

(12) 发明专利

(10) 授权公告号 CN 102301663 B

(45) 授权公告日 2013. 11. 06

(21) 申请号 201180001146. 2

(51) Int. Cl.

(22) 申请日 2011. 07. 06

H04L 12/70(2013. 01)

(85) PCT申请进入国家阶段日  
2011. 08. 27

(56) 对比文件

CN 1466340 A, 2004. 01. 07,  
CN 101834783 A, 2010. 09. 15,  
WO 2011/049135 A1, 2011. 04. 28,

(86) PCT申请的申请数据  
PCT/CN2011/076905 2011. 07. 06

审查员 王歆玥

(87) PCT申请的公布数据  
W02012/106869 ZH 2012. 08. 16

(73) 专利权人 华为技术有限公司  
地址 518129 中国广东省深圳市龙岗区坂田  
华为总部办公楼

(72) 发明人 林程勇

(74) 专利代理机构 深圳市深佳知识产权代理事  
务所(普通合伙) 44285  
代理人 彭愿洁 李文红

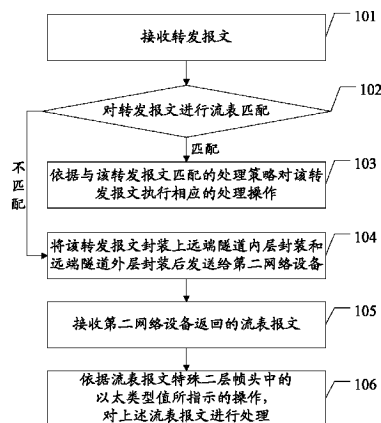
权利要求书3页 说明书13页 附图5页

(54) 发明名称

一种报文处理方法及相关设备

(57) 摘要

本发明实施例公开了一种报文处理方法及相关设备,其中,一种报文处理方法包括:第一网络设备接收转发报文;对所述转发报文进行流表匹配,若不匹配,则将所述转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便所述第二网络设备将所述转发报文转发给控制服务器处理;接收所述第二网络设备返回的流表报文,所述流表报文中携带所述转发报文的流表信息,且封装有特殊二层帧头;依据所述特殊二层帧头中的以太类型值所指示的操作,对所述流表报文进行处理。本发明提供的技术方案能有效解决网络设备控制平面的瓶颈限制问题,同时实现数据转发和路由控制的分离。



1. 一种报文处理方法,其特征在于,包括:  
第一网络设备接收转发报文;  
对所述转发报文进行流表匹配,若不匹配,则,  
将所述转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便所述第二网络设备将所述转发报文转发给控制服务器处理;  
接收所述第二网络设备返回的流表报文,所述流表报文中携带所述转发报文的流表信息,且封装有特殊二层帧头;  
依据所述特殊二层帧头中的以太类型值所指示的操作,对所述流表报文进行处理。
2. 根据权利要求1所述的方法,其特征在于,  
所述远端隧道内层封装具体为:多协议标签交换 MPLS 标签;  
所述远端隧道外层封装具体为:MPLS 标签,或通用路由封装协议 GRE 封装,或以太网协议安全协议 IPSec 封装。
3. 根据权利要求1或2所述的方法,其特征在于,  
所述发送给第二网络设备具体为:  
通过与所述第二网络设备建立的伪线 PW 隧道发送给第二网络设备。
4. 根据权利要求1或2所述的方法,其特征在于,  
所述对所述流表报文进行处理具体为:  
直接环回所述流表报文,或,透传所述流表报文,或对所述流表报文进行流表下发处理。
5. 根据权利要求1或2所述的方法,其特征在于,  
所述将所述转发报文封装上远端隧道内层封装和远端隧道外层封装的步骤还包括:  
在所述转发报文中携带所述转发报文的报文控制信息 PCI。
6. 一种报文处理方法,其特征在于,包括:  
网络设备接收转发报文;  
对所述转发报文进行流表匹配,若不匹配,则,  
将所述转发报文封装上二层帧头后,发送给控制服务器处理;  
接收所述控制服务器返回的流表报文,所述流表报文中携带所述转发报文的流表信息,且封装有特殊二层帧头;  
依据所述特殊二层帧头中的以太类型值所指示的操作,对所述流表报文进行处理。
7. 根据权利要求6所述的方法,其特征在于,  
所述对所述流表报文进行处理具体为:  
直接环回所述流表报文,或,透传所述流表报文,或依据所述流表报文生成流表。
8. 根据权利要求6或7所述的方法,其特征在于,  
所述发送给控制服务器处理的步骤之前包括:  
在所述转发报文中携带所述网络设备的标识。
9. 一种报文处理方法,其特征在于,  
控制服务器接收来自网络设备的转发报文;  
对所述转发报文进行解析,生成所述转发报文的流表信息;  
生成承载所述流表信息的流表报文,所述流表报文封装有特殊二层帧头,以便处理所

述流表报文的网络设备依据所述特殊二层帧头中的以太类型值指示的操作,对所述流表报文进行处理;

将所述流表报文发送给所述网络设备。

10. 根据权利要求 9 所述的方法,其特征在于,

所述转发报文中携带处理所述流表报文的网络设备的标识;

所述流表报文中携带处理所述流表报文的网络设备的标识。

11. 一种网络设备,其特征在于,包括:

接收单元,用于接收转发报文;

流表匹配单元,用于对所述接收单元接收到的转发报文进行流表匹配;

封装发送单元,用于当所述流表匹配单元的匹配结果不匹配时,将所述转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便所述第二网络设备将所述转发报文转发给控制服务器处理;

所述接收单元还用于接收所述第二网络设备返回的流表报文,所述流表报文中携带所述转发报文的流表信息,且封装有特殊二层帧头;

处理单元,用于依据所述流表报文的特殊二层帧头中的以太类型值所指示的操作,对所述流表报文进行处理。

12. 根据权利要求 11 所述的网络设备,其特征在于,

所述封装发送单元在所述转发报文上封装的远端隧道内层封装具体为多协议标签交换 MPLS 标签;

所述封装发送单元在所述转发报文上封装的远端隧道外层封装具体为 MPLS 标签,或通用路由封装协议 GRE 封装,或以太网协议安全协议 IPSec 封装。

13. 根据权利要求 11 或 12 所述的网络设备,其特征在于,

所述封装发送单元具体通过与所述第二网络设备建立的伪线 PW 隧道,向所述第二网络设备发送封装上远端隧道内层封装和远端隧道外层封装的转发报文。

14. 根据权利要求 11 或 12 所述的网络设备,其特征在于,

所述处理单元对所述流表报文进行的处理具体为:

直接环回所述流表报文,或,透传所述流表报文,或依据所述流表报文生成流表。

15. 一种网络设备,其特征在于,包括:

接收单元,用于接收转发报文;

流表匹配单元,用于对所述转发报文进行流表匹配;

封装发送单元,用于当所述流表匹配单元的匹配结果不匹配时,将所述转发报文封装上二层帧头后发送给控制服务器处理;

所述接收单元还用于接收所述控制服务器返回的流表报文,所述流表报文中携带所述转发报文的流表信息,且封装有特殊二层帧头;

处理单元,用于依据所述流表报文的特殊二层帧头中的以太类型值所指示的操作,对所述流表报文进行处理。

16. 根据权利要求 15 所述的网络设备,其特征在于,

所述处理单元对所述流表报文进行的处理具体为:

直接环回所述流表报文,或,透传所述流表报文,或依据所述流表报文生成流表。

17. 一种控制服务器,其特征在于,包括:

接收单元,用于接收来自网络设备的转发报文;

