第三届国家期刊奖百种重点期刊 | 中国科技核心期刊
工信部优秀科技期刊 | 中国五大文献数据库收录期刊

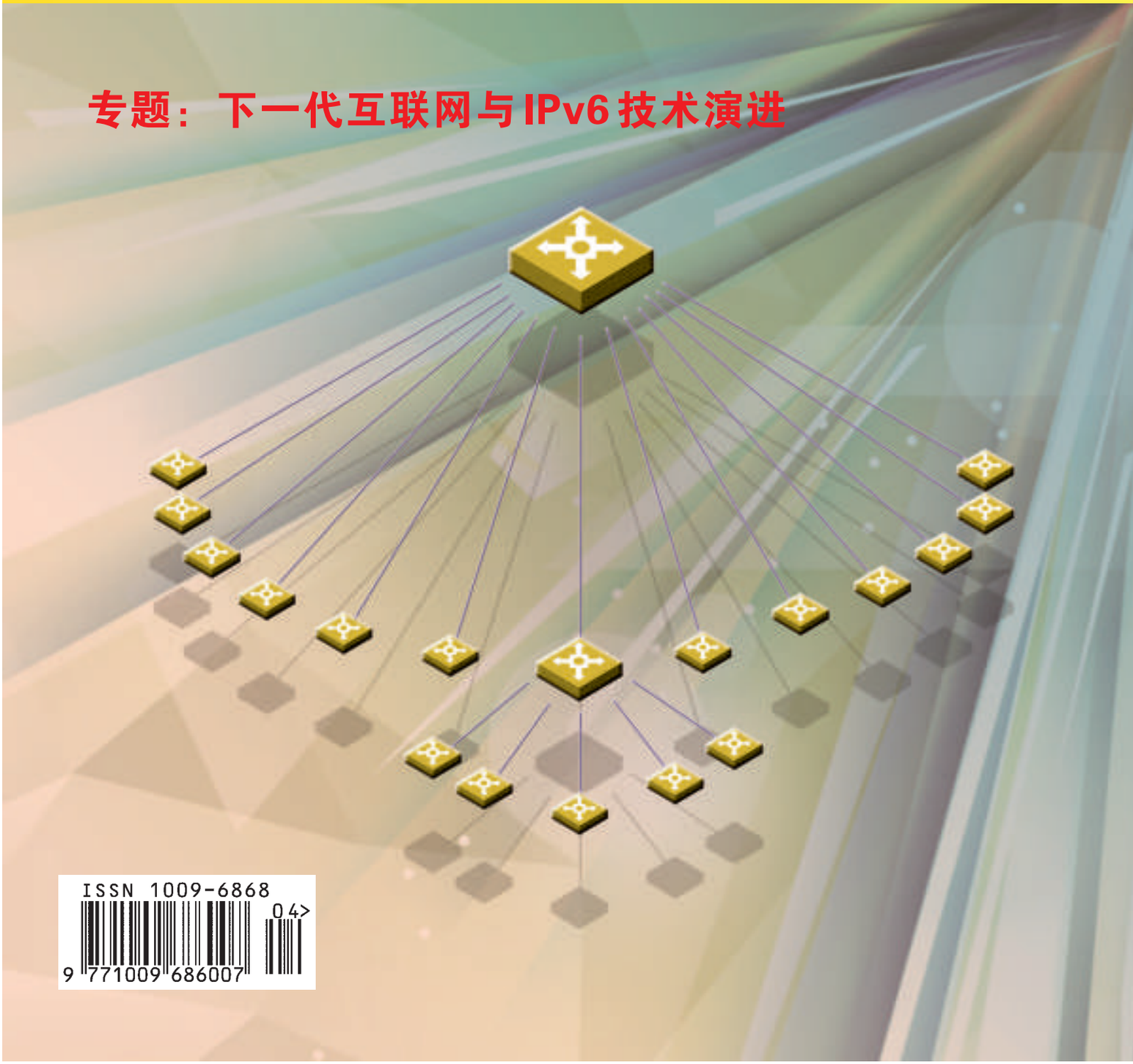
ISSN 1009-6868
CN 34-1228/TN

中兴通讯技术
ZTE TECHNOLOGY JOURNAL

www.zte.com.cn/magazine

2013年4月 • 第2期

专题：下一代互联网与IPv6技术演进



ISSN 1009-6868
04>
9 771009 686007

目次

中兴通讯技术　总第109期　第19卷　第2期　2013年4月

专题：下一代互联网与IPv6技术演进

02 IPv6过渡技术场景分析 …… 孙静文，孙琪，吴鹏
07 统一的IPv4/IPv6翻译与封装过渡技术——IVI/MAP-T/MAP-E …… 包丛笑，李星
12 PNAT——面向移动互联网的双重翻译过渡技术 …… 陈刚，邓辉，段晓东
16 面向NAT用户的IPv6过渡技术——Silkroad …… 刘敏
21 下一代互联网4over6软线隧道过渡技术 …… 崔勇，陈煜驰
25 域名系统对IPv6协议支持分析 …… 冷峰，王伟，张跃冬
29 ICP的IPv6演进 …… 侯金刚
33 IPv6-only网络互通技术 …… 刘笑寒，闫双舰，孙玉勤
36 IPv6测试研究与实践 …… 宋林健，李震

专家视点

40 计算通信：超量信息无线传输的深度探索 …… 王新兵，陶梅霞，刘辉

运营应用

44 无线通信系统的环境适应性 …… 郭丹旦，唐雄，崔卓

研究论文

49 GoTa 4G宽带多媒体集群系统的研究 …… 赵先明，徐云翔，朱伏生

开发园地

54 交换机中同步以太网技术的研究及其实现 …… 曾富前

系列讲座

58 大数据 (2) …… 韩晶，宋美娜

综合信息

中兴通讯创新方案落地欧洲 Mtel IPTV多屏融合平台商用(6)　中兴通讯荣获首届GTI Award融合方案大奖(20)　中兴通讯发布创新的全系列LTE小基站回传解决方案(32)　中兴通讯再破TD-LTE技术难关 全球首次实现跨频段载波聚合（CA）(39)　中兴通讯发布UMTS性能提升方案 频谱效率提升40% (57)　中兴通讯iCity智慧城市解决方案进入GSMA“全球移动大奖”短名单(62)

办刊宗旨

以人为本，荟萃通信技术领域精英；
迎接挑战，把握世界通信技术动态；
立即行动，求解通信发展疑难课题；
励精图治，促进民族信息产业崛起。

期刊基本参数：CN 34-1228/TN * 1995 * b * 16 * 64 * zh * P * ¥20.00 * 15000 * 14 * 2013-04

Contents

ZTE TECHNOLOGY JOURNAL Vol.19 No.2 Apr. 2013

Special Topic: Next-Generation Internet and the Evolution of IPv6 Technology

02 IPv6 Transition Technology …… SUN Jingwen, SUN Qi , WU Peng

