags	Fragment Offset
TTL	Protocol_number		Header Checksum	
Source IP				
Destination IP				
传输层报文头				
Payload				

表 1

后续片报文头不带传输层报文头，如表 2。

Version	IHL	ToS	Total Length	
Indentification			Flags	Fragment Offset
TTL	Protocol_number		Header Checksum	
Source IP				
Destination IP				
Payload				

表 2

如图1所示：N1路由器的入口接收到分片报文，该报文被分成3片，此时在入口配置基于传输层端口的ACL策略，结果由于frag1带有传输层端口号，frag1命中策略，并且根据策略进行了相应的动作（remark，car，mirror等），从N1路由器出去的时候，该报文已经被修改（修改优先级或者根据car配置丢弃），而由于从N1进来的frag2和frag3由于不带传输层报文头，没有传输层端口号，那么他们不会命中策略，结果他们没有进行策略动作处理（比如配置了基于端口的remark，只能对首片的ToS(Type of Service，服务类)域进行修改，而不能够修改后续片的ToS域），只是正常IP转发。这样对于同一条流，首片和后续片在整个转发过程中处理不同，对传输层处理造成了困难。

现有技术的缺点是：对分片报文的首片和后续片报文不能进行相同的策略动作处理。

发明内容

本发明为了克服现有技术中存在的复杂流分类分片报文中的后续片不能和首片一样实现基于传输层端口进行动作处理的问题，本发明提供了一种对分片报文进行复杂流分类的方

法及系统，具体方案如下：

本发明提供了一种对分片报文进行复杂流分类的方法，所述方法包括以下步骤：

步骤 A：收到分片报文后，判断该分片报文是否是首片报文，如果是，执行步骤 B，否则执行步骤 C；

步骤 B：提取报文中信息并存储，并根据报文携带的端口号在策略表中找到匹配的策略，执行动作操作；

步骤 C：将后续片报文的信息与存储的首片报文信息进行对比，找到对应的首片报文信息，按照对应的首片报文处理策略进行动作操作。

所述步骤 B 具体包括：

提取首片报文的报文信息查找 ACL 策略表判断是否命中，如果命中，将首片报文的报文信息和序列号进行存储，并根据命中的策略表项进行动作处理；否则，不作复杂流分类策略处理，直接转发报文。

所述步骤 C 具体包括：

提取后续片报文的报文信息和序列号判断是否命中存储的首片报文的报文信息和序列号，如果命中，与命中的首片报文执行相同的动作处理；否则，不作复杂流分类策略处理，直接转发报文。

所述步骤 C 还包括判断所述后续报文是否是分片报文的最后一片报文的步骤，如果是，删除存储的序列号；否则不删除该序列号。

所述步骤 B 还包括对每个序列号设置一个时间戳，

相应地，步骤 C 还包括定时检测时间戳是否超过了门限值，如果超过门限值，该时间戳对应的序列号设置为无效，否则更新时间戳，继续进行检测，当一个报文所有序列号都无效时，则将此报文的所有序列号删除。

本发明还提供了一种对分片报文进行复杂流分类的系统，所述系统包括以下模块：

判断模块，用于收到分片报文后，判断该分片报文是否是首片报文；

首片报文处理模块，用于当收到的报文是首片报文时，提取报文中信息并存储，然后根据报文携带的端口号在策略表中找到匹配的策略，执行动作操作；

后续片报文处理模块，用于当收到的报文是后续片报文时，将后续片报文的信息与存储的首片报文信息进行对比，找到对应的首片报文信息，按照对应的首片报文处理策略进行动作操作。

所述首片报文处理模块具体包括：

首片报文处理单元，用于提取首片报文的报文信息查找 ACL 策略表判断是否命中，如果命中，将首片报文的报文信息和序列号进行存储，并根据命中的策略表项进行动作处理；否则，不作复杂流分类策略处理，直接转发报文。