解析生成单元,用于对所述接收单元接收到的所述转发报文进行解析,生成所述转发报文的流表信息;

封装生成单元,用于生成承载所述流表信息的流表报文,所述流表报文封装有特殊二层帧头,以便处理所述流表报文的网络设备依据所述特殊二层帧头中的以太类型值指示的操作,对所述流表报文进行处理;

发送单元,用于将所述封装生成单元生成的流表报文发送给所述网络设备。

18. 根据权利要求 17 所述的控制服务器,其特征在于,

所述接收单元接收到的转发报文中携带有处理所述流表报文的网络设备的标识;

所述封装生成单元还用于在所述流表报文中携带处理所述流表报文的网络设备的标识。

## 一种报文处理方法及相关设备

### 技术领域

[0001] 本发明涉及通信领域,尤其涉及一种报文处理方法及相关设备。

### 背景技术

[0002] 随着网络架构的发展,基于网络扁平化及网络运营商快速部署业务需要,出现了各种网络开放技术。将对数据包的处理过程从路由器或交换机等网络设备分离出来,在专门的控制服务器上运行。这种技术将使网络更易于规划和管理,也更易于在现有网络格局上将网络开放。

[0003] Openflow 技术是将原来由交换机或路由器控制的报文转发过程转化为由 Openflow 交换机(或路由器)和控制服务器来共同完成的技术,其实现了数据转发和路由控制的分离。其主要过程如下:控制服务器通过传输控制协议(TCP,Transmission Control Protocol)或安全套接层(SSL,Secure Sockets Layer)与交换机(或路由器)的控制层建立连接,交换机(或路由器)的转发层在接收到数据包时,将该数据包上送到控制层的安全通道模块(即 Secure Channel),Secure Channel 将该数据包进行 TCP 及 Openflow 封装后(为便于描述,后续将经封装的数据包称为封装报文)发送给控制服务器,控制服务器对封装报文进行解析,将封装报文中的原始数据及端口信息转化为流表信息,并将流表信息通过 TCP 通道返回给该交换机(或路由器)的控制层,由交换机(或路由器)的控制层将该流表信息下发到该交换机(或路由器)的流表(即 flow table)中。

[0004] 上述方法虽然可实现数据转发和路由控制的分离,但由于 Openflow 是基于 TCP/SSL 上建立的连接,其性能完全受限于交换机(或路由器)上控制层与转发层的通信能力,而通常交换机(或路由器)上控制层与转发层之间的通信带宽很小,该带宽仅适用于控制层处理如开放式最短路径优先(OSPF,Open Shortest Path First)协议、多协议标签交换(MPLS,Multi-Protocol Label Switching)协议和边界网关协议(BGP,Border Gateway Protocol)等网络协议报文,可见,Openflow 技术存在着性能瓶颈。

### 发明内容

[0005] 本发明实施例提供了一种报文处理方法及相关设备,用于解决网络设备控制平面的瓶颈限制问题,实现数据转发和路由控制的分离。

[0006] 为解决上述技术问题,本发明实施例提供以下技术方案:

[0007] 一种报文处理方法,包括:

[0008] 第一网络设备接收转发报文;

[0009] 对上述转发报文进行流表匹配,若不匹配,则,

[0010] 将上述转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便第二网络设备将上述转发报文转发给控制服务器处理;

[0011] 接收上述第二网络设备返回的流表报文,上流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头;

- [0012] 依据上述特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理。
- [0013] 一种报文处理方法,包括:
- [0014] 网络设备接收转发报文;
- [0015] 对上述转发报文进行流表匹配,若不匹配,则,
- [0016] 将上述转发报文封装上二层帧头后,发送给控制服务器处理;
- [0017] 接收上述控制服务器返回的流表报文,上述流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头;
- [0018] 依据上述特殊二层帧头中的以太类型值所指示的操作,对上述流表报文进行处理。
- [0019] 一种网络设备,包括:
- [0020] 接收单元,用于接收转发报文;
- [0021] 流表匹配单元,用于对上述接收单元接收到的转发报文进行流表匹配;
- [0022] 封装发送单元,用于当上述流表匹配单元的匹配结果为不匹配时,将上述转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便上述第二网络设备将上述转发报文转发给控制服务器处理;
- [0023] 上述接收单元还用于接收上述第二网络设备返回的流表报文,上述流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头;
- [0024] 处理单元,用于依据上述流表报文的特殊二层帧头中的以太类型值所指示的操作,对上述流表报文进行处理。
- [0025] 一种网络设备,包括:
- [0026] 接收单元,用于接收转发报文;
- [0027] 流表匹配单元,用于对上述转发报文进行流表匹配;
- [0028] 封装发送单元,用于当上述流表匹配单元的匹配结果为不匹配时,将上述转发报文封装上二层帧头后发送给控制服务器处理;
- [0029] 上述接收单元还用于接收上述控制服务器返回的流表报文,上述流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头;
- [0030] 处理单元,用于依据上述流表报文的特殊二层帧头中的以太类型值所指示的操作,对上述流表报文进行处理。
- [0031] 一种控制服务器,包括:
- [0032] 接收单元,用于接收来自网络设备的转发报文;
- [0033] 解析生成单元,用于对上述接收单元接收到的转发报文进行解析,生成上述转发报文的流表信息;
- [0034] 封装生成单元,用于生成承载上述流表信息的流表报文,上述流表报文封装有特殊二层帧头,以便处理上述流表报文的网络设备依据上述特殊二层帧头中的以太类型值指示的操作,对上述流表报文进行处理;
- [0035] 发送单元,用于将上述封装生成单元生成的流表报文发送给上述网络设备。
- [0036] 由上可见,本发明实施例中通过在下发的流表报文上封装特殊二层帧头,使处理该流表报文的网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面

上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

### 附图说明

[0037] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0038] 图 1 为本发明提供了一种报文处理方法一个实施例流程示意图;

[0039] 图 2 为本发明提供了一种报文处理方法另一个实施例流程示意图;

[0040] 图 3 为本发明提供了一种报文处理方法再一个实施例流程示意图;

[0041] 图 4 为基于本发明一种报文处理方法下的转发报文在整个处理流程中的传输示意图;

[0042] 图 5 为本发明提供了一种应用场景下的报文处理方法再一个实施例流程示意图;

[0043] 图 6 为本发明提供了一种网络设备一个实施例结构示意图;

[0044] 图 7 为本发明提供了一种控制服务器一个实施例结构示意图。

### 具体实施方式

[0045] 本发明实施例提供了一种报文处理方法及相关设备。

[0046] 为使得本发明的发明目的、特征、优点能够更加的明显和易懂,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而非全部实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0047] 下面以第一网络设备为描述主体,对本发明实施例中一种报文处理方法进行描述,请参阅图 1,本发明实施例中的报文处理方法包括:

[0048] 101、第一网络设备接收转发报文;

[0049] 102、第一网络设备对转发报文进行流表匹配;

[0050] 第一网络设备在接收到转发报文后,首先对该转发报文进行流表匹配,如进行五元组匹配,以确定本地是否有与该转发报文匹配的处理策略,若匹配,则可执行步骤 103;若不匹配,则执行步骤 104。

[0051] 103、第一网络设备依据与该转发报文匹配的处理策略对该转发报文执行相应的处理操作。

[0052] 104、第一网络设备将该转发报文封装上远端隧道内层封装和远端隧道外层封装后发送给第二网络设备;

[0053] 需要说明的是,本发明实施例中的第一网络设备可以是部署于核心层或汇聚层的网络设备,也可以是部署于控制服务器远端的网络设备,其与控制服务器间有多台网络设备(如第二网络设备)间隔,因此,第一网络设备需通过其它网络设备将转发报文发送至控制服务器。

