



SD-WAN 隧道技术与组网模式

牛佳, 颜永明, 林志华

(中国电信股份有限公司上海分公司, 上海 200085)

摘 要: 介绍了 SD-WAN 技术现状, 比较了 SD-WAN 相对传统专线的优势, 描述了 SD-WAN 的基本技术特征, 对 SD-WAN 的软件定义网络控制器、安全加密与租户隔离、SD-WAN 业界标准进行了说明。通过分析 SD-WAN 主流厂商相关的技术, 对 SD-WAN 所用的隧道技术进行比较, 对不同厂商选用的私有 IPsec 隧道、VxLAN over IPsec、NvGRE over IPsec、SSL 隧道等组网方案做了分析比较, 给出了优选的隧道技术方案。并进一步针对如何应用隧道技术实现组网, 结合运营商组建 SD-WAN 的优势, 提出优化意见, 给出了 POP 点组网和点对点组网方案。最后对 SD-WAN 未来标准的制定、运营商与 SD-WAN 厂商的未来发展趋势进行了展望。

关键词: SD-WAN; VxLAN; IPsec; NvGRE

中图分类号: TN915.41

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021004

SD-WAN tunnel technology and network constuction

NIU Jia, YAN Yongming, LIN Zhihua

Shanghai Branch of China Telecom Co., Ltd., Shanghai 200085, China

Abstract: The status of SD-WAN technology was introduced, the SD-WAN advantages with traditional line were compared. The basic technical features of SD-WAN, the software definition network controllers of SD-WAN, security encryption and tenant isolation, the SD-WAN industry standards were described. The relevant technologies of SD-WAN main manufacturers were also analyzed, the tunnel technology used by SD-WAN was compared. And private IPsec tunnel, VxLAN over IPsec, NvGRE over IPsec, SSL tunnel network project used by the different companys was compared and analyzed, the optimized tunnel technology project was given, and optimizational suggestions on how to use tunnel technology to complete networking project were further put forwarded, combined with the advantages of Internet service product in establishing SD-WAN, the POP network and peer-to-peer networking solutions were given. Finally, it make future development prospect for SD-WAN standards, Internet service product and SD-WAN manufacturers trends.

Key words: SD-WAN, VxLAN, IPsec, NvGRE

1 引言

随着互联网数据中心业务和云业务的高速发

展, 越来越多的企业开始不满足于基本的网络解决方案, SD-WAN (software-defined wide area network, 软件定义广域网) 是软件定义网络



(SDN) 技术进一步在广域网中的应用, 近年逐渐进入人们的视线。在其出现之前, 传统企业多使用 Internet 加密或者专线, 搭建总部与分支机构的互联互通网络实现内部数据的同步与共享, 对网络层与应用层的交互联系没有那么密切的高要求。

2 SD-WAN 现状

2.1 SD-WAN 与传统专线的比较

2.1.1 传统专线特点

如今, 越来越多的用户选择最新的 SD-WAN 作为网络组网方案首选。因为与传统 MPLS-VPN 专线相比, SD-WAN 兼顾了企业用户对于成本低廉、组网灵活、质量优质的诉求。

传统的 MPLS-VPN 专线难以在网管系统/平台上实现集中配置下发, 缺乏对大规模设备进行远程管理的手段。其中, MPLS-VPN 专线必须在局端开通配置, 必须由运营商分配 RT (route target, 路由目标)、RD (route distinguisher, 路由区分符) 的值, 用户的配置必须和运营商 PE (provider edge, 边缘路由器) 的参数匹配, 运营商转发设备需全程支持 MPLS 标签化, 业务实现必须完全依赖运营商资源。

2.1.2 SD-WAN 与传统专线相比的优势

诸多限制因素, 导致 MPLS-VPN 没有办法满足配置简化、灵活性的需求。而 SD-WAN 能实现命令的集中下发, 采用网管侧预配置、终端侧零接触配置, 可免去对各分部终端的烦琐配置。在新增功能时, 迭代开发简单, 能在配置后台管理界面中集中查看设备的告警信息以及各个链路的状态。

同时, SD-WAN 可实现随选路由加载选择最优的分发路径。通过多网络随选和传送质量加强等功能, 解决了传统 Internet 传输不稳定, 时常出现的时延丢包、质量劣化等问题。另一方面, SD-WAN 继承了 Internet 组网的优势, 克服了专线租用独用通道成本高昂、配置开通时间周期较长等问题。

此外, SD-WAN 提供了对网络进行软件编程,

完整地实现了网络端到端、CPE 到 CPE 侧的实时监控、调整能力, 可实现网络和相关业务的快速部署和运维, 为用户提供高效快速的业务配置开通和配置变更能力。

2.2 运营商 SD-WAN 组网优势

2.2.1 运营商的运维优势

理想化的 SD-WAN 运维方式是实现一体化开通, 只需上门一次, 就能同时完成 Internet underlay 网络以及 SD-WAN 设备和 overlay 的线路开通。

