基于MPLS/BGP电子政务网L2VPN应用与实现

沈俊鑫[1]，沈 梅[2]
（1. 云南财经大学，云南 昆明 650208；2. 云南农业大学，云南 昆明 650201）

摘要：各厅局通过ATM，F/R，VPDN，Ethernet等方式接入电子政务网，通过分析比较各种基于MPLS的L2VPN技术，构建基于MPLS/BGP的L2VPN，分析改进的Kompella方式的实现过程和原理，实现Kompella和Martini方式的互联互通问题。

关键词：虚拟专用网；多协议标签交换；二层虚拟网；Kompell，Martini；CCC（电路交叉连接）；隧道组件；协议结构
中图分类号：TP 311 **文献标识码**：A **文章编号**：1004－390X（2008）01－0126－04

Implement and Application of L2VPN Based on MPLS/BGP E-Gov

SHEN Jun-xin[1]，SHEN Mei[2]
（1. Yunnan University of Finance and Economics，Kunming 650208，China；
2. Yunnan Agricultural University，Kunming 650224，China）

Abstract：The bureaus connect to the E－Governmental network platform via ATM/PVC，FR/PVC，VPDN or ETHERNET. In this article，three MPLS L2VPN technologies which were based on MPLS/BGP were compared and the principles of improved Kompella were analyzed to implement the connection between Kompella and Martini.

Key words：VPN；MPLS；L2VPN；Kompella；Martini；CCC（Circuit Cross Connect）；tunnel component；protocol structure

虚拟私有网络VPN（Virtual Private Network）基于隧道技术在公网上创建私有网络，将企业网的数据封装在隧道中进行传输。隧道协议可分为第2层隧道协议PPTP，L2F，L2TP和第3层隧道协议GRE，IPsec，通过某种格式对被传输数据包进行封装，例如L2TP封装后的数据格式：IP＋UPD＋L2TP＋PPP，即在原来传输协议（IP＋UPD）和数据（PPP Data）中间加入封装协议L2TP；GRE则在网络层进行封装，利用一种网络层协议封装另一种网络层协议，而IPSec则是在IP包头和IP数据净额中加入AH和ESP包头，实现对数据包的封装。

MPLS VPN基于标签交换的IP VPN，通过在IP包中嵌入Label的方式，利用标签进行IP包分发，通过LSP建立隧道，PE路由器根据FEC的虚拟路由表将IP包转发到相应的CE路由器[1]。MPLS L3VPN以MPLS/BGP为主，MPLS L2VPN主要基于VPWS和VPLS标准，其中VPWS标准采用Martini，Kompella方式，VPLS标准主要有CCC方式[2]。Martini方式使用LDP方式传递VC信息[3]，使用两层标签，第1层使用LDP建立Tunnel连接，第2层封装VC标签，因此L2VPN

收稿日期：2007－02－26 修回日期：2007－04－29
作者简介：沈俊鑫（1978－），男，福建漳州人，硕士，高级工程师，主要从事电子政务、网络工程等方面的研究。
E－mail：shenjunxin@ynnic.gov.cn

间可以共享 LSP[4]。Kompella 方式类似于 MPLS/BGP VPN，因此可以实现拓扑发现，组网灵活、部分解决了 n 方问题和跨域问题[5]；CCC 方式通过在 PE-CE 之间配置静态、透明隧道，只使用一层标签，单独使用 LSP[3]，这种方式配置不灵活。

云南省电子政务网在独立的 ATM 基础网上构建了 MPLS/BGP（如图 1 所示）。网络提供 Ethernet，FR，ATM 以及 VPDN 等接入方式，MPLS/BGP 支持 Full-Mesh，Hub-Spoke，Overlapping VPNs 等方式[7]。然而，现有 MPLS/BGP 虽在 Qos 上得到提高，但是无法满足 TE 方面的要求，同时，随着 VPN 数量的增加，PE 路由器的开销大，特别是内存开销（每个 VRF 在独立的内存空间中），VPN 路由的发现收敛慢等问题不断暴露。MPLS L2VPN 对 PE 设备的内存开销小，通过在 CE 之间直接扩展 VPN 路由，收敛快。