[0054] 在实际应用中,可通过命令行分别对第一网络设备和第二网络设备进行配置,在第一网络设备与第二网络设备间建立一条远端数据通道(如伪线(PW, Pseudo Wire)隧道),第一网络设备通过建立的远端数据通道,将上述转发报文发送至第二网络设备,由第二网络设备将转发报文转发至控制服务器。控制服务器在处理完该转发报文后,需要将处理完的转发报文(即流表报文)通过第二网络设备返回至第一网络设备,由于一台控制服务器可能管理多台具有流表处理功能的网络设备(如第一网络设备),为了保证控制服务器能将上述处理完的转发报文返回至第一网络设备,第二网络设备可在将转发报文转发至控制服务器前,在该转发报文中携带第一网络设备的标识。

[0055] 为保证转发报文可通过远端数据通道送达第二网络设备,第一网络设备将接收到的转发报文封装上远端隧道内层封装和远端隧道外层封装,具体的,远端隧道内层封装可以是MPLS标签,远端隧道外层封装可以是MPLS标签、也可以是通用路由封装(GRE, Generic Routing Encapsulation)封装或以太网协议安全协议(IPSec, Internet Protocol Security Protocol)封装。

[0056] 第一网络设备还可在上述转发报文中携带报文控制信息(PCI, Packet Control Info), PCI信息描述该转发报文的基本属性,如该转发报文的来源(如是上行报文或是下行环回而来的报文)、该转发报文的类型(如是http-get报文或是UDP报文)、该转发报文的上报方式(如是复制报文、或是引流重定向到控制服务器的报文)和第一网络设备的物理端口号等信息,使得控制服务器在接收到该转发报文后,可获知更多与该转发报文相关的信息,进而对该转发报文做出更为准确的判断及处理。

[0057] 105、接收第二网络设备返回的流表报文;

[0058] 在实际应用中,控制服务器接收到转发报文后,解析该转发报文并生成该转发报文的流表信息,其中,流表信息具体可包含如表1所示表项:

[0059] 表1

[0060]

流表属性	五元组	统计标识	带宽管理	动作	扩展保留
------	-----	------	------	----	------

[0061] 其中,流表属性:表明本流表是静态流表还是动态流表,含v4、v6标记;五元组包括源IP地址、目的IP地址、源端口号、目的端口号以及协议号;统计标识可按用户统计,也可按协议类型统计;带宽管理:可按用户进行带宽管理,可按协议类型进行管理,分上下行;动作:对匹配的流进行的动作处理,分上下行,动作包括丢弃、重定向、直行、复制等。

[0062] 在实际应用中,控制服务器根据转发报文的内容及处理此转发报文的业务来生成该转发报文的流表信息,比如流过滤控制业务对于非法的访问需要进行丢弃控制时,若该转发报文的访问属于非法访问(如可从该转发报文中的目的IP地址,即访问地址来判定该转发报文的访问是否属于非法访问),则控制服务器在生成的流表信息的动作中配置丢弃动作。

[0063] 控制服务器通过第二网络设备向第一网络设备下发携带上述流表信息的流表报文,并在流表报文中封装上特殊二层帧头,该特殊二层帧头中的以太网类型值可指示第一网络设备对该流表报文进行操作。

[0064] 在一种应用场景下,当一台控制服务器管理多台具有流表处理功能的网络设备



(如第一网络设备)时,为了保证控制服务器能将上述流表报文通过第二网络设备返回至第一网络设备,控制服务器还可在流表报文中携带第一网络设备的标识。

[0065] 106、依据上述特殊二层帧头中的以太类型值所指示的操作,对上述流表报文进行处理;

[0066] 第一网络设备的转发层通过对该流表报文二层报文解析,可解析出该流表报文的特殊二层帧头中的信息,依据该特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理,如,直接环回该流表报文,或,透传该流表报文,或依据该流表报文生成流表等。

[0067] 在实际应用中,上述特殊二层帧头中的以太类型值可由控制服务器根据转发报文携带的业务数据内容以及本地预置的与该业务相应的执行策略来确定。举例说明,假设转发报文中携带的业务数据内容属于家长控制业务,对于家长控制业务,控制服务器在接收到该转发报文时,会判断该转发报文访问的是否为非法网站(可通过转发报文中的目的 IP 地址确定),若是,则控制服务器可构造一个含有合法网址的 http 报文,封装入上述流表报文,将特殊二层头中的以太网类型值设置为指示透传动作的值,第一网络设备在接收到该流表报文后,将该流表报文直接透传给用户;如果该转发报文访问的是合法网站,则控制服务器可将特殊二层头中的以太网类型值设置为指示环回动作的值,第一网络设备在接收到该流表报文后,将该流表报文环回到上行查路由转发到目的网站。

[0068] 上述第一网络设备可以是具有流表处理功能的路由器、交换机或者其它网络设备,上述第二网络设备可以是路由器或交换机,上述控制服务器可以是可扩展业务平台服务器(SSP, Scalable Service Platform)。

[0069] 由上可见,本发明实施例中通过在下发的流表报文中封装特殊二层帧头,使第一网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

[0070] 上述实施例所描述的是带有流表处理功能的网络设备部署于核心层或汇聚层或控制服务器远端的情况,下面针对该网络设备部署于接入网或控制服务器近端的情况,对本发明实施例中一种报文处理方法进行描述,请参阅图 2,包括:

[0071] 201、网络设备接收转发报文;

[0072] 202、网络设备对转发报文进行流表匹配;

[0073] 网络设备在接收到转发报文后,首先对该转发报文进行流表匹配,如进行五元组匹配,以确定本地是否有与该转发报文匹配的处理策略,若匹配,则可执行步骤 203;若不匹配,则执行步骤 204。

[0074] 203、网络设备依据与该转发报文匹配的处理策略对该转发报文执行相应的处理操作。

[0075] 204、网络设备将该转发报文封装上二层帧头后,发送给控制服务器处理;

[0076] 由于该网络设备为接入层的网络设备,其离控制服务器的距离较近,因此可直接利用二层网络将转发报文发送给控制服务器处理。网络设备可将转发报文封装上二层帧头后直接发送给控制服务器。

[0077] 205、接收控制服务器返回的流表报文；

[0078] 在实际应用中，控制服务器接收到转发报文后，解析该转发报文并生成该转发报文的流表信息，其中，流表信息具体可包含如表 1 所示表项。

[0079] 在实际应用中，控制服务器根据转发报文的内容及处理此转发报文的业务来生成该转发报文的流表信息，比如流过滤控制业务对于非法的访问需要进行丢弃控制时，若该转发报文的访问属于非法访问（如可从该转发报文中的目的 IP 地址，即访问地址来判定该转发报文的访问是否属于非法访问），则控制服务器在生成的流表信息的动作中配置丢弃动作。

[0080] 控制服务器向该网络设备下发携带上述流表信息的流表报文，并在流表报文中封装上特殊二层帧头，该特殊二层帧头中的以太网类型值可指示该网络设备对该流表报文进行操作。

[0081] 在一种应用场景下，当一台控制服务器管理多台具有流表处理功能的网络设备（如第一网络设备）时，为了保证控制服务器能将上述流表报文返回至上述网络设备，控制服务器还可在流表报文中携带第一网络设备的标识。

