



(12) 发明专利

(10) 授权公告号 CN 102970227 B

(45) 授权公告日 2016. 03. 02

(21) 申请号 201210449232. 6

审查员 李孟

(22) 申请日 2012. 11. 12

(73) 专利权人 盛科网络(苏州)有限公司

地址 215021 江苏省苏州市工业园区星汉街
5号(腾飞工业坊)B幢4楼13/16单元

(72) 发明人 方沛昱 廖继平

(74) 专利代理机构 苏州慧通知识产权代理事务
所(普通合伙) 32239

代理人 安纪平

(51) Int. Cl.

H04L 12/741(2013. 01)

H04L 12/761(2013. 01)

(56) 对比文件

CN 101409685 A, 2009. 04. 15,

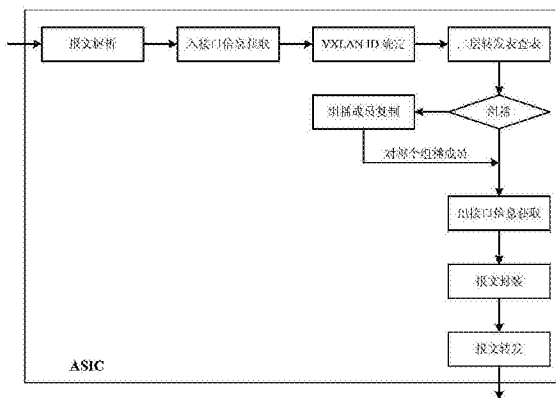
权利要求书1页 说明书4页 附图3页

(54) 发明名称

在 ASIC 中实现 VXLAN 报文转发的方法和装置

(57) 摘要

本发明揭示了一种在 ASIC 中实现 VXLAN 报文转发的方法和装置,所述 VXLAN 报文的转发由设置在服务器和 IP 网络间的 VXLAN 网关处理完成的,所述转发的 VXLAN 报文包括上行报文和下行报文,所述上行报文是从服务器侧发送至 IP 网络的报文,其处理过程包括对收到的报文进行解析,判断是否需要转发处理;对需要转发的报文,获取其对应的 VXLANID;使用 VXLAN ID 和报文目的 MAC 地址查表获得报文的处理信息和转发端口信息,对待转发报文封装处理后进行转发。本发明实现了 VXLAN 报文的单播、组播等行为,且通过逻辑固化在 ASIC 中,实现了 VLAN 的转发域向基于 VXLAN 的转发域的过渡。



1. 一种在 ASIC 中实现 VXLAN 报文转发的方法,其特征在于包括以下步骤:
 - a. 对 VXLAN 网关收到的报文进行解析,判断是否需要转发处理,所述 VXLAN 报文包括从服务器发送至 IP 网络的上行报文和从 IP 网络发送至服务器的下行报文;
 - b. 对需要转发的报文,获取其对应的 VXLAN ID,所述 VXLAN 报文是所述上行报文时,对需要转发的 VXLAN 上行报文,则先获得 VXLAN ID 的映射关系,所述 VXLAN ID 的映射关系通过 VLAN 或源 MAC 地址来标记 VXLAN ID,再根据获得的 VXLAN ID 的映射关系进行查表,在查找结果的输出中获得 VXLAN ID;
 - c. 根据报文的目的地 MAC 地址和 VXLAN ID 查二层转发表,获得报文的处理信息和转发端口信息;
 - d. 对待转发报文进行封装或编辑处理后进行转发。
2. 根据权利要求 1 所述的方法,其特征在于:所述 VXLAN 报文是从服务器发送至 IP 网络的上行报文,所述步骤 c 中的处理信息包括获得目的端口和封装信息;所述步骤 d 中对待转发报文进行的是封装处理。
3. 根据权利要求 1 所述的方法,其特征在于:所述 VXLAN 报文是从 IP 网络发送至服务器的下行报文,所述步骤 c 中 VXLAN ID 的获取是通过解封装 VXLAN 报文获得报头中所携带的信息。
4. 根据权利要求 3 所述的方法,其特征在于:所述步骤 c 中的处理信息包括获得目的端口和编辑信息;所述步骤 d 中对待转发报文进行的是编辑处理。
5. 根据权利要求 1 至 4 任意一项所述的方法,其特征在于:在步骤 c 中,如果查表结果是组播的话,则进行组播复制,并获得每一组播成员的处理信息。
6. 一种在 ASIC 中实现 VXLAN 报文转发的装置,其特征在于包括:
 - a. 报文解析及判断模块,用于对收到的报文进行解析,判断是否需要转发处理,所述 VXLAN 报文包括从服务器发送至 IP 网络的上行报文和从 IP 网络发送至服务器的下行报文;
 - b. VXLAN ID 确定模块,用于对需要转发的报文,获取其对应的 VXLAN ID,所述 VXLAN 报文是所述上行报文时,对需要转发的 VXLAN 上行报文,则先获得 VXLAN ID 的映射关系,所述 VXLAN ID 的映射关系通过 VLAN 或源 MAC 地址来标记 VXLAN ID,再根据获得的 VXLAN ID 的映射关系进行查表,在查找结果的输出中获得 VXLAN ID;
 - c. 转发表查表模块,用于通过 VXLAN ID 和报文目的 MAC 地址查表获得报文的处理信息和转发端口信息;
 - d. 报文转发模块,对待转发报文进行封装或编辑处理后进行转发。
7. 根据权利要求 6 所述的装置,其特征在于:所述 VXLAN 报文的转发由设置在服务器和 IP 网络间的 VXLAN 网关处理完成的,所述服务器上支持有多台虚拟机。
8. 根据权利要求 6 所述的装置,其特征在于:所述装置还包括获取 VXLAN 网关的报文入接口信息的入接口信息处理模块,获取出接口信息的出接口信息模块,对报文进行解封装的解封装模块,以及对报文进行编辑的编辑模块。