在运维方面, SD-WAN 运维需要具有全局各层级的视角, 能实现 overlay 和 underlay 网络的一体化监控维护, 可以同步检测 underlay、overlay 网络, 准确判断故障点及劣化点, 缩短故障处理时间, 确保用户网络始终处于高可用、高质量的水平。

2.2.2 运营商的网络覆盖优势

在网络覆盖范围方面, 使用遍布全国的云计算服务, 能快速实现全国几十个 SD-WAN POP 点的云化组网, 充分利用 SD-WAN 中的软件定义的控制面 (controller) 实现对控制和管理平面的统一编排, 并作为第一认证和注册点, 对 SD-WAN 的配置进行管理, 集中管理配置策略模板。全国组网的 POP 点能实现对各地的全覆盖。部署上线顺利运行后, 一支遍布全国的高水准、规范化、高效率的 7×24 h 数据网络维护团队也是 SD-WAN 稳定运维的基础。

在用户越来越关注网络组网高性价比的趋势下, 传统的 Internet 和网络专线已无法满足用户专网组网的需求。随着越来越多的用户优先选择 SD-WAN 组建专网, 电信运营商的传统专线业务也受到冲击。在此背景下, 各大电信运营商开始着手构建自己的 SD-WAN, 通过对 SD-WAN 技术深入研究, 对相关厂商设备开展测试, 形成自己的组网方案, 搭建体现电信运营商优势的 SD-WAN, 实现规模商用运营, 很好地满足了不同层次的用户需求。

2.3 SD-WAN 的技术特征

2.3.1 软件定义网络控制器

SD-WAN 最基本的技术特征在于软件定义、软件控制, 可以通过可视化的 Web 界面实现企业 WAN 的网络状态可视化, 提供页面进行智能自动化的能力, 对用户端 CPE 网络设备的性能状态做到实时性能监控。过去, 传统的网络部署架构比较依赖网络工程师对于网络知识的深度理解和对整体网络架构的规划, 网络配置也需要人工维护, 后期运维也需要人主动去判别故障处理, 对于每个网络节点都需要逐条命令行去配置下发。然而, 通过软件定义下发完全可以取代烦琐的预配置云维护, 网络控制器可实现网络业务抽象化, 网络配置模板可实现配置下发自动化, 对复杂的命令进行封装不对外开放, 对外屏蔽了具体的网络技术细节, 仅提供 RESTful API 和网络参数, 方便用户根据实际网络需求进行具体的网络业务编排、网络控制二次化开发、自动化智能化运维以及集中网络控制器配置下发。

2.3.2 安全加密与租户隔离

为了确保用户传输数据的可靠性和安全性, SD-WAN 的 overlay 网络需要对用户的数据报文进行加密, 满足了有安全性需求的企业用户, 同时保障了每个用户的安全隐私以防止数据泄露。通过加 VPN 租户标志的方式, 也满足了企业多租户的需求, 支持同一租户下多个部门的逻辑隔离。由于 CPE 是暴露在互联网上的, SD-WAN 也需要提供基本的安全策略保证 (如防攻击、基本的防火墙访问限制功能), 在用户内网与公网之间也需要提供不同的访问策略, 通过 ACL 保证访问控制。

2.3.3 SD-WAN 标准

根据 ONUG、Gartner、MEC 等组织定义的 SD-WAN 基本特征, 共性的看法有: SD-WAN 能实现用户分支的快速部署、快速上线, 部署效率大大提高; SD-WAN 能动态调整用户应用的流量

路径, 实现灵活智能化、便捷自动化的流量调度; SD-WAN 能集中管控, 通过可视化界面实现网络的编排运维功能; SD-WAN 提供网络安全、网络优化、业务快速发放等增值业务。

3 主流隧道/加密技术

SD-WAN 转发层面使用了隧道、加密技术。隧道技术是为了解决租户隔离甚至每个租户中各个部门的隔离问题, 实现能在 underlay 网络上区分用户。加密技术则是为了完成传输数据的加密, 在开放的 Internet 环境下保证用户业务数据的传输安全性。SD-WAN 隧道技术不依赖具体的 IP overlay 技术、运营商 WAN 组网技术, 可以通过现有的 IPSec、VxLAN、NvGRE 隧道技术进行组合搭建可用可扩展的隧道技术。下面将对传统隧道/加密技术进行分析。

3.1 IPSec

IPSec 是一个公开的网络层安全性框架协议。它不是一个简单孤立的协议, 而是由多个 IP 网络协议和安全服务组成的集合。IPSec 主要包括安全协议验证头部和封装安全载荷、密钥管理协议、Internet 密钥交换协议以及其他用于网络安全认证及加密的算法等。传统的 IPSec 隧道技术分为隧道 (tunnel) 封装模式和传送 (transport) 封装模式: tunnel 封装模式在头部封装了一个外网的 IP 地址, 实现对 LAN 侧的 IP 地址封装, transport 封装模式仅利用了 AH 协议和 ESP, 实现认证和加密。IPSec tunnel 模式可在 overlay 网络实现租户隔离, 并完成数据加密, 但在单独使用时需要建立较多 tunnel 实现租户级甚至部门级的隔离, 每个租户各使用一条隧道, 每条隧道都需要分配密钥、实现加密, 传统方式下, 密钥可以通过预分配或者电子证书的形式在每台终端上配置, 由于每个租户的密钥不同, 在租户较多的情况下, 会导致加密/解密资源大量消耗。IPSec 报文封装如图 1 所示。