所述后续片报文处理模块具体包括：

后续片报文处理单元，用于提取后续片报文的报文信息和序列号判断是否命中存储的首片报文的报文信息和序列号，如果命中，与命中的首片报文执行相同的动作处理；否则，不作复杂流分类策略处理，直接转发报文。

所述后续片报文处理单元具体还包括：

后续片报文处理子单元，用于判断所述后续报文是否是分片报文的最后一片报文，如果是，删除存储的序列号；否则不删除该序列号。

所述首片报文处理单元具体还包括：

设置时间戳子单元，用于对每个序列号设置一个时间戳；

相应地，后续片报文处理单元还包括：

定时检测子单元，用于定时检测时间戳是否超过了门限值，如果超过门限值，该时间戳对应的序列号设置为无效，否则更新时间戳，继续进行检测，当一个报文所有序列号都无效时，则将此报文的所有序列号删除。

应用本发明提供的技术方案带来的有益效果是：

(1) 通过对首片报文的报文信息进行存储，有效地解决了实际网络中对分片报文基于传输层进行复杂流分类策略的问题，使首片报文与后续片报文执行相同的动作处理；

(2) 通过对每片报文设置时间戳，节省了内存空间。

附图说明

图 1 是现有技术分片报文传输示意图；

图 2 是不同流分片报文传输示意图；

图 3 是本发明实施例 1 提供的对分片报文进行复杂流分类的方法流程图；

图 4 是本发明实施例 2 提供的分片报文定时检测流程图；

图 5 是本发明实施例 3 提供的对分片报文进行复杂流分类的系统示意图。

具体实施方式

下面结合附图和具体实施例对本发明作进一步说明，但不作为对本发明的限定。

本发明通过对分片报文的首片报文提取报文信息，后续片报文对比此首片报文提取的信息，根据对比结果进行相应的策略操作。从而实现首片报文与后续片报文进行相同的策略动作处理。

在应用本发明提供的技术方案时，上层为每个需要进行策略处理的数据流下发基于TCP/UDP 端口号的 ACL 策略，该策略对应一条 policy entry（策略表项），该表项记录在策略动作表（policy_action_table）中，并以 TCP/UDP 端口为掩码。同一数据流配置的动作相同。

对于普通报文，如果报文中的端口号与策略端口号匹配，则查策略动作表，作正常动作处理；对于分片报文，对首片和后续片报文进行不同的处理。

对分片报文具体处理方案如下：

实施例 1

本实施例采用每当收到分片报文的首片时，如果该报文内的 TCP/UDP 端口号匹配策略端口号，会命中该策略，此时首片报文需要将源 IP 地址（SIP），目的 IP 地址（DIP），协议号，分片标志位 frag_flag 写入分片信息 key 表，将动作表索引（action_index）写入分片信息表（frag_infor_table），同时将序列号 sequence_id 写入序列号 key 表。然后根据动作表索引（action_index），查找策略动作表（policy_action_table），根据策略动作表所配置动作（filter, remark, car 等）作相应的动作处理。

提取以上各信息的原因是：

由于后续片报文只带有 IP 头，不能匹配到策略所配置的传输层相关参数，所以对于后续片不应该按照普通报文构造相关 key 值查找策略动作表 policy_action_table，对于后续报文，需要判断后续片报文头和首片报文提取出的信息（上一段所描述的源 IP 地址（SIP），目的 IP 地址（DIP），协议号，分片标志位 frag_flag）是否匹配，来确认两者是否为同一条数据流，如果是同一条数据流，根据序列号 key 表中的序列号找到该后续报文对应的首片报文，这样就可以使用相应分片信息表项中的动作表索引来查找动作表，作相应的动作。