在 ASIC 中实现 VXLAN 报文转发的方法和装置

技术领域

[0001] 本发明涉及网络通信领域,尤其涉及在 ASIC(以太网交换芯片)设计中实现 VXLAN 报文转发的方法和装置。

背景技术

[0002] 随着虚拟化和云计算的逐步发展,未来的数据中心将大量部署虚拟机的应用。为了控制不同虚拟机和虚拟机之间,以及虚拟机和租户之间的通信,VLAN(Virtual Local Area Network,虚拟局域网)便顺理成章的成为隔离不同局域网的手段。但是由于虚拟技术的广泛应用,一个数据中心的可支持的租户数量也大大增加,二层 MAC 地址将出现爆发式的增长,并且在理论上不同虚拟局域网中的 MAC 地址是可以重叠的。因此,原有的 4KVLAN 数量已经捉襟见肘。

[0003] VXLAN(Virtual Extensible Local Area Network,可扩展虚拟网络)正是在这样的背景下,由 IETF 提出的一个解决 VLAN 空间不足的方案,由一个 24bit 的 VXLAN ID 来代替 VLAN 参与转发查表,大大扩展了虚拟局域网的数量。

[0004] VXLAN 的应用场景中,有一种 gateway 功能,可以实现非 VXLAN 的域和 VXLAN 域之间的互通,然而该功能需要交换机支持,来实现高带宽的转发。

[0005] VXLAN 目前随还在草案阶段,但是其清晰的网络层次和良好的扩展性,使得 VXLAN 成为一个很有可能被大众所认可的解决方案。然在网络交换芯片中,如何实现 VXLAN 报文的转发技术还未得到解决。

发明内容

[0006] 本发明的目的在于提供一种在 ASIC 中实现 VXLAN 报文转发的方法和装置,其能够对报文进行 VXLAN 封装和解封装,从而通过实现 VXLAN 的 gateway(网关)功能,实现基于 VLAN 的转发域向基于 VXLAN 的转发域的过渡。

[0007] 为实现上述目的,本发明提出如下技术方案:一种在 ASIC 中实现 VXLAN 报文转发的方法,包括以下步骤:

[0008] a. 对 VXLAN 网关收到的报文进行解析,判断是否需要转发处理;

[0009] b. 对需要转发的报文,获取其对应的 VXLAN ID;

[0010] c. 根据报文的源 MAC 地址和 VXLAN ID 查二层转发表,获得报文的处理信息和转发端口信息;

