

# MPLS 助“推”新一代互联网

爱立信（中国）通信有限公司 郭骏文

**摘要：**本文主要介绍了 MPLS 在 IPv6 网络中的运用及产品方案。首先简单分析了 IPv4/IPv6 混合组网时 MPLS VPN 的技术，包括 MPLS Tunneling、VPN over GRE、Carrier of Carrier VPN、L2 VPN 等；然后阐述了 MPLS DiffServ-Aware TE 的 QoS 方案；在最后部分介绍了爱立信公司的部分产品定位及应用。

**关键词：**IPv6 NGI MPLS VPN GRE 隧道技术 下一代网络 协议转换 L2TP

## 1. 引言

提起 IPv6，早年《数字财富》的一篇文章把这个技术称为“推拉之间的 IPv6”，“推拉”两字把当时其尴尬的地位描述得淋漓尽致。一方面，快速膨胀的互联网络、移动网络的庞大用户即将耗尽现有的 IP 地址，市场需求在“推”着它走；另一方面，发展了 20 多年的 IPv4 还在不断扩大它的市场份额，标准迁移带来的巨大成本和费用，新标准的成熟周期难以估量，这些又在“拉”着它不让前进。

但是，令人高兴的是，在今天，在多方面的努力下，笼罩着 IPv6 的迷雾逐渐散去，“推”的力量开始加速，下一代互联网的到来眨眼间已经是板上钉钉的事情了。

爱立信公司作为 IPv6 的主要倡导者，一直为推广和制定 IPv6 的相关标准，为建设下一代网络作着不懈的努力。两年多前的这一幕，便很好地体现了这一点：

2003 年 1 月，在斯德哥尔摩的一辆救护车上，一套名为“守护天使”的紧急医疗救护系统和爱立信公司的通讯设备在忙碌地工作着。在这些设备的协同工作下，大量的医疗数据和声音图像顺利地救护车和医疗中心之间进行传递……

这是爱立信公司利用 IPv6 技术，进行移动系统无缝漫游的一次演示。它成功地把下一代互联网的关键技术—IPv6 和 2G、3G、无线局域网、广域网结合在一起，完成了医疗多媒体信息的交互。

随着欧洲、美洲和亚洲的 IPv6 用户不断增长，中国政府八个部委也共同出台了 CNGI

项目来推动下一代互联网的发展。在这些项目里,爱立信公司更是积极参与了中国移动、电信、网通、铁通等运营商的网络建设,为其出谋划策,添砖加瓦,贡献自己的微薄之力。

因此,本文将根据在 CNGI 项目中的一些成功经验,介绍在下一代互联网中 MPLS 技术应用及产品方案。

## 2. IPv4/v6 的 MPLS VPN 混合组网

当然,在“推拉”之间,IPv6 不可能在一夜之间便替代了 IPv4,在未来很长的时间内,两个协议将共同存在、相互兼容。因此,如何使 IPv6/v4 和谐共存和发挥下一代互联网的优势一直是网络建设的难点。

而 MPLS 作为一种支持多协议的技术,其对二层、三层协议良好的支持,非常自然地成为 IPv6/v4 共存的桥梁。在这里,从这座桥梁开始,我们来了解一下在混合组网中的 MPLS VPN。

从原理上看,MPLS VPN 技术也是众多的隧道共存技术里的一种。当然,这种技术也有它适用的领域,它非常适合 IP 骨干网和城域核心网。

我们知道,大部分的隧道技术都是为了解决 IPv6 孤岛之间的互联问题,但是,这还远远不够;对于一个电信运营商,除了解决自身的互联问题外,还应该向广大的客户提供多种业务,其中包括目前比较热门的 VPN 业务。从这两方面出发,我们逐一分析 MPLS VPN 技术如何运用于 v4/v6 网络。

### 2.1 简单的 MPLS Tunneling

在目前 IPv4 的庞大网络里面,把 IPv6 孤岛连接起来,要解决两个问题:一、数据层面的传递和转发;二、路由信息的传递。MPLS Tunneling 对数据的转发当然是通过 LSP 路径,所以它需要经过的节点支持 MPLS;IPv6 路由信息的交换由 MPBGP 来完成,所以在其它地方也称为 BGP 隧道技术(或 6PE)。如下图所示:

