

文章编号: 2096-1618(2018)05-0503-06

PTN 中静态 L3VPN 的一种 IP 路由动态学习方法

孙捷, 姚尧, 任德昊

(成都信息工程大学通信工程学院, 四川 成都 610225)

摘要:PTN L3VPN 网络最重要的应用领域之一是 LTE 移动回传网络。静态 L3VPN 中对核心 PTN 节点都需要手工配置成千上万个 IP 路由表项,给网络配置和维护带来了复杂、大工作量及易出错问题。为解决该问题,提出一种创新方法,即 PTN 核心节点通过 DHCP 中继功能和 DHCP 消息侦听,从而可以在核心 PTN 节点上自动建立 IP 路由表。

关键词:通信与信息系统; 数字通信技术; 静态 L3VPN; DHCP 中继; DHCP 侦听; 分组传送网; LTE 回传网
中图分类号:TN915 **文献标志码:**A
doi:10.16836/j.cnki.jcuit.2018.05.004

0 引言

LTE 移动回传网络是基站 eNB 和网关 aGW/MME 之间的传输网络^[1-2]。LTE 移动回传主要有 PTN (packet transport network 包传输网)和 IP-RAN(基于 IP 的无线接入网)两种方式,其中 PTN 做 LTE 回传是被 CMCC 等国内外主流运营商广泛应用的方式^[3-4]。LTE 对 PTN 移动回传网络提出了新的承载需求:由于 X1 和 S2 接口业务都是靠 IP 地址来指示和区分,因此在 eNB 和 aGW/MME 之间及 eNB 之间需要采用基于 IP 地址的包转发策略。而现有的 PTN 设备都是基于 2.5 层的 MPLS TU(隧道)和 PW(伪线)标签进行包转发^[5],且由于 eNB 和 aGW/MME 的多归属关系,在 PTN 中给每个 eNB 建立多条单独 TU/PW 路径归属到不同的 aGW/MME 会导致 TU/PW 数量的急剧增加,因此在 PTN 网络中引入 IP 路由转发功能,即 L3VPN 功能,可以大幅度减少连接数量,避免移动回传组网的高复杂性和高成本。

L3VPN 功能一般考虑只部署在回传网络核心层,即核心 PTN 节点,而边缘接入 PTN 节点不部署 L3VPN,这样可以把 L3 路由域的规模控制得比较小,便于网络运维管理,降低移动回传网络建设和运维的综合成本,提高网络的可扩展性和安全性。由于考虑到 X2 带宽仅占 S1 带宽的 3%~5%^[6],因此 X2 可和 S1 共享 PTN 的 TU/PW 管道,利用核心 PTN 节点的 L3 IP 转发能力来支持。

1 基于 L3VPN 的 LTE 回传网络架构及 IP 转发路由表配置

L3VPN(三层虚拟专网)是一种基于 IP 的 VPN 技术,为用户提供三层可达性的虚拟专用网络,并实现不同用户业务之间的路由隔离。由于现有的 PTN 设备只支持 L2VPN^[7],并不支持 L3VPN,因此核心 PTN 节点需要改进以增加对 L3VPN 的支持^[8]。考虑到核心节点连接相对固定,为简化设计,核心 PTN 节点推荐采用静态 L3VPN 方式,用户人工配置所有 IP 路由,这样 PTN 节点无需支持 BGP 等复杂的路由协议。核心 PTN 网络节点实现 L2 到 L3 VRF 的桥接功能,即终结 L2 业务,识别 IP 报文,进行 L3 VRF(virtual route forwarding 虚拟路由转发)路由转发^[9]。

基于 L3VPN 的 LTE 回传网络架构如图 1 所示,分 PTN 接入层(L2VPN)和 PTN 核心层(静态 L3 VPN)两层。PTN 接入层连接 eNodeB,接入 PTN 节点(AN)是普通 MPLS 设备^[10],不需要支持 L3 VPN 功能。PTN 核心层连接 aGW/MME,核心 PTN 节点(HN)要支持静态 L3VPN 功能^[11]。图 1 还列出了从 eNodeB 到 aGW 路径在各段的包格式。具体的包转发过程如下:

(1)在接入 PTN 节点创建 EVPL 专线业务连接支持静态 L3VPN 功能的核心 PTN 节点,采用 PTN 标准的 TU/PW 封装:NNI 以太包头+TU/PW 标签+UNI 以太包头+IP 报头+LTE 报文。

(2)核心 PTN 节点终结 EVPL,删除 NNI 以太包头、TU/PW 标签、内部以太包头,得到 IP 包。查找相应的三层 IP 路由表,根据目的 IP 地址转发包。如果目的地址是本地 eNodeB,会再加上内部以太包头和

TU/PW 标签, NNI 以太包头, 形成标准的 PTN 包格式, 转发给相应接入 PTN 节点^[12]。

(3) 如果目的地址是本地 aGW/MME, IP 包直接转发到本地物理接口, 封装格式: UNI 以太包头+IP 报

头+LTE 报文。

(4) 如果目的地址是远端 eNodeB 或远端 aGW/MME, IP 包会加上 VRF 标签, TU/PW 标签, NNI 以太包头转发到远端核心 PTN 节点。

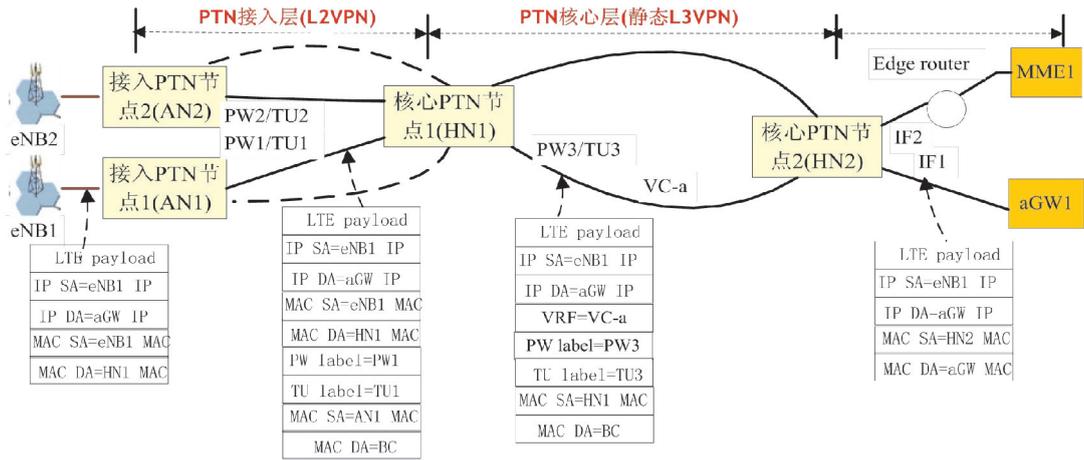


图1 基于 L3VPN 的 LTE 回传网络架构

2 静态 L3VPN 中动态学习 IP 路由表的方法

2.1 静态 L3VPN 中人工配置 IP 路由表的缺陷

虽然静态 L3VPN 避免了复杂的 IP 路由协议, 但有一个明显的缺陷。静态 L3VPN 中 IP 路由表是由人工配置到一个个核心 PTN 节点^[13]。一个 LTE 回传网络中每个 eNodeB 和 MME/aGW 的 IP 地址在 IP 路由表中都有一条记录。由于每个 LTE 回传网络可能有几十个核心 PTN 节点, 而每个核心 PTN 节点可连接几百个 eNodeB, 即需要对核心 PTN 节点用手工配置成千上万的 IP 路由表项, 增加了网络配置和维护的工作量, 使 PTN 网络配置和维护变得繁琐和容易出错。尤其在添加和移除 eNodeB 时, 每添加和移除一个 eNodeB, 每个核心 PTN 的路由表都要手工进行更新修改。