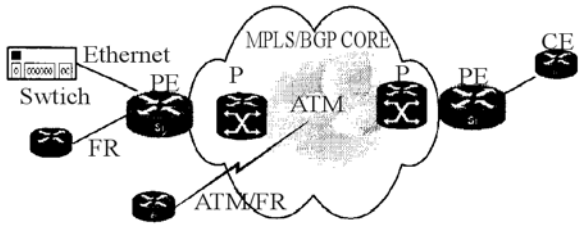

图 1　电子政务MPLS/BGP骨干网
Fig. 1　E-Gov MPLS/BGP core

1　系统设计与实现

由于各种 L2VPN 标准都是属于 Draft 阶段[8]，各设备厂商在实现方面也有所区别，因此系统除了考虑可扩展性、经济性、可管理性外，更重要的是在现有 MPLS/BGP VPN 基础上能实现不同设备协议间的兼容与互通性。

1.1　系统设计

PE 路由器上，需要对协议、报文结构进行重新设计，同时扩展 BGP NLRI。

1.1.1　协议结构

为了实现 L2VPN 功能，PE 路由器需要在协议结构上增加以下组件，分别实现连接控制、建立隧道、管理 VC 以及 VC 封装（如图 2 所示）。各组件需要分别实现以下功能：

（1）连接控制（Control Connection）模块：支持 LDP、MPLS/BGP、静态 LSP 方式实现 VC-Label 的协商、拆除及错误。

（2）传输组件（Transport Component）：支持 MPLS 标签或 GRE 隧道方式在入口和出口 PE 处获取 PDU。

（3）隧道组件（Tunnel Component）：实现通过 MPLS 的 Label 在 tunnel 上建立 VC Label（CCC 是静态 LSP）。

（4）二层协议数据单元（L2 PDU）：使用控制字仿真 VC 封装，32 位控制字。

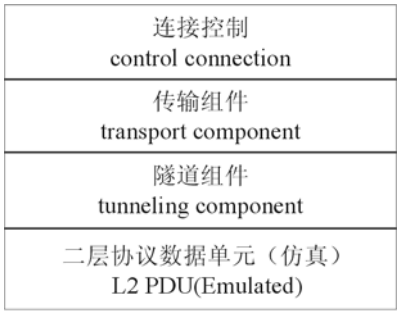

图 2　功能结构
Fig. 2　Function architecture

1.1.2　VC-Label 建立过程

VC-Label 是承载在 LDP 标记映射消息的通用标记 TLV 中（其值为 0x80）[8]，而 VC 的相关信息通过一个 LDP FEC 元素来承载。具体过程如下：

（1）入口 PE 启用 LDP 协议，利用第一层标签建立起 LSP：MPLS L2VPN 利用两层标签分别完成隧道建立和 VC 标记，第一层标签（外层标签）通过 LDP 建立入口 PE 到出口 PE 间的隧道 LSP（与 MPLS/BGP LSP 建立相同）。

（2）利用第二层标签实现 LDP 标记映射互换。

①首先入口 PE 为新接口分配 VC-Label，并绑定到配置的 VCID 中。

②入口 PE 向出口 PE 发送标签信息（包括 VC FEC TLV 和 VC Label TLV）。

③出口 PE 接受到标签信息后映射到本地的 VCID 中。

④出口 PE 以相同的过程向入口 PE 发送标签信息，完成出口 PE 和入口 PE 的 VC-Label 的相互学习。

⑤这样，对于 PE 路由器的 VCID，都保存了本地 Label 号和对端 Label 号，对于来自特定 PVC 的数据包都会在数据包增加对端 Label 号，然后转发。即新数据包增加了 VC Label 信息：Tunnel Label + VC Label + PDU。

（3）数据传输时，对端 PE 从收到的 Label-

Mapping 消息中提取 VC-ID，若与本地存储的 VC-ID 相匹配，则进行转发，同时把 VC Label 做为反方向传输的二层 PDUs 的内层标签；若 VC-ID 不匹配则忽略。