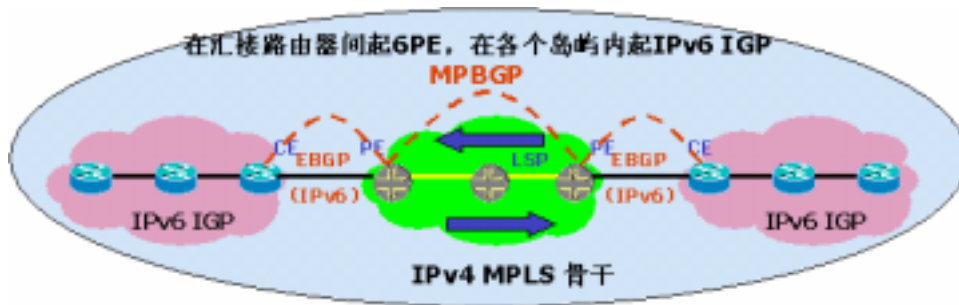


图 1 MPLS Tunneling 示意图

这里需要几个条件来满足 Tunneling 的要求：

- 两端的 PE ( Provide Edge ) 支持 v4/v6 双栈协议
- PE 与 CE ( Customer Edge ) 之间采用 EBGP 或 IGP 来交互 v6 路由
- 在一个 IPv6 的孤岛里，需要把 EBGP 收到的路由分发给所有的节点，那么可以采用 BGP 路由注入 IGP 的办法，也可以采用 FULL MESH IBGP 的方案
- 骨干网络所有节点支持 MPLS，运行 LDP 或 RSVP 协议

这种互联的方式简单明了，适合纯 IPv6 的孤岛进行通信。在整个通信的过程中，只需要在 PE 处封装两层 MPLS Label，然后在远端 PE 弹出 Label，并把 IPv6 的数据包进行 v6 的路由表查询，然后转发到相应的 CE。

它的优势比较明显，只要 PE 支持双栈，对 v6 网络里的路由器、主机均无特殊要求。但也有不足的地方，同一个 IPv4 MPLS 的骨干网络里只能支持这样的一个 IPv6 网络，如果还需要承载其它用户的 v6 网络并隔离它们间的路由信息，那这需要 v6 的 VPN 来完成；另外，如果这些 v6 的网络还向外提供 VPN 业务（对于 ISP 来说，这比较常见），那么要求采用 Carrier of Carrier 的技术。

## 2.2 引入 GRE tunnel

GRE ( Generic Routing Encapsulation ) 已经是运用广泛的隧道技术，把它和 MPLS 有机的结合起来，有时候会收到意想不到的效果。由于它对设备的要求较低，并且没有兼容性的问题，所以在实用性很强。

通过 GRE 或 IP-in-IP 的隧道把独立开来的一个个 v6 网络连接起来，形成连通的整体网络。由于这些隧道是点到点之间的连接，我们可以根据具体情况建立部分或者 FULL MESH 的连接关系。

同时，GRE 隧道支持的协议非常广泛，包括 IPv4、IPv6、ISO 和 MPLS 等。因此，可以在上面运行 OSPF、ISIS 路由协议，再结合 MPLS 及其相关的 LDP、RSVP 协议，便组成各种解决方案。