2.2 通过 DHCP 中继和侦听 DHCP 消息动态学习 IP 路由表

eNodeB 初始化时, 作为 DHCP server 的 aGW 通过 DHCP 协议给作为 DHCP client 的 eNodeB 动态分配 IP 地址。由于 DHCP server 和 DHCP client 不在同一个 IP 子网中, 因此 PTN 核心节点必须支持 DHCP 中继功能^[14], 确保 DHCP 包能穿越 PTN 网络到达 DHCP server。

提出一种创新方法, PTN 核心节点通过 DHCP 中继和侦听 DHCP 消息来获取 eNodeB 的 IP 地址和输出端口之间的映射关系, 从而可以在核心 PTN 节点上自动建立起 IP 转发表。

DHCP (dynamic host configuration protocol 动态主机配置协议) 使用 UDP 协议工作^[15]。DHCP server 使用 UDP port68 接收 client 的请求消息, UDP port67 向 client 发响应消息。DHCP 消息格式如图 2 所示。

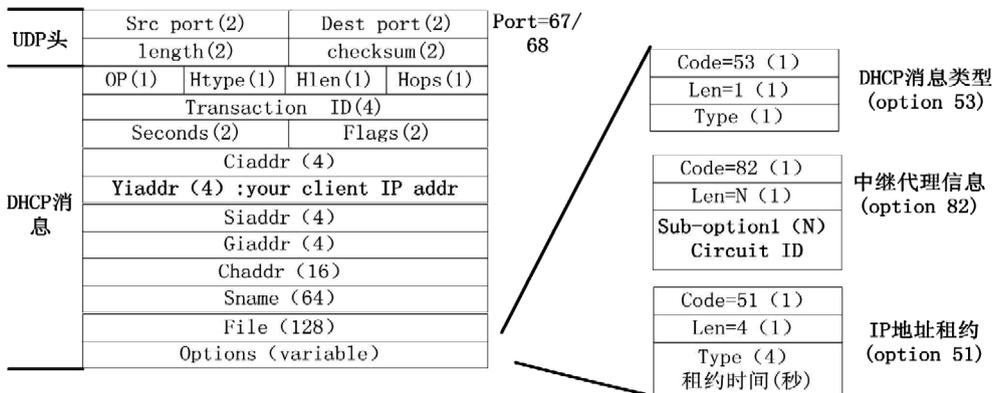


图2 DHCP 消息格式

OP:消息类型,为 1 时是 Client 发的请求,为 2 时是 Server 发的应答。Yiaddr (Your(client)IP address):即 server 分配给 client 的 IP 地址。DHCP 请求和应答报文中可带多个 option(选项)域。其中 option53(option 中 code=53)是所有 DHCP 消息都具备的,指示 DHCP 报文类型,比如 option53 中 type=1 表示 discovery 报文, type=2 表示 offer 报文等。获得 eNodeB 的 IP 地址和输出端口间映射关系用到的其它 DHCP 选项有 option 82 和 option51。

PTN 核心节点上各种 DHCP 消息处理及自动得到 IP 路由表的详细过程如下:

2.2.1 client 申请 IP 地址的 DHCP 中继过程及 IP 路由表建立

DHCP client 申请 IP 地址的 DHCP 中继消息流程图如图 3 所示,消息过程分为 Discover、Offer、Request、Ack 4 个阶段:

(1)DHCP client(eNodeB)没有 IP 地址,会发出一个 DHCP Discover 报文(option53 中 type=1),该报文

是广播报文,所有 DHCP Server 都会收到该报文。

(2)所有收到 DHCP Discover 报文的 server(aGW)都会从空闲 IP 池中选择 IP,并发送 DHCP Offer 报文(option53 中 type=2),向 client 提供该 IP 地址供其选择。DHCP Offer 报文也是广播报文,其中 Yiaddr 即服务器给客户端提供的 IP 地址。offer 报文中除 IP 地址外,还有掩码、网关和租约期限等信息。

(3)client 收到 server 提供的 IP 地址后从中选择一个,然后广播发送 DHCP Request 报文,告诉 server 选择了 IP 地址,并拒绝其他 server 的 IP 地址。没有被选中的 server 回收自己提供的 IP 地址。

