



(21) 申请号 201911426082.5

(22) 申请日 2019.12.31

(65) 同一申请的已公布的文献号
申请公布号 CN 112787923 A

(43) 申请公布日 2021.05.11

(66) 本国优先权数据
201911084038.0 2019.11.07 CN

(73) 专利权人 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 刘冰 彭书萍 李振斌 陈霞

(74) 专利代理机构 北京中博世达专利商标代理有限公司 11274
专利代理师 申健

(51) Int.Cl.

H04L 45/00 (2022.01)

(56) 对比文件

CN 101094155 A, 2007.12.26
s2-020729 "Dual-stack IPv4/IPv6
GSNs". 3GPP tsg_sa\WG2_Arch.2002, (TSGS2_23), 全文.

审查员 刘永喆

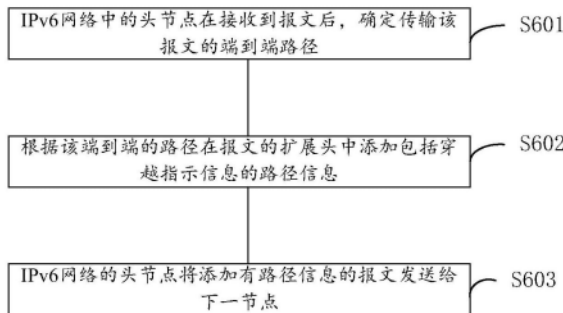
权利要求书3页 说明书18页 附图12页

(54) 发明名称

一种报文的处理方法、装置和系统

(57) 摘要

本申请实施例涉及一种报文处理方法,在该方法中,第一网络中的节点为报文确定传输路径,所述传输路径包括在第一网络中的传输路径以及穿越第二网络的路径。根据确定的传输路径,在报文中加入该路径信息;该路径信息中包括穿越指示信息,穿越指示信息指示第二网络的入口节点按照所述穿越指示信息中指定的第二网络的出口节点封装并传送报文。第一网络中的节点将添加有该穿越指示信息的报文经第一网络发往第二网络的入口节点。该方法通过在第一网络的节点实现选择第二网络的出口的能力,并通过穿越指示信息指示第二网络的入口节点按照指定出口节点路由,从而避免在第二网络的入口节点上独立部署路由策略,节约了资源开销。



1. 一种报文处理方法,其特征在于,所述方法包括:

第一网络中的头节点为报文确定传输路径,所述传输路径包括在第一网络中的路径以及穿越第二网络的路径;其中,所述第一网络为IPv6网络,所述第二网络为IPv4网络;

所述头节点根据确定的传输路径,在报文中加入路径信息,其中,所述路径信息中包括穿越指示信息,所述穿越指示信息指示了所述报文在所述第二网络中的入口节点和出口节点;

所述头节点根据所述路径信息,将所述报文经所述第一网络发往所述第二网络的所述入口节点,指示所述第二网络的所述入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式,并将封装后的所述报文路由至所述出口节点,所述出口节点用于传送所述报文。

2. 根据权利要求1所述的方法,其特征在于,所述路径信息包括所述第一网络中的路径上的中间节点的段标识SID以及所述第二网络的入口节点的SID,所述穿越指示信息包括在所述第二网络中的入口节点的SID中。

3. 根据权利要求2所述的方法,其特征在于,在确定传输路径前,所述方法还包括:

所述第一网络中的头节点接收所述第二网络的入口节点发布的至少一个SID;所述入口节点发布的不同SID携带不同的穿越指示信息;

所述第一网络中的头节点在报文中加入路径信息时,将穿越指示信息与所述第一网络中的头节点确定的穿越第二网络的路径相符的SID作为所述第二网络的入口节点的SID加入到所述路径信息中。

4. 根据权利要求1-3任一项所述的方法,其特征在于,所述穿越指示信息还指示了所述报文在穿越所述第二网络时采用的隧道封装类型;

所述报文到达所述第二网络的入口节点后,通过所述穿越指示信息指示所述入口节点按照所述隧道封装类型封装所述报文。

5. 根据权利要求4所述的方法,其特征在于,所述隧道封装类型为三层网络虚拟化NV03。

6. 根据权利要求5所述的方法,其特征在于,所述穿越指示信息还指示了封装时需拷贝的内容;

所述报文到达所述第二网络的入口节点后,通过所述穿越指示信息指示所述入口节点在封装报文时将所述需拷贝的内容,从所述第一网络所支持的协议的报文头中拷贝到封装后的所述第二网络所支持的协议的报文头中。

7. 一种报文处理方法,其特征在于,所述方法包括:

第二网络的入口节点接收来自第一网络的报文,所述报文中携带穿越指示信息,穿越指示信息中指示了所述报文在所述第二网络中的所述入口节点和出口节点;所述第一网络为IPv6网络,所述第二网络为IPv4网络;

所述入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式,其中,所述封装后的报文的地址为所述穿越指示信息中指示的出口节点;

所述入口节点将所述封装后的报文路由到所述出口节点。

8. 根据权利要求7所述的方法,其特征在于,所述报文中携带所述入口节点的段标识SID,所述穿越指示信息携带在所述入口节点的SID中。

9. 根据权利要求7所述的方法, 其特征在于, 所述入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式, 包括:

所述入口节点在所述报文外封装IPv4报文头, 将所述IPv4报文头的源地址设置为所述入口节点的IPv4地址或所述入口节点对应接口的IPv4地址, 并将所述IPv4报文头的目的地址设置为所述穿越指示信息中的出口节点的IPv4地址或所述出口节点对应接口的IPv4地址。

10. 根据权利要求7-9任一项所述的方法, 其特征在于, 所述穿越指示信息还指示了所述报文在穿越所述第二网络时采用的隧道封装类型;

所述入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式, 还包括:

所述入口节点按照所述穿越指示信息中指示的隧道封装类型, 建立与所述出口节点的传输隧道。

11. 根据权利要求10所述的方法, 其特征在于, 所述穿越指示信息中指示的隧道封装类型为三层网络虚拟化NV03。

12. 根据权利要求7-9任一项所述的方法, 其特征在于, 所述穿越指示信息还指示了封装时需拷贝的内容;

所述入口节点在将所述报文封装为所述第二网络支持的格式时, 还包括:

所述入口节点将所述穿越指示信息指示的需拷贝的内容拷贝到IPv4报文头中。

13. 根据权利要求7-9任一项所述的方法, 其特征在于, 在第二网络的入口节点接收来自第一网络的报文前, 所述方法还包括:

所述入口节点发布至少一个SID, 发布的不同SID携带不同的穿越指示信息。

14. 一种报文处理装置, 其特征在于, 所述装置为第一网络中的头节点, 所述装置包括:

确定单元, 用于为报文确定传输路径, 所述传输路径包括在第一网络中的路径以及穿越第二网络的路径; 其中, 所述第一网络为IPv6网络, 所述第二网络为IPv4网络;

报文处理单元, 用于根据所述确定单元确定的传输路径, 在报文中加入路径信息; 所述路径信息中包括穿越指示信息, 所述穿越指示信息指示了所述报文在所述第二网络中的入口节点和出口节点;

报文发送单元, 用于根据所述路径信息, 将所述报文经所述第一网络发往所述第二网络的所述入口节点, 指示所述第二网络的所述入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式, 并将封装后的所述报文路由至所述出口节点, 所述出口节点用于传送所述报文。

15. 根据权利要求14所述的装置, 其特征在于, 所述路径信息包括所述第一网络中的路径上的中间节点的段标识SID以及所述第二网络的入口节点的SID, 所述穿越指示信息包括在所述第二网络中的入口节点的SID中。

16. 根据权利要求15所述的装置, 其特征在于, 所述装置还包括:

接收单元, 用于接收所述第二网络的入口节点发布的至少一个SID; 所述入口节点发布的不同SID携带不同的穿越指示信息;

所述报文处理单元在报文中加入路径信息时, 将穿越指示信息与所述确定单元确定的穿越第二网络的路径相符的SID作为所述第二网络的入口节点的SID加入到所述路径信息

中。

17. 根据权利要求14-16任一项所述的装置,其特征在于,所述穿越指示信息还指示了封装时需拷贝的内容;

所述报文到达所述第二网络的入口节点后,通过所述穿越指示信息指示所述入口节点在封装报文时将所述需拷贝的内容,所述第一网络所支持的协议的报文头中拷贝到封装后的所述第二网络所支持的协议的报文头中。

18. 一种报文处理装置,其特征在于,所述装置为第二网络的入口节点,所述装置包括:

报文获取单元,用于接收来自第一网络的报文,所述报文中携带穿越指示信息,穿越指示信息中指示了所述报文在所述第二网络中的入口节点和出口节点;所述第一网络为IPv6网络,所述第二网络为IPv4网络;

报文处理单元,用于根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式;所述封装后的报文的地址为所述穿越指示信息中指示的出口节点;

报文发送单元,用于将所述报文处理单元封装后的报文路由到所述出口节点。

19. 如权利要求18所述的装置,其特征在于,所述报文中携带所述装置的段标识SID,所述穿越指示信息携带在所述装置的SID中。

20. 如权利要求18所述的装置,其特征在于,所述报文处理单元根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式,包括:

所述报文处理单元在所述报文外封装IPv4报文头,将所述IPv4报文头的源地址设置为所述入口节点的IPv4地址或所述入口节点对应接口的IPv4地址,并将所述IPv4报文头的目的地址设置为所述穿越指示信息中的出口节点的IPv4地址或所述出口节点对应接口的IPv4地址。

21. 如权利要求18-20任一项所述的装置,其特征在于,所述穿越指示信息还指示了所述报文在穿越所述第二网络时采用的隧道封装类型;

所述报文处理单元根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式,还包括:

所述报文处理单元按照所述穿越指示信息中指示的隧道封装类型,建立与所述出口节点的传输隧道。

22. 如权利要求21所述的装置,其特征在于,所述穿越指示信息中指示的隧道封装类型为三层网络虚拟化NV03。

23. 如权利要求18-20任一项所述的装置,其特征在于,所述穿越指示信息还指示了封装时需拷贝的内容;

所述报文处理单元在将所述报文封装为所述第二网络支持的格式时,还包括:

所述报文处理单元将所述穿越指示信息指示的需拷贝的内容拷贝到IPv4报文头中。

24. 如权利要求18-20任一项所述的装置,其特征在于,所述装置还包括:

发布单元,用于发布至少一个SID,所述SID携带穿越指示信息,在发布多个SID的情况下,不同SID携带不同的穿越指示信息。

25. 一种报文处理系统,其特征在于,包括如权利要求14-17任一项所述的装置以及如权利要求18-24任一项所述的装置。

一种报文的处理方法,装置和系统

技术领域

[0001] 本发明涉及通信领域,尤其涉及一种报文的处理技术。

背景技术

[0002] 段路由(segment routing,SR)是一种源路由技术,由SR隧道的头节点来为报文指定路径,并将路径转换成一个有序的段列表封装到报文头中,位于该路径的中间节点只需要根据报文头中指定的路径进行转发。其中,该SR隧道的头节点为报文进入SR隧道的第一个节点。当SR部署在第六版因特网协议(Internet Protocol version 6,IPv6)数据平面上时,称为IPv6段路由(Segment Routing over IPv6,简称SRv6)。为了在IPv6中支持段路由协议,在IPv6的基本首部后面新增了扩展头,即段路由头(Segment Routing Header,SRH),从而得到SRv6数据包。SRH扩展头部携带SL(segment list,段列表),每一个分段用一个128bit的IPv6地址来标示,且报文转发依赖IP地址编码的信息。在IPv6路由器中,段列表只能用IPv6地址形式的SID(Segment Identifier,段标识符)表示,否则不能转发。因此,SRH中的SL是一个显式的段标识符(SID)栈,通过中间节点不断的进行更新目的地址和偏移段标识栈的操作来完成逐跳转发。SRv6基于此实现在SR隧道的头节点上对转发路径的控制。

[0003] 但是,SRv6的转发面是IPv6,当网络中存在非IPv6网络(如,只支持IPv4转发的网络)时,SRv6则需要依赖其他技术穿越该网络。已有相关技术往往通过非IPv6网络与IPv6网络连接的边缘节点单独配置选路策略来实现对非IPv6网络的穿越。同一边缘节点处可存在多条路径。在边缘节点收到IPv6报文时,由边缘节点根据配置的选路策略选择穿越IPv4网络的路径。

[0004] 由于已有相关技术在穿越非IPv6网络时还需要在边缘节点部署独立的选路策略,导致资源浪费,而且无法实现SR隧道的头节点对端到端的转发路径的控制。

发明内容

[0005] 本申请实施例提供了一种报文处理方法、装置、以及系统,实现对IPv6报文端到端的传输控制。

[0006] 第一方面,本申请提供了一种报文处理方法。在该方法中,第一网络中的节点为报文确定传输路径,包括在第一网络中的传输路径以及穿越第二网络的路径。该第一网络中的节点根据确定的传输路径,在报文中加入路径信息。该路径信息中包括穿越指示信息,穿越指示信息指示了所述报文在第二网络中的出口节点。该第一网络中的节点将添加有所述穿越指示信息的报文经第一网络发往第二网络的入口节点,指示第二网络的入口节点按照所述穿越指示信息中指定的第二网络的出口节点封装并传送该报文。