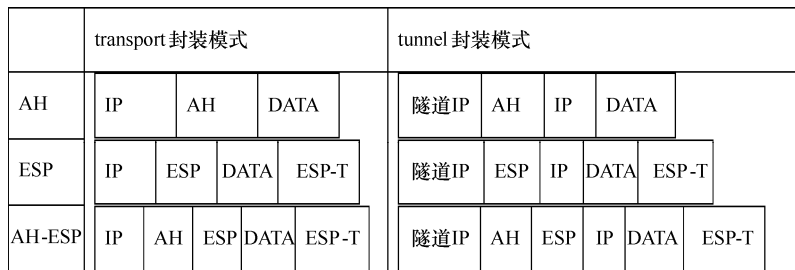


图1 IPSec 报文封装

3.2 VxLAN

国外专业机构定义了 VLAN 扩展方案——VxLAN（虚拟扩展局域网）。VxLAN 使用设备物理地址封装在用户数据报协议中的封装方式，是 NVO3（在第三层网络虚拟化）中的一种虚拟化网络技术。VxLAN 特性在本质上属于一种 VPN 技术，能在任意路由可达的网络上叠加二层虚拟网络，通过 VxLAN 网关实现 VxLAN 内部的互通，同时，也可以实现与传统的非 VxLAN 网络的互通。VxLAN 通过采用 MAC 地址封装在 UDP 中的形式扩展二层网络，将 Ethernet 报文封装在 IP 报文之上，通过路由协议在网络中进行传输，无须关注虚拟机的 MAC 地址。由于对路由的网络并无结构限制，因此具备了大规模的延伸能力。通过路由网络，虚拟机迁移不受原来运营商网络架构的制约。VxLAN 可以通过静态和动态的两种方式建立隧道，动态使用 EVPN/BGP 建立邻居关系。VxLAN 只具有隧道功能，不具备加密功能。VxLAN 报文封装如图 2 所示。

3.3 GRE

GRE 隧道是一种逻辑接口，通过 GRE 隧道可以将承载要发送的数据包整体封装在协议中传输。为 GRE 隧道设计的网络体系结构是为了构成点对点的传输封装服务。另外，GRE 隧道技术是一种完全无状态记录的设计，这表示每个 GRE 隧道的传输结束点从不记录关于其他 GRE 隧道结束点的信息状态。如果 GRE 隧道的结束对端不可达，在这样的设计之下，本地 GRE 隧道数量不会因为对端结束点状态变化而减少。也就是说，GRE 不

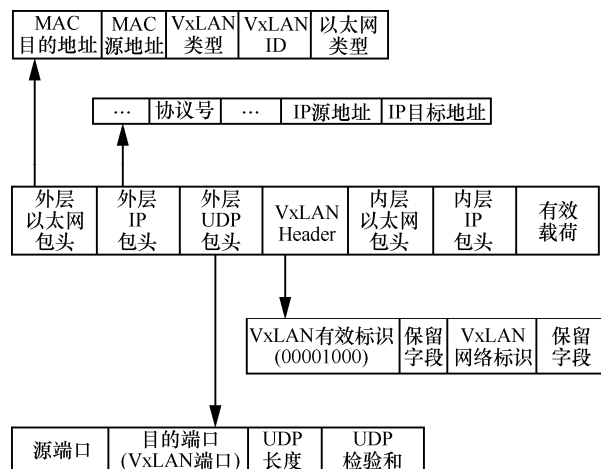


图2 VxLAN 报文封装

具备在链路对端不可达时将对该端的接口记录为 down，进而删除在路由表记录中所有使用这个接口节点作为对外转发接口的静态路由，为备选静态路由的安装或是为了基于策略的路由选择一个备选下一跳或接口做准备的功能。同时，GRE 不具备加密功能。

GRE 是一种 VPN 的第三层隧道协议。GRE 隧道是一种虚拟的点对点的连接，为需要在其他网络层协议传输的协议报文进行封装，被封装的协议报文可以在其他网络协议中传输。GRE 隧道接收报文后，对报文进行封装，在报文之前加入 GRE 头部，再将封装完成后的初始报文和 GRE 头部根据初始报文的 IP 报文通过传输协议进行转发。在解封装时，网络层发现外层头部报文协议号为 47，将 GRE 外层头部进行剥离，将剩下的 IP 报文和数据报文进行处理传输。GRE 报文封装如图 3 所示。