如图 2，通过 GRE 的连接，再运行 v6 IGP，IPv6 孤岛连成一片，完全可以形成一个独立的 AS：

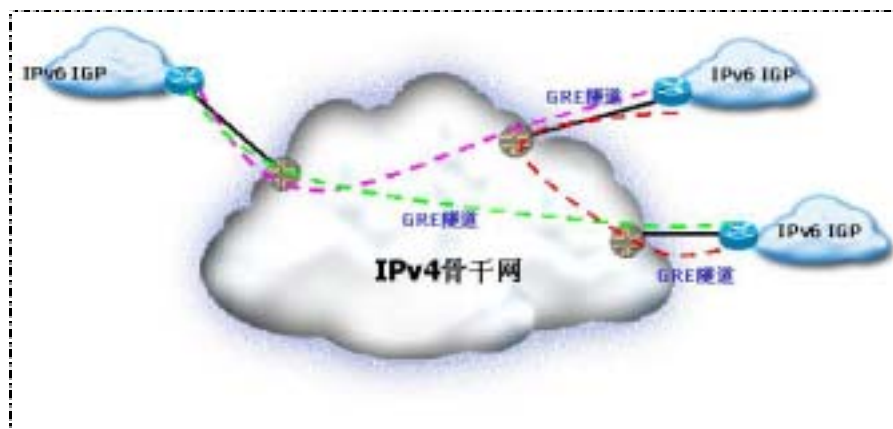


图 2 GRE Tunnel

仅仅是这样的连接，其效果其实与前面的 MPLS Tunneling 差别不大。如果我们在图 2 各个孤岛网络上再启用 IPv4 和 MPLS 的功能，就可以构成一个完整的 IPv4/v6 MPLS 的网络，这个网络便可以向客户提供 VPN 业务了。

在这个进一步优化的方案里：

- 对纯 IPv4 骨干网没有特殊要求；
- 在 v6 的孤岛里同时启用 IPv4/v6 双栈协议，运行 v4/v6 的 IGP，保证这些网络的 IGP 连通性；
- 在所有孤岛里启用 MPLS，并且需要 LDP 或 RSVP 建立相应的 LSP 路径，这些 LSP 有可能在 GRE 上面；
- 如图 3 里，与 IPv4 骨干网连接的 P 路由器，不但启用 IPv4/v6 双栈、MPLS，

还要求设备能够封装 v4/v6 和 MPLS 的数据包到 GRE Tunnel 里；

- 边缘的 PE 可以通过 MP-IBGP 部署 MPLS VPN。

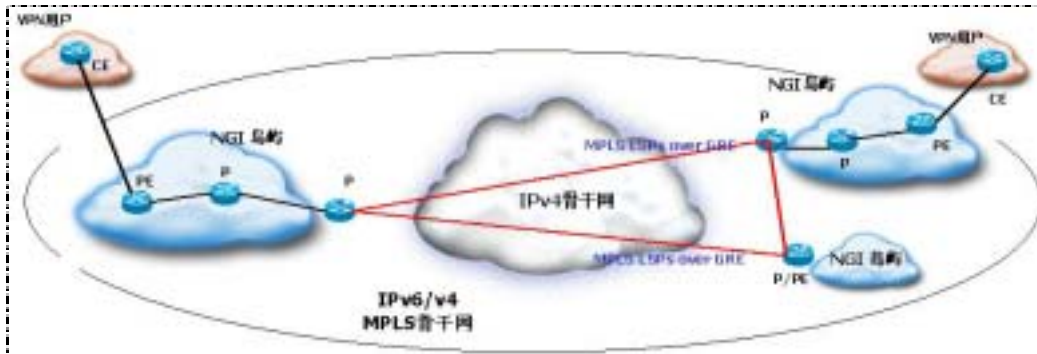


图 3 MPLS VPN over GRE

这样，这些原是 IPv6 的孤岛连接成为 IPv4/v6 的 MPLS 骨干网，一方面这些孤岛间可以进行 IPv6 的通信；另一方面，它们还可以向客户提供 VPLS、L2 VPN 和 L3 VPN（包括 IPv6 和 IPv4）。显然，它比单纯的 MPLS Tunneling 有更强的业务能力。

一般来说，MPLS VPN over GRE 要求孤岛里的路由器支持 MPLS 和双栈，这个要求比较容易满足。因为在运营商的新建网络里，大部分设备对这些协议在功能上都是支持的，只是性能上有所区别而已。

但是，这种结构对与 IPv4 骨干网连接的路由器（图 3 里发起 GRE tunnel 的 P 路由器）要求比较苛刻。它不但要支持以上提到的各种协议，而且负责所有数据包的 GRE 封装和解封装，负担非常重。如果仅利用路由器里的 CPU 来进行处理是基本不可行的，因此要添加专用的硬件加速模块来解决性能的不足。

