

4over6 基于非显式隧道的 IPv4 跨越 IPv6 互联机制

吴建平¹, 李 星², 崔 勇¹, 徐 恪¹, 王晓峰¹

(1 清华大学计算机科学与技术系, 北京 100084; 2 清华大学电子工程系, 北京 100084)

摘 要: IPv6 骨干网正在逐步建成并走向实用, 纯 IPv6 的骨干网 (如中国下一代互联网: CNGI) 已经出现, 而已有的大量互联网应用和服务仍存在于 IPv4 网络中, 因此出现了大规模复杂 IPv4 网络通过 IPv6 骨干网互联的需求, 然而目前没有满足这种需求的过渡机制. 本文提出了一种基于 IPv6 的 IPv4 网络互联过渡机制框架, 对 4over6 过渡机制框架进行了阐述和分析, 并通过原型系统实现, 对 4over6 机制框架进行了原理和功能验证测试. 为减小路由开销, 提出了 ISP 独立部署方案, 增强了 4over6 机制框架的可实施性. 4over6 过渡机制框架具备对网络和端系统透明性、轻型和可自适应动态选路特性, 网络管理维护负担小, 可适应于大规模、复杂的网络互联等优点.

关键词: 4over6; IPv6; IPv4; 过渡; 互联网

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2006) 03-0454-05

4over6 IPv4 Network Interconnection over IPv6 Backbone Without Explicit Tunneling

WU Jian-ping¹, LI Xing², CUI Yong¹, XU Ke¹, WANG Xiao-feng¹

(1 Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

(2 Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract As IPv6 networks rapidly deployed native IPv6 backbones (such as CNGI) are emerging. However vast applications and services still stay in IPv4 network. Thus interconnection of IPv4 networks over IPv6 backbones is required. Unfortunately, there are no suitable transition mechanisms available to satisfy this requirement. We propose a 4over6 framework for IPv4 network interconnection over IPv6 backbone without explicit tunneling. The 4over6 mechanism is discussed and analyzed. Prototype implementation verifies the theory of the framework. ISP independent deployment scheme then is proposed which reduces the routing overhead and increases the practicability of 4over6 deployment. Being transparent to IPv4 and IPv6 networks, the lightweight scalable 4over6 framework is fit for large-scale complexity network interconnection with automatic discovery and control.

Key words 4over6; IPv6; IPv4; transition; Internet

1 引言

随着 IPv6 技术的快速发展并逐步成熟和广泛被采用, IPv6 网络规模逐步扩大. 一些国家在政策上为了促进网络的过渡, 加速 IPv6 网络的发展, 推动 IPv6 及其相关技术的发展, 已经建成了纯 IPv6 骨干网络, 比如中国下一代互联网示范项目 (CNGI China Next Generation Internet). 然而, 已有的大量互联网应用和服务仍然存在于 IPv4 网络中, 将如此大量的应用和服务移植到 IPv6 上, 绝

不是短时期能够完成的工作, 因此, IPv4 和 IPv6 网络将长期共存, 相应的 IPv4/v6 过渡将具有长期性和复杂性.

IPv4/v6 过渡机制作为研究领域的一个热点问题, 研究者们已经提出了大量方案^[1, 2]. 现有研究成果可以分为双协议栈、隧道和协议翻译三类. 双协议栈技术同时对 IPv4 和 IPv6 网络、应用和服务提供了完全的兼容, 同时也是隧道和协议翻译等其他过渡技术的基础, 实现了较好的互联互通性, 然而由于需要同时维护两套协议基础设施, 维护管理开销大, 实现成本高, 并且由于 IPv4 地址缺乏而