图3 GRE 报文封装

3.4 NvGRE

NvGRE 是由微软公司提出，在 GRE 的基础上发展出来的。NvGRE 和 GRE 的封装包头格式基本一致，区别是在 NvGRE 的包头的 S 位和 C 位置零，所以 NvGRE 包头不含检验和与序列号。与 VxLAN 比较而言，NvGRE 并没有使用标准的传输协议，而是用通用的路由封装协议。NvGRE 使用 GRE 报文头部的前 24 bit 作为租户隔离的网络标识位，可以总共支持 2^{24} 个 VPN，这一点与 VxLAN 相同。NvGRE 传输网络使用了 GRE 报文头提供承载带宽利用率颗粒度的数据流，但是这样的设计使 NvGRE 不能提供传统的负载均衡，相对 VxLAN 而言，这是 NvGRE 的不足之处，也是和 VxLAN 有所区别的地方。NvGRE 报文封装如图 4 所示。



图4 NvGRE 报文封装

3.5 SSL

SSL（安全套接层）隧道技术是一种网络安全加密协议。它是一种通过 TCP/IP 通信传输协议实现的安全加密协议，并且使用公开的密钥技术。SSL 广泛支持各种类型的网络协议，同时提供 3 种基本的网络安全服务，这些服务都使用公开密钥技术。SSL 支持的服务通过网络进行通信却不损害安全性。它能在用户分支结构和总部之间创建一个安全连接，然后通过该 SSL 连接安全地发送任意数据量。

SSL 的初始目标是为两个传输应用之间提供私有的可靠连接。SSL 协议包括两层，在某些可靠传输协议的顶层是 SSL 记录层，SSL 记录层是用来封装其他更高层的协议。SSL 握手协议也是一种封装协议，允许服务器和客户端互相认证，在应用协议发送接收第一个数据前协商加密算法和密码。应用协议的独立性是 SSL 的优点之一，更高层次的协议显然可以使用在 SSL 协议之中。

4 隧道加密技术在 SD-WAN 的应用

对 5 家主流厂商的隧道加密技术进行了测试比较。SD-WAN 主流厂商中，C 厂商用了私有的 IPSec 隧道技术，H 厂商用了 NvGRE over IPSec 隧道技术和 VxLAN，V 厂商使用了 VxLAN over IPSec，D 厂商用了 SSL 技术，S 厂商用了 IPSec 隧道技术。

4.1 私有 IPSec

标准的 IPSec 隧道技术为了区分不同的租户，每个租户使用一个 tunnel，每个 tunnel 独立实现数据流的加密，在加密/解密的过程中，加密/解密的进程较多，将消耗大量计算资源。为解决这一问题，C 厂商在标准的 IPSec 隧道技术中引入私有字段，通过对 IPSec 报文中加一个 LBL 字段记录租户信息减少 tunnel 数量，以降低计算资源的消耗。私有 IPSec 封装格式如图 5 所示，格式中包含了源/目的 IP 地址、用户数据报协议、传输的源数据包，在加密的数据包外部加入了新的 IP 包头，与传统的公有 IPSec 相比，通过加入新的字段满足

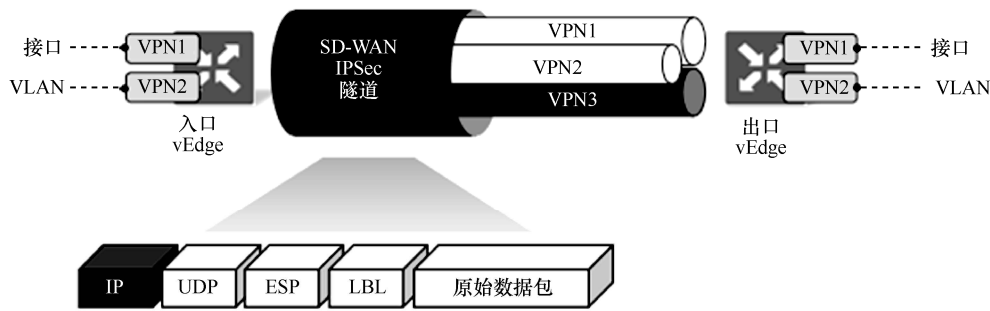


图5 私有 IPsec 封装

SD-WAN 对于多租户下的隔离问题, LBL 多占用 4 bit。

4.2 VxLAN over IPsec

单独的 VxLAN 只能解决租户隔离的问题, 无法实现加密传输, 为了解决单一隧道无法加密的问题, V 厂商使用了 VxLAN over IPsec 技术, 将 VxLAN 的报文封装在 IPsec 报文中。VxLAN 是明文报文在网络传输不安全, 通过 IPsec transport 模式, 可以实现安全传输。使用该技术, VxLAN 的外层 IP 地址暴露在外, IPsec 只对 VxLAN 数据包进行加密。因而, 无须使用多个 IPsec tunnel, 既减少了加密/解密资源开销, 同时也解决了租户隔离, 甚至租户内部跨部门隔离的问题。通过图 6 的抓包信息可以看到, VxLAN 协议中存在 Flags、Reserved、VNI 等字段信息, VxLAN Network Identifier 信息是为了给不同租户分配标识做到租户隔离, 每个租户的标识都是