### 2.3 Carrier of Carrier VPN

这已经不是什么新技术了，在 MPLS VPN 里，这是一种相对成熟的方案，我们希望在集成 IPv6/v4 的网络中，它能够为业界提供运营级的解决办法。

简而言之，IPv6 网络将作为 IPv4 骨干网的一个 VPN 客户网络，而在这个 IPv6 的 VPN 网络里还可以向外提供 VPN 业务。

这个方案主要采用 MPLS 封装来转发数据，用 MPBGP 来传递路由表。所以它要求

IPv4 骨干网和 IPv6 网络的设备都要支持 MPLS，在这点上它与前面的 MPLS VPN over GRE 的方案有所不同。

在 IPv4 骨干网络里，只需要支持普通 MPLS VPN 的功能即可，它透明地把孤岛网络的数据和路由信息传递过去，甚至其 PE 都不需要支持双栈协议。

由于采用 Carrier of Carrier 的方式，v6 的孤岛网络需要支持 v4 IGP 和 LDP 协议，由这些孤岛网络把所有 IPv4 IGP 路由（路由器的互联和环回接口网段）加上 MPLS Label 发送给 IPv4 骨干网络，由 IPv4 骨干网络透明传递到远端的孤岛网络。这样，每个孤岛网络的 CE 路由器都会收到各个网络的 IPv4 IGP 路由（带 MPLS Label），它可以通过 IPv4 IGP 重新分发给本岛里的所有路由器。

因此，v6 的孤岛里便有所有网络的 Labeled IPv4 Route，可以建立起到达所有节点的 MPLS LSP 路径，再利用前面的 MPLS tunneling 的技术，把 IPv6 的路由信息传递过来，孤岛之间的 IPv6 就可以畅通无阻了。

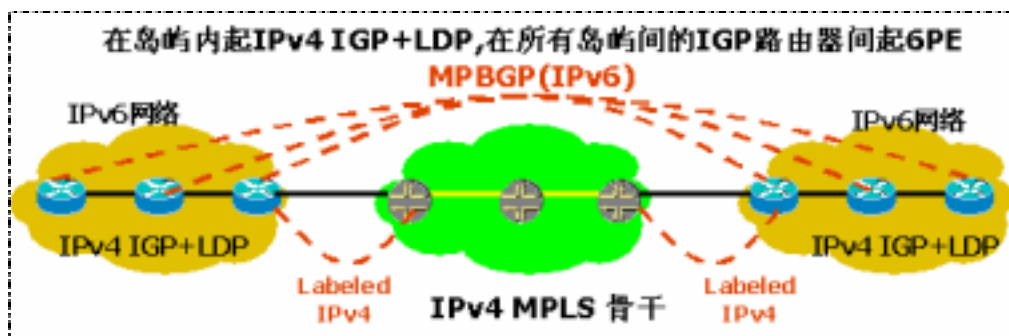


图 4 Labeled IPv4 Route + MPLS tunneling

至此，v6 的孤岛作为 VPN，已经完全连接起来了。那么，这些 v6 的网络作为运营网络，它向客户提供相应的 VPN 业务，也是没有问题的。如图 5 所示，在 PE 之间建立 MP-IBGP 的连接，传递相应 VRF 信息，建立 MPLS VPN。这个 VPN 可以二层 VPN，也可以是三层 VPN；而且，如果采用爱立信公司的路由设备，还可以支持双栈 VPN，也就是说这个 VPN 既支持 IPv4，也同时支持 IPv6，可以非常灵活地开展业务。

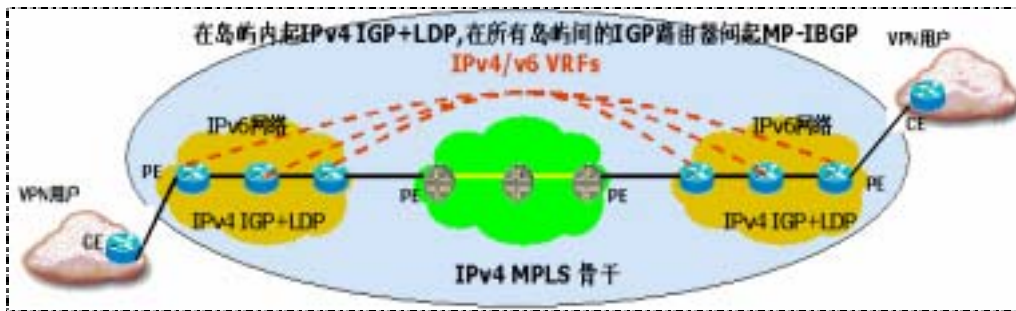


图 5 Carrier of Carrier MPLS VPN

由于在整个方案中嵌套了两层 PE，MPLS 标签堆栈中最多时存在 3 个标签：IPv4 MPLS 骨干网的 LSP Label、IPv6 网络里的 LSP Label、最终 VPN Label。这些 Label 的封装和解封装分别在这些 PE 路由器上完成。