无法全面部署。隧道过渡技术是目前被广泛采用的技术,适用于通信双方采用相同网络协议的情况,传输路径中可能需要穿越其他类型的网络,因此需要对分组进行封装处理。已经成为 IETF 正式标准的隧道技术有配置隧道技术、自动的 IPv4 兼容隧道^[3]、6to4 隧道^[4]、6over4 隧道^[5]和 GRE 隧道^[6],其他隧道技术如 ISATAP 隧道^[7]、Teredo^[8]/Silkroad^[9]隧道、基于 MPLS 的隧道^[10]等都还处在 IETF 标准草案阶段,上述隧道技术都是解决通过 IPv4 骨干网连接小规模、接入 IPv6 网络孤岛的问题。协议转换翻译技术主要有 SIIT^[11]、NAT-PT^[12]、TCP-UDP relay^[13]、BIS^[14]、SOCKS64^[15]和 BIA^[16],分别在网络层、传输层、应用层和主机应用程序接口进行 IPv4/v6 协议的翻译,其解决的主要问题是纯 IPv4 网络节点与纯 IPv6 网络节点之间的互联。

综上所述,目前大多数 IPv4/v6 网络过渡方案主要是解决基于 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,而针对已经出现的大量 IPv4 网络需要通过 IPv6 骨干网实现互联的需求,相关研究甚少。IPv6 通用隧道^[17]/GRE over IPv6 隧道、基于 MPLS 的隧道和 DSIM 技术^[18]虽然能够在一定程度上实现 IPv4 网络通过 IPv6 网络互联,然而由于存在可扩展性差、协议机制复杂或者需要对网络进行修改等问题,在大规模复杂网络中多个 IPv4 网络之间或者使用多路由器通过 IPv6 核心网互联情况下部署和实现这些机制是困难的,其管理维护负担已经超出了可控制的范围。

本文针对上述问题和实际需要,提出了一种基于 IPv6 的 IPv4 网络过渡机制框架,简称 4over6 机制。该机制通过 4over6 地址映射、4over6 路由和 4over6 分组处理的共同作用,实现了 IPv4 网络通过 IPv6 骨干网的互联,而不需要建立显式的隧道,并具备对网络和端系统透明特性、轻型、可自适应动态选路特性,具有较小的网络管理维护负担,能够适应于大规模、复杂网络的互联。

2 4over6 过渡机制设计原则

4over6 机制主要解决 IPv4 网络通过 IPv6 骨干网进行互联互通的问题。如图 1 所示的网络中,任意 IPv4 网络之间均应通过 IPv6 网络互联互通。为达到上述目标,所设计的 4over6 机制除了部署简单外,还需要满足以下要求。

- 透明性原则 (Transparency): 对 IPv4 网络和 IPv6 网络不需要做任何修改,所有的修改只在边界上完成,即对端系统、采用过渡机制的 IPv4 网络和 IPv6 网络透明。
- 轻型原则 (Lightweight): 充分利用现有的过渡机制和成果,如采用已经成熟并且已经获得广泛支持的数据分组的封装和协议转换技术,对路由器系统做尽量小的修改就可以实现;在网络中尽可能少的引入状态和协议交互负载。
- 可扩展原则 (Scalable): 具备与当前 IPv4 网络系统互

联一样的可扩展性,适应大规模的网络互联;适应不同规模的网络以及单一自治系统和多自治系统;在性能上同样具备可伸缩性。

- 动态自适应原则 (Automatic Discovery and Control): 具备自适应运行和动态选择路由特性,不需要人工干预路由或者建立显式的隧道,不需要或者只需要很少的网络管理和维护。

3 4over6 机制设计

4over6 机制原理如图 1 所示,其中 R1、R2 和 R3 是具备 4over6 机制的 4over6 边界路由器,IPv4 网络通过 4over6 边界路由器与 IPv6 网络互联。纯 IPv6 自治系统 (AS1) 内部运行域内路由协议,并具备至少一个标准的 IPv6 地址前缀。4over6 机制通过 4over6 地址映射、4over6 路由机制和 4over6 分组处理的共同作用,实现基于 IPv6 的 IPv4 网络互联,而不需要建立显式的隧道,同时能够很好的满足上述的所有功能需求透明、轻型、可扩展和自适应等原则。

