

ICS 33 040 40

M 32

YD

中华人民共和国通信行业标准

YD/T 1656-2007

采用边界网关协议多协议扩展（BGP-MP） 的基于 IPv6 骨干网的 IPv4 网络互联 （4 over 6）技术要求

Technical Requirement for Interconnection of IPv4 Network over IPv6
Backbone(4 over 6) Via Border Gateway Protocol Multi-Protocol
Extensions(BGP-MP)

2007-07-20 发布

2007-12-01 实施

中华人民共和国信息产业部 发布

目 次

前 言	II
1 范围	1
2 规范性引用文件	1
3 定义和缩略语	1
4 4 over 6机制概述	2
5 4 over 6数据平面处理	3
6 4 over 6控制平面处理	4
7 4 over 6机制示例	6
8 可扩展性	6
9 双栈接入	7
10 安全考虑	7

前 言

本标准主要依据最新研究成果制定完成，还参考了相关国内外标准。

本标准与《IPv4-IPv6过渡中互联互通技术要求》配套使用。

本标准由中国通信标准化协会提出并归口。

本标准起草单位：清华大学、比威网络技术有限公司、信息产业部电信研究院、华为技术有限公司、中国移动通信集团公司

本标准主要起草人：崔 勇、徐明伟、吴建平、李 星、徐 恪、付立政、何宝宏、蒋林涛、刘宏明、陈洪飞、段晓东

采用边界网关协议多协议扩展（BGP-MP）的 基于 IPv6 骨干网的 IPv4 网络互联（4 over 6）技术要求

1 范围

本标准对在 IPv4 网络向 IPv6 网络过渡中可能出现的基于 IPv6 骨干网进行 IPv4 网络互联互通的过渡技术 4 over 6 机制做了规范性描述。本标准中所涉及的过渡技术主要适用于 IPv6 网络逐步成熟的阶段，尤其是 IPv6 成为大规模骨干网络的情形下，大量 IPv4 接入网络通过大规模 IPv6 骨干网进行互联互通的情况。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

IETF RFC 2473

IPv4-IPv6过渡中互联互通技术要求
IPv6规范的通用分组隧道

3 定义和缩略语

下列定义和缩略语适用于本标准。

3.1 定义

封装（encapsulation）

把一种协议的完整数据分组作为内容数据，完整地放到另一协议数据分组内的过程称为封装。

解封装（decapsulation）

把封装过的数据分组内的内容数据恢复成原来的协议数据分组的过程称为解封装。

节点（node）

实现 IPv4 或 IPv6 的网络设备。

路由器（router）

负责转发最终目标地址不是它本身的数据分组的节点。

4 over 6路由器

支持 4 over 6 功能的路由器。

P路由器（P Router）

传输网中网络服务提供商所拥有的网络服务提供商内部骨干路由器。

PE路由器（PE Router）

传输网中网络服务提供商所拥有的提供网络接入功能的边缘路由器。

入口PE路由器（Ingress PE Router）

按照 4 over 6 封装规则对数据分组进行封装的 4 over 6 路由器。

出口PE路由器 (Egress PE Router)

对按照 4 over 6 封装规则封装过的数据分组进行解封装的 4 over 6 路由器。

CE路由器 (CE Router)

客户接入网络中的边缘路由器。

链路 (link)

相邻节点之间通信所依赖的数据链路层连接, 例如 Ethernet、PPP 链路、帧中继、ATM 网络。

接口(interface)

节点到链路的连接点。

4 over 6虚接口 (4 over 6 Virtual Interface)

位于 4 over 6 路由器上支持 4 over 6 功能的虚拟接口, 该接口需要配置 IPv4 地址和 IPv6 地址。

地址簇 (Address Family)

IPv4 和 IPv6 分别构成 IPv4 地址簇和 IPv6 地址簇。

地址簇边界路由器 (AFBR, Address Family Border Router)

运行 IPv4/v6 双协议栈的路由器, 该路由器同时分别连接 IPv4 地址簇网络和 IPv6 地址簇网络。在本标准所定义的 4 over 6 机制中, AFBR 具体实现在具有 4 over 6 功能的 PE 路由器上。