[0082] 206、依据上述特殊二层帧头中的以太类型值所指示的操作，对上述流表报文进行处理。

[0083] 网络设备的转发层通过对该流表报文进行二层报文解析，可解析出该流表报文的特殊二层帧头中的信息，依据该特殊二层帧头中的以太类型值所指示的操作，对该流表报文进行处理，如，直接环回该流表报文，或，透传该流表报文，或依据该流表报文生成流表等。

[0084] 在实际应用中，上述特殊二层帧头中的以太类型值可由控制服务器根据转发报文携带的业务数据内容以及本地预置的与该业务相应的执行策略来确定。举例说明，假设转发报文中携带的业务数据内容属于家长控制业务，对于家长控制业务，控制服务器在接收到该转发报文时，会判断该转发报文访问的是否为非法网站（可通过转发报文中的目的 IP 地址确定），若是，则控制服务器可构造一个含有合法网址的 http 报文，封装入上述流表报文，将特殊二层头中的以太网类型值设置为指示透传动作的值，网络上设备在接收到该流表报文后，将该流表报文直接透传给用户；如果该转发报文访问的是合法网站，则控制服务器可将特殊二层头中的以太网类型值设置为指示环回动作的值，网络设备在接收到该流表报文后，将该流表报文环回到上行查路由转发到目的网站。

[0085] 上述网络设备可以是具有流表处理功能的路由器、交换机或者其它网络设备，上述控制服务器可以是 SSP 服务器。

[0086] 由上可见，本发明实施例中通过在下发的流表报文中封装特殊二层帧头，使网络设备可直接在数据转发平面上，依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理，网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作，从而不受网络设备控制平面的瓶颈限制，也实现了数据转发和路由控制的分离。

[0087] 下面以控制服务器为描述主体，对本发明实施例中一种报文处理方法进行描述，请参阅图 3，包括：

[0088] 301、控制服务器接收来自网络设备的转发报文；

[0089] 302、对接收到的转发报文进行解析,生成该转发报文的流表信息;

[0090] 控制服务器在接收到转发报文后,可对该转发报文进行解析生成该转发报文的流表信息,其中,流表信息可包含如表 1 所示表项。

[0091] 303、生成承载该流表信息的流表报文;

[0092] 控制服务器对步骤 302 生成的流表信息进行封装,生成流表报文。

[0093] 为使得网络设备接收到上述流表报文后,可直接在转发层上对该流表报文进行处理,控制服务器在流表报文上封装特殊二层帧头,通过特殊二层帧头中的以太类型值指示上述网络设备对该流表报文进行处理。在实际应用中,控制服务器可根据转发报文携带的业务数据内容以及本地预置的与该业务相应的执行策略来确定特殊二层帧头中的以太类型值。举例说明,假设转发报文中携带的业务数据内容属于家长控制业务,对于家长控制业务,控制服务器在接收到该转发报文时,会判断该转发报文访问的是否为非法网站(可通过转发报文中的目的 IP 地址确定),若是,则控制服务器可构造一个含有合法网址的 http 报文,封装入上述流表报文,将特殊二层头中的以太网类型值设置为指示透传动作的值,上述网络设备在接收到该流表报文后,将该流表报文直接透传给用户;如果该转发报文访问的是合法网站,则控制服务器可将特殊二层头中的以太网类型值设置为指示环回动作的值,上述网络设备在接收到该流表报文后,将该流表报文环回到上行查路由转发到目的网站。

[0094] 可理解的是,由于上述特殊二层帧头中的以太类型值为自定义的参数值,控制服务器需在封装了特殊二层帧头后,再在流表报文上封装普通二层帧头,以保证流表报文能在二层网络中正常传输。

[0095] 进一步的,当一台控制服务器管理多台网络设备的情况时,控制服务器还可在流表报文中携带上述网络设备的标识,如可将该网络设备的标识携带在上述普通二层帧头中。

[0096] 304、将上述流表报文发送给上述网络设备;

[0097] 若上述网络设备与流表报文通过二层网络直连,则控制服务器可直接通过二层网络将流表报文返回给该网络设备;

[0098] 若上述网络设备是部署在核心网或汇聚网或控制服务器远端的设备,则控制服务器可将该流表报文下发给近端网络设备(具有转发功能的网络设备,如第二网络设备),由该近端网络设备将流表报文转发至该网络设备。

[0099] 上述控制服务器可以是 SSP 服务器。

[0100] 由上可见,本发明实施例中控制服务器通过在下发的流表报文中封装特殊二层帧头,使网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

[0101] 下面以第一网络设备、第二网络设备和控制服务器为描述主体,以一具体应用场景对本发明实施例中一种报文处理方法进行描述。

[0102] 首先请参阅图 4,图 4 所示为转发报文在整个处理流程中的传输示意图,其中,图中的 R1 表示第一网络设备接收到的转发报文,L2 表示封装的普通二层帧头,L2' 表示封装

的特殊二层帧头, R2 表示控制服务器生成的该转发报文的流表信息, 具体地报文处理流程请参阅图 5, 包括:

[0103] 401、第一网络设备接收转发报文, 并对接收到的转发报文进行流表匹配;

[0104] 第一网络设备在接收到转发报文后, 对该转发报文进行流表匹配, 如进行五元组匹配, 以确定本地是否有与该转发报文匹配的处理策略, 若匹配, 则执行步骤 402; 若不匹配, 则执行步骤 403。

[0105] 402、第一网络设备依据与该转发报文匹配的处理策略对该转发报文执行相应的处理操作。

[0106] 403、第一网络设备将该转发报文进行封装后发送给第二网络设备;

[0107] 在实际应用中, 可在第一网络设备与第二网络设备间建立 PW 隧道, 第一网络设备通过该 PW 隧道将封装后的转发报文发送给第二网络设备, PW 隧道的建立可通过复用静态 PW 流程来实现, 具体的, 假设第一网络设备的 IP 地址为 192. 2. 2. 2, 第二网络设备的 IP 地址为 192. 1. 1. 1, 则可分别在第一网络设备和第二网络设备中进行如下配置来建立上述 PW 隧道。

[0108] 在第一网络设备进行如下命令行配置: `mpls static-l2vc destination 192. 1. 1. 1 transmit-vpn-label 200 receive-vpn-label 100 DPI_DPI;`

[0109] 在第二网络设备进行如下命令行配置: `mpls static-l2vc destination 192. 2. 2. 2 transmit-vpn-label 100 receive-vpn-label 200 ip-interworking.`

[0110] 假设第一网络设备和第二网络设备进行了上述命令行配置, 转发报文的下发处理流程由第一网络设备上的 DPI 单板和网络处理器 (NP, Network Processor) 来实现, 则首先, 第一网络设备内部资源可申请一个标签 (简称 M1 标签) 下发到 DPI 单板中, 并在 NP 中建立两个转发表项: insegment 表和 outsegment 表。其中, 在 insegment 表中: 动作 = 标签交换, 远端隧道内层封装 = transmit-vpn-label, 目的板和目的端口 (简称 TB/TP) = LSP 隧道的入口板 TB/TP, outsegment 表索引值 = LSP 隧道的对应的 outsegment 表索引值; outsegment 表中包含远端隧道外层封装信息, 封装方式为 MPLS 封装。上述 outsegment 表中的远端隧道外层封装信息、上述 insegment 表中的 LSP 隧道的入口板 TB/TP 以及该 LSP 隧道的对应的 outsegment 表索引值, 可通过第二网络设备的 IP 地址来获取到 (假设第一网络设备与第二网络设备间已建立 LSP 隧道)。