本实施例采用表格的形式存储首片报文的信息，具体应用上不局限于表格形式，可以根据需要自行设计首片报文信息与后续片报文信息间的对应关系。

参见图 3，一种对分片报文进行复杂流分类的方法，具体包括以下步骤：

步骤 101：分片报文到达路由器后，判断该分片报文是否是首片报文，如果是，进行步骤 102；否则进行步骤 107。

具体判断报文是否为首片报文是通过判断该片是否有传输层报文头（带有传输层端口号），如果有则说明是首片报文，否则是后续报文。

本实施例以图 2 所示的 2 个不同数据流报文的分片报文为例，其中 2 个不同数据流报文分别用 pkt1, pkt2 表示，pkt1 包括首片 pkt1_frag1、后续片 pkt1_frag2、pkt1_frag3，pkt2 包括首片 pkt2_frag1、后续片 pkt2_frag2、pkt2_frag3。

如图 2 所示，最先达到路由器 N 的报文是首片 pkt1_frag1。

步骤 102：提取首片报文中的报文信息（源 IP 地址（SIP），目的 IP 地址（DIP），协议号，分片标志位 frag_flag）查找 ACL 策略表判断是否命中。如果命中，执行步骤 103，否则执行步骤 113。

策略 key 表的数据结构如表 3 所示：

Rule_id	flag	tos	sip	dip	protocol	Tcp/Udp端口号	传输层flag
	掩掉	掩掉	掩掉	掩掉			

表3

由于策略只配置了基于传输层的 ACL 策略表，提取首片报文中的报文信息，需要将其他的部分掩掉，仅根据 Rule_id 和报文中的传输层端口号和传输层其他信息进行策略表的查找。

步骤 103：策略命中，将首片报文的信息按照表 4 的格式写入分片信息 key 表。

分片信息 key 表的数据结构如表 4：

Rule_id	Valid_flag	sip	dip	protocol	Frag_flag	ttl	tos
---------	------------	-----	-----	----------	-----------	-----	-----

表4

具体字段表示的含义为：

Valid_flag: key 表有效标志位，该标志位等于 0，表示该表项无效；该标志位等于 1，表示该表项有效。

Rule_id: ACL 配置规则 id。

其他字段从报文头中取出，和 IP 报文头定义一样。

步骤 104：将首片报文查到的策略动作表索引（策略动作表的地址）写入分片信息 key 表所对应的分片信息表。

分片信息表的数据结构如表 5：

Action_index

表5

该表中字段Action_index的含义是：它是ACL 策略表的索引，通过该索引查找ACL策略表，根据策略表内容作相应的动作（filter, remark, car...）。

步骤 105：将首片报文头中的序列号写入序列号 key 表。

序列号key表的数据结构如表6：

Rule_id	Valid_flag	Sequence_id
---------	------------	-------------

表6

其中，Rule_id：表示 ACL 配置规则 id；

Valid_flag：该标志位等于 1，表示该表项有效；该标志位等于 0，表示该表项无效；

Sequence_id：确认后续片与首片是否为同一报文的唯一标志。

步骤 106：将首片报文此时的时间戳写入该序列号 key 表对应的序列号表，然后执行步骤 112。

序列号表的数据结构如表 7：

Pkt_timestamp

表 7

Pkt_timestamp：该值表示首包匹配动作后，写入的时间戳，用于定时检测。

步骤 107：后续片报文，提取 IP 报文头中信息，查找分片信息 key 表，如果命中，说明后续片报文与表中命中的报文属于同一数据流报文，会返回该分片信息 key 表项对应的分片信息表项。执行步骤 108。否则，表明后续片报文在分片信息 key 表中没有相同的数据流，即没有匹配的 ACL 策略，执行步骤 113。

步骤 108：根据后续片报文中的序列号构造 key 值，查找序列号 key 表，看是否可以命中该表项，如果命中，执行步骤 109，如果没有命中，执行步骤 113。