1.1.3 报文结构

根据协议结构，很容易得知经过 PE 路由器上的二层报文头需要包含 Tunnel，VC 以及控制字等信息。具体结构如图 3 所示：

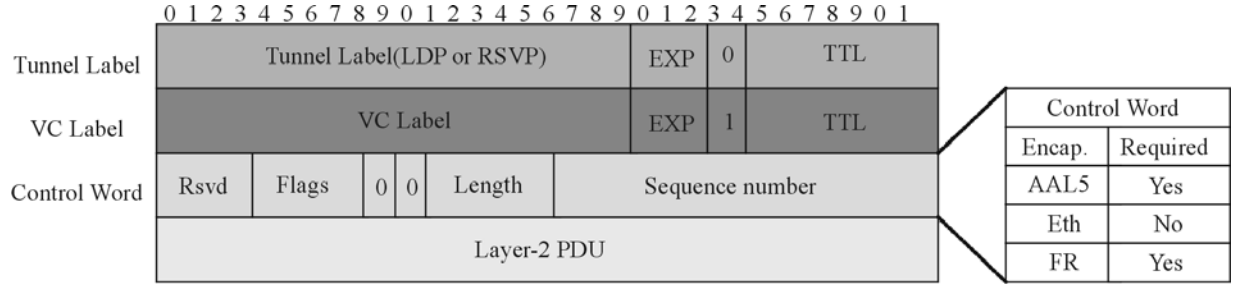

图 3 报文结构格式

Fig. 3 Generic packet format

1.1.4 主要代码

```
#include " l2vpn. h"
/ * 主要常量： * /
#define LDP_ TLV_ FEC 0x0100
#define LDP_ TLV_ LABEL_ REQUEST_ MSG_ ID 0x0600
#define LDP_ FEC_ KOMPELLA_ VC0x70 [6] / * KOMPELLA _ VC 在 LDP FEC 元素的 TLV 值 * /
#define LDP_ FEC_ MARTINI_ VC 0x80 [3] / * MARTINI _ VC 在 LDP FEC 元素的 TLV 值 * /
#define VC_ TYPE_ VALUES…
/ * 主要数据结构：（通过扩展 LDP 实现，通过 LDP FEC 来携带 VC 信息） * /
static const struct tok ldp _ fec _ values [ ] = {…}; / * ldp_ fec 值 * /
static const struct tok tunnel_ label_ values [ ] = {…}; / * tunnel_ label 值 * /
static const struct tok vc _ label _ values [ ] = {…}; / * vc_ labe 值：如 Martini VC 和 Kompella_ VC * /
/ * 产生封装 Tunnel Lable + VC Label 主要程序 * /
static int decode_ labeled_ vpn_ l2 (const u_ char * pptr, char * buf, u_ int buflen)
{/ * * pptr 指针存储 TLV 类型 * /
…
strlen = snprintf (buf, buflen," RD：%s, CE-ID：%u, Label-Block Offset：%u, Label Base %u",
bgp_ vpn_ rd_ print (pptr),
EXTRACT_ 16BITS (pptr + 8),
EXTRACT_ 16BITS (pptr + 10),
EXTRACT_ 24BITS (pptr + 12) > > 4); / * the label is offsetted by 4 bits so lets shift it right * /
while (tlen > 0) {/ * 获取 TLVs * /
tlv_ type = * pptr + +;
tlv_ len = EXTRACT_ 16BITS (pptr);
ttlv_ len = tlv_ len;
pptr + = 2;
switch (tlv_ type) {/ * 根据 TLV 类型计算 tlv_ len * /
case 1：…
case 2：…
default：
snprintf (buf + strlen, buflen-strlen," \ n \ t \ tunknown TLV #%u, length：%u",
tlv_ type,
tlv_ len);
break;
}
tlen - = (tlv_ len < < 3); / * the tlv-length is expressed in bits so lets shift it tright * /
}
return plen + 2;
}
```

1.2 协议处理过程

在 MPLS/BGP VPN 中，P 路由器只维护到达

PE 路由器的路由，负责 MPLS 标签转发，不需要参与 VPN 路由，即不维护 VRF，同时在数据转发时不需要查找 IP 路由，而采用标签交换即可[1]。在 L2VPN 中，PE 的协议处理过程如下：