Carrier of Carrier 的 MPLS VPN 这种网络共存方案，从性能上看是比较理想的，扩展性很好，IPv4 MPLS 骨干网可以为多个 IPv6 网络服务，IPv6 网络也可以同时为多个 IPv4/v6 的网络创建 VPN。虽然，它要求 IPv4/v6 的网络都要支持 MPLS，但这在如今的运营网络里不是很困难的事情。因此，这是一个值得推广的 VPN 解决方案。

## 2.4 二层 VPN

这个方案与 GRE Tunnel 非常类似，其思路是通过 MPLS L2VPN 模拟出二层的链路，以供 VPN 使用，不管这个 VPN 是 IPv6 的还是 IPv4 的网络。

它要求 IPv4 骨干网络支持 L2 VPN。目前普及面最广的是 Martini-L2-VPN，它依靠 LDP LSP 在两个 PE 间传递二层链路数据帧，配置相对简单。

其优点是对 CE 没有特殊要求，CE 可以支持 IPv6/IPv4/MPLS 的任何一种。它的主要问题是，需要互联的 CE 间二层链路类型是相同的，如 FR 和 FR，VLAN 和 VLAN，但在现实网络里，很难做到全网完全统一。

## 3. MPLS DiffServ-aware 流量工程

如何提高服务质量，是互联网的一个棘手但又必须解决的问题。作为下一代的互联

网络，而且长期是 IPv4 和 IPv6 混合组网的状况，我们没有理由忽视 QoS 的实施：一、保证 IPv6 流量与 IPv4 流量各自的正常运行；二、在未来给客户id提供端到端的服务质量。

MPLS DiffServ-Aware TE 的出现，给我们展现了一个新的方向。再结合 MPLS VPN、MPLS FRR 等技术，它非常有可能为 IPv4/v6 共存网络，甚至下一代网络提供端到端的 QoS 方案。

为了解决在 IP 互联网络上承载语音、视频等实时业务，IETF 组织在 1998 年定义了基于 DSCP 的 DiffServ 解决方案，它是一个基于“类”的 QoS 技术，主要是针对在骨干网上部署的 QoS 解决方案。DiffServ 虽然具备了良好的扩展性，但缺乏了有效的端到端部署机制，运营商必须在域中根据分组的 DSCP 定义所有节点设备。另外一个问题是 DiffServ 简单地把各种业务用简单的几个“类”来定义，但往往各种业务的流量模型和业务模型不尽相同，把各种业务叠加在一起后，其流量模型和业务模型将会是非常复杂。这些问题都使 DiffServ 至今都没有被运营商在骨干网上广泛采用。

MPLS TE (MPLS 流量工程) 撇开了使用传统利用 SPF 计算方式建立路由而是通过 RSVP (资源预留) 协议建立 LSP，使网络流量根据带宽需求合理地被引导，达到类似传统利用 ATM 技术实现流量工程的效果。但 MPLS TE 并没有定义服务类型，即没有解决如何在 LSP 中传送 EF、AF 等不同等级的业务，这也成为了 MPLS TE 最大的一个缺陷。

由于 DiffServ 和 MPLS TE 都存在不完善的地方，因此 2002 年业界提出了一种 MPLS DiffServ-Aware TE (简称 MPLS DS - TE) 的解决方案，它在原有的 MPLS TE 技术基础上增加了基于类别区分的功能。