唯一的, Flags 为 VxLAN 的标志位, 第一个 Reserved 为保留位, 后一个 Reserved 为保留字段。

4.3 NVGRE over IPsec

NVGRE 只实现了租户隔离的隧道化功能, 没有解决加密问题。H 厂商主推 NVGRE over IPsec 方案, 就是将 NVGRE 的报文封装在 IPsec 报文中。由于 NVGRE 是明文报文, 在网络传输不安全, 通过 IPsec transport 模式, 可以实现加密传输, 保证私密性。使用 NVGRE 的原因和 VxLAN 大致相同, 也是为了解决 IPsec tunnel 模式下, 实现租户隔离需消耗大量计算资源的问题, 同时, 也解决了单 NVGRE 隧道下无法做到安全加密的问题。但 NVGRE 相与 VxLAN 相比存在一点不足, NVGRE 封装包头采用的是 GRE 协议, 许多网络设备和防火墙无法对 GRE 头部进行处理, 所以 NVGRE 协议不支持负载均衡功能, 而 VxLAN 使用的协议是 UDP, 天然地支持网络负载均衡。H 厂商隧道

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	192.168.2.2	61.151.158.115	VxLAN	154	
2	0.011994	192.168.2.2	61.151.158.115	VxLAN	190	
3	0.011997	192.168.2.2	61.151.158.115	VxLAN	146	
4	0.035999	192.168.2.2	61.151.158.115	VxLAN	146	
5	0.060001	192.168.2.2	61.151.158.115	VxLAN	126	
6	0.087998	192.168.2.2	61.151.158.115	VxLAN	158	
7	0.091997	61.151.158.115	192.168.2.2	VxLAN	166	
8	2.875994	192.168.2.2	61.151.158.115	VxLAN	134	
9	2.879994	61.151.158.115	192.168.2.2	VxLAN	146	
10	2.891997	61.151.158.115	192.168.2.2	VxLAN	146	

> Frame 1: 154 bytes on wire (1232 bits), 154 bytes captured (1232 bits)	
> Ethernet II, Src: fa:16:3e:82:c9:e0 (fa:16:3e:82:c9:e0), Dst: fa:16:3e:e7:12:bc (fa:16:3e:e7:12:bc)	
> Internet Protocol Version 4, Src: 192.168.2.2, Dst: 61.151.158.115	
> User Datagram Protocol, Src Port: 4790, Dst Port: 4790	
Virtual eXtensible Local Area Network	
> Flags: 0x1c Reserved: 32771 Next Protocol: Unknown (252) VxLAN Network Identifier (VNI): 27921 Reserved: 128	
> Data (104 bytes) Data: 0200757a00003ec700003ec7000000003e0c775ad370b3b4... [Length: 104]	

图6 V 厂商隧道技术抓包信息

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	50.1.2.2	50.1.1.2	ESP	446	ESP (SPI=0xcb348a32)
2	0.009989	50.1.2.2	50.1.3.2	ESP	446	ESP (SPI=0x5855e3f7)
3	0.019983	50.1.2.2	50.1.4.2	ESP	446	ESP (SPI=0x29fecfe1)
4	0.823440	50.1.2.2	50.1.4.2	ESP	190	ESP (SPI=0x29fecfe1)
5	0.999967	50.1.2.2	50.1.1.2	ESP	446	ESP (SPI=0xcb348a32)
6	1.009941	50.1.2.2	50.1.3.2	ESP	446	ESP (SPI=0x5855e3f7)
7	1.019972	50.1.2.2	50.1.4.2	ESP	446	ESP (SPI=0x29fecfe1)
8	1.082098	HuaweiTe_d0:90:10	LLDP_Multicast	LLDP	492	MA/60:d7:55:d0:90:10 IN/GigabitEthe
9	1.999961	50.1.2.2	50.1.1.2	ESP	446	ESP (SPI=0xcb348a32)
10	2.009938	50.1.2.2	50.1.3.2	ESP	446	ESP (SPI=0x5855e3f7)

> Frame 1: 446 bytes on wire (3568 bits), 446 bytes captured (3568 bits)

> Ethernet II, Src: HuaweiTe_d0:10:15 (60:d7:55:d0:10:15), Dst: HuaweiTe_56:06:66 (38:4c:4f:56:06:66)

> Internet Protocol Version 4, Src: 50.1.2.2, Dst: 50.1.1.2

▼ Encapsulating Security Payload

ESP SPI: 0xcb348a32 (3409218098)

ESP Sequence: 38191

图 7 H 厂商隧道技术抓包信息

技术抓包信息如图 7 所示, 在 NVGRE over IPSec 协议中, IPSec 隧道使用了协议 ESP, 包含 ESP SPI、ESP Sequence (ESP 序列), SPI 是用来确定唯一的安全联盟、Sequence 是为了保护接收侧防止受到攻击。ESP 的尾部字段 padding 用来隐藏加密数据的真实长度, padlength 代表需要填充的字节数, Next Header 表示下一个被加密的头部数据类型。NVGRE 中的 Checksum 表示 GRE 报文中所有数据的校验和通常置零不用, Sequence Number 表示数据包的序号也在 NVGRE 报文中置零不用, Key 表示密钥信息, 一般置 1 表示包含 VSID, 另外还包括版本号、数据协议类型, 再使用 24 bit