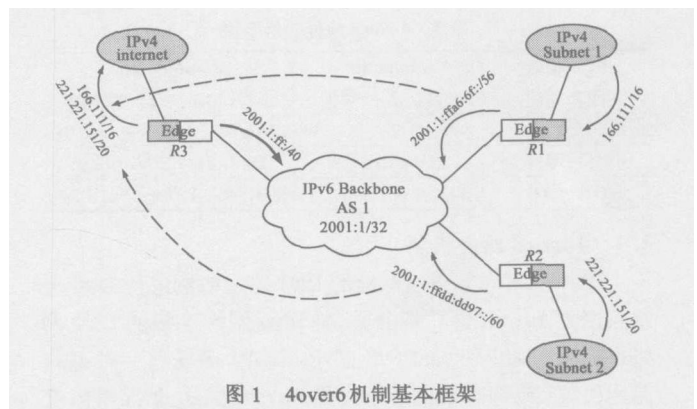


图 1 4over6 机制基本框架

3.1 4over6 边界路由器

4over6 边界路由器是 4over6 机制定义的新概念,是该机制的基本组成部分,也是网络中定义的新实体,用于同时连接 IPv4/v6 网络。其基本特征是对一个双协议栈路由器的功能进行扩展,具备下述 4over6 路由功能和 4over6 分组处理功能。

3.2 4over6 地址映射

4over6 地址映射是 4over6 机制的基础。基于 4over6 地址映射,IPv4 分组可以很容易的被翻译到 IPv6 分组,而不需要在网络中维护软状态;同样,IPv4 路由也可以很方便的翻译成一个 IPv6 路由,反之亦然。具体的方法是将 IPv4 地址 (含前缀,下同) 嵌入 IPv6 地址并基于该前缀建立一一对应关系,即给定的 IPv4 地址能够唯一对应并确定一个 IPv6 地址,反之亦然。具体方法如图 2 所示,其中 V4ADDR 为映射前的 IPv4 地址,4over6prefix 为 4over6 地址前缀。4over6prefix 有两种定义方式,一种是 IANA 分配的方式,即全局范围内有效的特殊前缀 4over6 地址映射,类似于

6to4 机制中的“2002:”前缀;另一种是 ISP或自治系统从其所拥有的 IPv6地址段中自行分配的可标识 IPv4地址的 4over6地址前缀,即局部或者本地 4over6地址映射,这样的前缀可能存在多个.用 4over6prefix代表 4over6单播地址前缀,其长度为 len 在特定的 ISP或者自治域系统内,可以同时存在上述两种映射.

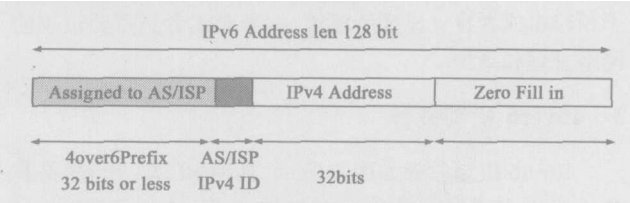


图 2 4over6地址映射

IPv4网络地址前缀,形如 a b c d M,其对应的 IPv6地址前缀为 4over6prefix a b c d /len+ M.对确定的自治系统内部,其 4over6prefix前缀长度是已知的,因此可以进行双向的地址映射.假设全局分配的 4over6地址前缀为 2004: /16,自治系统获得的 IPv6前缀为 2001: da8: /32,其 IPv4地址标识为 0xFF,在上述情况下表 1是一些地址和前缀映射的示例.