[0007] 通过在该第一网络中的节点实现选择第二网络的隧道出口的能力,并通过穿越指示信息指示第二网络的入口节点按照指定出口节点路由,从而避免在第二网络的入口节点上独立部署路由策略,节约了资源开销,而且还实现了第一网络中的节点对端到端的转发路径的控制。

[0008] 一种具体的实施方式中,该路径信息可包括第一网络中的路径上的中间节点的段标识SID以及所述第二网络的入口节点的SID,穿越指示信息包括在所述第二网络中的入口节点的SID中。通过改造SID,将穿越指示信息携带在在第二网络中的入口节点的SID中,避免对设备进行大量改造,节约成本。

[0009] 一种具体的实施方式中,在确定传输路径前,第一网络中的节点还接收第二网络的入口节点发布的至少一个SID。其中,不同SID携带不同的穿越指示信息,第一网络中的节点在报文中加入路径信息时,将穿越指示信息与第一网络中的节点确定的穿越第二网络的路径相符的SID作为所述第二网络的入口节点的SID加入到所述路径信息中。通过不同SID携带不同的穿越指示信息,来表示不同的穿越路径,使第一网络中的节点通过选择SID来实现对穿越路径的选择,以一种代价较小的方式实现了头节点对端到端的转发路径的控制。

[0010] 一种具体的实施方式中,穿越指示信息还指示了所述报文在穿越所述第二网络时采用的隧道封装类型。报文到达所述第二网络的入口节点后,通过该穿越指示信息指示入口节点按照所述隧道封装类型封装所述报文。通过穿越指示信息,第一网络中的节点不仅实现了对穿越路径的控制,还进一步实现了对穿越隧道类型的选择,使得第一网络中的节点可以根据业务情况,进行灵活的选路,增强了第一网络中的节点的灵活性。

[0011] 一种具体的实施方式中,穿越指示信息还可以指示封装时需拷贝的内容。在所述报文到达所述第二网络的入口节点后,通过所述穿越指示信息指示所述入口节点在封装报文时将所述需拷贝的内容,从所述第一网络所支持的协议的报文头中拷贝到封装后的所述第二网络所支持的协议的报文头中。通过在穿越指示信息中实现拷贝能力,可以使报文在原网络中的能力能够迁移到另一网络中,保障穿越过程中业务能力不丢失。

[0012] 一种具体的实施方式中,所述需拷贝的内容为随流操作、管理和维护in-situ operation administration and maintenance(iOAM)信息,或inbound OAM(iOAM)带内操作、管理和维护inband OAM。

[0013] 本实施方式中,通过将第一网络所支持的协议的报文头中的iOAM信息,拷贝至封装后新增加的第二网络所支持的协议的报文头中,可实现端到端路径上iOAM数据的采集,从而能够更加准确的获取整个网络OAM相关信息。

[0014] 一种具体的实施方式中,所述穿越指示信息中指示的隧道封装类型为三层网络虚拟化NV03,所述NV03包括虚拟可扩展局域网Virtual eXtensible Local Area Network(VXLAN),虚拟扩展局域网的通用协议扩展Virtual eXtensible Local Area Network-Generic Protocol Extension(VXLAN-GPE),或通用网络虚拟化封装Generic Network Virtualization Encapsulation(GENEVE)。

[0015] 一种具体的实施方式中,所述第一网络为IPv6网络,所述第二网络为IPv4网络。

[0016] 第二方面,本申请提供了一种报文处理方法。在该方法中,第二网络的入口节点接收来自第一网络的报文,该报文中携带穿越指示信息,穿越指示信息中指示了所述报文在所述第二网络中的出口节点。入口节点根据该穿越指示信息将所述报文封装为第二网络支持的格式。其中,封装后的报文的地址为所述穿越指示信息中指示的出口节点。入口节点将封装后的报文路由到该出口节点。

[0017] 第二网络的入口节点根据第一网络发来的报文中的穿越指示信息来路由报文,从而避免在第二网络的入口节点上独立部署路由策略,节约了资源开销,而且还实现了由第

一网络中的第一网络中的节点对端到端的转发路径的控制。

[0018] 一种具体的实施方式中,该报文中携带入口节点的段标识SID,所述穿越指示信息携带在该入口节点的SID中。通过改造SID,将穿越指示信息携带在在第二网络中的入口节点的SID中,避免对设备进行大量改造,节约了成本。

[0019] 一种具体的实施方式中,该方法可以应用在IPv6网络中穿越IPv4网络的场景中,其中,第一网络为IPv6网络,第二网络为IPv4网络,入口节点根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式的具体实现过程如下:入口节点在报文外封装IPv4报文头,将所述IPv4报文头的源地址设置为所述入口节点的IPv4地址或所述入口节点对应接口的IPv4地址,并将所述IPv4报文头的目的地址设置为所述穿越指示信息中的出口节点的IPv4地址或所述出口节点对应接口的IPv4地址。通过穿越指示信息指示IPv4网络中的入口节点按照指定出口节点路由,从而在不对入口节点上部署独立路由策略的情况下,实现了在IPv6网络中穿越IPv4网络,节约了资源开销。

[0020] 一种具体的实施方式中,穿越指示信息还指示了所述报文在穿越所述第二网络时采用的隧道封装类型。入口节点将所述报文封装为所述第二网络支持的格式时,按照所述穿越指示信息中指示的隧道封装类型,建立与所述出口节点的传输隧道。通过穿越指示信息,第一网络中的节点不仅实现了对穿越路径的控制,还进一步实现了对穿越隧道类型的选择,使得第一网络中的节点可以根据业务情况,进行灵活的选路,增强了第一网络中的节点的灵活性。

[0021] 一种具体的实施方式中,所述穿越指示信息中指示的隧道封装类型为NV03。

[0022] 一种具体的实施方式中,所述穿越指示信息还指示了封装时需拷贝的内容。所述入口节点在将所述报文封装为所述第二网络支持的格式时,还将所述穿越指示信息指示的需拷贝的内容拷贝到IPv4报文头中。通过在穿越指示信息中实现拷贝能力,可以使报文在原网络中的能力能够迁移到另一网络中,保障穿越过程中业务能力不丢失。

[0023] 一种具体的实施方式中,第二网络的入口节点接收来自第一网络的报文前,所述入口节点发布至少一个SID,发布的不同SID携带不同的穿越指示信息。通过不同SID携带不同的穿越指示信息,来表示不同的穿越路径,使第一网络中的节点通过选择SID来实现对穿越路径的选择,以一种代价较小的方式实现了第一网络中的节点对端到端的转发路径的控制。

[0024] 第三方面,本申请实施例提供一种报文处理装置,包括执行如上述第一方面或者第一方面任意具体实施方式中所描述方法的单元。

[0025] 该报文处理装置及其各实施例的技术效果可参考上述第一方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0026] 第四方面,本申请实施例提供一种报文处理装置,包括执行如上述第二方面或者第二方面任意具体实施方式中所描述方法的单元。

[0027] 该报文处理装置及其各实施例的技术效果可参考上述第二方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0028] 第五方面,本申请实施例提供一种报文处理装置,包括处理器、通信接口以及存储器;所述存储器用于存储指令,所述处理器用于执行所述指令,所述通信接口用于接收或者发送报文;其中,所述处理器执行所述指令时执行如上述第一方面或者第一方面的任意具

体实施方式中所描述方法。

[0029] 该报文处理装置及其各实施例的技术效果可参考上述第一方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0030] 第六方面,本申请实施例提供一种报文处理装置,包括处理器、通信接口以及存储器;所述存储器用于存储指令,所述处理器用于执行所述指令,所述通信接口用于接收或者发送报文;其中,所述处理器执行所述指令时执行如上述第二方面或者第二方面的任意具体实施方式中所描述方法。

[0031] 该报文处理装置及其各实施例的技术效果可参考上述第二方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0032] 第七方面,本申请提供一种计算机存储介质,所述计算机介质存储有用于网络故障定位的计算机程序,所述计算机程序被处理器执行时实现如上述第一方面或者第一方面的任意具体实施方式中所描述方法。

[0033] 该报文处理装置及其各实施例的技术效果可参考上述第一方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0034] 第八方面,本申请提供一种计算机存储介质,所述计算机介质存储有用于网络故障定位的计算机程序,所述计算机程序被处理器执行时实现如上述第二方面或者第二方面的任意具体实施方式中所描述方法。

[0035] 该报文处理装置及其各实施例的技术效果可参考上述第二方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0036] 第九方面,本申请实施例提供一种芯片,包括:处理器以及接口电路,所述接口电路用于接收指令并传输至处理器;处理器用于接受接口电路的指令,执行上述第一方面或第一方面的任一种具体的实施方式中的方法,或实现上述第二方面或第二方面的任一种具体实施方式中的方法。

[0037] 可选地,该芯片系统中的处理器可以为一个或多个。该处理器可以通过硬件实现,也可以通过软件实现。当通过硬件实现时,该处理器可以是逻辑电路、集成电路等。当通过软件实现时,该处理器可以是一个通用处理器,通过读取存储器中存储的软件代码来实现。

[0038] 可选地,该芯片还包括一个或多个存储器。该存储器可以与处理器集成在一起,也可以和处理器分离设置。

[0039] 该报文处理装置及其各实施例的技术效果可参考上述第一方面以及各可选实施方式中的具体描述,或者上述第二方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

[0040] 第十方面,本申请实施例提供一种报文处理系统,该系统包括上述第三或第五方面所描述的报文处理装置,以及上述第四或第六方面所描述的报文处理装置。

[0041] 该报文处理系统的各实施例的技术效果可参考上述第一至二方面以及各可选实施方式中的具体描述,这里不再逐一赘述。

附图说明

[0042] 图1所示为本申请实施例中的一种SRH扩展头格式示意图;

[0043] 图2所示为本申请实施例中的一种SRH扩展头抽象格式示意图;

- [0044] 图3所示为本申请实施例中的一种IPv6 DA变换示意图；
- [0045] 图4所示为本申请实施例一种SRv6 SID结构示意图；
- [0046] 图5所示为本申请实施例一种应用场景示意图；
- [0047] 图6所示为本申请实施例一种第一网络中的节点对报文的处理流程图；
- [0048] 图7所示为本申请实施例中的一种IPv4网络的入口节点的处理流程图；
- [0049] 图8所示为本申请实施例中采用VXLAN-GPE穿越IPv4的应用场景示意图；
- [0050] 图9所示为本申请实施例中SRv6采用VXLAN-GPE穿越IPv4的方法流程图；
- [0051] 图10所示为本申请实施例中的一种采用VXLAN-GPE封装后的报文格式示意图；
- [0052] 图11所示为本申请实施例中的一种采用GENEVE作为封装方式进行封装后的报文格式示意图；
- [0053] 图12所示为本申请实施例中的一种采用VXLAN作为封装方式进行封装后的报文格式示意图；
- [0054] 图13所示为本申请实施例中的一种采用NVGRE作为封装方式进行封装后的报文格式示意图；
- [0055] 图14所示为本申请实施例中的一种报文处理装置结构示意图；
- [0056] 图15所示为本申请实施例中的一种报文处理装置结构示意图；
- [0057] 图16所示为本申请实施例中的一种报文处理装置结构示意图；
- [0058] 图17所示为本申请实施例中的一种报文处理装置结构示意图；
- [0059] 图18所示为本申请实施例中的一种报文处理装置结构示意图；
- [0060] 图19所示为本申请实施例中的一种报文处理装置结构示意图。

具体实施方式

[0061] 下面结合附图和具体实施例对本发明展开详细说明。但需要注意，下面的这些实施例，仅为帮助理解技术方案所举的例子，并不用于限定本发明。

[0062] 本申请实施例主要应用于采用穿越技术来传输报文的场景。通常情况下，在传输报文的第一网络中存在一段不同于第一网络的第二网络时，需要采用穿越技术来穿越第二网络，以完成报文的传输。本申请实施例在需要穿越网络时，由第一网络中的节点确定报文传输路径，包括在第一网络中的传输路径以及穿越第二网络的路径。本申请实施例中的第一网络中的节点指的是第一网络中建立的SR隧道的头节点。第一网络中的节点根据确定的路径，在报文中加入端到端的路径信息，其中，该路径信息中包括穿越指示信息，用于指示报文在第二网络中的入口节点和出口节点。第二网络中的入口节点在接收到报文后，根据报文中的穿越指示信息，将所述报文封装为所述第二网络支持的格式，并将封装后的报文路由到指定的出口节点，从而完成穿越。这样，通过在第一网络中的节点实现选择第二网络的隧道出口的能力，并通过穿越指示信息指示第二网络的入口节点按照指定出口节点路由，从而避免在第二网络的入口节点上独立部署路由策略，节约了资源开销，而且还实现了第一网络中的节点对端到端的转发路径的控制。