[0011] d. 对待转发报文进行封装或编辑处理后进行转发。

[0012] 更进一步地,所述 VXLAN 报文是从服务器发送至 IP 网络的上行报文。

[0013] 所述步骤 b 中 VXLAN ID 是通过 VLAN 或源 MAC 地址进行标记的。

[0014] 所述步骤 c 中的处理信息包括获得目的端口和封装信息;所述步骤 d 中对待转发报文进行的是封装处理。

[0015] 所述 VXLAN 报文是从 IP 网络发送至服务器的下行报文,所述步骤 c 中 VXLAN ID

的获取是通过解封装 VXLAN 报文获得报头中所携带的信息。

[0016] 所述步骤 c 中的处理信息包括获得目的端口和编辑信息；所述步骤 d 中对待转发报文进行的是编辑处理。

[0017] 在步骤 c 中，如果查表结果是组播的话，则进行组播复制，并获得每一组播成员的处理信息。

[0018] 本发明还提供了一种在 ASIC 中实现 VXLAN 报文转发的装置，其包括：

[0019] a. 报文解析及判断模块，用于对收到的报文进行解析，判断是否需要转发处理；

[0020] b. VXLAN ID 确定模块，用于对需要转发的报文，获取其对应的 VXLAN ID；

[0021] c. 转发表查表模块，用于通过 VXLAN ID 和报文目的 MAC 地址查表获得报文的处理信息和转发端口信息；

[0022] d. 报文转发模块，对待转发报文进行封装或编辑处理后进行转发。

[0023] 更进一步地，所述 VXLAN 报文的转发由设置在服务器和 IP 网络间的 VXLAN 网关处理完成的，所述服务器上支持有多台虚拟机。

[0024] 所述装置还包括获取 VXLAN 网关的报文入接口信息的入接口信息处理模块，获取出接口信息的出接口信息模块，对报文进行解封装的解封装模块，以及对报文进行编辑的编辑模块。

[0025] 与现有技术相比，本发明提出的在 ASIC 中实现 VXLAN 报文转发的方法和装置实现了 VXLAN 报文的单播、组播和广播。同时，本发明作为以太网交换芯片的一个子特性，可通过逻辑固化在 ASIC 中，充分利用以太网交换芯片高带宽、低成本的优势，为将来的数据中心进行 VLAN 扩展降低部署的成本。

附图说明

[0026] 图 1 是本发明在 ASIC 中实现 VXLAN 报文转发的网络示意图；

[0027] 图 2 是本发明转发上行 VXLAN 报文的转发流程图；

[0028] 图 3 是本发明转发下行 VXLAN 报文的转发流程图。

具体实施方式

[0029] 下面将结合本发明的附图，对本发明实施例的技术方案进行清楚、完整的描述。

[0030] 如图 1 所示，本发明揭示的在 ASIC 中实现 VXLAN 报文转发，是在 VXLAN 的网关中实现的，在 IP 网络和服务器间设有多个 VXLAN 网关，如图中的 VXLAN gateway 1, VXLAN gateway 2, VXLAN gateway 3 及 VXLAN gateway 4, 以进行 IP 网络与服务器间的数据传输。每一 VXLAN 网关上连接有多台服务器，每一服务器上设置有多台虚拟机。各 VXLAN 网关能够接收从相应的服务器发送至 IP 网络的上行报文及从 IP 网络发送至对应的服务器的下行报文，其对所收到的报文进行处理后则转发至目的地址，VXLAN 网关对上行报文的处理和对下行报文的处理流程分别介绍如下：

[0031] 一、上行报文处理

[0032] 上行报文是指 VXLAN 网关收到的从服务器侧发来需要通过 VXLAN 封装的报文。当 VXLAN 网关收到从服务器侧发过来的上行报文时，如图 1 中的 VXLAN gateway1, 其接收到的

从服务器 1 或服务器 2 上发送至 IP 网络的报文时,该报文在 VXLAN gateway 1 中处理后转发至 IP 网络。VXLAN 网关中对该上行报文的处理过程包括解析报文,获取入接口信息,确定报文中的 VXLAN ID,查找二层转发表表,组播后,获取出接口信息,再进行报文封装和报文转发。对应地,其上行报文处理装置包括报文解析模块,入接口信息获取模块,VXLAN ID 模块,二层转发查表模块,组播模块,组播成员复制模块,出口信息获取模块,报文封装模块以及报文转发模块。