表 1 4over6地址映射示例

| IPv4地址 或者前缀 | 4over6prefix= 2004 /16的映射 | 4over6 prefix= 2001: da8: /32的映射 |
|-------------------|------------------------------|-------------------------------------|
| 218 106 172 0 /24 | 2004 DA6A: AC /40 | 2001 da8: F1DA: 6AAC /64 |
| 166 111 0 0 /16 | 2004 A66F /32 | 2001 da8 FFA6: 6F /56 |
| 166 111 68 2 | 2004 A66E: 4402 : | 2001 da8 FFA6: 6F44 02 : |

3.3 4over6路由机制

IPv4网络内部通过标准的 IPv4路由机制来使 IPv4分组路由到 4over6边界路由器.与 IPv6网络连接的 IPv4网络可以采用如 RIP,OSPFv2,IS-IS,BGP4等任何一种 IPv4路由协议(或静态配置)与边界 4over6路由进行路由交互,通告 IPv4网络的路由信息,使得 4over6边界路由器获得与之连接的 IPv4网络的路由.

4over6边界路由器采用 4over6地址映射转换的方法将 IPv4路由映射到对应的 IPv6路由,并注入到 IPv6网络中,是 4over6路由机制的关键部分.实现的方式是采取路由信息映射的方式,即对路由信息中的 IPv4地址进行 4over6地址映射,实现 IPv4路由信息转换为 IPv6路由信息,这种转换得到的 IPv6路由称为伪 IPv6路由.路由大部分属性信息可直接映射,如路由度量等,而某些属性的直接映射并没有任何意义,如下一跳地址,因此可以根据实现者的需求重新定义.

IPv6网络内部通过标准的 IPv6路由机制使伪 IPv6分组(经过 4over6分组处理的 IPv4分组,参见 3.4节)路由到适当的 4over6边界路由器.与 IPv4网络连接的 IPv6网络可以采用如 RIPng,OSPFv3,IS-IS,BGP4+等任何一种 IPv6路由协议(或者静态配置)与边界 4over6路由进行路由交互,通告 IPv6网络的路由信息(其中可能包含伪 IPv6

路由),使 4over6边界路由器获得与之连接的 IPv6网络的路由,同时,4over6边界路由器也向 IPv6网络通告本身的 IPv6路由信息,其中包含了 4over6路由映射转换得到的伪 IPv6路由.

源 IPv4网络需要使用一定的路由信息,从而将发送到目标 IPv4网络的分组能够通过适当的路由传送到 4over6边界路由器,进而再通过 IPv6骨干网传送到目标端 IPv4网络.为了向源 IPv4网络通告 IPv6网络所连接的所有或者部分 IPv4网络的可达性信息,4over6机制可以使用 4over6路由映射转换的逆过程-4over6路由映射还原来实现.即将伪 IPv6路由还原称为 IPv4路由.通过 4over6路由映射转换和还原,所有与 IPv6网络相连的 IPv4网络的路由都被注入到 IPv6网络中,并以伪 IPv6路由形式存在;

IPv4路由可以透明的穿越 IPv6网络.管理员应当可以通过配置来允许或者禁止 4over6路由映射转换和还原功能,还可以配置 4over6路由映射转换和还原只针对特定的 IPv4地址前缀以及范围.为了避免 IPv4路由信息进入 IPv6网络,上述 4over6路由还原并不是 4over6机制的必须部分.另外一种可行的方法是引入隧道代理机制,从而向通信双方一方告知目标地址或其封装地址.对隧道代理机制的详细讨论超出了本文的范围.

3.4 4over6分组处理

IPv4数据分组穿越 IPv6网络时,需要对数据分组进行特殊的处理,即 4over6分组变换和还原.目前对数据分组的处理方式主要有两种:封装和翻译.分组封装是在原 IPv4分组上增加一个额外的 IPv6分组头,将 IPv4分组封装到 IPv6的载荷内,封装后的数据分组在 IPv6网络内进行传输.分组翻译是将 IPv4分组头转换成 IPv6分组头,载荷内容不变化,转换后的分组头加上原先的载荷,原先的 IPv4分组头丢弃,组成新的分组,在 IPv6网络中传输.