(4)被选中的 server 收到 DHCP Request 后广播回送 DHCP ACK 报文,确认 client 申请 IP 地址成功。ACK 报文同样包含了分配的 IP 地址、掩码和租约期限等信息。option51(IP Address Lease Time)即 IP 租约期限,若 IP 租约超期未续租,server 会回收该 IP 放入空闲 IP 池。

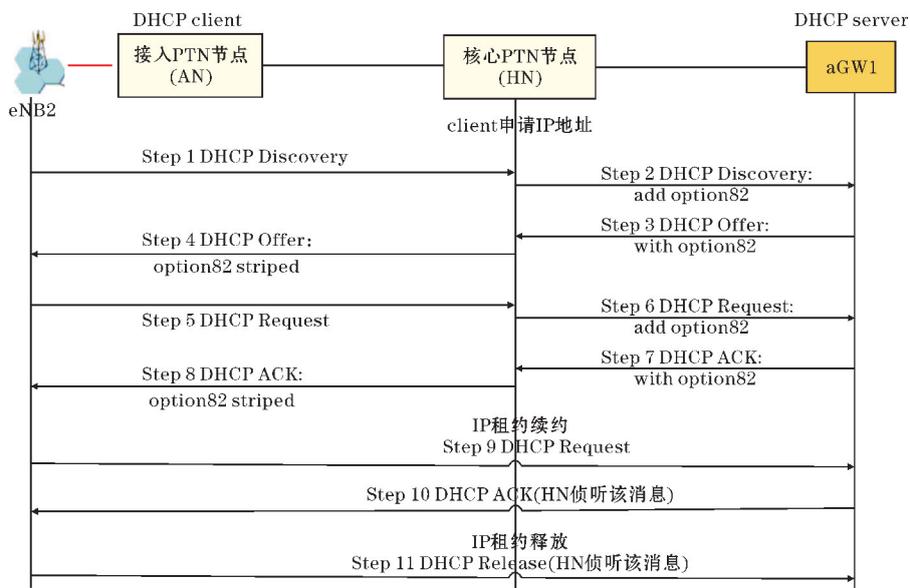


图 3 DHCP IP 申请/续约/释放的消息流程图

option 82 是 DHCP 报文中的中继代理信息选项 (relay agent information option)。核心 PTN 节点支持 DHCP 中继功能,收到 client 发的 DHCP 请求报文后添加 option 82 选项(其中包含了客户端的接入物理端口和接入设备标识等信息),然后再把该报文转发给 server。server 发的 DHCP 应答报文也回带 option82,PTN 核心节点收到 server 应答报文后,剥离 option 82 选项,并根据选项中的物理端口信息转发应答报文给 client。

Option 82 包含多个子选项 sub-option,其中子选项

1 是必备的,定义了电路 ID(Circuit ID),它表示接收到的 DHCP 请求报文来自的链路标识^[7],这个链路标识即核心 PTN 转发到该 DHCP client (eNodeB)的 IP 包的输出 port,也就是 IP 路由表中对应该 eNodeB IP 表项的“输出口”。该链路标识只在中继代理节点内部有意义,server 不解析其含义,只是简单地把收到的 option82 拷贝回应答报文中。

因此,PTN 核心节点在 step7 收到 server 的 DHCP ACK 消息后,取出 Yiaddr,即得到 IP 路由表中的“IP/subnet”,解析 option82 中 sub-option1 的电路 ID(Cir-

cuit ID)字段,即得到 IP 路由表中“输出口”。此外,PTN 核心节点还要解析 IP 租约期限 option51,将初始租约时间写入 IP 路由表中‘剩余租约时间’域,即完成了 IP 路由表中 eNodeB 对应表项的建立。

此外,PTN 核心节点收到 step7 的 DHCP ACK 消息的源 IP 地址即 server(aGW)的 IP 地址,而 PTN 核心节点知道自己从哪个输入口收到 DHCP ACK 消息,这样就可建立 IP 路由表中 aGW 对应的表项。该表项“剩余租约时间”域为 forever,即永不过期。