07 IVI/MAP-T/MAP-E: Unified IPv4/IPv6 Stateless Translation and Encapsulation Technologies …… BAO Congxiao, LI Xing

12 PNAT: Double-Translation Transition Technology for Mobile Internet …… CHEN Gang, DENG Hui, DUAN Xiaodong

16 Silkroad: An IPv6 Transition Technology for NAT Users …… LIU Min

21 4over6 Softwire Tunneling Transition Technology for Next-Generation Internet …… CUI Yong, CHEN Yuchi

25 Making DNS Compatible with IPv6 Protocol …… LENG Feng, WANG Wei, ZHANG Yuedong

29 IPv6 Transition of ICP Networks …… HOU Jin'gang

33 Preliminary Research on IPv6-only Network Interoperability Technology …… LIU Xiaohan, YAN Shuangjian, SUN Yuqin

36 IPv6 Testing: Research and Practice …… SONG Linjian, LI Zhen

Expert View

40 Computing Communications: Wireless Transmission of Mass Information …… WANG Xinbing, TAO Meixia, LIU Hui

Operational Application

44 Environmental Adaptability of Wireless Communication Systems …… GUO Dandan, TANG Xiong, CUI Zhuo

Research Paper

49 GoTa 4G Wideband Multimedia Trunk System …… ZHAO Xianming, XU Yunxiang, ZHU Fusheng

Development Field

54 Synchronous Ethernet Technologies in Ethernet Switch …… ZENG Fuqian

Lecture Series

58 Big Data (2) …… HAN Jing, SONG Meina

敬告读者

本刊享有所发表文章的版权，包括英文版、电子版、网络版和优先数字出版版权，所支付的稿酬已经包含上述各版本的费用。

未经本刊许可，不得以任何形式全文转载本刊内容；如部分引用本刊内容，须注明该内容出自本刊。

邮购须知

本刊常年办理邮购订阅业务，欢迎订阅。订阅方法：从邮局汇款至编辑部，在汇款单上将订阅者的详细地址、收件人姓名及联系电话填写清楚，并在汇款单附言栏注明所购杂志期次及数量。