[0111] 当步骤 401 中转发报文不匹配时, DPI 单板将该转发报文打上 M1 标签及相应的 PCI 信息, 封装成 MPLS 报文后发送到 NP, NP 通过解析发现转发报文上存在 M1 标签, 则查询 insegment 表, 获得 TB/TP、动作、远端隧道内层封装信息和 outsegment 表索引值, 通过获得的 outsegment 表索引值获得远端隧道外层封装信息, 进行标签交换处理, 即先弹掉原有的 MPLS 标签, 然后依据获得的远端隧道内层封装信息和远端隧道外层封装信息分别将转发报文封装上远端隧道内层封装和远端隧道外层封装, 最后封装二层帧头 (为普通的二层帧头) 后从 PW 隧道发送出去, 完成转发报文中送的过程。

[0112] 404、第二网络设备将接收到的转发报文转发至控制服务器;

[0113] 第二网络设备接收到转发报文, 去掉封装在转发报文上的远端隧道内层封装和远端隧道外层封装, 重新封装二层帧头后发送给控制服务器;

[0114] 由于第二网络设备为控制服务器的近端设备, 故可直接利用与控制服务器间的二

层连接将转发报文发送到控制服务器。

[0115] 具体地,可事先在第二网络设备上进行如步骤 403 中的命令行配置: `mpls static-l2vc destination 192.2.2.2 transmit-vc-label 100 receive-vc-label 200 ip-interworking`,其中 `ip-interworking` 表明传输为异种介质属性。第二网络设备收到上述转发报文后去掉封装在转发报文上的远端隧道内层封装和远端隧道外层封装,由于是异种介质,第二网络设备查找地址解析 (`arp`, `address resolution`) 表,重新为转发报文封装二层帧头,进一步还可在二层帧头中携带第一网络设备的标识 (如 `router id`),将该转发报文从出接口发送至控制服务器。

[0116] 405、控制服务器对接收到的转发报文进行解析,生成该转发报文的流表信息;

[0117] 控制服务器接收到转发报文后,可对该转发报文进行解析生成该转发报文的流表信息,其中,流表信息可包含如表 1 所示表项。

[0118] 406、控制服务器生成承载该流表信息的流表报文;

[0119] 控制服务器对步骤 405 生成的流表信息进行封装,生成流表报文。

[0120] 为使得网络设备接收到上述流表报文后,可直接在转发层上对该流表报文进行处理,控制服务器在流表报文上封装特殊二层帧头,通过特殊二层帧头中的以太类型值指示上述网络设备对该流表报文进行处理。在实际应用中,控制服务器可根据转发报文携带的业务数据内容以及本地预置的与该业务相应的执行策略来确定特殊二层帧头中的以太类型值。举例说明,假设转发报文中携带的业务数据内容属于家长控制业务,对于家长控制业务,控制服务器在接收到该转发报文时,会判断该转发报文访问的是否为非法网站 (可通过转发报文中的目的 IP 地址确定),若是,则控制服务器可构造一个含有合法网址的 `http` 报文,封装入上述流表报文,将特殊二层头中的以太网类型值设置为指示透传动作的值,上述网络设备在接收到该流表报文后,将该流表报文直接透传给用户;如果该转发报文访问的是合法网站,则控制服务器可将特殊二层头中的以太网类型值设置为指示环回动作的值,上述网络设备在接收到该流表报文后,将该流表报文环回到上行查路由转发到目的网站。

[0121] 可理解的是,由于上述特殊二层帧头中的以太类型值为自定义的参数值,控制服务器需在封装了特殊二层帧头后,再在流表报文上封装普通二层帧头,以保证流表报文能在二层网络中正常传输。

[0122] 进一步的,当一台控制服务器管理多台网络设备的情况时,控制服务器还可在流表报文中携带上述网络设备的标识,如可将该网络设备的标识携带在上述普通二层帧头中。

[0123] 407、控制服务器将上述流表报文发送给第二网络设备;

[0124] 408、第二网络设备将接收到流表报文转发给第一网络设备;

[0125] 第二网络设备接收到流表报文后,可通过步骤 403 中建立的 PW 隧道向第一网络设备返回该流表报文,首先去掉封装在流表报文上的上述普通二层帧头,通过查询本地的 `insegment` 表和 `outsegment` 表,将流表报文封装上远端隧道内层封装、远端隧道外层封装及二层帧头后发送给第一网络设备。

[0126] 409、第一网络设备依据特殊二层帧头中的以太类型值的指示,对流表报文进行处理;

[0127] 第一网络设备在接收到第二网络设备返回的流表报文后,去掉封装在该流表报文上的普通二层帧头、远端隧道内层封装和远端隧道外层封装,查看特殊二层帧头,依据该特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理,如,直接环回该流表报文,或,透传该流表报文,或依据该流表报文生成流表等。

[0128] 由上可见,本发明实施例中通过在下发的流表报文上封装特殊二层帧头,使第一网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

[0129] 下面对本发明实施例中的一种网络设备进行描述,请参阅图 6,本发明实施例中一种网络设备 500 包括:

[0130] 接收单元 501,流表匹配单元 502,封装发送单元 503 和处理单元 504。

[0131] 其中:

[0132] 接收单元 501,用于接收转发报文;

[0133] 流表匹配单元 502,用于对接收单元 501 接收到的转发报文进行流表匹配;

[0134] 流表匹配单元 502 对该转发报文进行流表匹配,如进行五元组匹配,以确定本地是否有与该转发报文匹配的处理策略,若不匹配,则触发封装发送单元 503。

[0135] 封装发送单元 503,用于当流表匹配单元 502 的匹配结果为不匹配时,将接收到的转发报文封装上远端隧道内层封装和远端隧道外层封装后,发送给第二网络设备,以便上述第二网络设备将该转发报文转发给控制服务器处理;

[0136] 在实际应用中,可通过命令行分别对网络设备 500 和第二网络设备进行配置,在网络设备 500 与第二网络设备间建立一条远端数据通道(如伪线(PW, Pseudo Wire)隧道),网络设备 500 通过建立的远端数据通道,将上述转发报文发送至第二网络设备,由第二网络设备将转发报文转发至控制服务器。

[0137] 为保证转发报文可通过远端数据通道送达第二网络设备,封装发送单元 503 将接收到的转发报文封装上远端隧道内层封装和远端隧道外层封装,具体的,远端隧道内层封装可以是 MPLS 标签,远端隧道外层封装可以是 MPLS 标签,也可以是 GRE 封装或 IPSec 封装。

[0138] 封装发送单元 503 还可在上述转发报文中携带 PCI 信息,PCI 信息描述该转发报文的基本属性,如该转发报文的来源(如是上行报文或是下行环回而来的报文)、该转发报文的类型(如是 http-get 报文或是 UDP 报文)、该转发报文的上报方式(如是复制报文、或是引流重定向到控制服务器的报文)和网络设备 500 的物理端口号等信息,使得控制服务器在接收到该转发报文后,可获知更多与该转发报文相关的信息,进而对该转发报文做出更为准确的判断及处理。

[0139] 进一步的,接收单元 501 还用于接收上述第二网络设备返回的流表报文,该流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头。其中,上述流表信息可如表 1 所示,上述特殊二层帧头中的以太网类型值可指示处理单元 504 对该流表报文进行操作。

[0140] 处理单元 504,用于依据接收单元 501 接收到的流表报文的特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理。

[0141] 处理单元 504 通过对该流表报文进行二层报文解析,可解析出该流表报文的特殊

二层帧头中的信息,依据该特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理,如,直接环回该流表报文,或,透传该流表报文,或依据该流表报文生成流表等。

[0142] 网络设备 500 可以是部署于核心层或汇聚层或控制服务器远端,可以是具有流表处理功能的路由器、交换机或者其它网络设备,此处不作限定。

[0143] 由上可见,本发明实施例中控制服务器通过在下发的流表报文中封装特殊二层帧头,使网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