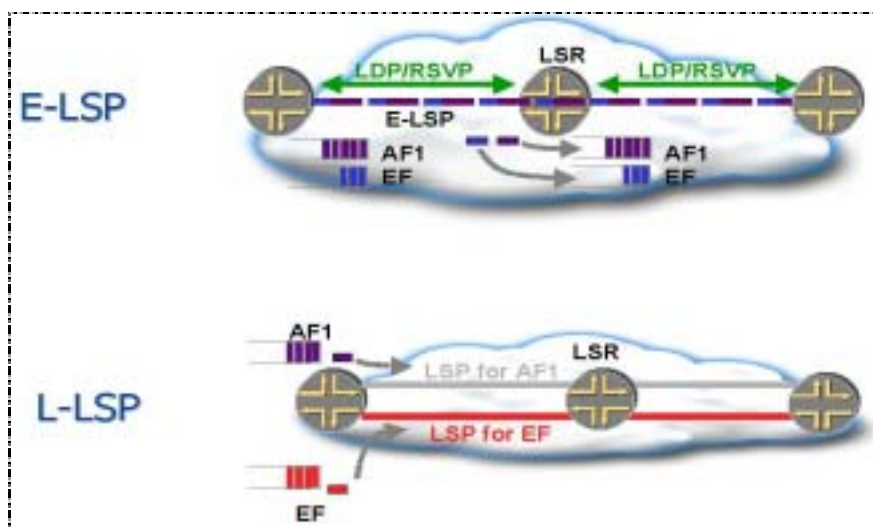


图6 E-LSP and L-LSP 原理

在 MPLS DS-TE 中,有两种方法可以确保业务类型信息映射到特定 DiffServ-TE LSP 的流量,映射到适当调度队列中。第一种方法是在 LSP 入口适当设置 EXP 位 (E-LSP), 通过 EXP 确定 PHB 的调度行为,因此,该解决方案适用于少于 8 个 PHB 的网络。第二种方法是通过 L-LSP 中的标签来映射调度行为,并使用 EXP 位来决定流量的丢弃优先级。与该 L-LSP 相关的调度行为在 LSP 建立时由信令协议(如 RSVP 等)进行传递。这种方法将支持任意数量的 PHB。图 6 分别说明了 MPLS DS - TE 的 E-LSP 和 L-LSP 实现原理。

通过这些方法,不同的 LSP 对应了不同的业务等级,也就对应了不同的服务质量;而且各种 MPLS 的应用如 VPN、CCC、FRR 等均基于 LSP 来开展。因此把它们结合起来,便非常容易地确保了不同业务执行不同的服务质量。

可以预料,随着 MPLS 在下一代互联网中的普遍应用,MPLS DiffServ-Aware TE 将成为我们手中的 QoS 利器。

#### 4. 专注的 IPv6 产品方案

所谓巧妇难为无米之炊,前面介绍的 MPLS/IPv6 技术方案,同样需要强有力的产品来支持。为此,爱立信公司自从 1995 年率先发布 IPv6 商用路由器以来,一直致力于为业界提供端到端的 MPLS/IPv6 产品解决方案。

## 4.1 核心路由平台

近几年，爱立信和 Juniper 公司共同推出了整个 IPv6 核心网络和骨干网络的解决方案，这其中包括大家并不陌生的 T640/T320/M320/M40e 等路由产品。这些设备对 IPv6 和 MPLS 的良好支持，使它们适合运行在骨干网络上，可以十分出色地扮演 MPLS VPN 里 P 或 PE 的角色。

这些路由设备在系统设计开发和软件设计开发中，就已经把 IPv6/MPLS 的功能内置。随着新产品的不断推出，目前，爱立信/Juniper 的全系列产品一视同仁地支持所有运营商网络所需要的 IPv6 功能，并与 IPv4 一样，采用硬件来支持 IPv6 的转发和进行路由选择，因此其 IPv6 性能与 IPv4 没有任何区别，都能以线速转发数据包。

我们的 IPv6 核心路由器采用专门的 ASIC 芯片来支持 IPv6 网络，使得核心路由器在支持多种 IPv6 功能和大量安全过滤功能的情况下，各种速率接口，从 2Mbps 到 10Gbps，都能以线速转发 IPv6 数据包，或者 IPv6/IPv4 混合数据包，并且具有非常小的延迟 (<150us)。目前，爱立信/Juniper 路由器支持所有运营商网络所需的路由、寻址、传输、安全、应用等协议；同时，随着 IPv6 技术和标准的发展，爱立信/Juniper 将及时支持最新的功能和协议，只需要对软件版本进行升级即可。

## 4.2 接入核心 E320

面对 IPv4/IPv6 共存的网络，接入时可以利用 GRE、MPLS 等隧道接入技术；但是，在网络的边缘，我们认为利用双栈 BRAS 或者 L2TP Tunnel 的接入方式更加利于未来网络的迁移，并与现有接入方式保持一致性：