[0033] 具体来说,如图 2 所示的上行报文处理的流程图,VXLAN 网关对收到的上行报文通过报文解析模块进行解析,得到后续处理所需要的各种信息如外层 IP 头、以太网头,以及 VXLAN 报头等;入接口信息获取模块根据报文接收端口的配置信息来决定是否需要进行 VXLAN 上行报文处理;如果需要进行 VXLAN 上行报文的处理,则还将获得 VXLAN ID 的映射关系,该 VXLAN ID 的映射关系可通过 VLAN 或源 MAC 地址来标记 VXLAN ID。

[0034] VXLAN ID 确定模块根据之前获得的 VXLAN ID 的映射关系进行查表,如 VXLAN ID 用 VLAN 标记时,用入接口 ID 和 VLAN ID 作为关键字进行查找,在查找结果的输出中可以获得转发需要的 VXLAN ID。

[0035] 二层转发表查表模块根据报文的目的地 MAC 地址和 VXLAN ID 共同作为关键字进行查表,获得该报文的目的地以及编辑方式。如果是组播报文,则针对每一个组播组成员,获得对应的目的地以及编辑方式,后续处理将针对每一个组播组成员依次处理。在二层转发表查表模块中,通常需要设置一条默认条目,这样在查表的时候必然可以返回结果。通常情况下,广播、组播和未知单播报文都可以使用该条目,其目的是使当前报文洪泛到虚拟转发域中的所有成员。

[0036] 出接口信息获取模块根据转发表决定的目的地获取出接口的信息。

[0037] 报文封装模块根据转发表决定的封装信息,进行相应的封装,包括 VXLAN 头、UDP 头、外层 IP 头、外层以太网头的封装,然后送给报文转发模块从相应的目的端口发送出去,从而完成 VXLAN 上行报文的处理过程。

[0038] 二、下行报文处理

[0039] 下行报文是指 VXLAN 网关收到从 IP 网络发来的带有 VXLAN 封装的报文。VXLAN 网关中对下行报文的处理包括接收到下行报文后,然后将其解封装,并且使用内部 MAC 地址查表并转发给本地服务器的处理过程,如图 1 中的 VXLAN gateway 1 接到的从 IP 网络发送至服务器 1 或服务器 2 上的虚拟机的报文。VXLAN 网关中的下行报文处理装置中包括报文解析模块,入接口信息获取模块,解封装模块,二层转发表查表模块,组播模块,组播成员复制模块,出接口信息获取模块,报文编辑模块以及报文转发模块。

[0040] 具体来说,如图 3 所示的下行报文的处理流程图,报文解析模块对收到的报文进行解析,得到后续处理所需要的各种信息如外层 IP 头、以太网头,以及 VXLAN 报头等;入接口信息获取模块根据报文接收端口的配置信息以及当前报文的封装结构,来决定是否需要进行 VXLAN 下行报文处理,只有当报文具有 VXLAN 封装,并且外层目的 IP 是本接入点需要处理的时候,才会进入后续的解封装模块;否则仅仅执行普通的路由处理。解封装模块根据入接口信息,使用外层目的 IP、源 IP 和封装在报文中的 VXLAN ID 一起作为关键字进行查表,获得解封装信息。然后去掉报文的外层 IP 头、以太网头,以及 VXLAN 报头,并且获得 VXLAN 报头中所携带的虚拟转发域信息。

[0041] 二层转发表查表模块根据内层的目的 MAC 地址和封装在报文中的 VXLAN ID 共同作为关键字进行查表,得到报文的处理信息,该处理信息包括目的端口和编辑信息,由于此时不需要对报文进行封装,因此只可能有一些编辑信息,如 VLAN。当查表结果是组播的话,则需要对组播复制,然后获取每一个组播成员的目的端口和编辑信息。

[0042] 出接口信息获取模块根据查表的结果,得到需要转发出去的端口信息。

[0043] 报文编辑模块中,根据查表结果和目的端口的配置,对内层报文进行必要的编辑,然后交给报文转发模块进行转发。

[0044] 在报文编辑模块中,最主要的作用是根据 VXLAN 虚拟转发域映射出在目的端口虚拟机上所对应的 VLAN 信息,以便使服务器可以识别,发送给对应的虚拟机。

[0045] 在二层转发表查表模块中,和上行报文处理一样,需要设置一条默认条目,使报文洪泛到本地服务器中虚拟转发域中的所有成员。

[0046] 本发明的技术内容及技术特征已揭示如上,然而熟悉本领域的技术人员仍可能基于本发明的教示及揭示而作种种不背离本发明精神的替换及修饰,因此,本发明保护范围应不限于实施例所揭示的内容,而应包括各种不背离本发明的替换及修饰,并为本专利申请权利要求所涵盖。