[0063] 在具体实施例时，第一网络中的节点不仅可以用穿越指示信息来指示第二网络的入口节点和出口节点，还可以用穿越指示信息来指示需采用的隧道封装类型，第二网络的入口节点在接收到报文后，按照穿越指示信息指定的隧道封装类型对报文进行封装。第一

网络中的节点还可以利用穿越指示信息来实现封装前的报文和封装后的报文间的信息拷贝。具体的,第一网络中的节点可以在穿越指示信息中指示需要拷贝的信息,第二网络的入口节点在封装报文时,按照穿越指示信息指定的需要拷贝的信息进行拷贝。

[0064] 下面以IPv6网络中存在一段需穿越的IPv4网络为例对本申请实施例的实施过程进行详细说明。

[0065] 由于本申请实施例在实施过程中,涉及到对报文的更改,因此,为了便于理解,先对IPv6网络中传输的报文的格式进行简单介绍。

[0066] IPv6网络中传输的报文是采用的IPv6协议的IPv6报文或者是基于IPv6报文扩展得到的SRv6报文。本申请实施例提供的方案可基于SRv6报文来实施。IPv6报文是由IPv6标准头+扩展头(0...n)+负载Payload组成。为了基于IPv6转发平面实现SRv6,新增加一种IPv6扩展头,称作段路由头(Segment Routing Header,SRH),该扩展头指定一个IPv6的显式路径,存储的是IPv6的段列表(Segment List)信息。SRv6隧道的头节点在IPv6报文增加一个SRH扩展头(即得到SRv6报文),中间节点就可以按照SRH扩展头里包含的路径信息转发。图1为本申请实施例提供的一种SRH扩展头的格式。

[0067] 其中,图1所示的各SRH扩展头字段解释如下

[0068] Next Header字段:长度8比特,用于标识紧跟在SRH之后的报文头的类型。

[0069] Hdr Ext Len字段:长度8比特,用于表示SRH头的长度。主要是指从Segment List[0]到Segment List[n]所占用长度。

[0070] Routing Type字段:长度8比特,用于标识路由头部类型,SRH Type是4。

[0071] Segments Left字段:长度8比特,用于表示到达目的节点前仍然应当访问的中间节点数。

[0072] Last Entry字段:长度8比特,用于表示在段列表中包含段列表的最后一个元素的索引。

[0073] Flags字段:长度8比特,数据包的一些标识。

[0074] Tag字段:长度16比特,标识同组数据包。

[0075] Segment List[n]字段:长度128*n比特,由该字段组成段列表,段列表从路径的最后一段开始编码。Segment List是IPv6地址形式。

[0076] 为了便于叙述转发原理,图1所示SRH扩展头可以抽象成为图2的形式。

[0077] 图2中各个字段解释如下:IPv6 Destination Address:IPv6报文的目地址,简称IPv6 DA。在普通IPv6报文里,IPv6 DA是固定不变的。在SRv6中,IPv6 DA仅标识当前报文的下一个节点,是不断变换的。

[0078] <Segment List[0],Segment List[1],Segment List[2],...,Segment List[n]>:SRv6报文的段列表,在入节点生成。Segment List[0]是SRv6路径上第一个需要处理的SRv6 Segment,Segment List[1]是第二个,Segment List[2]是第三个,...,Segment List[n]是第n+1个。

[0079] 如图3所示,在SRv6中,每经过一个SRv6节点,Segments Left(SL)字段减1,IPv6 DA信息变换一次。Segments Left和Segments List字段共同决定IPv6 DA信息。如果SL值是n(即n-0),则IPv6 DA取值就是Segments List[0]的值。

[0080] • 如果SL值是n-1,则IPv6 DA取值就是Segments List[1]的值。

[0081] • 如果SL值是n-2,则IPv6 DA取值就是Segments List[2]的值。

[0082] • ...

[0083] • 如果SL值是0(n-n=0),则IPv6 DA取值就是Segments List[n]的值。

[0084] 需要说明的是,以上实施例中的字段或者报文结构仅是为了说明SRv6报文及其规则,本申请并不限于此,具体实践中,可以根据不同的实现场景,做不改变基本原理的一些具体实现层面的变化。

[0085] 段列表中的Segment list是IPv6地址形式,通常也可以称为SID(Segment Identifier)。如图4所示,图4为本申请实施例提供的SRv6 SID的格式示意图,其中,SID包括定位(Locator)和功能(Function),相应的格式为Locator:Function,其中Locator占据IPv6地址的高比特位,Function部分占据IPv6地址的剩余部分。

[0086] Locator具有定位功能,所以要在SR域内唯一。节点配置Locator之后,系统会生成一条Locator网段路由,并且通过IGP在SR域内扩散。网络里其他节点通过Locator网段路由就可以定位到本节点,同时本节点发布的所有SRv6 SID也都可以通过该条Locator网段路由到达。Function代表设备的指令(Instruction),这些指令可由设备预先设定,Function部分用于指示SID的生成节点进行相应的功能操作。Function部分还可以分出一个可选的参数段(Arguments),此时SRv6 SID的格式变为Locator:Function:Arguments,Arguments占据IPv6地址的低比特位,通过Arguments字段可以定义一些报文的流和服务等信息。Function和Arguments都是可以定义的。

[0087] SRv6 SID有很多类型,不同类型的SRv6 SID代表不同的功能。

[0088] 本申请实施例提出了一种新的SRv6 SID的功能,即提出一种新类型的SID,用以指示一种新的网络节点的行为。本申请实施例定义的新的SID类型可以表示为:End.T4.Encaps,其全称为Endpoint encapsulates the original SRv6 packet into an IPv4 tunnel。在具体实现时,可以通过在SID的Function部分定义新的比特位来代表新的网络行为。新定义的比特位以及该比特位所代表的含义,可由控制面发布到转发面的各节点,使转发面的节点可以使用和识别该新类型的SID。本申请实施例中新定义的SID所定义的网络行为为穿越网络的行为。其中可包括穿越指示信息,用于指示在SRv6报文穿越IPv4网络时,该SID所对应的IPv4网络边缘节点在接收到该SRv6报文时的行为。具体的,该SID中的穿越指示信息可以指示穿越IPv4时的出口节点,也可以进一步指示穿越IPv4时采用的隧道的封装类型或指示穿越IPv4时,需要拷贝到封装成IPv4格式的报文中的信息。

[0089] 下面结合应用场景对该新增SID的应用进行详细说明。图5为应用本申请实施例的一种场景示意图。如图5所示,网络中存在三段网络,两段IPv6网络由一段IPv4网络相连。其中,IPv6网络为本申请实施例中的第一网络,IPv4网络为本申请实施例中需要穿越的第二网络。由R1发出的SRv6报文需要穿过中间的IPv4网络到达R7。R3、R4和R5三个网络设备是IPv4网络的边缘节点,同时支持IPv4、IPv6双栈。R3、R4和R5之间存在隧道,当SRv6报文到达R3时,可以选择R4或者R5作为隧道出口。

[0090] 本申请实施例中,边缘节点在发布自身的SID时,可以采用本申请实施例中新定义的SID。例如,图5中的R3可发布两个SID:SID1和SID2。其中,SID1的function部分指示的穿越IPv4时的出口节点为R5。SID2的function部分指示的穿越IPv4时的出口节点为R4。在具体实现时,边缘节点还可以在SID的function部分指示穿越时所采用的隧道封装类型或穿

越时需要拷贝到封装后的格式中的内容等。

[0091] 下面以图5所示场景为例,对本申请实施例提供的SRv6报文穿越IPv4网络的方法进行说明。图6为本申请实施例提供的IPv6网络中的SRv6隧道的头节点对报文的处理流程图。如图6所示,该头节点的处理过程包括:

[0092] S601:IPv6网络中的头节点在接收到报文后,确定传输该报文的端到端路径。

[0093] 其中,IPv6网络是部署了SR的SRv6网络。IPv6网络中的头节点是该IPv6网络中建立的SRv6隧道的头节点,即报文进入SRv6隧道的第一个节点(即该报文在SRv6网络的入口节点),如图5中的节点R1。

[0094] 头节点在接收到报文后,为该报文选择一条端到端的传输路径,其中,该端到端路径包括在SRv6网络中的传输路径以及穿越IPv4网络的路径。其中,穿越IPv4网络的路径可以是确定了IPv4网络入口节点和出口节点的一条隧道。以图5为例,R1可以选择R3作为IPv4网络的入口节点,R5作为IPv4网络的出口节点。

[0095] 需要说明的是,IPv4网络的入口节点和出口节点是IPv4网络与IPv6网络连接的边缘节点,支持IPv4和IPv6双栈。

[0096] S602:IPv6网络的头节点根据该端到端的路径在报文的扩展头中添加路径信息。其中,该路径信息中包括穿越指示信息,该穿越指示信息用于指示报文在IPv4网络中的入口节点和出口节点。

[0097] 该路径信息可以以SRv6报文格式中的段列表的形式存在。具体的,头节点可以将路径信息添加到报文的SRH扩展头中,在SRH扩展头的段列表中存放该端到端的路径中的中间节点的SID。其中,在报文中携带的IPv4网络入口节点的SID可采用本申请实施例提供的具有新增功能的SID。穿越指示信息则携带于在该SID中,用于指定IPv4网络入口节点的行为。

[0098] 需要说明的是,IPv4网络入口节点发布的SID可能有多个,不同SID携带不同的穿越指示信息。例如,图5中,R3发布SID1和SID2。SID1中的穿越指示信息指示的穿越IPv4时的出口节点为R5。SID2中的穿越指示信息指示的穿越IPv4时的出口节点为R4。头节点在报文中加入端到端的路径信息时,可以将穿越指示信息与头节点确定的穿越第二网络的路径相符的SID作为IPv4网络入口节点的SID加入到端到端的路径信息中。例如,本实施例中头节点确定的穿越IPv4网络的路径为R3作为IPv4网络的入口节点,R5作为IPv4网络的出口节点。因此,头节点可选择SID1作为IPv4网络的入口节点的SID加入到路径信息中。

[0099] 本实施例中,穿越指示信息还可以进一步指示穿越IPv4网络时的隧道类型,例如,隧道类型为NV03。

[0100] S603,IPv6网络的头节点将添加有路径信息的报文发送给下一节点。

[0101] 如图5中的头节点R1将添加有路径信息的报文发送给中间节点R2。中间节点R2接收报文后,按照报文中的段列表中的节点地址继续将报文发送给下一节点R3。到该步骤可采用已有的处理过程,这里不再赘述。

[0102] 以上过程为传输报文的头节点的处理过程,在该过程中,IPv6网络中的头节点除了具备选择IPv6网络中路径的能力(由SRv6 SRH实现),还要具备选择穿越IPv4网络的隧道出口的能力,通过穿越指示信息指示IPv4网络的入口节点按照指定出口节点路由,从而避免在IPv4网络的入口节点上独立部署路由策略,节约了资源开销,而且还实现了头节点对

端到端的转发路径的控制。

[0103] 下面进一步对IPv4网络中的入口节点的处理过程进行说明。如图7所示,图7为本申请实施例提供的一种IPv4网络的入口节点的处理流程图。该处理过程包括:

[0104] S701:IPv4网络的入口节点接收到来自IPv6网络的报文后,对IPv6中的剩余段数(Segment Left,SL)字段以及IPv6目的地址(IPv6 DA)字段进行更新。

[0105] 入口节点在接收到报文后,判断报文中下一报文头是SRH,且SL大于0,则执行该更新操作。该更新操作包括:将SL字段的值减1,并使用SRH[SL]替换IPv6 DA字段中的地址。该更新过程可参见图3部分的说明,这里不再赘述。

[0106] 需要说明的是,如果SL等于0,则表示该报文已到达目的节点,报文传输流程终止。而在本实施例中该入口节点不是目的节点。

[0107] 同样以图5所示场景为例,入口节点R3在接收到后,将SL字段的值修改为2,并将IPv6 DA字段中的地址修改为SRH[2]对应的地址。

[0108] S702:入口节点将报文封装为IPv4网络支持的格式,并在封装过程中按照穿越指示信息将报文的源地址设置为对应入口节点的地址,将报文的地址设置为对应穿越指示信息中的出口节点的地址。

[0109] 具体的,入口节点在更新IPv6报文后,会在IPv6报文外层封装IPv4报文头,以便该报文在IPv4网络中传输。在封装过程中,入口节点根据SID中的穿越指示信息将IPv4报文头的源地址设置为该入口节点的IPv4地址或该入口节点对应接口的IPv4地址,IPv4报文头的目的地址设置为穿越指示信息中的出口节点的IPv4地址或该出口节点对应接口的IPv4地址。