3.2 缩略语

BGP	Border Gateway Protocol	边界网关协议
BGP-MP	Multi-Protocol extension to BGP	边界网关协议多协议扩展
IP	Internet Protocol	互联网协议
IPv4	Internet Protocol Version 4	互联网协议版本 4
IPv6	Internet Protocol Version 6	互联网协议版本 6

4 4 over 6 机制概述

随着 IPv6 技术的快速发展和纯 IPv6 骨干网络的建成, 大量 IPv4 网络需要通过 IPv6 骨干网来实现互联互通的需求日益增多。如图 1 所示, 在 IPv4/v6 互联的网络拓扑上, 一些只运行 IPv6 协议栈的 P 路由器构成了纯 IPv6 骨干网。然而, 由于现有大量 IPv4 应用将在一定时期内仍然广泛使用, 因此该纯 IPv6 骨干网需要向边界网络提供 IPv4 协议栈接入。

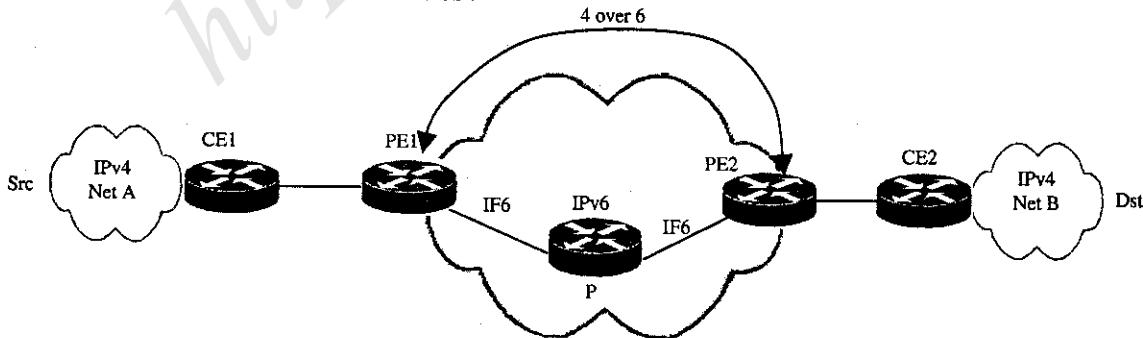


图 1 使用 4 over 6 机制传输的网络拓扑

PE 路由器作为地址簇边界路由器 AFBR, 同时运行 IPv6 和 IPv4 双协议栈。PE 路由器在骨干网内使用 IPv6 协议连接纯 IPv6 骨干网, 对接入网络则使用 IPv4 协议栈与边缘网络的 IPv4 单协议栈 CE 路由器连接, 从而对已有 IPv4 网络提供接入服务。因此, 需要设计一种机制, 使得在 CE 路由器上只运行 IPv4

协议，而 IPv6 骨干网作为传送 IPv4 分组的传输网络。

目前大多数 IPv4/IPv6 网络过渡方案都主要是解决基于 IPv4 骨干网的小规模 IPv6 接入网络之间的互联问题以及 IPv4/IPv6 网络之间的互联互通性问题。IPv6 通用隧道、基于 MPLS 的隧道和 DSTM 技术等虽然能够在一定程度上实现 IPv4 网络通过 IPv6 主干网进行互联，但由于存在可扩展性差、协议机制复杂或者对网络核心改动很大等问题，部署和实现这些机制是困难的，其管理维护负担已超出了可控范围。本标准提出了基于 IPv6 的 IPv4 网络互联机制（简称 4 over 6 机制），解决 IPv4 网络通过纯 IPv6 骨干网络实现互联的问题。

与其他类似机制相比，4 over 6 机制不需要手工配置端到端隧道，具备对网络和端系统的透明特性，能够自适应的动态选路，是一种 IPv4/v6 异构网络自动传输机制，具有较小的网络管理及维护负担，适应于大规模复杂网络拓扑。本标准提出的 4 over 6 机制主要解决 mesh 问题，而不涉及对 Hub&Spoke 场景的支持。