2.2.2 IP 租约续约过程及 IP 路由表更新

IP 租约续约过程的消息流程图如图 3 所示。DHCP client 在 IP 租约期限过半时,会自动向 server 发送单播 DHCP request 报文,续延 IP 租约信息。如果得不到 server 确认可继续使用该 IP,当租约期到 87.5% 时,会重复一次申请续约的过程。server 收到 DHCP Request 续租报文后,根据报文中的 ciaddr 信息看该 IP 如果可续租,给 client 回应单播的 DHCP ACK 报文。

这个过程中 DHCP 消息的目的 IP 是 server 或 client 的单播 IP 地址,对 PTN 核心节点的 DHCP 中继功能是透明的,PTN 核心节点只是透传该消息,不会在消息上添加 option82。此时 PTN 核心节点要侦听 server 发出的 IP 续约的 DHCP ACK 消息(step10):若 IP DA 和 DHCP ACK 中的 Ciaddr(client IP 地址)指示的 IP 相同,且该 IP DA 在该 PTN 核心节点的 IP 路由表中,则解析 DHCP ACK 中 IP 租约期限 option51,在 IP 路由表中该 IP 对应的“剩余租约时间”域加上新的租约时间。

根据目的 IP 地址的不同可以对 step7 DHCP ACK 和 step10 的 DHCP ACK 的处理区分开:step7 的 DHCP ACK 是广播 IP 或 IP 就是 PTN 核心节点自己的 IP,而 step9 的 DHCP ACK 是单播 IP 且 IP 地址不是 PTN 核心节点自己的 IP。

2.2.3 IP 释放过程及 IP 路由表删除

IP 释放过程的消息流图如图 3 所示。client 在成功获取 IP 地址后,随时可以通过发送 DHCP Release 报文释放自己的 IP 地址,server 收到 DHCP Release 报文后,回收相应的 IP 地址重新分配,Server 无需回应 DHCP ACK。

DHCP Release 报文是 server IP 的单播 IP 报文。对 PTN 核心节点的 DHCP 中继功能是透明的,PTN 核心节点只是透传该消息,不会在消息上添加 option82。此时 PTN 核心节点要侦听 DHCP RELEASE 消息(step11):若 DHCP RELEASE 消息中的 Ciaddr(client

IP 地址)指示的 IP 在该 PTN 核心节点的 IP 路由表中,PTN 核心节点在 IP 路由表中删除该 IP 地址的表项。

2.2.4 PTN 核心节点完整的 IP 路由表动态建立和维护过程

通过上述过程,PTN 核心节点就动态建立了 IP 路由表中 eNodeB 和 aGW 对应的表项,并可在 eNodeB 续租或释放 IP 时动态更新 IP 路由表。对回传网络中 MME 和其他 aGW,包括一个 eNB 归属多个 MME/aGW 的情况,其 IP 路由表项仍需要人工配置到核心 PTN 节点,不过一个回传网络中由于只有 2-4 个 MME 和 aGW,这种需要人工配置的表项只剩 1-3 项,相对之前需要人工配置成千上万项的 IP 路由大为简化,且一般 aGW 和 MME 配置好后比较固定,不像 eNodeB 经常有添加和移除的变化,也大幅减少了维护工作量。

两个 eNB 之间通信的情况,如果两个 eNB 属于同一个核心 PTN 节点,比如图 1 中 eNB1 和 eNB2,通过文中方法,在 DHCP 申请 IP 地址过程中已经在 HN1 路由表中分别建立了自己的 IP 转发表项,这样 eNB1 和 eNB2 就可以在 IP 层互通。比如 HN1 收到 eNB1 发给 eNB2 的包,根据目的 IP=eNB2 IP 查路由表,得到其输出口是 PW2,则知悉是 eNB 之间报文,直接剥掉原来报文中 TU1 和 PW1 标签,加上 TU2 和 PW2 标签,在 MPLS 层转发出去。如果两个 eNB 归属不同的核心 PTN 节点,需要在核心 PTN 节点上手工配置非属于自己的 eNB 的静态路由,不过根据现有移动组网情况,这种情形很少,需要人工配置的表项也很少。