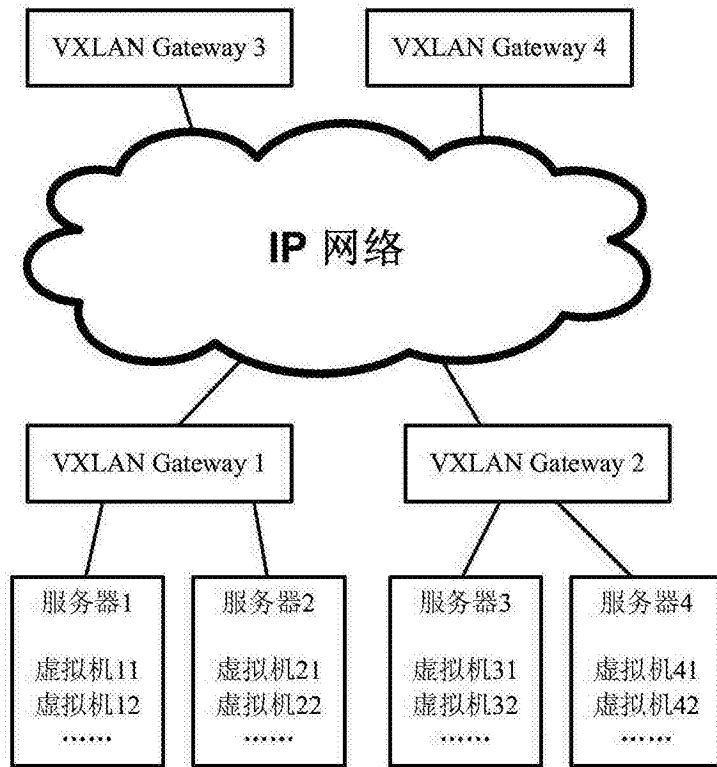


图 1

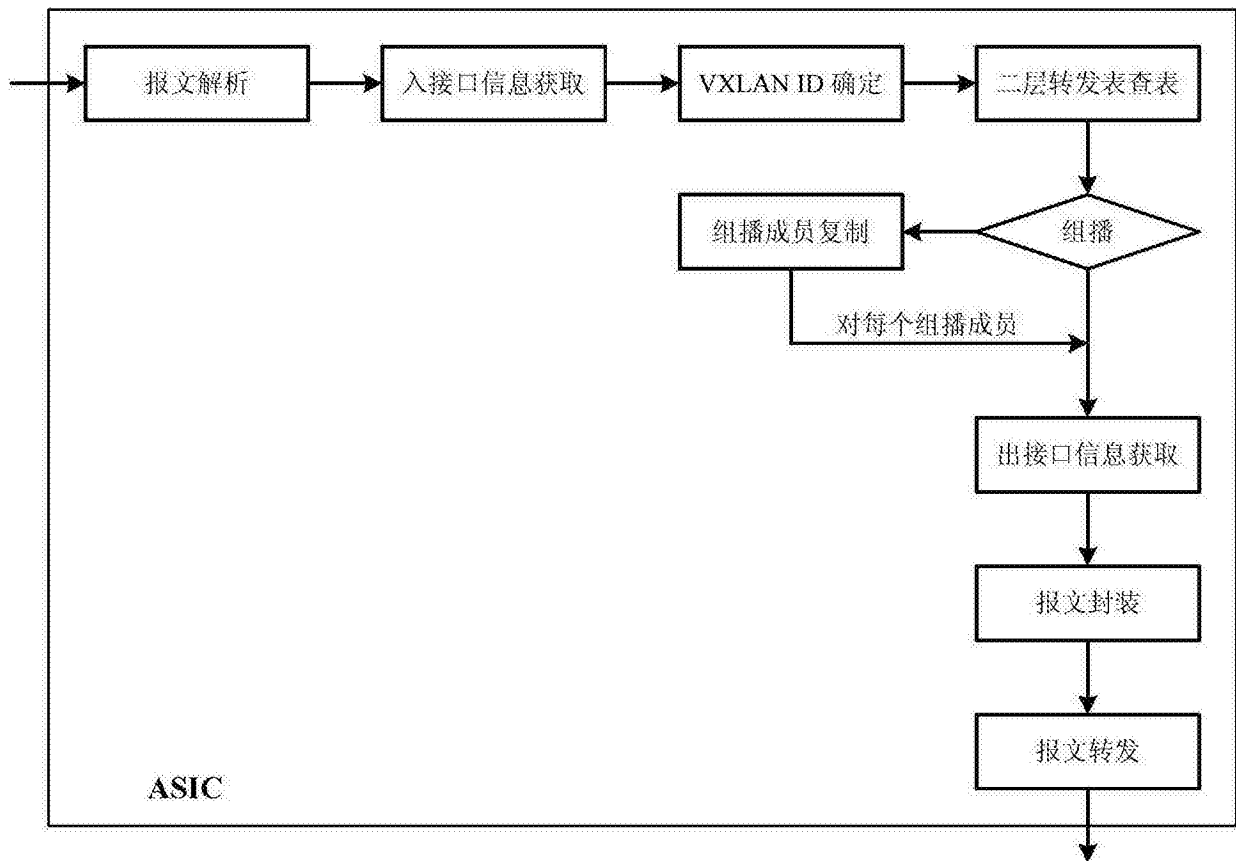