步骤 109：判断是否该后续片报文是报文的最后一块，如果是，执行步骤 110，否则执行步骤 111。

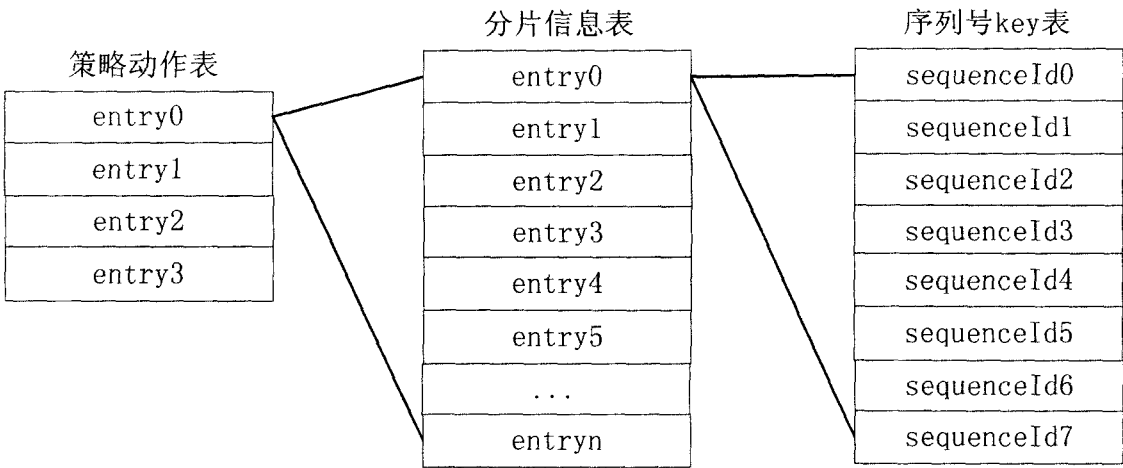
步骤 110：将对应的序列号 key 表中的 valid 标志位清零，表示该序列号表项已经释放，然后执行步骤 112。

步骤 111：更新序列号表中对应的时间戳，然后执行步骤 112。

步骤 112：根据策略表索引查 ACL 策略表对首片报文进行相应的动作处理并结束。

步骤 113：不作复杂流分类策略处理，走正常转发流程。

策略表、分片信息表和序列号 key 表的表项对应关系如下：



考虑到同一个流的分片报文在 8 个报文间隔后应该已经能够到达，所以设计的时候取经验值，每个流可以同时存放 8 个分片报文（一个流存放多少个报文可以根据实际情况制定）。

其中上述分片信息 key 表的表项和分片信息表中的表项是一一对应的，并且序列号 key 表的表项和序列号表的表项也是一一对应的，在实现过程中，通过构造 key 值查找相应 key 表，如果命中，会返回 key 表中表项对应的信息表或者序列表的表项。

实施例 2

在网络堵塞的时候可能会发生分片报文中片或者最后一块报文丢失的情况。对于中间片报文丢失的情况，由于报文分片后，在路由器上是感知不到的，只能在报文到达终端后，在高层重组报文时，才能感知，然后根据该层的一些机制（比如重传机制）处理。对于最后一块报文丢失的情况，由于分片信息表（frag_info_table）和序列号表（sequenceID_table）一直等不到最后一块报文，资源一直被占用，不能够被释放。这样对于其他流量，不能分配到空间，所以本发明为序列号表中每一个表项定义一个时间戳，如果一定时间以后，该序列号还没有被释放，那么认为该报文已经被丢弃，老化相应序列号和分片信息表中对应的表项。

参见图 4，针对网络堵塞的情况，对分片信息表的具体检测更新过程如下：

步骤 201：定时轮循检测有效的分片信息 key 表（valid 标志置 1，表示该表项有效），对于有效的分片信息 key 表，检测每个分片信息表下挂的序列号表（通过设置序列号 key 值结构中的 valid 标志位为 1，看是否能和序列号 key 匹配），判断是否命中，如果命中，说明该表项有效，执行步骤 202，如果不能命中，则表示分片信息表下挂的所有的序列号表项无效，执行步骤 203。