专题：下一代互联网和IPv6技术演进

专|题|导|读

互联网技术已成为人类生活中不可或缺的一部分。作为互联网协议簇的中流砥柱，IPv4取得了令人难以置信的成功。近些年来，随着互联网规模爆炸式地增长，互联网对IP地址的需求变得极为迫切。然而，IPv4地址消耗已经成为不争的事实，严重制约了全球互联网的发展。大量智能终端的涌入，使得有限的IPv4地址空间显得更加狭小。在这样的历史关口下，下一代互联网IPv6技术成为了亿万用户、大量ICP、众多运营商和设备商的聚焦点，IPv6的过渡技术也显得尤为重要。

IPv6过渡场景繁多复杂，但技术方案缺乏总体架构，各类IPv6过渡单元技术也优缺点并存，这些因素都使得IPv6过渡技术面临着各方面巨大的挑战：如何保障运营商同时稳定地提供IPv4和IPv6服务；采用何种方案，便于运营商的运营管理，且能够减轻配置、维护和排障的负担；如何保证终端用户和跨层应用的透明性，让用户无需感知过渡的存在，使之拥有良好用户体验；如何在保证对现有网络最小影响下，做好逐步过渡，平滑演进。为了解决这些技术难题，需要在各方面广泛而深入合作，既包括基础设施运营商、服务提供商（ISP）、网络内容服务商（ICP），又包括大量终端用户、各类操作系统乃至应用程序开发者。

做好下一代互联网的平滑演进，具有重要的社会价值和现实意义。目前各个科研机构和相关部门的互联网专家锐意进取，开拓创新，不断推动着IPv6过渡技术向前发展。那么IPv6过渡背景和场景如何？主要IPv6过渡技术是如何实现的？目前IPv6过渡技术已经发展到什么地步？又都涌现出怎样的新技术？这些技术又是如何得到ISP、ICP和用户等的支持和协作？这些问题不仅包括技术问题，还与政治、经济和文化等方面息息相关。

本期专题主要反映了当前IPv6平滑过渡中的技术层面。每篇论文都凝聚着作者对过渡技术研究的心血和汗水，希望这些成果与观点可以让读者受到启发，拓宽思路，开拓视野。在此，也对各位作者的积极支持和辛勤工作表示衷心的感谢！

崔勇

本期专题策划人

崔勇

清华大学计算机科学与技术系教授、博士生导师、网络技术研究所副所长，教育部新世纪人才和北京市科技新星计划获得者，IETF IPv6过渡Softwire工作组主席；研究领域包括计算机网络体系结构、IPv6与下一代互联网、无线移动计算等；先后主持了多项国家和省部级科研项目，包括国家自然科学基金重大国际合作项目、中国下一代互联网示范工程CNGI项目IPv6过渡课题、国家高科技发展计划“863”项目、国家重点基础研究“973”计划子课题等；获得40余项国家发明专利授权，荣获国家科学技术进步奖二等奖等；在互联网领域的重要国际刊物上发表了100余篇学术论文。

2013年第1—6期专题计划

1 自组织网络技术与应用
陈前斌 重庆邮电大学通信与信息工程学院院长

2 下一代互联网与IPv6技术演进
崔勇 清华大学计算机系网络所副所长

3 单波长Tb/s以上超高速光通信系统技术与器件
张成良 中国电信北京研究院副总工

4 大数据技术与应用
刘鹏 解放军理工大学教授

5 软件定义网络
王文东 北京邮电大学网络技术研究院副院长

6 移动互联网的发展趋势和技术方向
蒋林涛 工信部电信研究院科技委主任

DOI: 10.3969/j.issn.1009-6868.2013.02.001
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130204.1017.001.html