[0110] 在穿越IPv4网络时,可采用隧道技术进行穿越。不同的隧道技术,有不同的封装方式。也就是说,在封装报文的时候,可以采用不同的隧道封装类型。该隧道封装类型可以在穿越指示信息中指定。本实施例中穿越指示信息中指示采用NV03隧道封装类型。采用三层网络虚拟化Network Virtualization over Layer 3(NV03)隧道封装类型,IPv6报文可以通过叠加网络Overlay network,如二层租户网络,穿越IPv4网络。它通过在IPv4网络的入口节点和出口节点之间建立起Overlay隧道,利用Overlay隧道进行穿越。NV03隧道封装类型的封装方式可包括:虚拟可扩展局域网Virtual eXtensible Local Area Network(VXLAN),虚拟扩展局域网的通用协议扩展Virtual eXtensible Local Area Network-Generic Protocol Extension(VXLAN-GPE),或通用网络虚拟化封装Generic Network Virtualization Encapsulation(GENEVE)等。

[0111] 为了实现NV03隧道封装,入口节点可进一步在报文外层封装UDP报头,并设置Overlay协议报头对应的UDP目的端口和Overlay报头。该过程具体可包括:

[0112] 设置UDP端口号,目的端口号为Overlay报头对应的端口号,例如,VXLAN(4789), VXLAN-GPE(4790),或GENEVE(6081)。

[0113] 根据SID的指示填充虚拟网络ID字段,如图8中的VNI=200。

[0114] 设置Overlay报头中的Next-header字段,如图8中的next protocol=IPv6。

[0115] 上述封装UDP报头及Overlay报头的过程均可采用已有的实现方式来实现,这里不再赘述。

[0116] 需要说明的是,在封装过程中涉及到的已有技术中的一些处理,由于本申请实施

例中不做更改,在这里就不一一列举了,例如,封装时将内层IPv6报文的hop limit减1等。

[0117] S703,入口节点将报文通过建立的隧道将报文路由到穿越指示信息中指示的IPv4网络的出口节点。

[0118] 报文到达出口节点后,出口节点对报文进行解封装,暴露原始IPv6报文,然后在IPv6网络中继续按现有的IPv6或SRv6转发流程进行转发,直到到达目的节点。

[0119] 例如,在图5中,报文到达出口节点R5后,出口节点R5将报文解封后,继续转发给下一节点R6,直至到达目的节点R7。

[0120] 下面分别以通过VXLAN,VXLAN-GPE,GENEVE这几种不同的封装方式来穿越IPv4网络为例进行说明。

[0121] 图8为本申请实施例提供的一种采用VXLAN-GPE穿越IPv4的应用场景图。下面通过图9对图8所示的应用场景的处理流程进行说明。图9为本申请实施例提供的一种SRv6采用VXLAN-GPE穿越IPv4的方法流程图,该方法包括:

[0122] S901,IPv6网络中的头节点(节点E)将IPv4网络的入口节点(节点2)发布的SID封装在SRv6的SRH报头中。本实施例中,IPv6网络中的头节点是该IPv6网络中建立的SRv6隧道的头节点,即报文进入SRv6隧道的第一个节点(即该报文在SRv6网络的入节点),如图8中的节点E,或图5中的节点R1。

[0123] 其中,节点2的SID采用本申请实施例提供的新的SID类型,该SID的示例如图8所示。节点E对该SID的封装过程可参见图6所示实施例来实现,这里不再赘述。

[0124] S902,报文经非边缘节点(如图8中的节点1)时,按照普通IPv6或SRv6转发流程进行转发。

[0125] 需要说明的是,在本申请实施例中,边缘节点是支持IPv6与IPv4双栈的节点,连接着IPv6网络与IPv4网络。报文在传输过程中到达的第一个边缘节点即为IPv4网络的入口节点。

[0126] S903,判断报文是否到达边缘节点(图8中的节点2),若未到达边缘节点,则继续执行步骤S902;若到达边缘节点(即到达IPv4网络的入口节点),则执行步骤S904-S908;

[0127] 步骤S904,该边缘节点按照SID的指示,在IPv6报头外封装IPv4报头,UDP报头和VXLAN-GPE报头。

[0128] 如图8所示,报文在到达节点2时,按照SRH中对应的SID的指示(即SID中的穿越指示信息的指示),将原始报文使用VXLAN-GPE进行封装,在VXLAN-GPE头之前添加IPv4头和UDP头,以穿越节点2和节点3之间的IPv4网络。其中,IPv4报文头的源地址和目的地址可设为分别设备2和3对应的接口地址,UDP目的端口号为4790,VNI按照SID的指示进行设定(图8中举例VNI=200),VXLAN-GPE中的next protocol字段设为IPv6对应的“0x2”。图10为采用本实施例的方式封装后的报文格式。封装过程的具体实现可参考图7所示实施例中的步骤S702,这里不再赘述。

[0129] S905,封装后的报文在IPv4网络中按照现有IPv4转发流程转发。

[0130] 具体实现方式参考已有的IPv4转发处理方式来实现,这里不再赘述。

[0131] S906,报文到达IPv4目的地址所指示的对端边缘节点(如图8中节点3)时,对端边缘节点对IPv4报头,UDP报头和VXLAN-GPE报头进行解封装,得到原始IPv6报文。

[0132] S907,解封后的报文继续在IPv6网络中按照现有IPv6或SRv6转发流程进行转发。

[0133] S908,判断报文是否到达目标节点(图8中的节点F),若未到达,则继续执行步骤S907,若到达,则报文转发流程结束。

[0134] 以上是采用VXLAN-GPE穿越IPv4网络的实施例。

[0135] 在另一实施例中,也可以采用GENEVE作为封装方式穿越IPv4网络。采用GENEVE作为封装方式穿越IPv4网络与采用VXLAN-GPE穿越IPv4网络的区别在于,执行步骤S904中的封装时,Overlay报头为GENEVE,UDP报头的目的端口为6081,GENEVE报头中的protocol type字段设为0x86DD。图11为采用GENEVE作为封装方式进行封装后的报文格式。

[0136] 在另一实施例中,还可以采用VXLAN作为封装方式穿越IPv4网络。采用VXLAN作为封装方式穿越IPv4网络与采用VXLAN-GPE穿越IPv4网络的区别在于,执行步骤S904中的封装时,Overlay报头为VXLAN;UDP报头的目的端口为4789;

[0137] VXLAN中没有Next protocol字段,默认其后紧跟的报文头为以太报头,源MAC和目的MAC地址可分别设为设备2和3对应的接口MAC地址。图12为采用VXLAN作为封装方式进行封装后的报文格式。

[0138] 在另一实施例中,也可以采用NVGRE作为封装方式穿越IPv4网络。采用NVGRE作为封装方式穿越IPv4网络与采用VXLAN-GPE穿越IPv4网络的区别在于,执行步骤S904中的封装时,Overlay报头为NVGRE;无UDP报头。图13为采用NVGRE作为封装方式进行封装后的报文格式。

[0139] 本申请实施例通过扩展一个新类型的SID来承载穿越指示信息,使头节点不仅具备选择IPv6网络中路径的能力(由SRv6 SRH实现),还要具备选择IPv4网络隧道出口的能力,从而实现了端到端路径的控制。

[0140] 本申请实施例扩展该新类型的SID不仅能实现头节点对端到端路径的控制,还能够实现IPv6报文与IPv4报文之间的拷贝。下面以拷贝随流操作、管理和维护in-situ operation administration and maintenance(iOAM)报头(iOAM有时候也称为带内操作、管理和维护inband OAM,本申请实施例对此不做限制)和/或拷贝优先级为例进行说明。

[0141] 在IPv6报文和IPv4报文间拷贝iOAM报头的实施例中,IPv6报文中携带有iOAM报头。该实施例与图9所示实施例的区别在于,在步骤S902的报文转发过程中,还要进一步按照iOAM报头的指示随路收集信息。在步骤S904的报文封装过程中,边缘节点(节点2)还将IPv6扩展报头中的iOAM报头拷贝到IPv4扩展报头或者Overlay报头中。在步骤S905的转发过程中,进一步根据iOAM报头的指示随路收集信息。在步骤S906的解封过程中,还进一步将IPv4扩展报头或者Overlay报头中的iOAM报头拷贝到IPv6扩展报头中,并在步骤S907的转发过程中,继续根据iOAM报头的指示随路收集信息。

[0142] 上述实施例在IPv6/SRv6报头被封装进Overlay隧道前,可将部分信息拷贝到IPv4/Overlay报头中。比如,IPv6/SRv6中的iOAM信息可以拷贝到IPv4或者Overlay报头中;在IPv4域转发时,网络设备将要采集的信息添加到iOAM字段中;在Overlay隧道的出口处可将iOAM采集的数据拷贝回IPv6/SRv6报头中。如此可实现端到端路径上iOAM数据的采集。

[0143] 在拷贝优先权的实施例中,IPv6报文中携带有APP-ID。该APP-ID可通过映射到不同的SR policy等保证不同应用的服务等级协议Service Level Agreement(SLA)。在IPv6/SRv6报文在IPv4网络中转发时,为了在IPv4域中保证应用的SLA,可以将APP-ID映射为IPv4报文头中的DSCP字段,以指示在IPv4域中转发的优先级。具体的,在封装报文的过程中,

IPv4网络的入口节点可将IPv6 Hop-by-hop报头或者SRH报头中携带的APP-ID信息设置在IPv4报头的DSCP字段。报文在IPv4网络中转发时,根据DSCP匹配优先级。

[0144] 通过本申请实施例提供的拷贝功能,使得报文在IPv6网络中的特性在穿越IPv4网络时得以保留。

[0145] 需要说明的是,本申请实施例中的节点均为网络设备,其可以是网络中的交换机,路由器等。

[0146] 图14是本申请实施例提供的一种报文处理装置的结构示意图,该报文处理装置可以为上述任一实施例中第一网络中的节点或第二网络的入口节点。本申请实施例中的第一网络中的节点指的是第一网络中建立的SR隧道的头节点。该报文处理装置1400可以为交换机,路由器或者其他转发报文的网路设备。在该实施例中,该报文处理装置1400包括:主控板1410、接口板1430和接口板1440。多个接口板的情况下可以包括交换网板(图中未示出),该交换网板用于完成各接口板(接口板也称为线卡或业务板)之间的数据交换。

[0147] 主控板1410用于完成系统管理、设备维护、协议处理等功能。接口板1430和1440用于提供各种业务接口(例如,POS接口、GE接口、ATM接口等),并实现报文的转发。主控板1410上主要有3类功能单元:系统管理控制单元、系统时钟单元和系统维护单元。主控板1410、接口板1430以及接口板1440之间通过系统总线与系统背板相连实现互通。接口板1430上包括一个或多个处理器1431。处理器1431用于对接口板进行控制管理并与主控板上的中央处理器进行通信,以及用于报文的转发处理。接口板1430上的存储器1432用于存储转发表项,处理器1431通过查找存储器1432中存储的转发表项进行报文的转发。

[0148] 所述接口板1430包括一个或多个网络接口1433用于接收上一跳网络节点发送的探测报文,并根据处理器1431的指示向下一跳网络节点发送处理后的探测报文。具体实现过程请参考图6所示实施例中步骤S601以及S603,图7所示实施例中的步骤S701和S703,或图9所示实施例中的步骤S902、S905以及S907。

[0149] 所述处理器1431用于执行上述方法实施例中所描述的第一网络中的节点或第二网络中的入口节点的处理步骤和功能,具体可以参考上述图6所示实施例中的步骤S601以及S602,图7所示实施例中的步骤S701至步骤S702,或图9所示实施例中的步骤S901至S908,这里不再逐一赘述。

[0150] 需要说明的是,如图14所示,本实施例中包括多个接口板,采用分布式的转发机制,这种机制下,接口板1440上的操作与所述接口板1430的操作基本相似,为了简洁,不再赘述。

[0151] 此外,需要说明的是,图14中的接口板1430中的处理器1431和/或1441可以是专用硬件或芯片,如网络处理器或者专用集成电路(application specific integrated circuit)来实现上述功能,这种实现方式即为通常所说的转发面采用专用硬件或芯片处理的方式。采用网络处理器这一专用硬件或芯片的具体实现方式可以参考下面图15所示的实施例的描述。在另外的实施方式中,所述处理器1431和/或1441也可以采用通用的处理器,如通用的CPU来实现以上描述的功能,通用CPU的实现方式可以参考下面图18和19所示的实施例的描述。

[0152] 此外,需要说明的是,主控板可能有一块或多块,有多块的时候可以包括主用主控板和备用主控板。接口板可能有一块或多块,该设备的数据处理能力越强,提供的接口板越