PTN 核心节点有定时器,定期去减少 IP 路由表中 eNodeB 表项的“剩余租约时间”域,“剩余租约时间”减到 0 后会自动删除该表项。不过对 aGW/MME 对应的 IP 路由表项,其“剩余租约时间”为 forever,即永不过期,定时器不会减少 forever 的表项,这些表项要在 aGW/MME 的 IP 变化时人工删除和重建。PTN 核心节点完整的 IP 路由表动态建立和维护流程图如图 4 所示。

3 实验验证

搭建验证平台如图 1 所示,在 HN1 和 HN2 上运用算法自动建立核心 PTN 节点 HN1 和 HN2 的静态 IP 路由表如表 1 所示。核心 PTN 节点根据该路由表可以正确完成 L3VPN 包的转发。

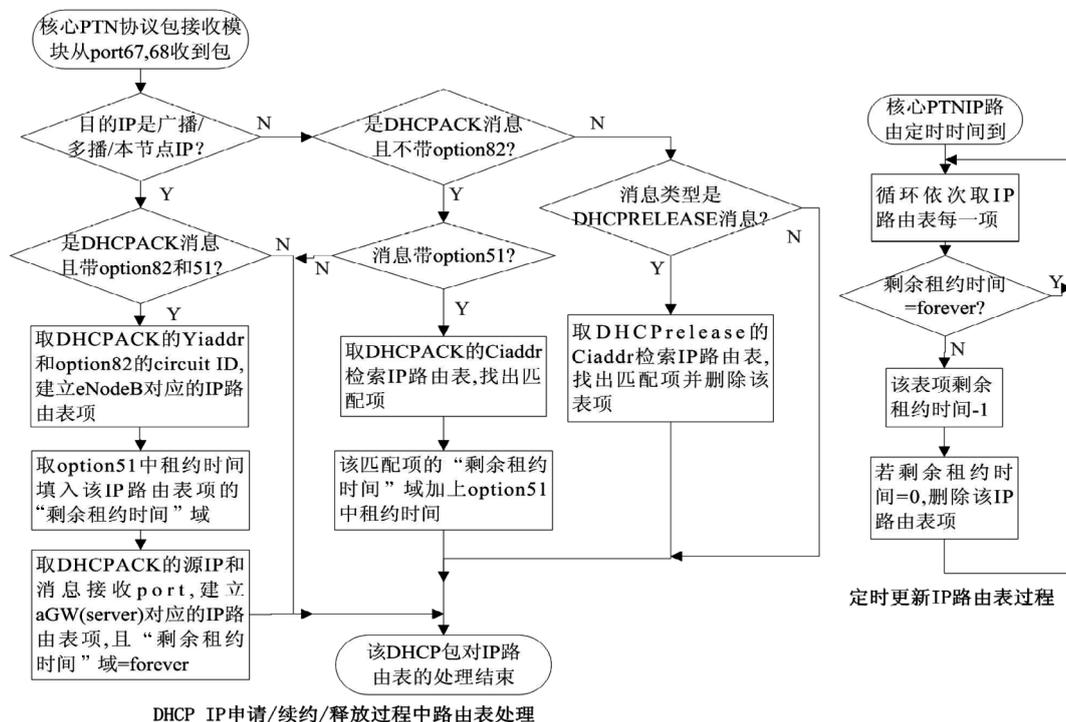


图 4 PTN 核心节点完整的 IP 路由表动态建立和维护流程

表 1 PTN 核心节点静态 IP 路由表配置

HN1 节点静态 IP 路由表			HN2 节点静态 IP 路由表		
IP/subnet	输出口	剩余租约时间	IP/subnet	输出口	剩余租约时间
eNB1 IP	PW1	4	eNB1 IP	VC_a	4
eNB2 IP	PW2	4	eNB2 IP	VC_a	4
MME1 IP	VC_a	forever	MME1 IP	IF2	forever
aGW1 IP	VC_a	forever	aGW1 IP	IF1	forever