IPv6过渡技术场景分析

IPv6 Transition Technology

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0002-005

孙静文/SUN Jingwen
孙琪/SUN Qi
吴鹏/WU Peng
（清华大学下一代互联网实验室，北京 100084）
(Next Generation Internet Lab, Tsinghua University, Beijing 100084, China)

摘要： 分析了当前比较成熟的隧道和翻译过渡技术，主要包括主干网隧道过渡技术、接入网隧道过渡技术、接入网翻译过渡技术。针对过渡场景及其解决方案各不相同，认为在过渡过程中应当充分考虑网络的特征与现状。以主干网和接入网为主线，给出了各场景适用的技术方案，包括主干网的IPv6过渡场景、接入网的IPv6过渡场景、互联网内容提供商过渡场景。

关键词： 下一代互联网；过渡技术；过渡场景

Abstract: In this paper, we discuss more mature IPv6 transition technologies related to the backbone tunnel, access network tunnel, and access network translation. We suggest that it is necessary to consider the characteristics and status quo of the network in the transition process. We analyze the technology applicable to a backbone IPv6 transition scenario, an access network IPv6 transition scenario, and an Internet content provider transition scenario.

Keywords: next generation Internet; transition technology; transition scenarios

1 IPv6过渡技术的背景

IPv4互联网已在全球范围取得了巨大的成功，但随着互联网规模的持续增长，IPv4地址资源日益短缺。采用私有IP地址的网络地址转换（NAT）技术虽然可以缓解地址短缺压力，但是破坏了互联网端到端特性，影响了许多互联网应用的推广。IPv6具有地址空间巨大、改进了地址格式和路由汇聚特性等诸多优点，是下一代互联网中取代IPv4的最佳选择。然而IPv6与IPv4不兼容，因而在IPv4向IPv6过渡时期，IPv4与IPv6网络共存，带来了IPv4/IPv6互联与穿越等基本需求，产生了异构路由、地址映射、状态维护与可扩展性等问题。如何平稳实现互联网向IPv6过渡，已经成为全球下一代互联网发展的重大技术挑战之一。

针对IPv6过渡问题，目前国际上提出了许多过渡方案。这些技术方案针对不同的网络场景、用户的需求和过渡阶段提出，其技术思想也不尽相同。过渡技术从整体上可以划分为双栈技术、隧道技术和翻译技术。双栈过渡技术需要节点上并行维护IPv4和IPv6协议栈，代价高，不能解决IPv4地址耗竭的问题，因此对隧道和翻译过渡技术的研究在互联网工程任务组已经成为了重点和热点。

隧道技术通过对报文的封装、解封装，使得两个同构网络能够在一个异构网络的两边桥接起来从而实现相互通信。通过隧道技术使底层承载网络保持单栈，仅将隧道端点的地址簇边界路由器升级成双栈，即可实现数据报文跨异构网络的透传。隧道技术具有保持端到端特性，对终端上层应用透明，对底层网络独立的优势。早期的隧道方案包括手工配置隧道、6to4[1]、6PE[2]、ISATAP[3]、Teredo[4]等。近年来，IETF Softwire工作组针对主干网和接入网研究并制订了一系列的隧道过渡标准。

翻译技术是指将IP数据包在IPv4和IPv6两种协议簇之间的转换，其可以分为有状态和无状态两种。有状态翻译是指在翻译设备中需要保存动态信息或是在接收转发报文时需要新建或者修改翻译设备中的数据结构，反之，如果都不需要则称为无状态翻译。翻译技术能够直接、高效地实现IPv4/IPv6的互通。早期的翻译技术包括SIIT[5]、NAT-PT[6]等，目前翻译相关标准的制订主要在behave工作组进行。现阶段比较成熟的翻译技术包括NAT64[7]、IVI[8]等。由于翻译技术本身的限制以及IPv6地址空间远大于IPv4地址空间等因素，这些翻译技术对过渡场景的适用有着一定的局限性。