总体来说，4 over 6 机制包括控制平面和数据平面两方面的问题。控制平面需要解决的问题是如何通过隧道端点发现机制来建立 4 over 6 隧道。由于多个 PE 路由器连接到 IPv6 传输网上，为了准确地封装 IPv4 分组并转发到某个出口 PE 路由器，入口 PE 路由器需要知道具体哪个 PE 路由器是出口路由器。本标准扩展 BGP-MP 协议，在 IPv6 骨干网上携带 IPv4 目的网络的信息和隧道端点信息并发送到 IPv6 骨干网的另一端，以此来在 PE 路由器上建立无状态的 4 over 6 隧道。PE 与 CE 之间可以通过域内或域间 IPv4 路由协议来交互 IPv4 路由，也可以由 CE 路由器配置缺省路由到 PE 路由器，视具体使用场景而定。

在建立 4 over 6 隧道的基础上，数据平面主要关注包括封装和解封装的分组转发处理。在入口 PE 路由器找到恰当的出口路由器后，入口路由器需要采用某一特定的封装机制来封装并转发原始 IPv4 分组。出口路由器从 IPv6 传输网络收到封装分组后，出口路由器对分组进行解封装，并转发到相应的 IPv4 目的网络。由于 4 over 6 机制主要运行在 PE 路由器上，且只涉及对 IPv4 分组最外层头部的封装处理，因此也同样适用于 IPv4 网络中使用 NAT 机制的场景。

5 4 over 6 数据平面处理

4 over 6 数据平面主要包括下面 3 个部分：

- (1) 入口 PE 路由器将接收到的 IPv4 分组用 IPv6 头部封装；
- (2) 封装后的分组在 IPv6 骨干网中传输；
- (3) 出口 PE 路由器把分组解封装为原来的 IPv4 分组格式。

由于封装后的 4 over 6 数据分组在 IPv6 骨干网传输时，是对 IPv6 骨干网透明的，所以这类封装的数据分组就如同正常的 IPv6 数据分组一样在骨干网当中传输。基于 4 over 6 的分组传输特点，数据分组的封装和解封装需要在双栈路由器 AFBR 上完成，因此，在本标准所提出的 4 over 6 机制中，由 ISP 提供的接入路由器 PE 实现 AFBR 的分组封装和解封装的功能。

在具体实现中，每个 4 over 6 PE 路由器维护一个 4 over 6 虚接口，该虚接口通过维护 4 over 6 封装表，用于处理分组的封装和解封装。该封装表的每个表项中包含了目的 IPv4 网络的地址和掩码，以及需要转发到的出口 PE 路由器的 4 over 6 虚接口 IPv6 地址。

如图 2 所示，IPv4 分组到达 4 over 6 入口 PE 路由器，入口 PE 路由器在查找 IPv4 转发表后，发现应进入本地的 4 over 6 虚接口处理；在 4 over 6 虚接口处理中，通过查找 4 over 6 封装表对该分组进行封装，封装的目的 IPv6 地址是 4 over 6 出口 PE 路由器的虚接口 IPv6 地址，而源 IPv6 地址是 4 over 6 入口 PE

路由器的虚接口 IPv6 地址；封装后的 IPv6 分组通过 IPv6 骨干网发送到出口 PE 路由器；出口路由器对分组进行解封装后，再通过查找 IPv4 转发表将原始 IPv4 分组转发到相应的目的 IPv4 网络。因此，为了进行正确的 4 over 6 数据处理，需要将目的 IPv4 网络地址的转发接口置为 4 over 6 虚接口，并且需要自动维护 4 over 6 封装表，这些在 4 over 6 控制平面来完成。