由于 HN1 和 HN2 都支持 DHCP 中继功能,而关于多级 DHCP 中继并没有标准定义。算法采用一种简单的处理方式,核心 PTN 节点若从其他核心 PTN 节点收到带 option 82 的 DHCP Discovery 和 Request 消息则透明转发(例如图 1 中 HN2 收到 HN1 发来的 DHCP Discovery 和 Request 消息),并记录收发端口。若收到 DHCP Offer 和 ACK 消息,则看出口,若出口连接的是其它核心 PTN 节点,则透明转发(例如图 1 中 HN2 收到 aGW1 的 DHCP Offer 和 ACK 消息),并记录收发端口;若出口连接的是接入 PTN 节点,则剥掉 option82。

4 结束语

针对 PTN LTE 回传网络中静态 L3VPN 中需要手工配置成千上万的 IP 路由表项给网络配置和维护带来大工作量及易出错的问题,提出一种创新方法:PTN 核心节点通过 DHCP 中继功能和侦听 DHCP 消息在核心 PTN 节点上自动建立动态 IP 转发表。该方法可极

大减少静态 L3VPN 部署和维护中配置 IP 路由的工作量,使 LTE 回传网络维护和配置简单和不易出错。

文中提出的方法还有进一步改进的空间,比如一个接入 PTN 节点如果以双归属保护方式连接到两个核心 PTN 节点,此时如何自动配置备用核心 PTN 节点的 IP 路由表及发生保护倒换时如何自动更新 IP 路由表都需要进一步研究。

参考文献:

[1] 赵杰. 浅谈 PTN 技术与网络架构[J]. 通讯世界,2017(24):133-133.
 [2] 方伟津,黄剑华,陈鑫雄. 面向 TD-LTE 的 PTN 接入层演进策略研究[J]. 电信工程技术与标准化,2015(7):17-20.
 [3] 陈艳,黄建. 基于 PTN 技术的传输网络构建与应用[J]. 通讯世界,2016(2):41-41.
 [4] 王会敏. 浅析 LTE 时代下的 PTN 网络优化[J].

- 通讯世界,2015(15):49-50.
- [5] YD/T 2336.1-2016 分组传送网(PTN)网络管理技术要求 第1部分[S].北京:2016.
- [6] 徐宗标.面向TD-LTE移动回传的PTN部署策略及演进方向[J].电信技术,2014(3):95-98.
- [7] 官向民.PTN技术在二层专网中的应用[J].电信快报,2017(6):8-10.
- [8] 兰可.基于PTN技术的TD-LTE光传输组网模式研究[J].通信世界,2018(2):114-115.
- [9] 王牧云.PTN支持L3VPN技术的研究与验证[J].电信科学,2012(4):119-123.
- [10] 程伟强,王敏学,张晴.城域POTN应用、设备和组网需求及分析[J].邮电设计技术,2015(3):6-10.
- [11] PTN网管系统中静态L3VPN业务的设计与实现[D].武汉:武汉理工大学,2013.
- [12] Bocci M, Bryant S, Frost D, et al. A framework for MPLS in transport networks[R]. draft-ietf-mpls-tp-framework-10. 2010.
- [13] 覃晓霞,卢灵宣.EPC Pool组网下PTN网络配置方案[J].邮电设计技术,2016(4):72-75.
- [14] 何钰,李瑞祥.配置DHCP中继管理地址[J].网络运维与管理,2016(11):52-53.
- [15] W. Richard Stevens. TCP/IP详解[M].北京:机械工业出版社,2016.

A Method to Dynamic Learning IP Routing in PTN Static L3VPN

SUN Jie, YAO Yao, REN De-hao

(College of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: LTE backhauling is one of the most important applications of PTN L3VPN. Normally, thousands of IP route table items is manually configured, which related to PTN core nodes in static L3VPN. This will cause problems such as complexity and trivial steps in the proceedings of both network configuration and maintenance, so it is difficult for administrator to correctly configure. In order to solve this issue, a novel method about automatically building IP routing table is proposed. By utilizing DHCP relay function and DHCP messages snoop, this method can be implemented in core PTN node to solve above issue.

Keywords: communication and information system; digital communication technology; static L3VPN; DHCP relay; DHCP snooping; PTN; LTE backhauling