多。多块接口板的情况下,该多块接口板之间可以通过一块或多块交换网板通信,有多块的时候可以共同实现负荷分担冗余备份。在集中式转发架构下,该设备可以不需要交换网板,接口板承担整个系统的业务数据的处理功能。在分布式转发架构下,该设备包括多块接口板,可以通过交换网板实现多块接口板之间的数据交换,提供大容量的数据交换和处理能力。所以,分布式架构的网络设备的数据接入和处理能力要大于集中式架构的设备。具体采用哪种架构,取决于具体的组网部署场景,此处不做任何限定。

[0153] 具体的实施例中,存储器1432可以是只读存储器(read-only memory,ROM)或可存储静态信息和指令的其它类型的静态存储设备,随机存取存储器(random access memory, RAM)或者可存储信息和指令的其它类型的动态存储设备,也可以是电可擦可编程只读存储器(electrically erasable programmable read-only memory,EEPROM)、只读光盘(compact disc read-only Memory,CD-ROM)或其它光盘存储、光碟存储(包括压缩光碟、激光光碟、光碟、数字通用光碟、蓝光光碟等)、磁盘或者其它磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其它介质,但不限于此。存储器1432可以是独立存在,通过通信总线与处理器1431相连接。存储器1432也可以和处理器1431集成在一起。

[0154] 其中,存储器1432用于存储程序代码,并由处理器1431来控制执行,以执行上述实施例所提供的报文处理方法。处理器1431用于执行存储器1432中存储的程序代码。程序代码中可以包括一个或多个软件模块。这一个或多个软件模块可以为下面图16或图17任一实施例中提供的软件模块。

[0155] 具体实施例中,所述网络接口1433,可以是使用任何收发器一类的装置,用于与其它设备或通信网络通信,如以太网,无线接入网(radio access network,RAN),无线局域网(wireless local area networks,WLAN)等。

[0156] 图15是本申请实施例提供的另一种报文处理装置的结构示意图,该报文处理装置可以为上述任一实施例提供的第一网络中的节点或者所述第二网络中的入口节点。本申请实施例中的第一网络中的节点指的是第一网络中建立的SR隧道的头节点。该报文处理装置1500可以为交换机,路由器或者其他转发报文的网路设备。在该实施例中,该报文处理装置1500包括:主控板1510、接口板1530、交换网板1520和接口板1540。主控板1510用于完成系统管理、设备维护、协议处理等功能。交换网板1520用于完成各接口板(接口板也称为线卡或业务板)之间的数据交换。接口板1530和1540用于提供各种业务接口(例如,POS接口、GE接口、ATM接口等),并实现数据包的转发。控制平面由主控板1510的各管控单元及接口板1530和1540上的管控单元等构成。主控板1510上主要有3类功能单元:系统管理控制单元、系统时钟单元和系统维护单元。主控板1510、接口板1530和1540,以及交换网板1520之间通过系统总线与系统背板相连实现互通。接口板1530上的中央处理器1531用于对接口板进行控制管理并与主控板上的中央处理器进行通信。接口板1530上的转发表项存储器1534用于存储转发表项,网络处理器932通过查找转发表项存储器1534中存储的转发表项进行报文的转发。

[0157] 所述接口板1530包括一个或多个网络接口1533用于接收上一跳网络节点发送的探测报文,并根据处理器1531的指示向下一跳网络节点发送处理后的探测报文。具体实现过程请参考图6所示实施例中步骤S601以及S603,图7所示实施例中的步骤S701和S703,或

图9所示实施例中的步骤S902、S905以及S907,这里不再逐一赘述。

[0158] 所述处理器1532用于执行上述方法实施例中所描述的第一网络中的节点或第二网络中的入口节点的处理步骤和功能,具体可以参考上述图6所示实施例中的步骤S601以及S602,以及图7所示实施例中的步骤S701至步骤S702,或图9所示实施例中的步骤S901至S908,这里不再逐一赘述。

[0159] 需要说明的是,如图15所示,本实施例中包括多个接口板,采用分布式的转发机制,这种机制下,接口板1540上的操作与所述接口板1530的操作基本相似,为了简洁,不再赘述。此外,如上所述,图15中的网络处理器1532以及1542的功能可以用专用集成电路(application specific integrated circuit)替换来实现。

[0160] 此外,需要说明的是,主控板可能有一块或多块,有多块的时候可以包括主用主控板和备用主控板。接口板可能有一块或多块,该设备的数据处理能力越强,提供的接口板越多。接口板上的物理接口卡也可以有一块或多块。交换网板可能没有,也可能有一块或多块,有多块的时候可以共同实现负荷分担冗余备份。在集中式转发架构下,该设备可以不需要交换网板,接口板承担整个系统的业务数据的处理功能。在分布式转发架构下,该设备可以有至少一块交换网板,通过交换网板实现多块接口板之间的数据交换,提供大容量的数据交换和处理能力。所以,分布式架构的网络设备的数据接入和处理能力要大于集中式架构的设备。具体采用哪种架构,取决于具体的组网部署场景,此处不做任何限定。

[0161] 参见图16,本申请实施例还提供了一种报文处理装置1600,该装置可以应用于网络中的节点,该节点为该网络中建立的SR隧道的头节点,例如该节点可以实现图6或9所示SR隧道的头节点的功能。在一个具体实施例中,报文处理装置包括:确定单元1602、报文处理单元1604,以及报文发送单元1606。

[0162] 其中,所述确定单元1602用于执行图6所示实施例中步骤S601,为报文确定端到端的传输路径,包括在第一网络中的传输路径以及穿越第二网络的路径。

[0163] 所述报文处理单元1604,用于执行图6所示实施例中步骤S602,根据确定的传输路径,在报文中加入端到端的路径信息;所述路径信息中包括穿越指示信息,穿越指示信息指示了所述报文在所述第二网络中的出口节点。

[0164] 所述报文发送单元1606具体用于发送报文处理单元1604处理后的报文,具体可以将添加有所述穿越指示信息的报文经所述第一网络发往所述第二网络的入口节点,指示所述第二网络的入口节点按照所述穿越指示信息中指定的所述第二网络的出口节点封装并传送所述报文。具体实现过程,可参考图6所示实施例中步骤S601以及图9所示实施例中步骤S901,这里不再赘述。

[0165] 在一个具体的实施方式中,所述报文处理装置1600可以进一步包括:接收单元1608,用于接收所述第二网络的入口节点发布的至少一个SID;所述入口节点发布的不同SID携带不同的穿越指示信息;

[0166] 所述报文处理单元1604在报文中加入端到端的路径信息时,将穿越指示信息与所述确定单元确定的穿越第二网络的路径相符的SID作为所述第二网络的入口节点的SID加入到所述端到端的路径信息中。

[0167] 上述接收单元1608以及所述报文处理单元1604的具体实现过程,可参考图5所示实施例中关于R3节点的处理过程的描述,以及图9所示实施例中步骤S901的描述,这里不再

赘述。

[0168] 参见图17,本申请实施例还提供了一种报文处理装置1700,该装置可以应用于网络的边缘节点,该节点可以实现图7或9所示入口节点的功能。

[0169] 在一个具体实施例中,报文处理装置包括:报文获取单元1702、报文处理单元1704,以及报文发送单元1706。

[0170] 其中,所述报文获取单元1502,用于接收来自第一网络的报文,所述报文中携带穿越指示信息,穿越指示信息中指示了所述报文在所述第二网络中的出口节点。具体实现过程可以参考上述图7以及图9所示实施例中的具体描述,此处不再赘述。

[0171] 所述报文发送单元1704,用于根据所述穿越指示信息将所述报文封装为所述第二网络支持的格式;所述封装后的报文的地址为所述穿越指示信息中指示的出口节点。具体实现过程可以参考上述图7所示实施例中的步骤702。

[0172] 所述报文发送单元1706,用于将所述封装后的报文路由到所述出口节点。

[0173] 在一个具体的实施方式中,所述报文处理装置1700可以进一步包括:发布单元1708,用于发布至少一个SID,所述SID携带穿越指示信息,在发布多个SID的情况下,不同SID携带不同的穿越指示信息。具体实现过程,请参考图5所述实施例关于R3节点的处理过程的描述,以及图9所示实施例中步骤S901的描述。

[0174] 参见图18,本申请实施例提供了一种报文处理装置1800,该报文处理装置1800可以实现上述图6和图9所示实施例中SR隧道的头节点的功能。该报文处理装置包括至少一个处理器1802和至少一个通信接口1804。

[0175] 此外还可以进一步包括存储器1806,用于存储指令。所述存储器1806既可以是处理器内的缓存存储器,也可以是处理器之外的存储器。在图16实施例中所描述的各单元为通过软件实现的情况下,执行图16中的报文处理单元1604和报文发送单元1606功能所需的软件或程序代码存储在存储器1806中。

[0176] 处理器1802,用于执行存储器1806中的指令,执行上述应用于图6或图9所示实施例中的SR隧道的头节点的处理步骤。

[0177] 通信接口1804,用于进行通信。

[0178] 存储器1806(处理器外置情况下)、处理器1802和通信接口1804通过总线1808相互连接;总线1808可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示,图18中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0179] 在具体实施例中,处理器1802用于为报文确定端到端的传输路径,并根据确定的传输路径,在报文中加入端到端的路径信息;将添加有穿越指示信息的报文通过通信接口1804发往第二网络的入口节点。具体的过程请参考上述图6以及图9中的具体描述,此处不再赘述。

[0180] 参见图19,本申请实施例提供了一种报文处理装置1900,该报文处理装置1900可以实现上述图7以及图9所示实施例中入口节点的功能。该报文处理装置包括至少一个处理器1902和至少一个通信接口1904。

[0181] 此外还可以进一步包括存储器1906,用于存储指令。所述存储器1906既可以是处

理器内的缓存存储器,也可以是处理器之外的存储器。在图17实施例中所描述的各单元为通过软件实现的情况下,执行图17中的报文处理单元1704和报文发送单元1706功能所需的软件或程序代码存储在存储器1906中。

[0182] 处理器1902,用于执行存储器1906中的指令,执行上述应用于图7以及图9所示实施例中入口节点的处理步骤。

[0183] 通信接口1904,用于进行通信。

[0184] 存储器1906(处理器外置情况下)、处理器1902和通信接口1904通过总线1908相互连接;总线1908可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示,图19中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0185] 在具体实施例中,处理器1902用于接收来自第一网络的报文,根据报文中的穿越指示信息将所述报文封装为第二网络支持的格式,并将封装后的报文的地址设置为所述穿越指示信息中指示的出口节点,然后将所述封装后的报文路由到所述出口节点。具体的过程请参考上述图7以及图9所示实施例中的具体描述,此处不再赘述。

[0186] 通信接口1904用于收发报文。具体的过程请参考具体的过程请参考上述图7以及图9所示实施例中的具体描述,此处不再赘述。

[0187] 上述存储器1806和存储器1906可以是随机存取存储器(random-access memory, RAM)、闪存(flash)、只读存储器(read only memory, ROM)、可擦写可编程只读存储器(erasable programmable read only memory, EPROM)、电可擦除可编程只读存储器(electrically erasable programmable read only memory, EEPROM)、寄存器(register)、硬盘、移动硬盘、CD-ROM或者本领域技术人员知晓的任何其他形式的存储介质,也可以是处理器中的缓存。

[0188] 上述处理器1802和处理器1902例如可以是中央处理器(central processing unit, CPU)、Network-Processor、通用处理器、数字信号处理器(digital signal processor, DSP)、专用集成电路(application-specific integrated circuit, ASIC)、现场可编程门阵列(field programmable gate array, FPGA)或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件、处理电路或者其任意组合。其可以实现或执行结合本申请公开内容所描述的各种示例性的逻辑方框,模块和电路。处理器也可以是实现计算功能的组合,例如包含一个或多个微处理器组合,DSP和微处理器的组合等等。

[0189] 上述通信接口1804和通信接口1904例如可以是接口卡等,可以为以太(ethernet)接口或异步传输模式(asynchronous transfer mode, ATM)接口,以及接口电路等。

[0190] 本申请实施例还提供了一种芯片,芯片设置在如图16或18所示的报文处理装置中,该报文处理装置为方法实施例中的SR隧道的头节点。该芯片包括处理器和接口电路。

[0191] 接口电路用于接收指令并传输至处理器。

[0192] 处理器用于接受接口电路的指令,执行上述应用于图6以及图9所示实施例中SR隧道的头节点的处理步骤。

[0193] 在具体实施例中,处理器用于为报文确定端到端的传输路径,并根据确定的传输路径,在报文中加入端到端的路径信息;将添加有穿越指示信息的报文通过通信接口1804