4over6分组封装可以存在两种基本的形式-IP-IP封装和 GRE封装方法.IP-IP封装方法类似 IPv6通用隧道封装^[15].IPv6通用隧道封装方式定义了包括将 IPv4分组封装在 IPv6分组中的方法,如图 3(b)的分组 IP-IP封装.4over6机制的 IP-IP封装处理与该方法的关键区别在于,IPv6封装分组头的各个字段信息是从 IPv4分组头中的相关字段中直接映射得到,如源地址和目的地址是经过 4over6地址映射得到的,并且不需要在路由器中维护关于隧道的软状态,如图 3(a)所示的 4over6分组头映射.IPv6头部的相关字段直接对对应颜色的 IPv4分组头部映射,在 IPv4分组头中没有的信息,则按照特定的实现和已有的协议进行创建.其他处理与隧道机制相同,如最大传输单元(MTU)、LMP处理等.GRE封装方法类似标准的 GRE封装^[15],定义了包括将 IPv4分组封装在 IPv6分组中的方法,如图 3(c)的分组 GRE封装.4over6机制的 GRE封装处理与该方法的关键区别在于 IPv6封装分组头的各个字段信息由 4over6分组头映射获得而不需要在路由器中维护关

于隧道的软状态 (图 3 (a)).

4over6 分组翻译采用类似无状态 IP/ICMP 分组翻译协议的分组处理方法^[9], 其中的主要区别在于, IPv6 封装分组头的各个字段信息是从 IPv4 分组头中的相关字段中直接映射得到, 而不需要在路由器中维护关于地址映射表等软状态信息. 详细分组翻译以及相关分组的处理参考 RFC2765^[9] 相关部分.

分组封装可以保证端到端的透明传输并支持任意业务, 因此适用范围广泛, 其缺点是传输开销增加. 分组封装机制主要用于单播数据的处理和传输, 组播数据的处理和传输. 分组翻译传输比分组封装具有更小的传输开销, 具有对上层协议语义透明的特点和较高的传输效率. 由于对上层协议语义透明的特点, 使得 4over6 机制分组处理中的分组翻译机制具备一个很有趣的特性就是该机制提供了一种 IPv4/v6 网络之间透明互访的可能性. 然而分组翻译方式由于 IPv4 分组头转换到 IPv6 分组头的过程存在信息丢失, 例如会丢失 IPv4 分片分组信息、IPv4 选项信息等, 因此二者各有优缺点, 无法互相取代. 如果要求保证端到端的透明性, 或者可能需要传输 IPv4 选项信息时, 则建议使用分组封装方式; 在大多是使用环境下, 由于分组翻译具有相对封装方式较高的传输效率, 因此建议无特殊要求时使用分组翻译进行传输; 实现者应该允许管理员配置使用何种传输分组处理方式.

4 ISP 独立部署方案

通过对路由转换机制的分析, 4over6 机制可能产生 IPv6 路由表规模扩大问题. 目前大多数主流的骨干网路由器都能够支持大容量的路由表和高性能的路由查找, 使用硬件路由查找 (如 TCAM 技术) 复杂度一般为 $O(1)$, 使用软件路由查找的边缘路由器一般不需要处理整个表, 因此, 在性能上的降低和性能恶化并不明显. 此外, 考虑到 IPv4 系统和网络在未来可能逐步缩减, 而 4over6 机制是一种过渡技术, 是网络过渡时期临时采用的解决方案, 因此 IPv4 路由表引入 IPv6 网络在技术上是可行的. 尽管如此, 本文仍研究了 4over6 机制的路由表规模抑制策略, 提出了一种在单独的 ISP 中独立部署 4over6 机制的方案, 并可以增强 4over6 机制的渐进部署特性. 通过图 4 示例阐述 ISP 独立部署方案. 假设在 ISP1 中部署了 4over6 机制而在其他 ISP 中没有部署 4over6 机制. (1) 在 IPv6 路由器 R4 上