图 2

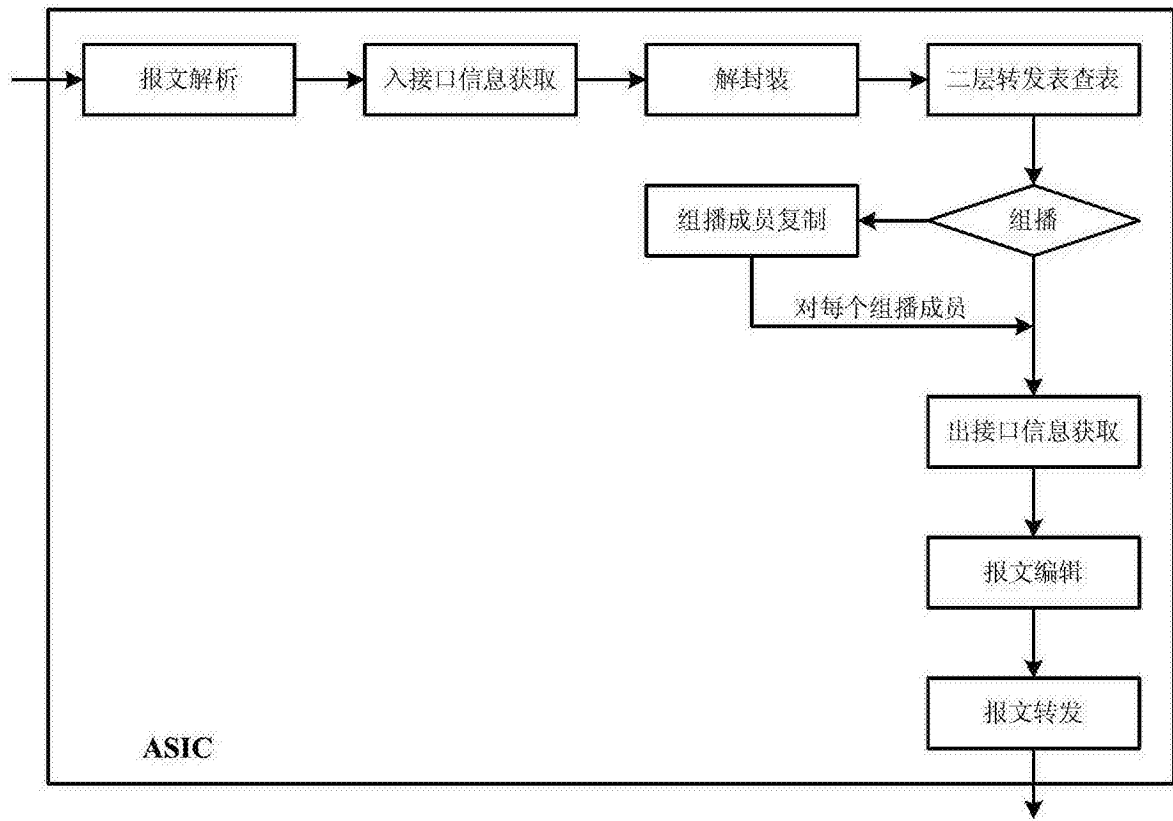


图 3