[0144] 需要说明的是,本实施例的网络设备 500 可以如上述方法实施例中的第一网络设备,可以用于实现上述方法实施例中的全部技术方案,其各个功能模块的功能可以根据上述方法实施例中的方法具体实现,其具体实现过程可参照上述实施例中的相关描述,此处不再赘述。

[0145] 下面对本发明实施例中的另一种网络设备进行描述,包括:

[0146] 接收单元,用于接收转发报文;

[0147] 流表匹配单元,用于对接收单元接收到的转发报文进行流表匹配;

[0148] 流表匹配单元对该转发报文进行流表匹配,如进行五元组匹配,以确定本地是否有与该转发报文匹配的处理策略,若不匹配,则触发封装发送单元。

[0149] 封装发送单元,用于当上述流表匹配单元的匹配结果为不匹配时,将该转发报文封装上二层帧头后发送给控制服务器处理;

[0150] 封装发送单元可将转发报文封装上二层帧头后,通过二层网络直接发送给控制服务器。进一步的,封装发送单元还可在该转发报文中携带本网络设备的标识。封装发送单元还可在上述转发报文中携带 PCI 信息,PCI 信息描述该转发报文的基本属性,如该转发报文的来源(如是上行报文或是下行环回而来的报文)、该转发报文的类型(如是 http-get 报文或是 UDP 报文)、该转发报文的上报方式(如是复制报文、或是引流重定向到控制服务器的报文)和本网络设备的物理端口号等信息,使得控制服务器在接收到该转发报文后,可获知更多与该转发报文相关的信息,进而对该转发报文做出更为准确的判断及处理。

[0151] 上述接收单元还用于接收上述控制服务器返回的流表报文;该流表报文中携带上述转发报文的流表信息,且封装有特殊二层帧头。其中,上述流表信息可如表 1 所示,上述特殊二层帧头中的以太类型值可指示处理单元对该流表报文进行操作。

[0152] 处理单元,用于依据接收单元接收到的流表报文的特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理。

[0153] 处理单元通过对该流表报文进行二层报文解析,可解析出该流表报文的特殊二层帧头中的信息,依据该特殊二层帧头中的以太类型值所指示的操作,对该流表报文进行处理,如,直接环回该流表报文,或,透传该流表报文,或依据该流表报文生成流表等。

[0154] 上述网络设备部署于接入层或控制服务器近端,可以是具有流表处理功能的路由器、交换机或者其它网络设备,此处不作限定。

[0155] 需要说明的是,本实施例的网络设备可以如上述方法实施例中的网络设备,可以用于实现上述方法实施例中的全部技术方案,其各个功能模块的功能可以根据上述方法实

施例中的方法具体实现,其具体实现过程可参照上述实施例中的相关描述,此处不再赘述。

[0156] 下面对本发明实施例中的控制服务器进行描述,请参阅图 7,本发明实施例中的控制服务器 600 包括:

[0157] 接收单元 601,用于接收来自网络设备的转发报文;

[0158] 解析生成单元 602,用于对接收单元 601 接收到的转发报文进行解析,生成该转发报文的流表信息;

[0159] 其中,流表信息可包含如表 1 所示表项。

[0160] 封装生成单元 603,用于生成承载解析生成单元 602 生成的流表信息的流表报文;

[0161] 为使得网络设备接收到上述流表报文后,可直接在转发层上对该流表报文进行处理,封装生成单元 603 在流表报文上封装特殊二层帧头,通过特殊二层帧头中的以太类型值指示上述网络设备对该流表报文进行处理。在实际应用中,封装生成单元 603 可根据转发报文携带的业务数据内容以及本地预置的与该业务相应的执行策略来确定特殊二层帧头中的以太类型值。举例说明,假设转发报文中携带的业务数据内容属于家长控制业务,对于家长控制业务,控制服务器 600 在接收到该转发报文时,会判断该转发报文访问的是否为非法网站(可通过转发报文中的目的 IP 地址确定),若是,则封装生成单元 603 可构造一个含有合法网址的 http 报文,封装入上述流表报文,将特殊二层头中的以太网类型值设置为指示透传动作的值,上述网络设备在接收到该流表报文后,将该流表报文直接透传给用户;如果该转发报文访问的是合法网站,则封装生成单元 603 可将特殊二层头中的以太网类型值设置为指示环回动作的值,上述网络设备在接收到该流表报文后,将该流表报文环回到上行查路由转发到目的网站。

[0162] 可以理解的是,由于上述特殊二层帧头中的以太类型值为自定义的参数值,封装生成单元 603 需在封装了特殊二层帧头后,再在流表报文上封装普通二层帧头,以保证流表报文能在二层网络中正常传输。

[0163] 进一步的,当控制服务器 600 管理多台网络设备的情况时,封装生成单元 603 还可在流表报文中携带上述网络设备的标识,如可将该网络设备的标识携带在上述普通二层帧头中。

[0164] 发送单元 604,用于将封装生成单元 603 生成的流表报文发送给上述网络设备。

[0165] 若上述网络设备与流表报文通过二层网络直连,则发送单元 604 可直接通过二层网络将流表报文返回给该网络设备;

[0166] 若上述网络设备是部署在核心网或汇聚网或控制服务器远端的设备,则发送单元 604 可将该流表报文下发给近端网络设备(具有转发功能的网络设备,如第二网络设备),由该近端网络设备将流表报文转发至该网络设备。

[0167] 控制服务器 600 可以是 SSP 服务器。

[0168] 由上可见,本发明实施例中控制服务器通过在下发的流表报文上封装特殊二层帧头,使网络设备可直接在数据转发平面上,依据上述特殊二层帧头中的以太类型值所指示的操作对该流表报文进行相应的处理,网络设备和控制服务器直接在数据转发平面上完成转发报文的引流工作和流表报文的下发工作,从而不受网络设备控制平面的瓶颈限制,也实现了数据转发和路由控制的分离。

[0169] 需要说明的是,本实施例的控制服务器可以如上述方法实施例中的控制服务器,



可以用于实现上述方法实施例中的全部技术方案,其各个功能模块的功能可以根据上述方法实施例中的方法具体实现,其具体实现过程可参照上述实施例中的相关描述,此处不再赘述。

[0170] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0171] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0172] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0173] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0174] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0175] 以上对本发明所提供的一种报文处理方法及相关设备进行了详细介绍,对于本领域的一般技术人员,依据本发明实施例的思想,在具体实施方式及应用范围上均会有改变之处,综上,本说明书内容不应理解为对本发明的限制。