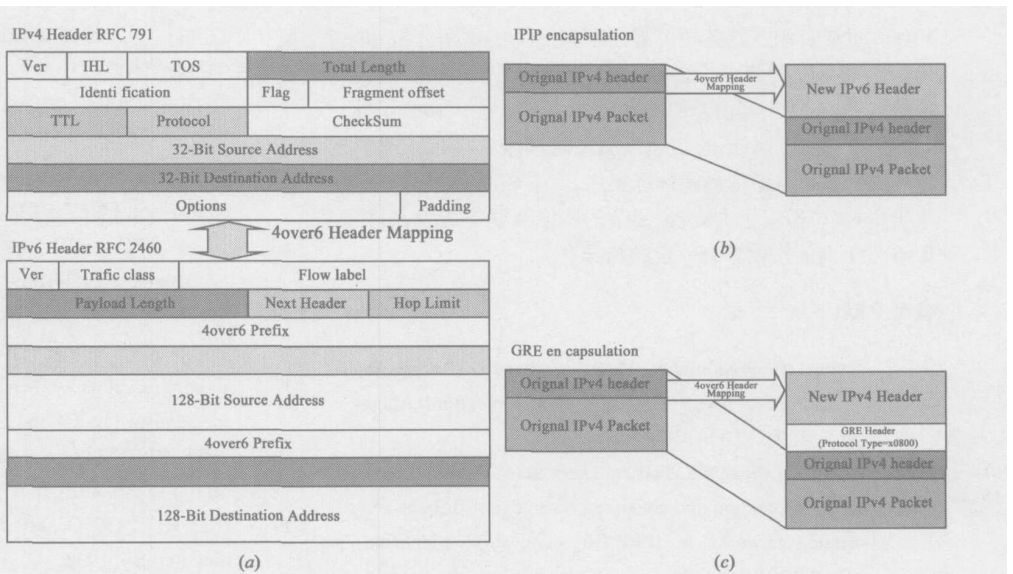


图 3 4over6 分组处理

设置禁止 ISP1 中的伪 IPv6 路由向其他 IPv6 ISP 网络中传播; (2) 在 4over6 路由器 R1、R2 上限制所有其他 IPv4 ISP 网络中的 IPv4 路由信息向 IPv6 网络中转换并传播. 通过采用上述 4over6 机制的 ISP 独立部署方案, ISP 的 IPv6 网络中仅有本 ISP 所属 IPv4 网络相应的伪 IPv6 路由, 大大限制了其所维护的伪 IPv6 路由数量, 保证了该 ISP 内部网络节点的性能. 与此同时, 该 ISP 在使用 4over6 机制的过程中, 该 ISP 由于采用 4over6 机制而引入的路由信息不会传播到其他 ISP, 因此实现了与其他 ISP 的完全独立性. 这种独立性保证了各个 ISP 可以渐进地、独立地采用 4over6 机制, 也可以独立地废除 4over6 机制, 而不会影响到其他 ISP, 保证了网络的可扩展性.

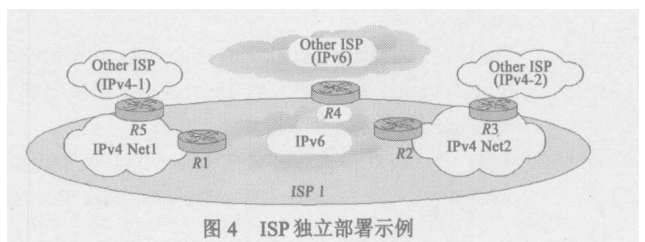


图 4 ISP 独立部署示例

5 结论

本文针对 IPv6 骨干网已经开始逐步部署的情况, 提出了 IPv4/v6 过渡的 4over6 机制框架. 该机制通过 4over6 地址映射、4over6 路由机制、4over6 分组处理三个步骤, 实现了基于 IPv6 骨干网的 IPv4 网络互联, 克服了目前类似机制中普遍存在的管理维护困难、机制复杂、可扩展性差等缺点, 适用于大规模复杂 IPv4 网络之间通过 IPv6 骨干网进行互联互通的情形, 满足 IPv4 网络通过 CNG I 等纯 IPv6 骨干网络实现互联的要求. 本实验室基于自主研发的 THU 4000 系列双栈路由器, 实现了 4over6 机制的原型系统, 对 4over6 机制进行了原理和功能验证测试. 本文对