2 主干网隧道过渡技术

6to4用于解决IPv6孤岛穿越IPv4

网络出版时间：2013-02-04
收稿日期：2013-01-15
基金项目：国家“973”计划项目（2009CB320501）

网络实现相互通信，以及与IPv6 Internet通信。这是通过6to4边界路由器（连接IPv6子网和IPv4网络）采用无状态自动隧道实现。6to4将固定前缀2002::/16划分成多个IPv6网络前缀，而这些网络前缀在中继处无法聚合，将极大地扩张路由器路由表规模，从而带来IPv6可扩展性问题。

网状软线Softwire Mesh[9-11]采用的也是一种网状穿越场景下路由器到路由器的隧道机制，它应用的场景是内部IP（I-IP）主干网下外部IP（E-IP）客户网间的互联，包括IPv4承载IPv6（IPv4-over-IPv6）和IPv6承载IPv4（IPv6-over-IPv4）两类场景。Softwire Mesh通过扩展MP-BGP协议，将E-IP目的前缀与I-IP封装目的地址形成映射，实现跨异构网络路由信息传递。映射表的大小不会超过E-IP转发表，所以查表时间在可以接受的范围之内。网状软线保持了4/6独立编址，具有良好的可扩展性，可适用于大规模网络；支持多种隧道类型，包括IP-IP、GRE、MPLS、L2TP、IPsec等；通过扩展边界网关协议（BGP）提供了通用的信令机制支持；为主干网的过渡提供了重要的解决方案。

3 接入网隧道过渡技术

6RD[12-13]技术面向IPv6网络穿越IPv4接入网与IPv6互联网通信的场景，将6to4隧道应用到星型IPv6-over-IPv4场景中。6RD采用无状态封装，并且在6to4技术的基础上扩展，使用因特网业务提供商（ISP）实际可变前缀（NSP），按照NSP+{IPv4}::/的格式组成IPv6前缀，再生成IPv6地址。6RD具有无状态和简单的特性，但由于IPv6地址与IPv4地址耦合，6RD同样存在IPv6可扩展性的问题。

近年来，IPv6接入网逐渐增多，但用户和应用仍大量留在IPv4中。由于IPv6接入网中的IPv4应用或终端仍需要访问IPv4互联网，IPv4-over-IPv6场景逐渐增多。为缓解IPv4地址紧缺的压力，该场景下增加了IPv4地址共享的考虑。基于地址共享方式和有无状态维护两个维度，IETF提出3类IPv4-over-IPv6过渡方案[14]。

Dual-Stack Lite[15]方案采用运营商级别的网络地址转换（CGN）和有状态封装技术。在该机制中公有地址在CGN处集中控制，本地网络需要使用私有IPv4地址，而隧道汇聚点即CGN负责将封装在IPv6中的IPv4报文解封装并进行NAT转换。该方案能有效地缓解IPv4地址不足的问题，但处于较高位置的NAT会带来状态维护上的巨大开销，并且由于NAT的阻隔，外部网络无法主动发起对IPv6域内IPv4主机的访问。

4over6机制将公有地址的分配、共享和有状态封装有机地结合起来。Public 4over6[16]方案通过DHCPv4-over-IPv6技术跨越IPv6接入网动态地分配完整的公有IPv4地址[17]。用户获取公有地址后可以直接将IPv4报文经4over6隧道发送到运营商侧的隧道汇聚点，由汇聚点进行解封装和IPv4转发。由于该机制用户侧无NAT，因此支持外部网络对域内主机或服务器的访问。

4over6机制的状态维护从每流的规模降到每用户的规模，从而减轻隧道汇聚点的性能压力；将公有地址下放至用户侧，可以保持端到端的特性；没有应用层翻译的问题，并且可以保持通信的双向透明性[18]。

为了进一步提高IPv4地址的复用率，Public 4over6的扩展方案lightweight 4over6[19]方案可以根据实际需求分配适当的端口范围，通过划分端口段的方式来共享提高IPv4地址利用率。

IPv4-over-IPv6的第三类解决方案是4RD[20]和MAP-E[21]，采用的是包含地址共享的无状态封装方式。支持该类方案将IPv4地址和端口段信息内嵌入IPv6地址，通过IPv4/IPv6地址映射算法[22]实现双向地址解析。通过计算获取的地址进行路由和封装/解封装。通过将IPv4与IPv6地址进行结合，该机制实现了无状态维护，这使得该类方案在状态维护上优于DS-Lite和4over6方案。然而，IPv4地址与IPv6地址的耦合性将会给IPv4/IPv6网络的规划和运行带来较大的限制。