发往第二网络的入口节点。具体的过程请参考上述图6以及图9中的具体描述,此处不再赘述。

[0194] 本申请实施例还提供了一种芯片,芯片设置在如图17或19所示的报文处理装置中,该报文处理装置为方法实施例中的入口节点。该芯片包括处理器和接口电路。

[0195] 接口电路用于接收指令并传输至处理器。

[0196] 处理器用于接受接口电路的指令,执行上述应用于图7以及图9所示实施例中入口节点的处理步骤。

[0197] 在具体实施例中,处理器用于接收来自第一网络的报文,根据报文中的穿越指示信息将所述报文封装为第二网络支持的格式,并将封装后的报文的地址设置为所述穿越指示信息中指示的出口节点,然后将所述封装后的报文路由到所述出口节点。具体的过程请参考上述图7以及图9所示实施例中的具体描述,此处不再赘述。

[0198] 本申请实施例还提供了一种网络系统,网络系统包括第一网络中的设备和第二网络中的设备,第一网络中的设备包括节点,第二网络中的设备包括与第一网络连接的边缘节点,边缘节点同时位于第一网络中,支持第一网络和第二网络的传输协议,第一网络中的节点发出的报文经第二网络的边缘节点,穿越过第二网络到达目的节点。本申请实施例中的第一网络中的节点指的是第一网络中建立的SR隧道的头节点。

[0199] 所述网络系统中的节点可以执行图6以及图9所示实施例中的SR隧道的头节点的步骤,此处不再赘述,或者相应的,所述网络系统中的节点为图16或18所示实施例的报文处理装置1600或1800。

[0200] 所述网络系统中的边缘节点在可以执行图7以及图9所示实施例中的入口节点的步骤,此处不再赘述,或者相应的,所述网络系统中的边缘节点为图17或19所示实施例的报文处理装置1700或1900。

[0201] 本申请实施例还提供了一种计算机可读存储介质,包括指令,当其在计算机上运行时,使得计算机执行上述应用于报文处理装置1600的报文处理方法。

[0202] 本申请实施例还提供了一种计算机可读存储介质,包括指令,当其在计算机上运行时,使得计算机执行上述应用于入口节点1800的报文处理方法。

[0203] 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。

[0204] 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,单元的划分,仅仅为一种逻辑业务划分,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。

[0205] 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。

[0206] 另外,在本申请各个实施例中的各业务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,也可以采用软件业务单元的形式实现。

[0207] 集成的单元如果以软件业务单元的形式实现并作为独立的产品销售或使用时,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,本申请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)执行本申请各个实施例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。

[0208] 本领域技术人员应该可以意识到,在上述一个或多个示例中,本发明所描述的业务可以用硬件、软件、固件或它们的任意组合来实现。当使用软件实现时,可以将这些业务存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。计算机可读介质包括计算机存储介质和通信介质,其中通信介质包括便于从一个地方向另一个地方传送计算机程序的任何介质。存储介质可以是通用或专用计算机能够存取的任何可用介质。