步骤 202：检测表项中时间戳是否超过了门限，如果超过了门限，执行步骤 204；否则执行步骤 205。

步骤 203: 将对应的分片信息表释放。

步骤 204: 设置对应的序列号 key 表中的 Valid_flag 等于 0, 即该序列号表项无效。

步骤 205: 更新时间戳 (每个报文的发送过程中, 硬件上都有一个时钟计时器, 通过该时钟计时器的显示进行时间戳的更新), 当一个分片信息表中所下挂的所有序列号 key 表中的表项都已经 invalid(无效), 则将此信息表释放, 具体释放是通过将分片信息 key 表中的 valid 标志位清零。

本发明的上述方法对于网络中 QoS 的扩展应用, 有非常重要的意义。

实施例 3

本发明还提供了一种对分片报文进行复杂流分类的系统, 包括以下模块:

判断模块, 用于收到分片报文后, 判断该分片报文是否是首片报文;

首片报文处理模块, 用于当收到的报文是首片报文时, 提取报文中信息并存储, 然后根据报文携带的端口号在策略表中找到匹配的策略, 执行动作操作;

后续片报文处理模块, 用于当收到的报文是后续片报文时, 将后续片报文的信息与存储的首片报文信息进行对比, 找到对应的首片报文信息, 按照对应的首片报文处理策略进行动作操作。

其中, 首片报文处理模块具体包括:

首片报文处理单元, 用于提取首片报文的报文信息查找 ACL 策略表判断是否命中, 如果命中, 将首片报文的报文信息和序列号进行存储, 并根据命中的策略表项进行动作处理; 否则, 不作复杂流分类策略处理, 直接转发报文。

后续片报文处理模块具体包括:

后续片报文处理单元, 用于提取后续片报文的报文信息和序列号判断是否命中存储的首片报文的报文信息和序列号, 如果命中, 与命中的首片报文执行相同的动作处理; 否则, 不作复杂流分类策略处理, 直接转发报文。

为了节约内存, 后续片报文处理单元具体还包括:

后续片报文处理子单元, 用于判断所述后续报文是否是分片报文的最后一块报文, 如果是, 删除存储的序列号; 否则不删除该序列号。

首片报文处理单元具体还包括:

设置时间戳子单元, 用于对每个序列号设置一个时间戳;

相应地, 后续片报文处理单元还包括:

定时检测子单元, 用于定时检测时间戳是否超过了门限值, 如果超过门限值, 该时间戳

对应的序列号设置为无效，否则更新时间戳，继续进行检测，当一个报文所有序列号都无效时，则将此报文的所有序列号删除。

以上所述的实施例，只是本发明较优选的具体实施方式的一种，本领域的技术人员在本发明技术方案范围内进行的通常变化和替换都应包含在本发明的保护范围内。

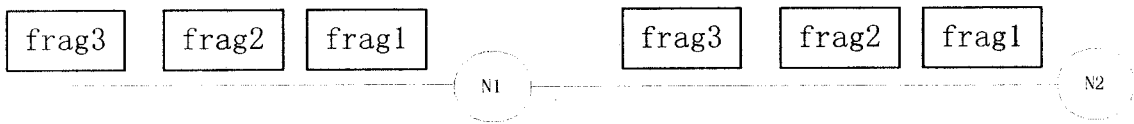


图 1

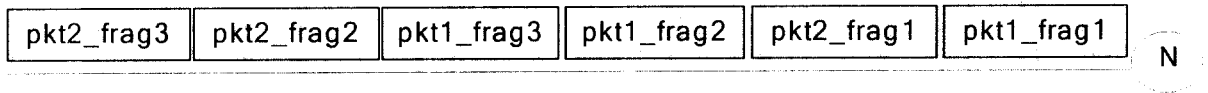


图 2

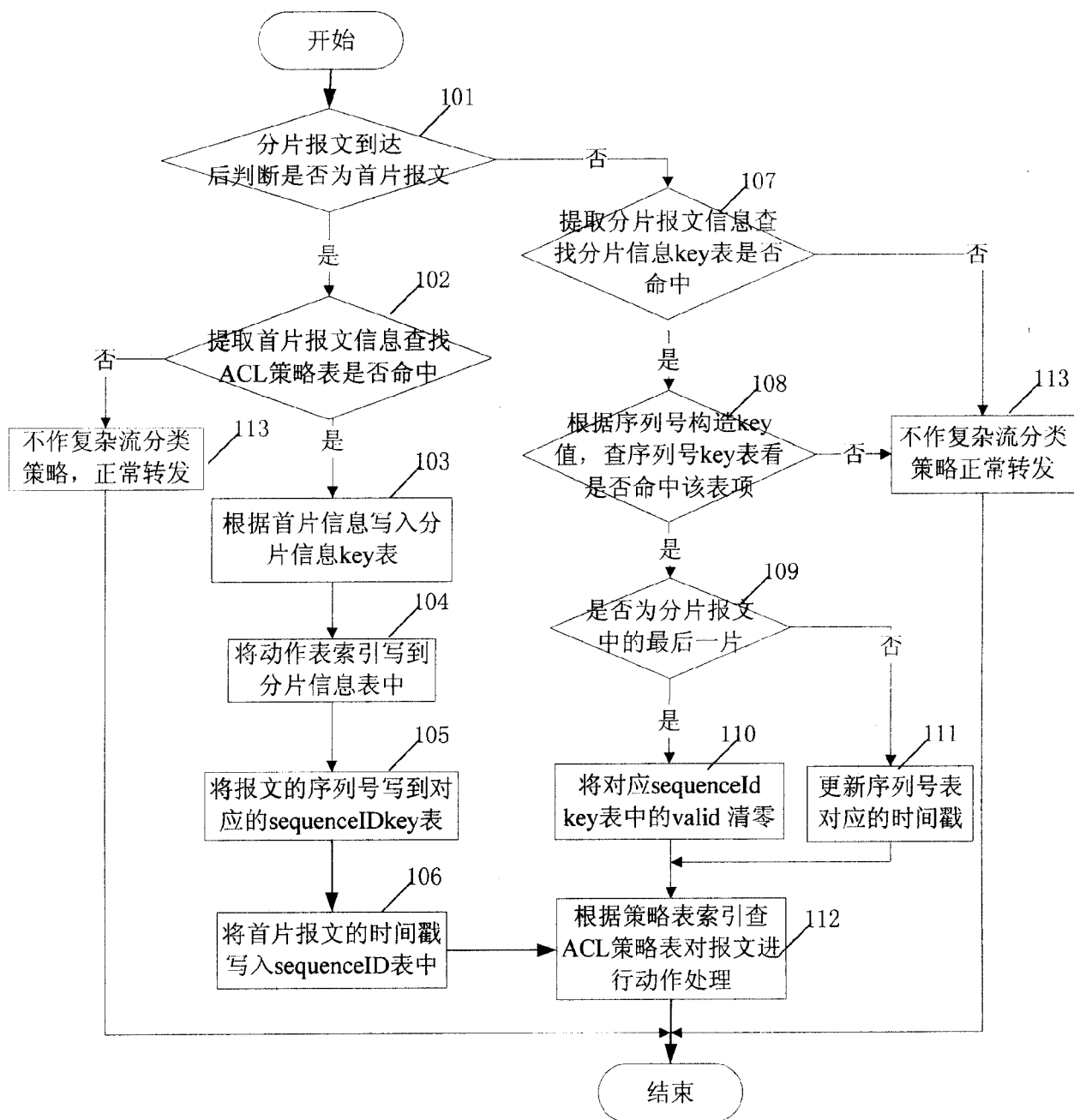


图 3

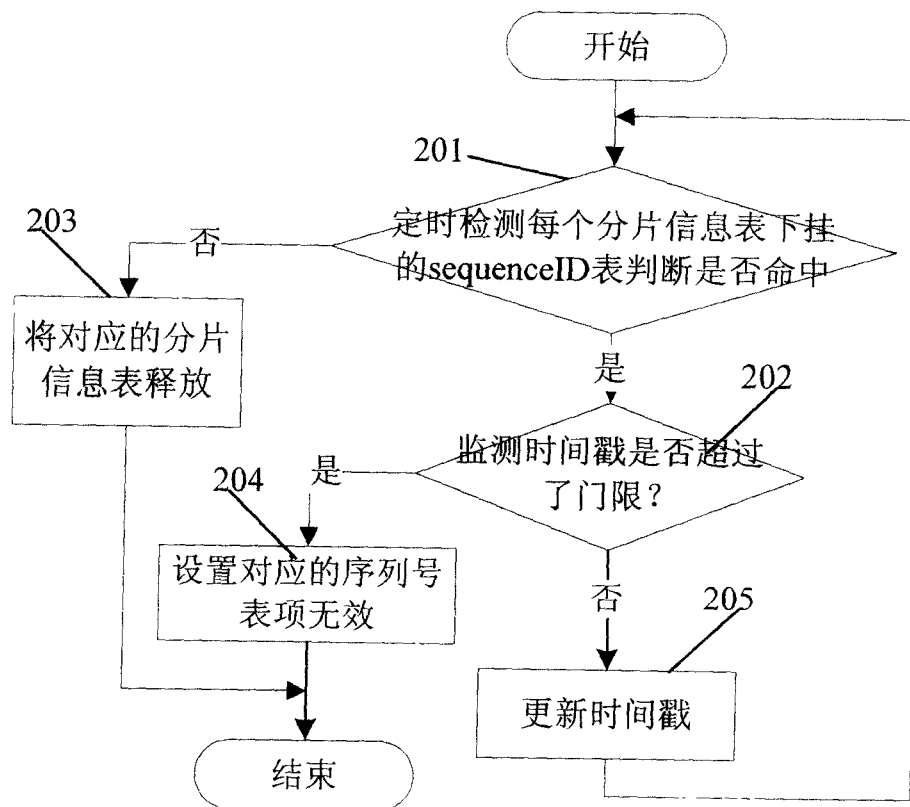


图 4

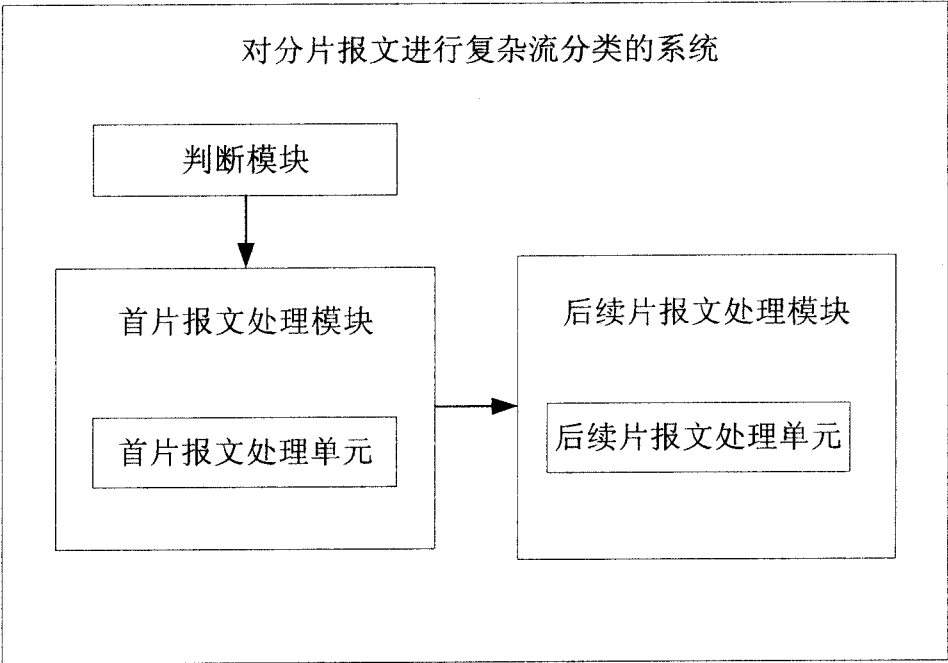


图 5