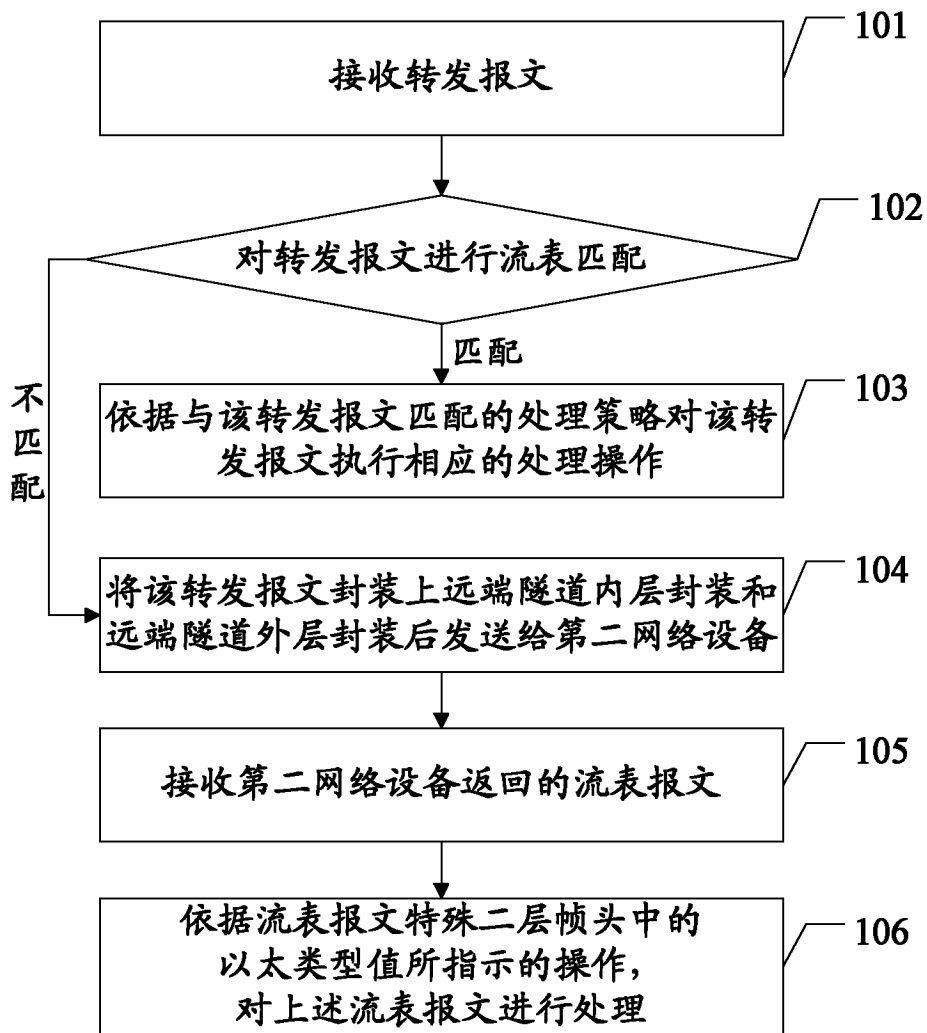


图 1

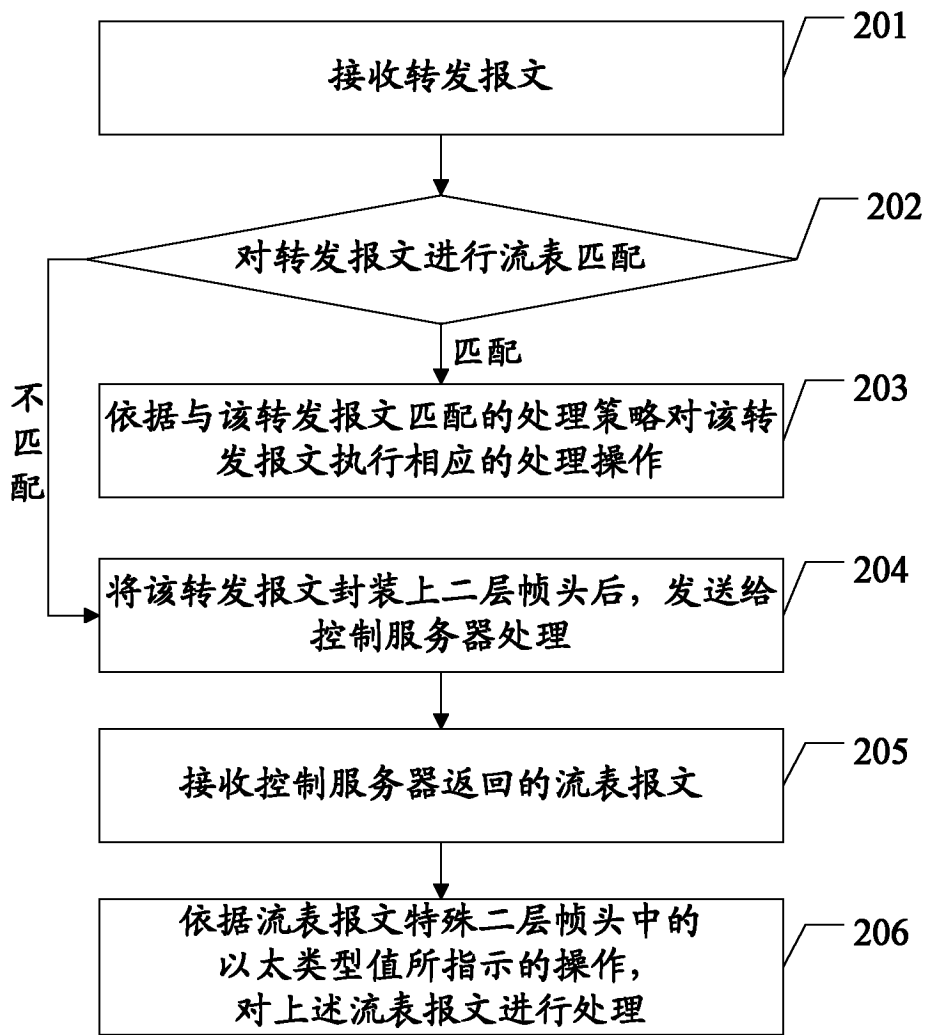


图 2

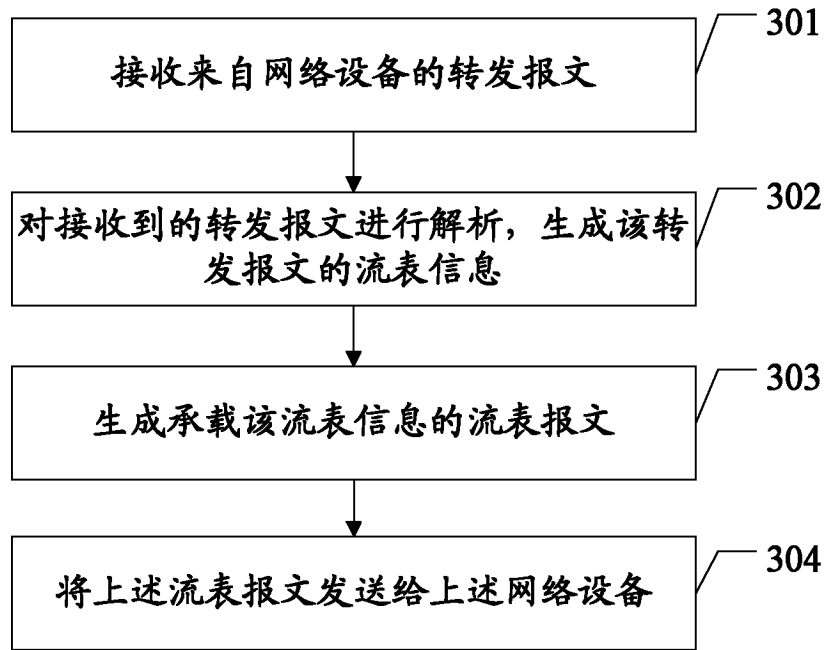


图 3

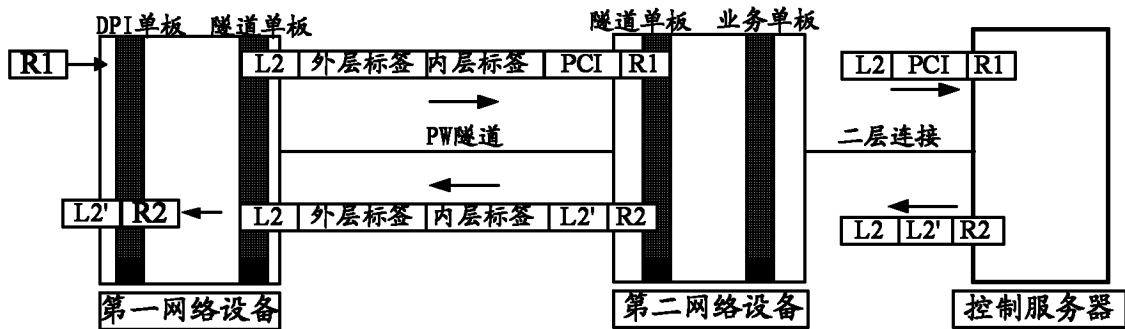