4over6过渡机制及其各项关键技术进行了阐述和分析,并讨论了 4over6机制中的路由表规模问题,提出了 ISP独立部署方案,在减小路由开销的同时进一步增强了 4over6机制的可实施性。4over6过渡机制具备对网络和端系统透明特性、轻型、可自适应动态选择路由,具有较小的网络管理维护负担、可适应于大规模、复杂的网络互联等优点,对纯 IPv6 骨干网的建设具有一定的指导意义。

参考文献:

- [1] DanielGWaddington, Fangzhe Chang Realizing the transition to IPv6 [J]. IEEE Communications Magazine, 2002, 6(3): 138- 148
- [2] Mallik Tatipanula, Patrick Grossetete, Hiroshi Esaki IPv6 integration and coexistence strategies for next-generation networks[J]. IEEE Communications Magazine, 2004, 42(1): 88- 96
- [3] IETF RFC2893, Transition Mechanisms for IPv6 Hosts and Routers[S]. 2000
- [4] IETF RFC 3056, Connection of IPv6 Domains via IPv4 Clouds[S]. 2001
- [5] IETF RFC2529, Transmission of IPv6 over IPv4 Domains Without Explicit Tunnels[S]. 1999
- [6] IETF RFC 1701, Generic Routing Encapsulation (GRE) [S]. 1994
- [7] FTamplin, TGleson, M Talwar, D Thaler Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [S]. IETF Draft draft-ietfngtrans-isatap-22.txt May, 2004
- [8] C Huitema, Teredo Tunneling IPv6 over UDP Through NATs[S]. IETF draft draft-huitema-v6ops-teredo-02.txt 2004
- [9] Li Min, Wu Xianguo et al Tunneling IPv6 with Private IPv4 Addresses Through NAT Devices[S]. IETF draft draft-lim-irv6ops-silkroad-01.txt 2004
- [10] J De Clercq, D Ooms et al Connecting IPv6 Islands over IPv4 MPLS using IPv6 Provider Edge Routers

(6PE) [S]. IETF Draft draft-ietfngtrans-6ops-bgp-tunnel-03.txt 2004

- [11] IETF RFC2765, Stateless IP/ICMP Translation Algorithm (SIIT) [S]. 2000
- [12] IETF RFC2766, Network Address Translation-Protocol Translation (NAT-PT) [S]. 2000
- [13] IETF RFC 3142, An IPv6-to-IPv4 Transport Relay Translation [S]. 2001
- [14] IETF RFC 2767, Dual Stack Hosts Using the Bump-In-the-Stack Technique (BIS) [S]. 2000
- [15] IETF RFC 3089, A SOCKS-based IPv6/IPv4 Gateway Mechanism (SOCKS64) [S]. 2001
- [16] IETF RFC 3338, A Durand Dual Stack Hosts Using Bump-in-the-API (BIA) [S]. 2002
- [17] IETF RFC2473, Generic Packet Tunneling in IPv6 Specification[S]. 1998
- [18] J Bound et al Dual Stack Transition Mechanism (DSTIM) [S]. IETF draft draft-ietfngtrans-dstm-08.txt 2002

作者简介:



吴建平 男, 1953年 10月生于山西太原, 教授, 博士生导师, 主要研究领域为计算机网络体系结构、下一代互联网路由体系结构和形式化方法等, 近年来在国内外学术刊物和学术会议上发表学术论文二百余篇。

E-mail jianping@cernet.edu.cn



李 星 男, 1956年 4月生于北京, 教授, 博士生导师, 清华大学电子工程系教授, 中国教育和科研计算机网 CERNET网络中心副主任; 美国 Signa X 科学学会会员, 亚太网络工作组 APNG主席, 亚太网络信息中心 APNIC执行委员会委员, 发表 120多篇论文, 著作 1本。

E-mail xing@cernet.edu.cn