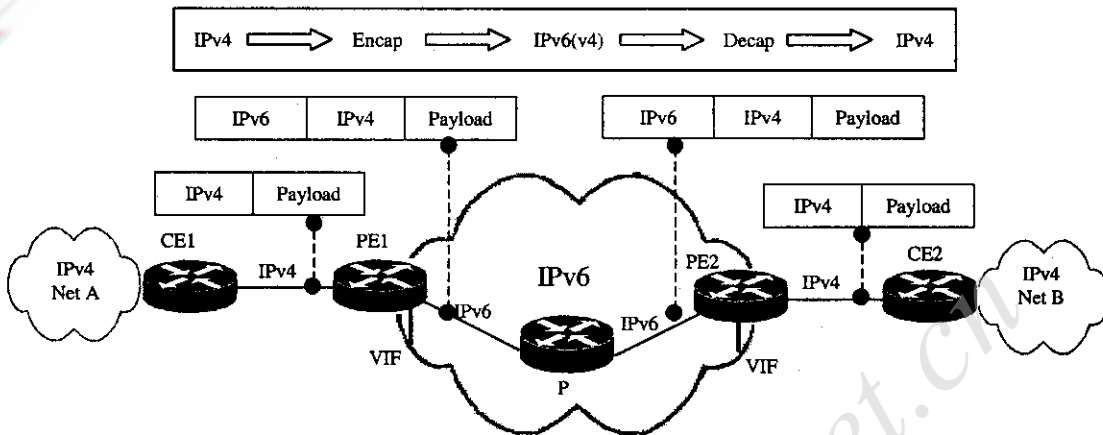


图 2 4 over 6 数据平面处理

数据分组的封装和解封装采用 IPv4-in-IPv6 的方法，如图 3 所示，具体封装、解封装以及分组分片和 ICMP 分组的相应处理可以参见 RFC 2473 和《IPv4-IPv6 过渡中互联互通技术要求》。除了规范性引用文件所描述的 4 over 6 的封装方法，对其他已有封装方法如 GRE 隧道方法、IPSec、MPLS 隧道以及 PPVPN 隧道等，4 over 6 机制将来可以通过进一步扩展进行支持。此外，对于将来出现的 IPv4 over IPv6 封装技术，本标准也可扩展采用。

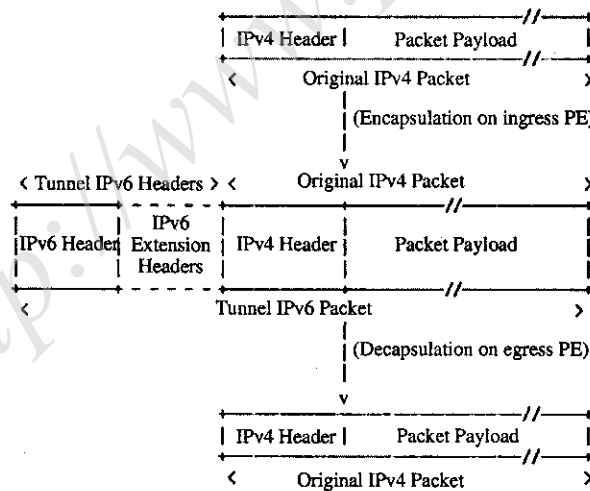


图 3 IPv4-in-IPv6 的封装和解封装方法

6 4 over 6 控制平面处理

4 over 6 机制中，控制平面主要解决隧道端点发现和网络可达性信息传输的问题。在本标准所定义的 4 over 6 机制中，4 over 6 控制平面通过对边界网关协议多协议扩展 BGP-MP 再进行 4 over 6 扩展来实现。作为 AFBR，PE 路由器之间基于 IPv6 骨干网建立 BGP 的邻居关系，通过 BGP 的多协议扩展携带不同种类的路由信息。

BGP-MP 协议规定，在建立 BGP 邻居关系时使用 OPEN 消息的下列两个参数的组合标识 BGP 实体的能力：地址簇标识符 AFI 和后续地址簇标识符 SAFI。由于传输上述 4 over 6 隧道端点和目的地址是 BGP-MP 实体的一个新能力，即 4 over 6 能力，因此需要定义新的 {AFI,SAFI} 组合。由于 4 over 6 所需要传输的目的地址是 IPv4 网络地址，因此 AFI 使用已经定义的 AFI_IP4=1。而基于申请定义 SAFI 的先来先服务原则，本标准申请定义 SAFI_4 over 6 = 67。基于上述定义，4 over 6 的 BGP-MP 协议在 OPEN 消息中使用 {AFI,SAFI} = {AFI_IP4=1, SAFI_4 over 6=67} 标识发出该 BGP-MP 消息的 BGP 实体的 4 over 6 能力。