[0209] 以上的具体实施方式,对本发明的目的、技术方案和有益效果进行了进一步详细说明,所应理解的是,以上仅为本发明的具体实施方式而已。

[0210] 以上,以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的范围。

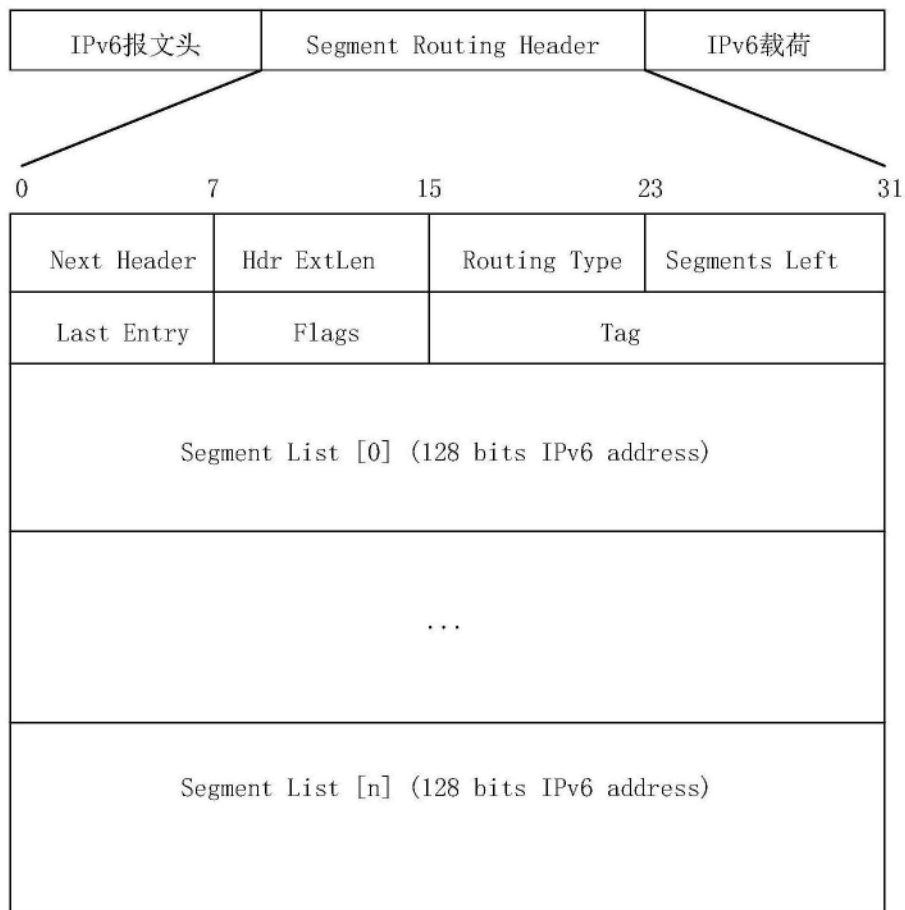


图1

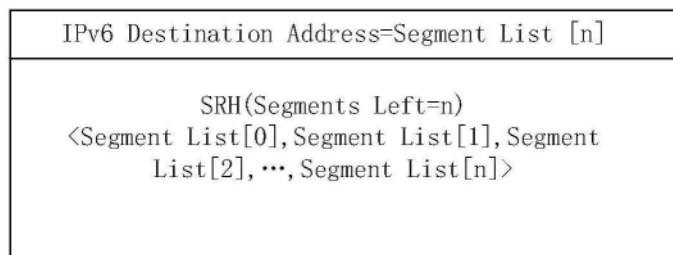


图2

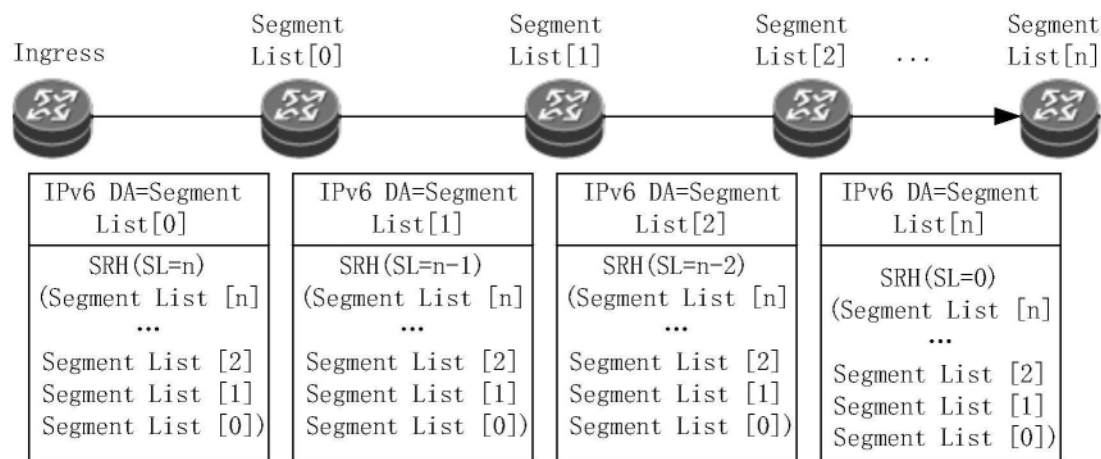


图3

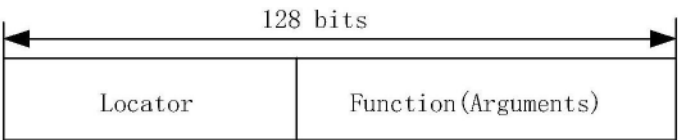


图4

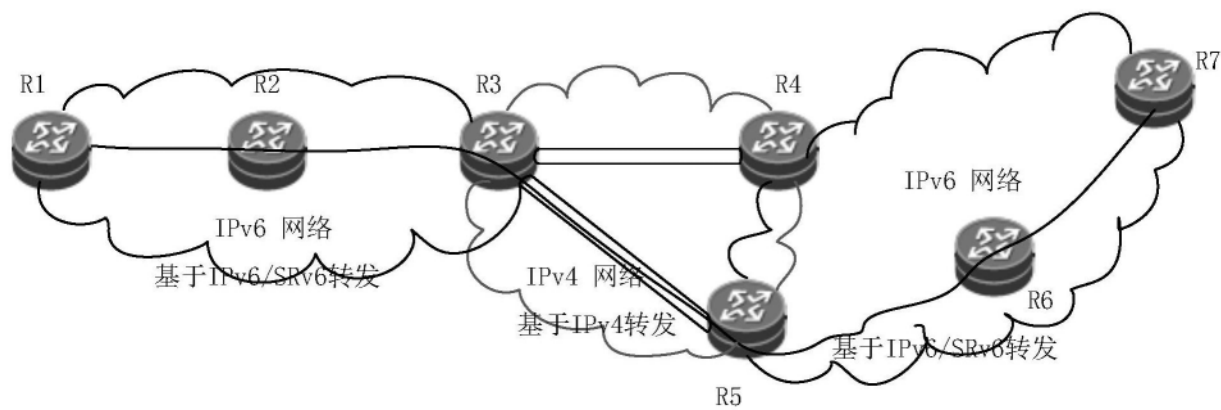


图5

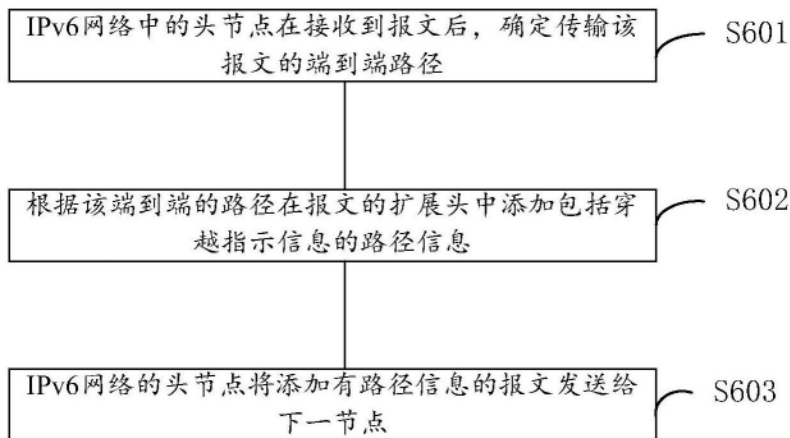


图6

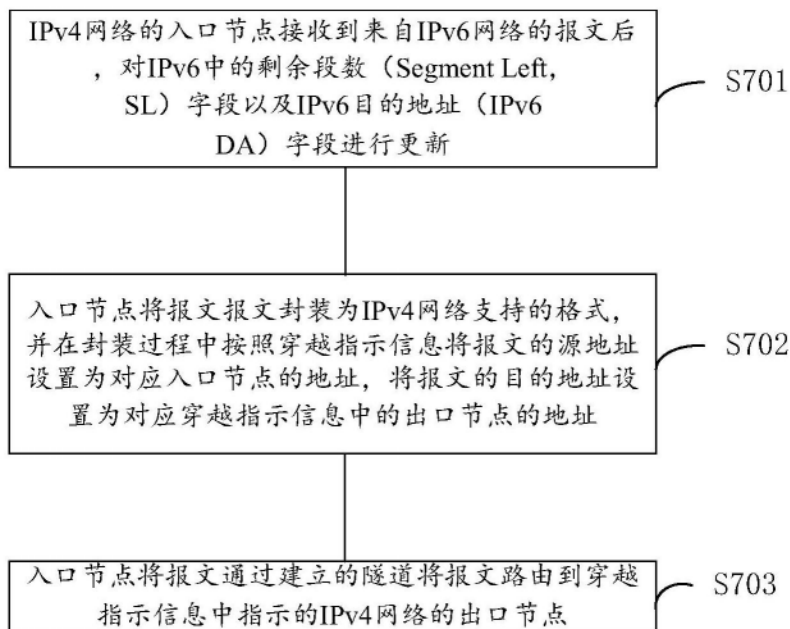


图7

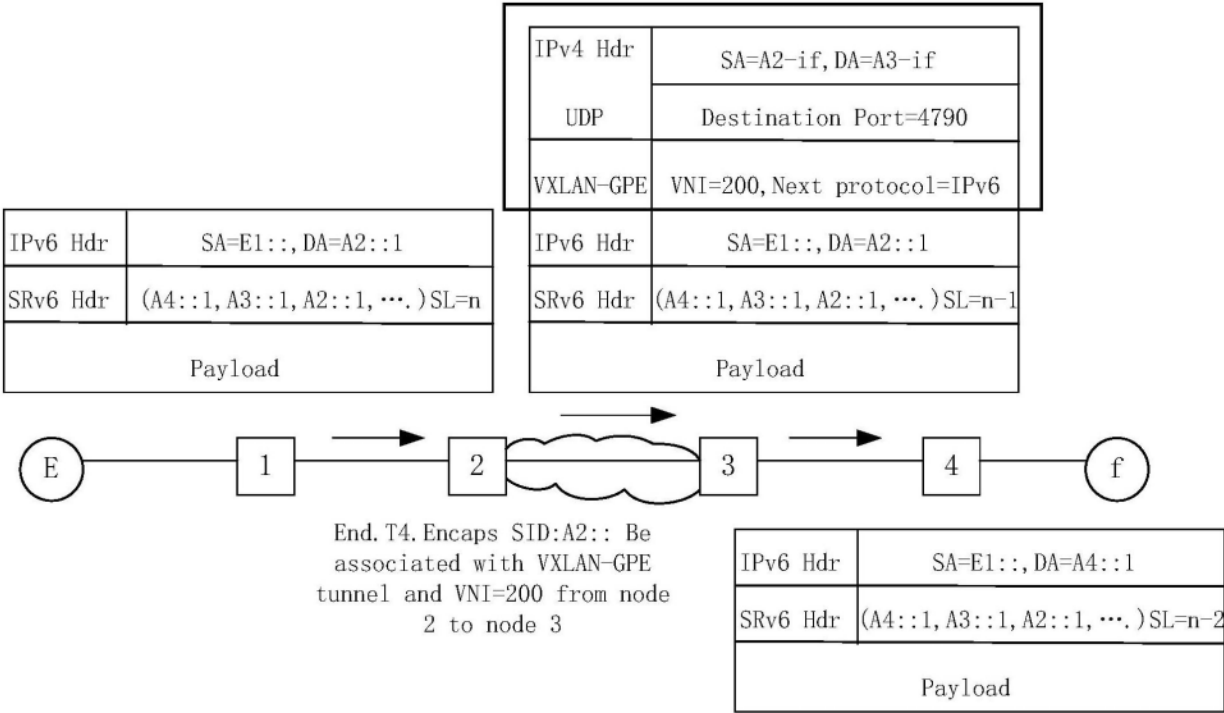


图8

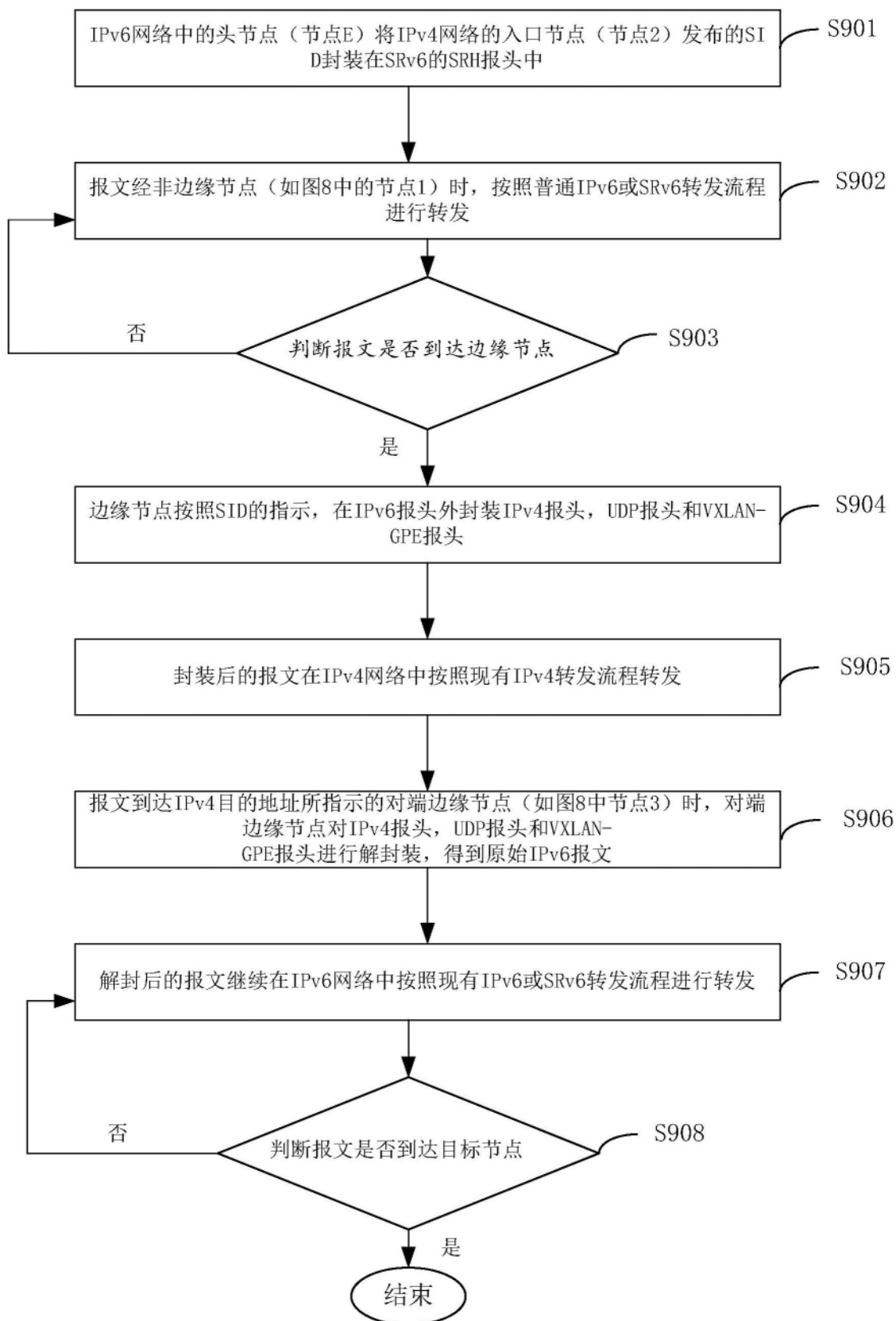


图9

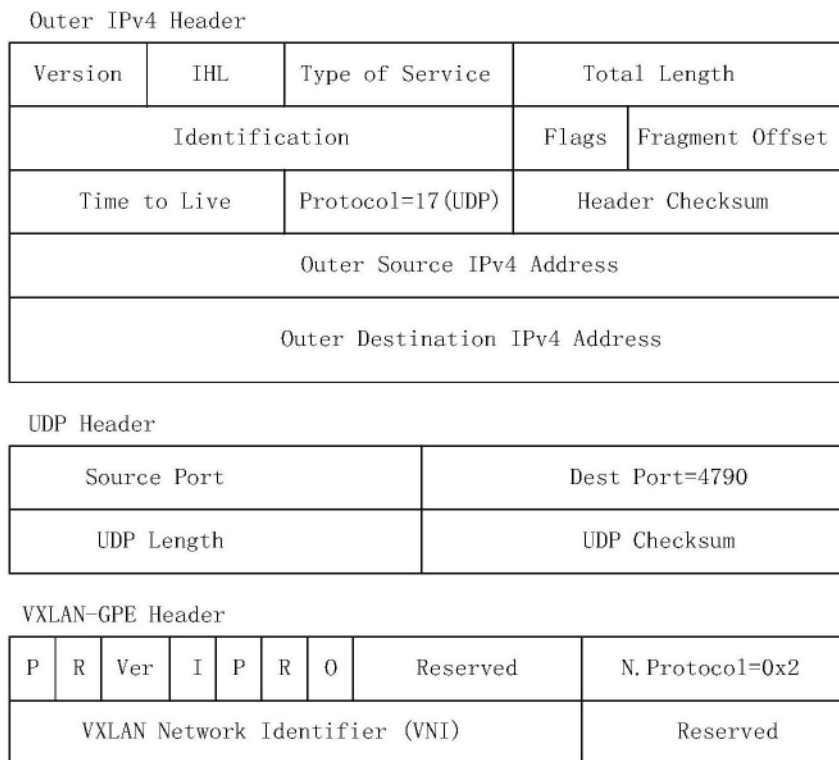


图10

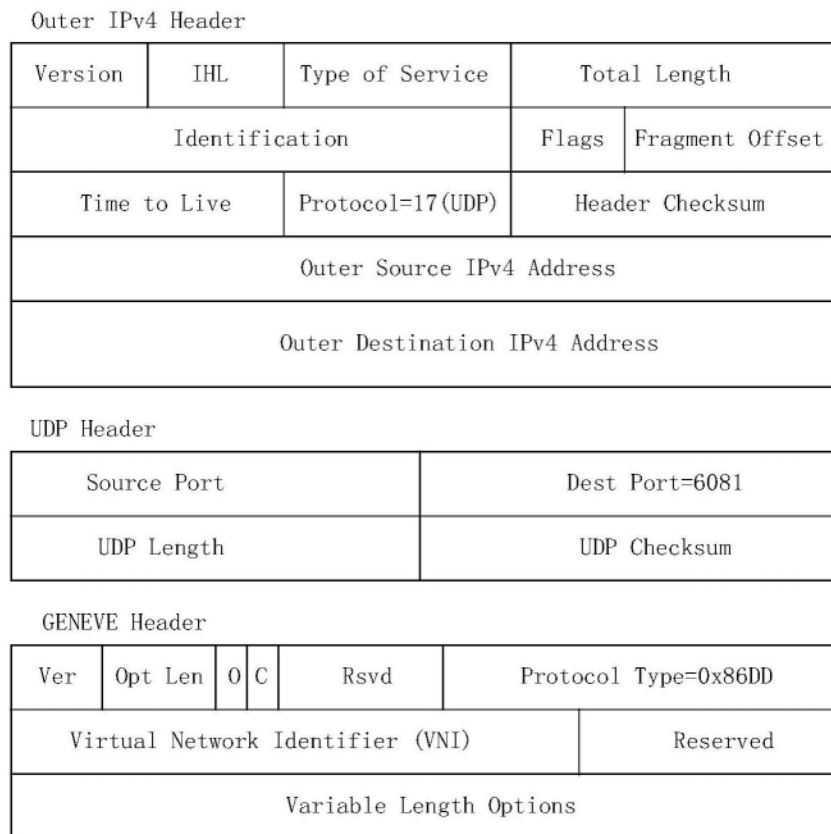


图11

Outer IPv4 Header

Version	IHL	Type of Service	Total Length	
Identification			Flags	Fragment Offset
Time to Live		Protocol=17 (UDP)	Header Checksum	
Outer Source IPv4 Address				
Outer Destination IPv4 Address				

UDP Header

Source Port	Dest Port=4789
UDP Length	UDP Checksum

VXLAN Header

R	R	R	R	I	R	R	R	Reserved	
VXLAN Network Identifier (VNI)								Reserved	

Inner Ethernet Header

Inner Destination MAC Address	
Inner Destination MAC Address	Inner Source MAC Address
Inner Source MAC Address	
OptnlEthertype=C-Tag 802.1Q	Inner VLAN Tag Information
Ethertype=0x86DD	