- 双栈 BRAS。BRAS 同时支持 IPv6 和 IPv4 的宽带接入，并把这两部分流量分别转发至不同的网络，开通简单，管理也比较方便。
- L2TP Tunnel。双栈或者 IPv6 的用户通过 PPP 连接到 LAC 服务器，然后 LAC 建立到双栈 LNS 的 L2TP 隧道，把用户 PPP 会话终结在 LNS 服务器。这样，用户和双栈 LNS 之间建立起 IPv6 的 PPP 会话，完成 IPv6 的接入。

这些 GRE、MPLS、双栈 BRAS 和 L2TP 等接入技术，在我们的新款产品 E320 上已

经得到非常完美的实现了。E320 结合了 IPv6、IPv4 及 MPLS 的互联技术，融合了原有的 T/M 系列和 ERX 系列产品的优点，将成为下一代互联网主流的接入设备。

这个产品的交换能力高达 320G，完全可以满足中国互联网高密度接入的需求；并且它优化了整个体系结构，提高板卡的重用性和灵活性，进一步保护了运营商的设备投资。



图 7 新一代接入核心 E320

E320 和其它 ERX 系列路由器一样，在创新的可编程 ASIC 体系结构的支持下，都是在运营商边缘部署 IPv6/v4 双栈业务的理想平台，可支持运营商级的路由选择和宽带用户管理。此外，利用 SDX-300 业务部署系统，服务供应商还可以进一步提高控制能力，以便迅速地为用户定义并激活基于策略的 IP 业务。

### 4.3 协议转换网关

为了实现 IPv4 和 IPv6 网络的相互通讯，我们建议在骨干网层面部署协议转换网关（NAT-PAT）来完成。它对 IPv4 和 IPv6 的数据包进行 NAT 转换，这些转换包括对 IP 包头以及有效负载的翻译。对于一些内嵌地址信息的高层协议（如 FTP、DNS 等），协议转换网关需要运用相应的应用层网关（FTP ALG、DNS ALG 等）协议来完成翻译。

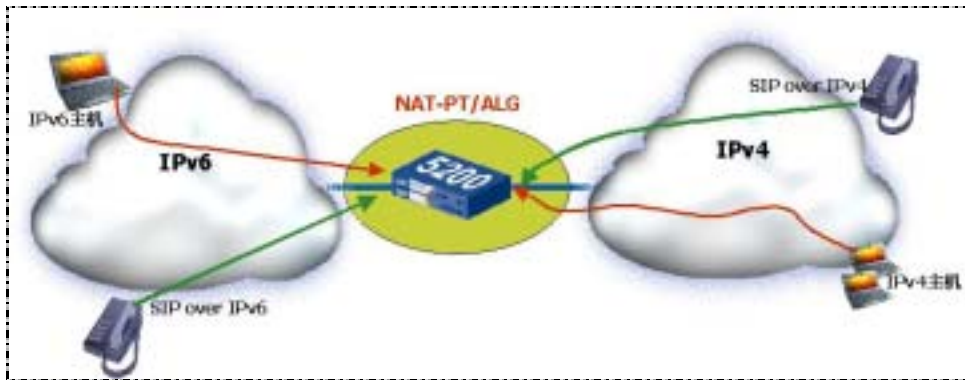


图 8 协议转换网关

为此，爱立信公司采用 NS5200/NS5400 设备进行协议转换。这些设备可以无阻塞地完成 4G 到 12G 的流量转换，完全满足当今一个大型网络互访流量的需求。

而且，该设备结合 NAT-PT 功能与防火墙功能一起来做流量过滤，可以充分利用防火墙的安全策略，严格限制外部主机对内部网络的访问。通过它们，顺利过滤掉目前在互联网里横行的网络病毒和攻击，达到保护 IPv6 网络的目的。

## 5. 结束语

下一代互联网涉及的技术、产品领域非常多，其中包括了路由协议、地址规划、QoS、VPN、安全等等，本文仅简单介绍了 MPLS VPN、DS-TE 的技术和部分产品方案，并且大部分观点均以建设运营网络为出发点，不免有很大的局限性，希望大家进一步斟酌。

另外，这些方案都有不同的优缺点，不同的适用空间，用户可以根据自己的具体情况进行具体的部署。