图 4

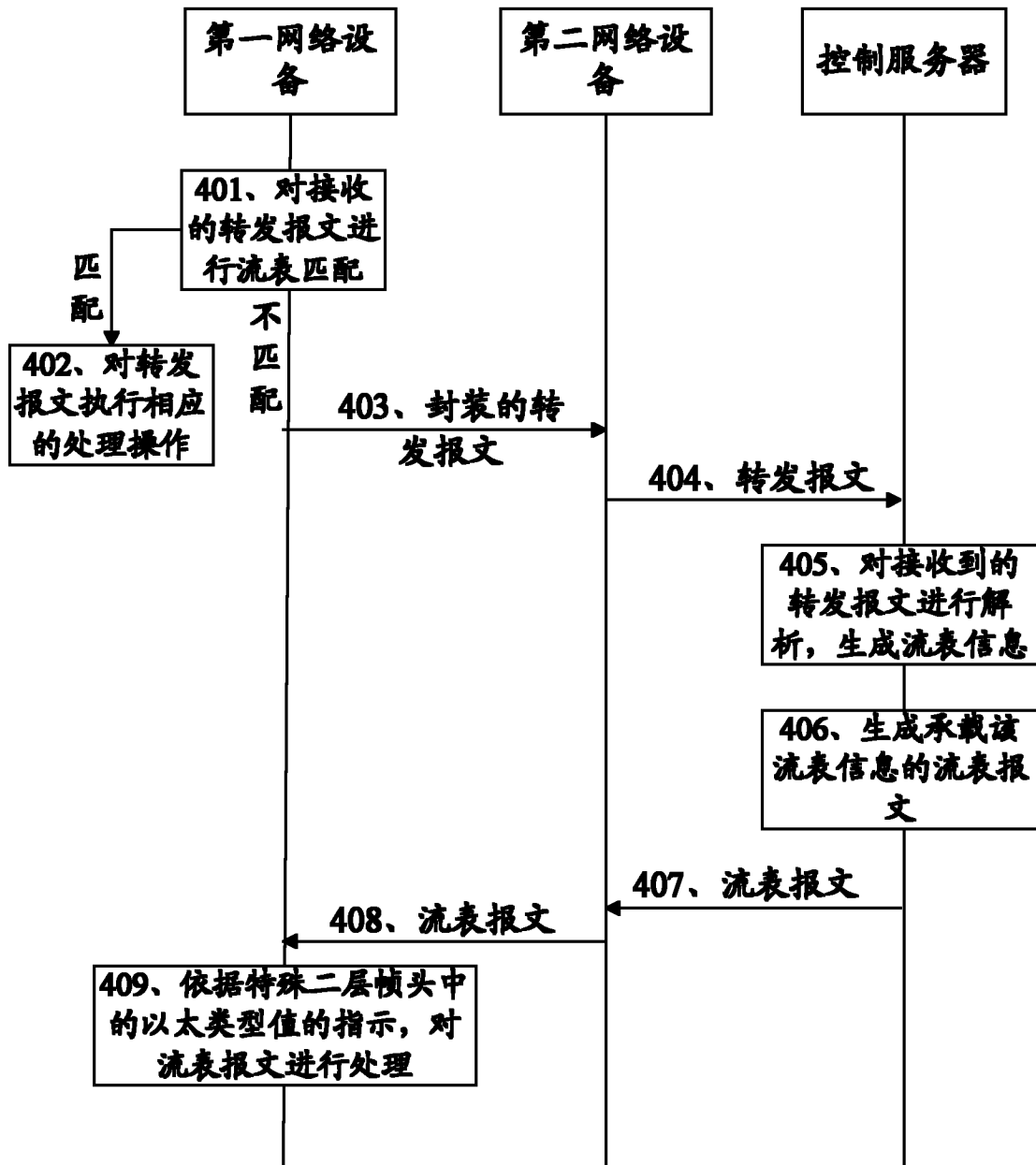


图 5

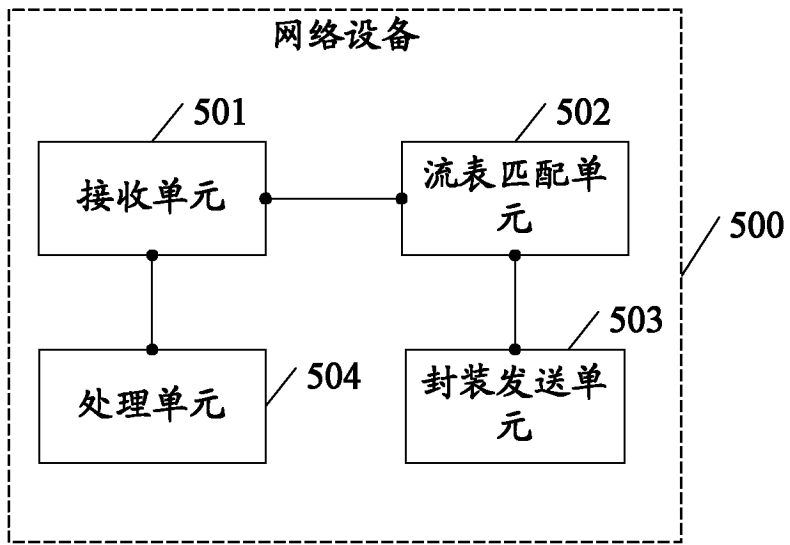


图 6

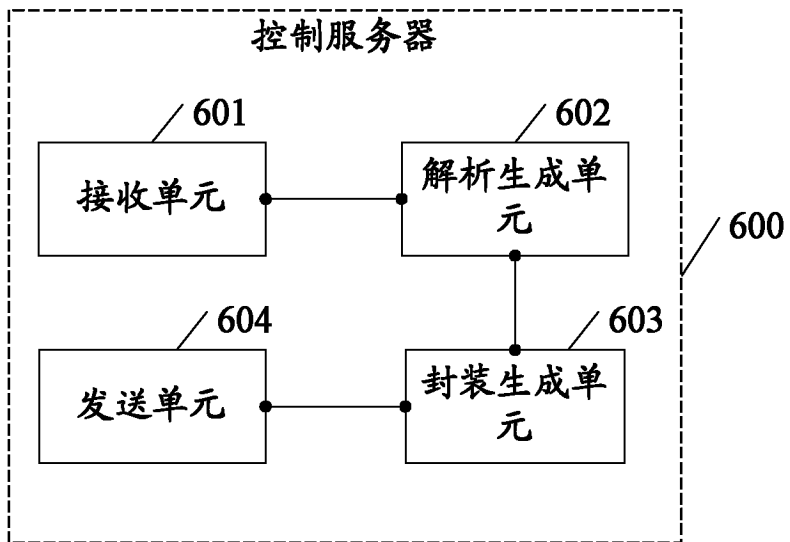


图 7