的虚拟子网标识符来标记 NVGRE 网络。

4.4 SSL

D 厂商采用 SSL 技术实现 SD-WAN。但是 SSL 主要在网页的应用上使用得较多, 功能和多租户隔离方面的能力较弱, 能否完美支持未来 SD-WAN 的 POP 点组网模式, 值得商榷。SSL 不支持多 VPN 业务隔离, 只具备加密功能, 如果需要多租户, 必须使用 SSL 多进程, 对租户隔离的支持能力较弱。D 厂商隧道技术抓包信息如图 8 所示, 从图 8 的抓包信息可以看到, SSL 协议内部使用了 TLS 协议, 消息两端的加密通过非对称或公钥加密算法, 协商过程非常安全, 无法被监听者观察。

No.	Time	Source	Destination	Protocol	Length	Info
1	0.000000	7.7.7.2	7.7.7.1	DTLSv1.0	67	Application Data
2	0.000385	7.7.7.1	7.7.7.2	TLSv1.2	107	Application Data
3	0.000473	7.7.7.1	7.7.7.2	DTLSv1.0	67	Application Data
4	0.000857	7.7.7.2	7.7.7.1	DTLSv1.0	67	Application Data
5	0.001039	7.7.7.1	7.7.7.2	DTLSv1.0	67	Application Data
6	0.001099	7.7.7.2	7.7.7.1	TLSv1.2	107	Application Data
7	0.001158	7.7.7.1	7.7.7.2	TCP	66	35842 → 443 [ACK] Seq=42 Ack=42 Win=49 Len=0 TSval
8	1.002472	7.7.7.1	7.7.7.2	TLSv1.2	107	Application Data
9	1.002566	7.7.7.1	7.7.7.2	DTLSv1.0	67	Application Data
10	1.002598	7.7.7.2	7.7.7.1	DTLSv1.0	67	Application Data
11	1.002852	7.7.7.1	7.7.7.2	DTLSv1.0	67	Application Data
12	1.003024	7.7.7.2	7.7.7.1	DTLSv1.0	67	Application Data

> Frame 3: 67 bytes on wire (536 bits), 67 bytes captured (536 bits)

> Ethernet II, Src: Fuzhoux_05:2e:12 (68:91:d0:d5:2e:12), Dst: Fuzhoux_05:2f:05 (68:91:d0:d5:2f:05)

▼ Internet Protocol Version 4, Src: 7.7.7.1, Dst: 7.7.7.2

0100 = Version: 4

.... 0101 = Header Length: 20 bytes (5)

> Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)

Total Length: 53

Identification: 0x8ba6 (35750)

> Flags: 0x0000

Fragment offset: 0

Time to live: 64

Sequence number: 0

Window size: 0

Checksum: 0x0000

Urgent pointer: 0

Options: 0x0000

0000 68 91 d0 d5 2f 05 68 91 d0 d5 2e 12 08 00 45 00 h.../..E-

0010 00 35 8b a6 00 00 40 11 d3 01 07 07 01 07 07 -5....@.....

0020 07 02 b8 29 01 bb 00 21 1c 43 17 fe ff 00 02 00 ...)...!..C.....

0030 00 00 00 0f 5a 00 0c f3 f4 a7 5a 48 9c 71 e5 3fZ...ZH-q?

0040 2d 18 8a

图 8 D 厂商隧道技术抓包信息



4.5 方案比较

SD-WAN 各厂商使用的隧道及加密技术比较见表 1。

对于 SD-WAN 场景下的隧道/加密技术,目前主流厂商采用的技术中最优的方案是 C 厂商的私有化 IPSec 技术,私有的 IPSec 技术满足隧道、加密支持业务隔离,配置复杂性低、报文开销小,支持负载均衡,但是由于 C 厂商为私有协议,无法全面地在其他厂商中全面兼容使用,希望 C 厂商私有 IPSec 技术能成为国际标准组织的标准,实现大规模推广。退而求其次, V 厂商的 VxLAN over IPSec 技术是隧道技术的次优方案,由于采用了标准技术,它能在所有的厂商中进行推广,既满足了数据加密和租户隔离的功能,也能解决 NVGRE 隧道存在的负载均衡问题,但封装机制较为复杂,相较于 C 厂商的私有 IPSec 技术, VxLAN over IPSec 技术配置复杂性较高,报文开销一般但比私有 IPSec 技术开销要大,多出 20 bit 的开销。与 VxLAN over IPSec 技术相比,由于 NVGRE over IPSec 技术的 NVGRE 特性不支持负载均衡,为了实现负载均衡需要使用多个 IP 地址,增加了网络地址的额外开销。SSL 隧道技术不支持业务隔离,需要通过开启多进程手段实现业务隔离,并且 SSL 隧道技术更适用于 Web 应用,所以也不建议推广适用。