图12

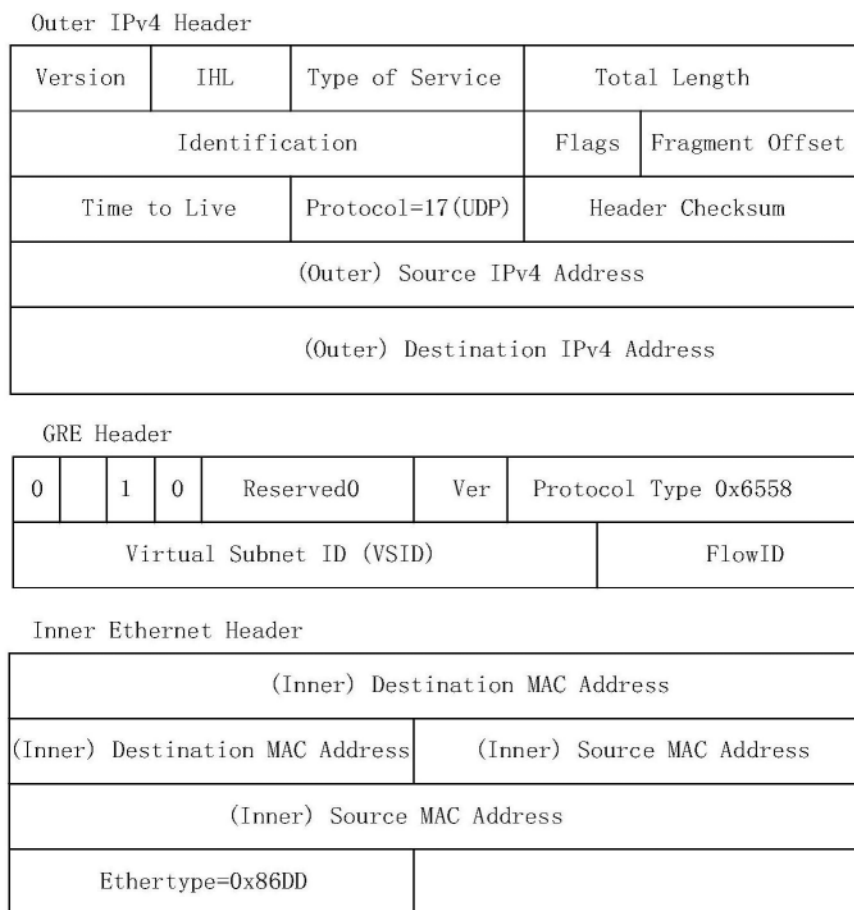


图13

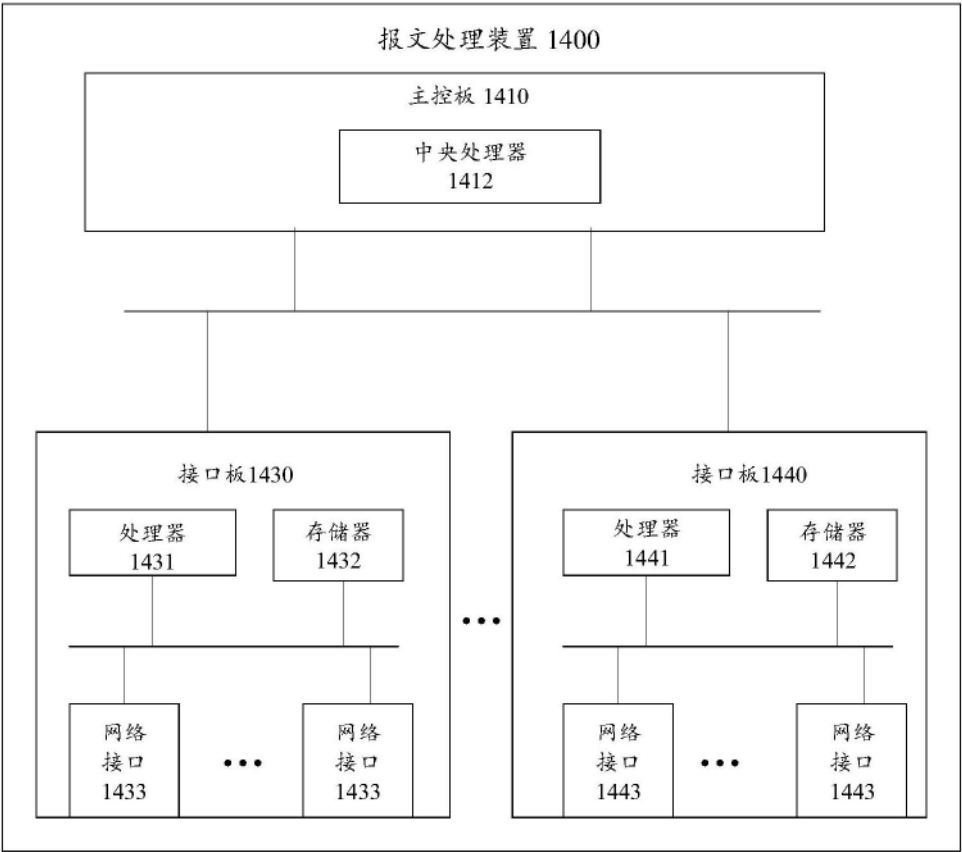


图14

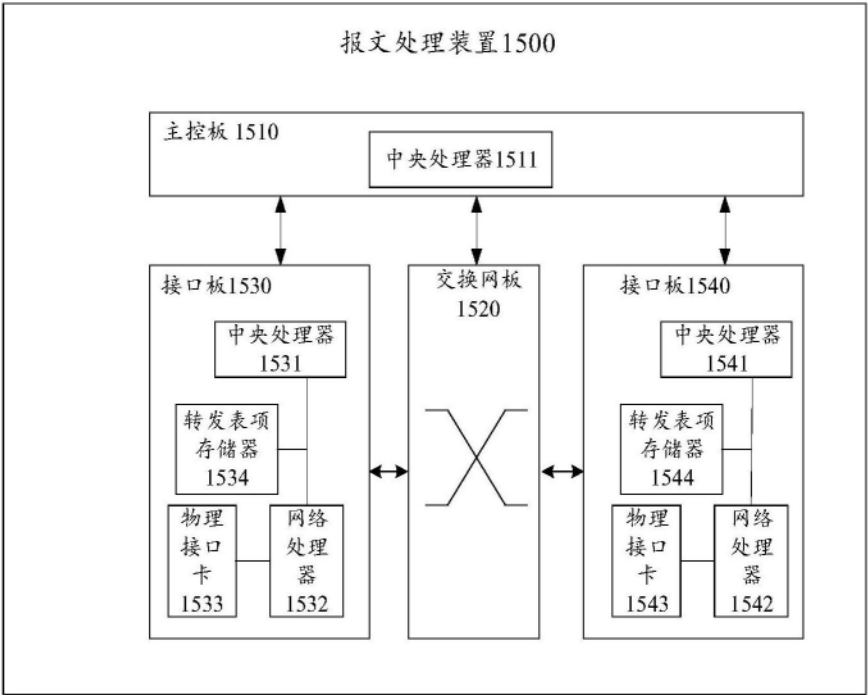


图15

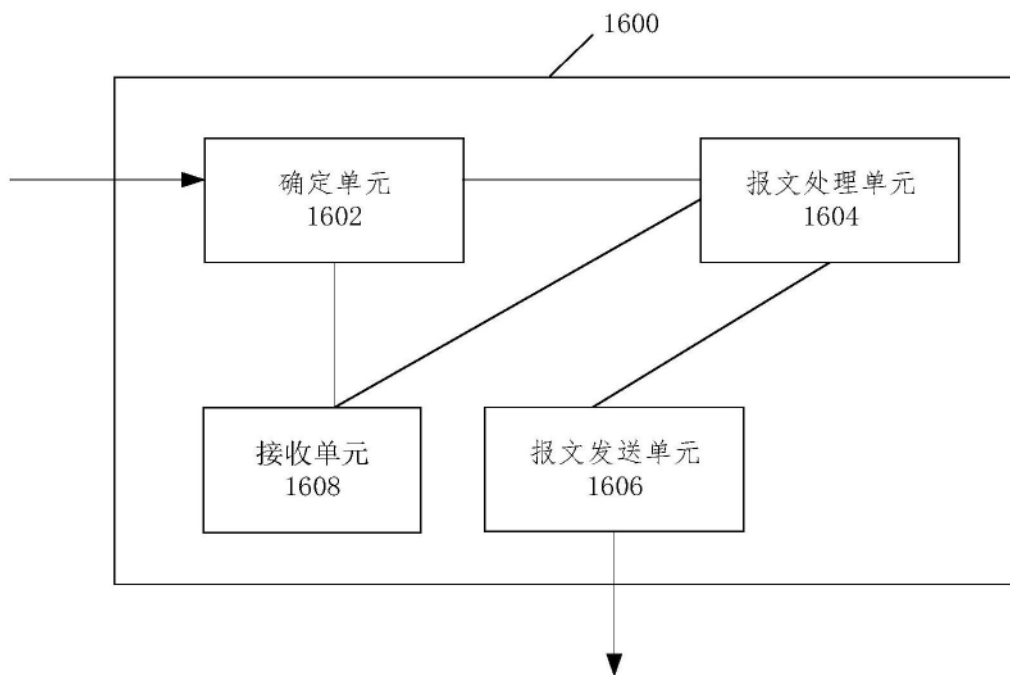


图16

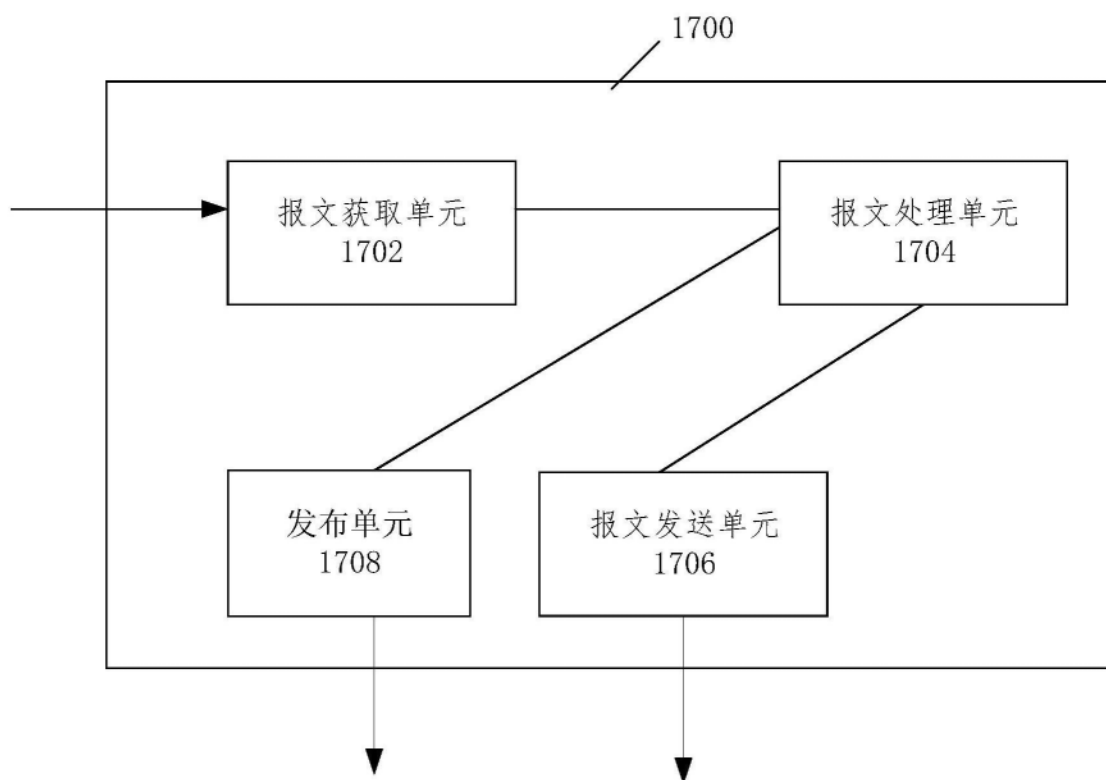


图17

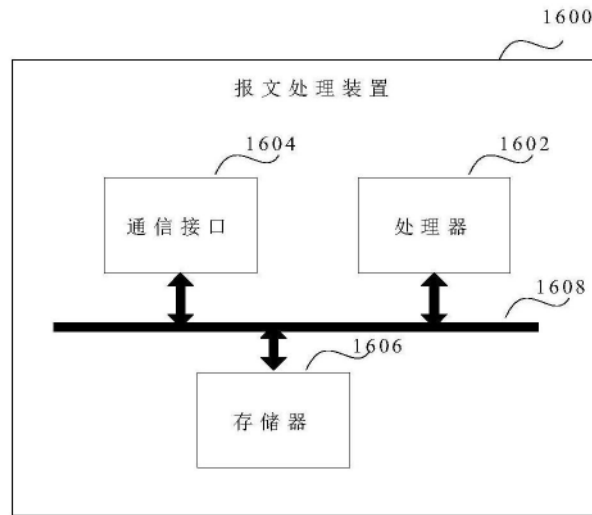


图18

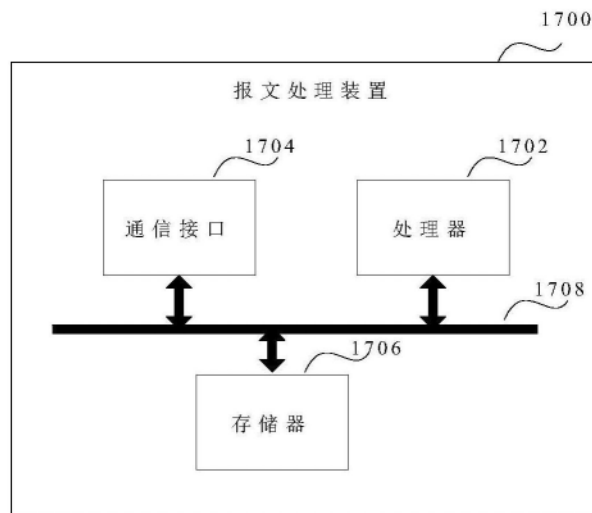


图19