在 BGP 建立了邻居关系的基础上，在 BGP UPDATE 消息中的路径属性包含下一跳的 AFI、SAFI 以及下一跳地址和网络层可达性信息。在使用 BGP-MP 的 UPDATE 消息传输隧道端点和网络可达性信息时，仍然使用上述 AFI 和 SAFI。与此同时，在路径属性中的下一跳网络地址必须是 4 over 6 虚接口的 IPv6 地址，而网络可达性信息 NLRI 包含了目的 IPv4 网络地址和掩码。该 Update 分组如图 4 所示。

Address Family Identifier (2 octets): AFI_IP4=1
Subsequent AFI (1 octet): SAFI_4OVER6 = 67
Length of Next Hop (1 octet): 16
Next Hop: IPv6 Address of 4over6 Virtual Interface
Number of SNPAs (1 octet)
Length of first SNPA (1 octet)
First SNPA (variable)
Length of second SNPA (1 octet)
Second SNPA (variable)
...
Length of Last SNPA (1 octet)
Last SNPA (variable)
NLRI (variable): Destination IPv4 Network Address

图 4 BGP 协议的 4 over 6 扩展

当 4 over 6 路由器的 IPv4 路由信息发生变化时，路由管理模块通知 BGP 协议，启动了 4 over 6 功能的 BGP 协议发送 Update 分组到其他 Peer 路由器。对端 4 over 6 路由器收到 Update 分组后，更新本地维护的封装表的同时，更新本地 IPv4 路由表，把相应目的 IPv4 地址的出接口置为本地 4 over 6 虚接口。

对于配置有路由反射器（Route Reflection 简称为 RR）的场景，PE 路由器分为 RR 和非 RR 两种。每个 RR 连接多个非 RR，这些非 RR 称为该 RR 的客户端路由器。所有 RR 之间建立全连接的 Full-mesh 结构。因此每个 RR 的 IBGP Peer 分为两种，即 RR 和客户端路由器，其中客户端路由器把所连接的 IPv4 网络中的 IPv4 路由变化通过 IBGP 报文通告对应的 RR。如果一个 RR 收到其他 RR 通过 IBGP 发来的 4 over 6 路由，将其转发给所有的客户端路由器。而如果一个 RR 收到客户端路由器通过 IBGP 发来的 4 over 6 路由，则将其转发给所有的 IBGP Peer。任何一个远端 IPv4 网络中的路由变化，可以经过最多 3 步传到整个网络，即远端出口客户端 PE 路由器、RR 以及网络中任意入口客户端 PE 路由器。因此，本标准所定义的 4 over 6 机制能够支持配置有路由反射器的场景。

7 4 over 6 机制示例

基于如上定义的 4 over 6 数据平面和控制平面的处理过程，下面以图 5 为例说明 4 over 6 机制的实际工作过程。PE2 从 CE2 学习到到达目的 IPv4 网络 Net B 和 Net C 的路由信息，更新本地的 IPv4 路由表 (PE2 IPv4 Routing table)，进而产生本地封装表 (PE2 Encapsulation table)，并将本地封装表发送给 BGP-MP 模块。PE2 的 BGP-MP 模块把目的 IPv4 网络 Net B 和 Net C 的可达性信息以及隧道端点信息 (PE2 的 4 over 6 VIF) 通过 BGP Update 消息发送到 PE1。PE1 在接收到 BGP 更新分组后，把 Net B 和 Net C 的网络地址以及隧道端点 (PE2 的 4 over 6 VIF) 存储到封装表 (PE1 Encapsulation table)，并在本地 IPv4 路由表中增加相关信息，把目的地址为 Net B 和 Net C 的转发接口置为本地 4 over 6 VIF，即 PE1 VIF。