5 SD-WAN 组网模式

目前,上海电信 SD-WAN 组网包括 POP 点组

网和点对点直接组网两种方式。主流厂商中的 H 厂商、V 厂商、C 厂商是主推互联网点对点组网,由于市场选择改进使用 POP 点的组网方式,而 D 厂商和 S 厂商主要使用 POP 点的组网。

5.1 SD-WAN POP 点组网方式

POP 点组网示意图如图 9 所示。POP 点组网方式采用分段隧道化,POP 点与 POP 点之间由 MPLS-VPN 建立专用通道,底层使用的是覆盖范围、网络容量、业务能力各项指标最强的网络——中国电信 CN2 骨干网。

POP 点的组网过程中,在 POP 点与 POP 点之间使用 MPLS 进行传送,由于 MPLS-VPN 通道的封闭性,使传输过程中不需要使用额外的加密技术。此外,POP 点之间的 MPLS 使用 Full-Mesh 的结构,资源消耗较少,Full-Mesh 结构只需要配置一次 RT 值就可以实现导入导出、配置简化,每增加/减少一个用户不需改变 underlay 网络的配置,underlay 网络只需初始配置,能减少配置量。

DCI 网络中采用 MPLS-VPN 建立的专用封闭大通道,没有必要做加密处理,可以去掉 IPSec 封装,使用的 MPLS 标签,网络资源开销较小,无须额外的资源来对 tunnel 进行加密/解密,计算量大幅减少。

与 SD-WAN 互联网厂商的 DCI 网络相比,跨地域 SD-WAN 线路的质量优势明显。POP 点的组网可以使用 POP 点双挂方案,通过双上联解决设备冗余性问题。

表 1 SD-WAN 各厂商使用的隧道及加密技术比较

隧道/加密	厂商	隧道	加密	报文开销	配置复杂性	负载均衡	单加密进程实现多 VPN 业务隔离
私有 IPSec	C	支持	支持	低	低	支持	支持
VxLAN over IPSec	V	支持	支持	一般	高	支持	支持
NVGRE over IPSec	H	支持	支持	一般	高	不支持	支持
SSL	D	支持	支持	低	低	支持	不支持
标准 IPSec	S	支持	支持	低	低	支持	不支持

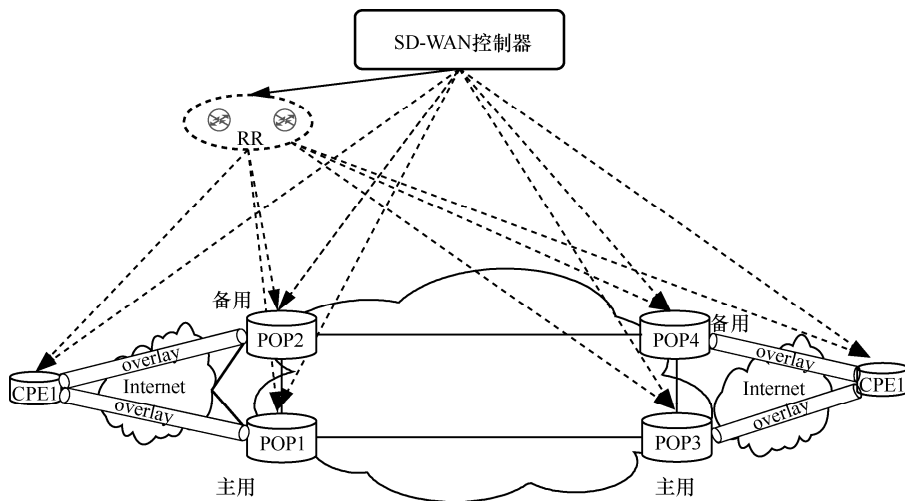


图9 POP点组网示意图

5.2 SD-WAN Internet 点对点方式

点对点技术直接在两个终端之间进行加密实现内部传输，缺点是在 underlay 网络上，基于 Internet 跨地址的传输质量无法保证，点对点技术适合用户节点较少、网络结构简单的场景，在多终端场景下配置较复杂。点对点组网方式示意图如图 10 所示。

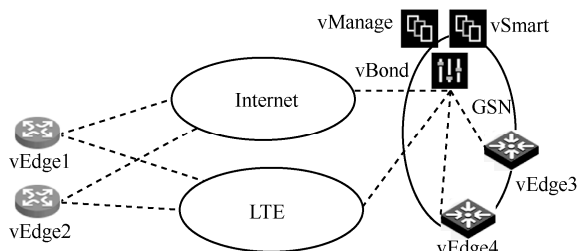


图10 点对点组网方式示意图

5.3 SD-WAN 引流技术

目前可实现的POP点方案解决了点对点方案中 Internet 跨地域网络质量无法保证的问题，缺点在于需分段配置，用户侧和POP点之间、POP点与POP点之间需要两次3段配置，单独配3段overlay工作量较大。为了将这3段配置简化为1次配置，建议进一步开发引入定向引流技术，实现一次性配置能跨越POP点全程开通的功能，保留POP点DCI MPLS VPN通道的同时，使用IPSec或者NvGRE隧道完成端到端配置的下发。引流技