（1）入口 PE 接收到来自 CE 的 L2 数据帧。

（2）如果到目的出口 PE 上还没有建立 LDP 或 MPLS/BGP，则建立 LDP 或 MPLS/BGP。

（3）入口 PE 在 Tunnel 上分配 VC Label，绑定 VC-ID 到相应的子接口上；PEi 用 PEj 标签块的首值 + i 来作为 PEj 报文的 VC Label，例如 PE1 的标签首值是 1000，PE2 的标签首值是 2000，那么 PE1 发给 PE2 的 VC Label 就是 2001，而 PE2 发给 PE1 的 VC Label 就是 1002。

（4）入口 PE 发送已经绑定的信息（包括 VC FEC TLV 和 VC Label TLV）到出口 PE；信息结构如下：Tunnel Label + VC Label + L2 PDU

（5）出口 PE 匹配入口 PE 发送已经绑定的信息，建立匹配接口。

1.3　PE 层次结构的改进

在 MPLS/BGP VPN 基础上提供 L2VPN 服务，对 PE 路由器性能要求比较高，不仅要维护 MPLS/BGP VPN 路由即 VRF，LSP 标签，而且还要维护 L2 VPN 的 Tunnel 标签和 VC 标签，信令开销大[9]。因此实际部署时应该考虑采用分层 PE 方式，里层 PE 除连接外层 PE 只实现 3 层 VPN，而外层 PE 只实现 L2VPN。

2　结束语

2.1　MPLS L2VPN 在电子政务应用情况

云南、黑龙江、陕西、江西、河南、安徽等省电子政务网络都已部署了基于 MPLS/BGP VPN。在云南省电子政务网络，已成功部署了基于 MPLS/BGP 骨干网 L2VPN，为部分“金”字号工程（例如检查系统、工商系统、卫生、金保）提供二层的 VPN 服务。网络运行情况正常，MPLS/BGP L3VPN 相比，省、州市、县三级 PE 路由器的内存、信令协议开销明显下降，三层路由的收敛明显加快，但是采用 MPLS L2VPN 接入的单位，需要自己维护路由（例如云南省金盾工程）。

同时，准备在 MPLS L2VPN 网络上部署、运行 IPv6，方案正在规划中。

2.2　MPLS L2VPN 不足

由于 L2VPN 各标准仍处于 Draft 阶段，仍在不断的完善当中，因此大部分情况都是只实现 Draft 文档的部分功能[9]。

L2VPN 是点对点方式，而且在 PE 路由器配置过程需要大量的手工配置，因此整个网络系统对于二层 Tunnel 和 VC 的管理显得尤为复杂。

和 MPLS/BGP 三层 VPN 不同，L2VPN 不支持 VPN 嵌套，即无法支持 Hub-Spoke、Full-Mesh 以及 Overlapping 等 VPN。

[参考文献]

[1] 沈俊鑫. MPLS VPN 技术及产品探究 [A]. 崔汝贤编. 云南发展与改革论文集 [C]. 昆明：云南大学出版社，2006.

[2] Marc Lasserre [EB/OL]. http://www.ietf.org/, draft-ietf-l2vpn-vpls-ldp. IETF. 2006.

[3] Kireeti Kompella [EB/OL]. http://www.ietf.org/, draft - martini - l2circuit - trans - mpls - 19. IETF. 2006.

[4] Luca Martini [EB/OL]. http://www.ietf.org/, Draft - martini - l2circuit - encap - mpls - 12. IETF. 2006.

[5] Kireeti Kompella [EB/OL]. http://www.ietf.org/, Draft - kompella - ppvpn - l2vpn - 02. IETF. 2003.

[6] W. Augustyn [EB/OL]. http://www.ietf.org/, draft - ietf - ppvpn - l2 - framework. IETF. 2003.

[7] 沈俊鑫. MPLS/VPN 与 IPSec VPN 融合在电子政务中的应用 [A]. 崔汝贤编. 云南发展与改革论文集 [C]. 昆明：云南大学出版社，2006.

[8] VAN PEPLIJAK. MPLS 和 VPN 体系结构 CCIP 版 [M]. 北京：人民邮电出版社，2003.

[9] CISCO SYSTEM. Layer 2 VPN Architectures [M]. USA：Cisco Press，2005.