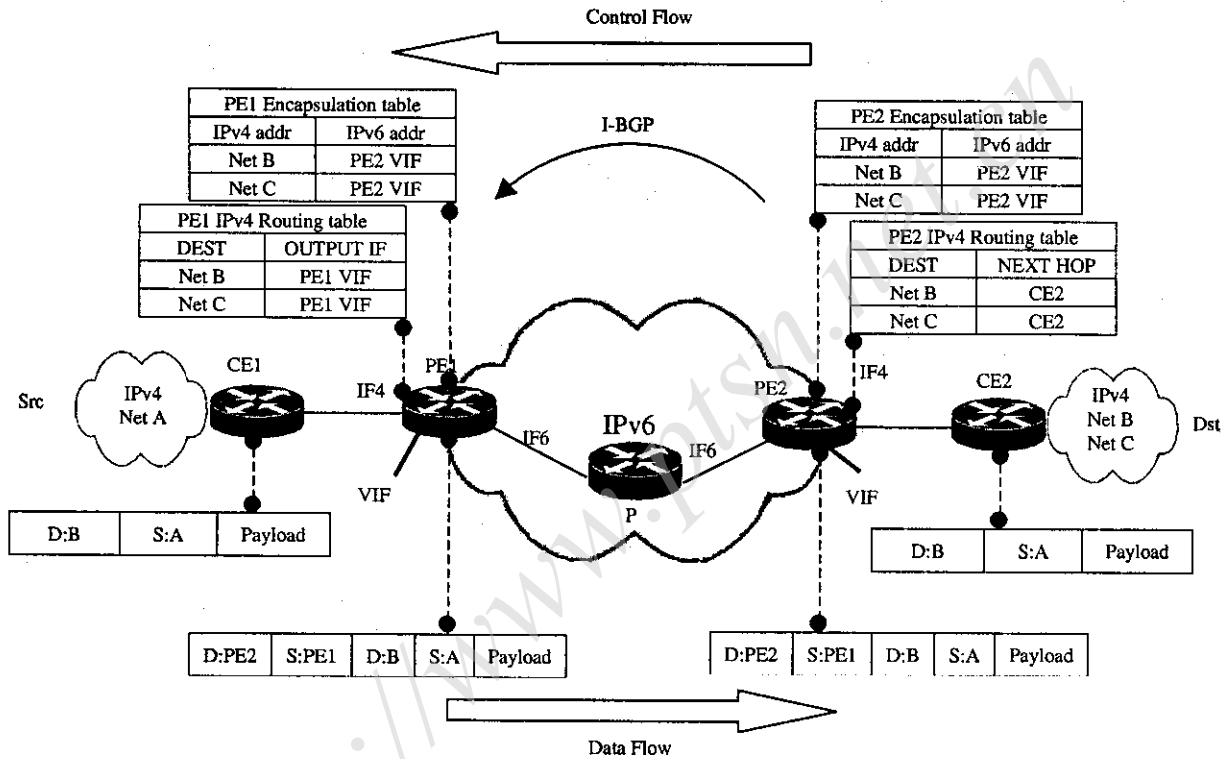


图 5 4 over 6 机制示例

在完成上述路由信息交互后，PE1 和 PE2 分别维护了如图所示的封装表和路由表。来自源网络 Net A 的目的地址在网络 Net B 内部的 IPv4 分组到达 PE1 后，PE1 在 IPv4 转发表中进行路由查找，发现对应转发接口为本地 VIF 接口，交由本地 VIF 接口处理。VIF 接口处理中，查找封装表，找到对应的 IPv6 地址 (即 PE2 的 4 over 6 虚接口的 IPv6 地址)，将该地址作为目的 IPv6 地址并将本地 4 over 6 虚接口的 IPv6 地址作为源地址进行封装，进而发送到 IPv6 网络。PE2 收到封装后的分组后，进行解封装，并根据原始 IPv4 分组的地址查找 IPv4 转发表，转发到目的网络 Net B。