术可以使用多层头部，即多层隧道、多次封装实现。

首先，在隧道外层封装到POP点的IP地址，在第二层内部封装租户的信息，实现租户隔离。通过识别外层的头部IP信息从用户侧到达第一个POP点之后，替换外部隧道头部IP地址内容，内层内容保持不变，仍旧能实现租户隔离，将外层头部IP地址替换为下一个POP点的地址，用来实现到下一个POP点的寻址转发，到达最后一个与用户目的地址直连的POP点后，通过内层的IP地址，找到用户目的设备。通过引流技术可以简化多个POP点之间的配置，大大减少了配置的工作量。也可以使用segment routing（分段路由）、flowspec等技术解决接入侧的引流问题。

6 结束语

就目前网络组网方案的发展现状来说，SD-WAN已形成了替代中低端专线业务的趋势。从2019年起，上海电信在SD-WAN相关领域技术进行了研究和探索，并实现了规模化商用组网。通过比较发现，在SD-WAN隧道技术方面C厂商的私有化IPSec技术具有优势，但在还未实现标准化的前提下，现阶段无法跨厂商推广，其他厂商可优先选择VxLAN over IPSec技术。SD-WAN



POP 点组网方式中, 现有的解决方案中配置量和工作量较大, 可以结合引流方案通过多层头部、多次封装来进行优化, 端到端的一次配置经由 POP 点转发。与互联网厂商相比, 电信运营商具有 overlay 和 underlay 兼顾的监控维护优势。相信在不久的将来, 在运营商和厂商的实践努力下, SD-WAN 的相关标准会更加完善, 为用户打造更好的服务。由于多云融合的推进, 公有云、私有云、混合云的开放接入, SD-WAN 可以有更多与云服务对接的机会, 相信云计算和 IDC 网络之间会有更深入的融合。在面对企业用户日益增长的新需求(如视频会议优化、国际加速业务、网络直播视频等), SD-WAN 会有更加精准的网络服务解决方案。

参考文献:

- [1] RISDIANTO A C, SHIN J, LING T C, et al. Leveraging ONOS SDN controllers for OF@ TEIN SD-WAN experiments[C]//Proceedings of the Asia-Pacific Advanced Network. [S.l.:s.n.], 2015: 1-6.
- [2] 颜永明, 左良, 许斌, 等. LSN DCI EVPN VxLAN 组网架构研究及实现[J]. 电信科学, 2017, 33(6): 171-178.
YAN Y M, ZUO L, XU B, et al. LSN DCI research and implementation of EVPN VxLAN network architecture[J]. Telecommunications Science, 2017, 33(6): 171-178.
- [3] 王霞. Openstack 云平台多租户网络隔离证据收集研究与分析[D]. 北京: 北京工业大学, 2017.
WANG X. Research and analysis of forensic collection of multi-tenant network isolation in openstack cloud platform [D]. Beijing: Beijing University of Technology, 2017.
- [4] 柴瑶琳, 穆域博, 马军锋. SD-WAN 关键技术 [J]. 中兴通讯技术, 2019(3): 7-9.
CHAI Y L, MU Y B, MA J F. Key technology in SD-WAN[J]. ZTE Communications, 2019(3): 7-9.
- [5] 鲁子奕, 杨文斌. 运营商 SDN 云网协同架构和关键技术研究 [J]. 中兴通讯技术, 2019(3).
LU Z Y, YANG W B. Infrastructure and key technologies of SDN cloud network for service provider [J]. ZTE Communications, 2019(3).
- [6] IETF. Virtual eXtensible local area network, a framework for overlaying virtualized layer 2 networks over layer 3 networks:RFC 7348[S]. 2019.
- [7] IEFT. MPLS-based Ethernet VPN-EVPN, BGP: RFC7432 [S]. 2019.
- [8] 穆域博, 柴瑶琳, 宋平, 等. SD-WAN 产业发展与关键技术研究 [J]. 信息通信技术与政策, 2019(11).
MU Y B, CHAI Y L, SONG P, et al. Research on industry developments and key technologies of SD-WAN [J]. Information and Communications Technology and Policy, 2019(11).
- [9] 中国互联网协会. 软件定义广域网(SD-WAN)研究报告[R]. 2018.
Internet Society of China. SD-WAN research report [R]. 2018.
- [10] 杨帅. 传统广域网向软件定义广域网(SD-WAN)架构转型浅析[J]. 中国信息化, 2019(5).
YANG S. Elementary analysis of traditional WAN to software defined WAN (SD-WAN) architecture transformation[J]. In-formalization Research, 2019(5).

[作者简介]



牛佳 (1995-), 女, 中国电信股份有限公司上海分公司助理工程师, 主要研究方向为数据网络、SDN 等。



颜永明 (1978-), 男, 中国电信股份有限公司上海分公司教授级高级工程师、综合运营监控中心经理, 主要研究方向为数据网络、SDN 等。



林志华 (1981-), 男, 中国电信股份有限公司上海分公司工程师, 主要研究方向为数据网络、SDN 等。