8 可扩展性

4 over 6 机制不需要更改 P 路由器，只需要对 PE 路由器的功能进行扩展，即所有 AFBR 的操作全部在 PE 路由器上完成，因此具有较高的可扩展性。此外，4 over 6 路由器不需要维护每流状态，只需要增加 4 over 6 的路由信息。

由于 AFBR 之间需要使用 BGP 进行邻居交互，因此对 PE 路由器的数量可扩展性与自治系统边界路

由器 ASBR 的数量可扩展性相同。如果采用了路由器反射 Router Reflector (RR) 技术, 则理论上对 PE 路由器的数量是没有限制的。

9 双栈接入

边缘网络可能通过不同的地址簇接入骨干网, 如图 6 所示。在这种双栈接入的情况下, CE 通过配置适当的默认路由、策略路由或者优先级设置来选择具体采用哪个骨干网的接入方式。对于 PE 从 CE 学习路由的方式, 可以通过手工配置的方法或者使用路由协议与 CE 交互, 但需要使用路由过滤机制保证 IPv4 接入网络是末端网络而不是穿通网络, 即过滤掉通过 CE 所学习到的目的地址不属于该接入网络的路由。

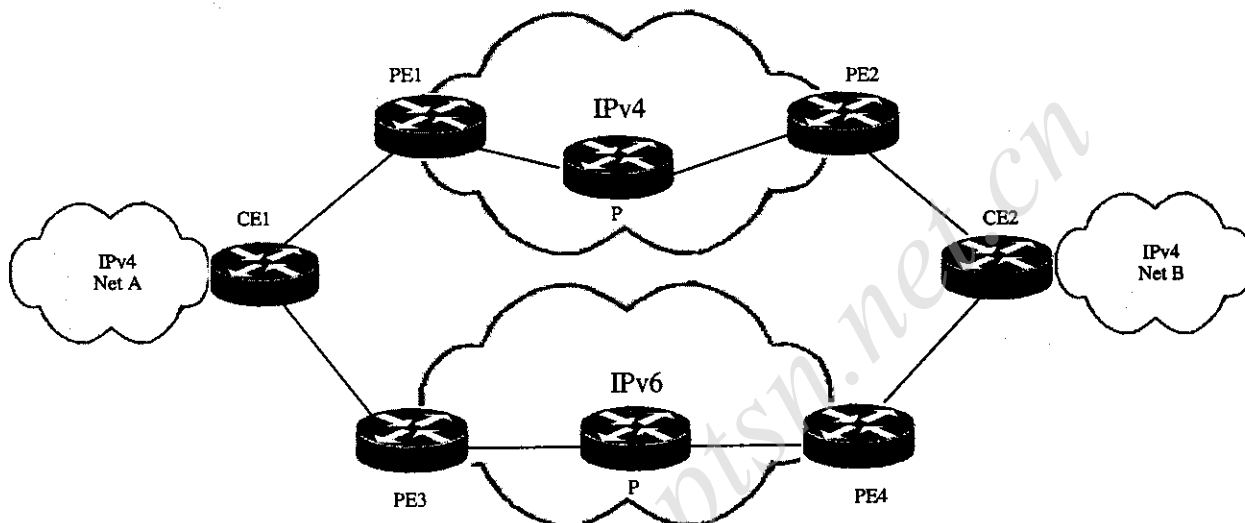


图 6 IPv4 网络通过 IPv4/v6 骨干网多点接入互联网

10 安全考虑

虽然 4 over 6 数据分组在 4 over 6 边界路由器中需要处理分组的封装, 使得 4 over 6 边界路由器的处理负载增加, 但由于 4 over 6 边界路由器不需要进行资源分配和状态维护, 因此具有较小的负载, 并具有一定抗 DDoS 攻击的能力。此外, 由于分组封装机制较为简单, 因此如果采用硬件实现, 将进一步增强 4 over 6 边界路由器的抗攻击能力。

BGP Peers 之间的控制会话和隧道分组传输还可以进一步扩展采用 IPSec 来实现以增加安全性。