4 接入网翻译过渡技术

接入网翻译技术主要包括有状态和无状态翻译。NAT64采用有状态翻译方式，明确了IPv4主机地址映射到IPv6时使用固定前缀64:FF9B::/96。NAT64细化了DNS寻址的方法，将DNS翻译的功能独立为DNS64设备。NAT64方案的适用场景包括IPv6网络访问IPv4互联网、IPv6互联网访问IPv4网络、IPv6网络访问IPv4网络。状态维护的特性决定了它不可能适用于IPv6互联网访问IPv4互联网的访问场景。而对IPv4访问IPv6的4种场景，现有的一些机制并不可行。

IVI机制为无状态翻译技术采用网络专用的可变前缀（NSP），用NSP+IPv4地址+后缀的形式组建IPv6地址。IVI机制的路由可扩展和寻址问题都得到了解决，同时具有高性能、支持双向无状态通信的优点，适用于IPv6网络与IPv4网络互通的场景。然而IVI机制本身仍需要消耗IPv6用户量级的IPv4地址，因此其并不适合大规模部署。对此IVI又给出了支持IPv4地址复用的扩展方案，每个用户只占用1个IPv4地址的部分端口空间。这就需要主机端或用户终端设备（CPE）端进行相应的端口映射。

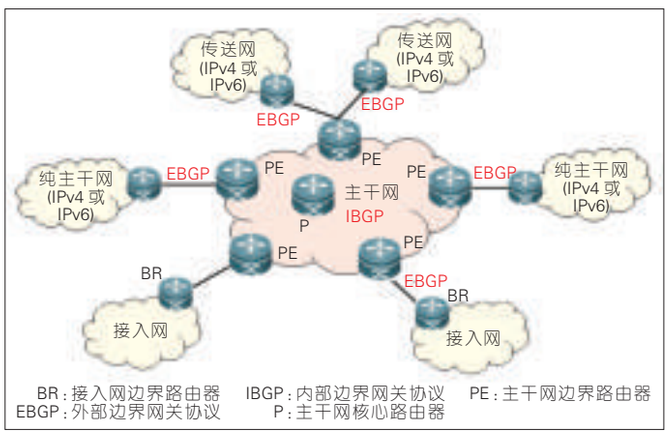
5 场景分析与适用技术方案

IPv6过渡符合下一代互联网的发展方向，但过渡场景及其解决方案各不相同，在过渡过程中应当充分考虑网络的特征与现状。本文以主干网和接入网为主线，对各种IPv4/IPv6

过渡的场景进行描述和分析，供ISP和ICP进行实际的决策和部署时作参考。

5.1 主干网的IPv6过渡场景

主干网向上连接为自己提供传输服务的上级供应商网络，向下可连接各个边缘网络，同时主干网还与同级的其他主干网相连[23]。主干网基本过渡场景如图1所示。主干网的主要任务是提供IPv4和IPv6的接入和传输服务，无论是IPv4分组还是IPv6分组都能够通过所接入的主干网转发到正确的目的地。为达到这一目的，可以将主干网升级成双栈，但是开销巨大；对于单栈的主干网，则需要借助过渡技术来同时提供IPv4和IPv6的传输服务。



◀图1 主干网基本过渡场景

主干网的过渡场景主要包括IPv6客户网穿越IPv4主干网与IPv6 Internet互通和IPv4客户网穿越IPv6主干网与IPv4 Internet互通两类。现有的一些IPv4主干网可能需要提供IPv6传输服务，出现IPv6-over-IPv4过渡场景，此时可以只将主干网边界路由器升级成双栈，而核心路由器仍保持IPv4单栈，使用6to4或Softwire Mesh技术等来提供IPv6传输服务。

随着下一代互联网的发展，纯IPv6主干网逐渐建成。但是IPv4客户网以及IPv4 Internet仍会长期存在，为促进IPv6主干网的利用率，IPv4-over-IPv6过渡场景越来越多，可以利用Softwire Mesh技术使IPv6主干网能够提供IPv4的传输服务。这样可以将IPv4流量引入到IPv6主干网中，减轻IPv4网络的负担，并逐渐取代IPv4网络，也可以积累较多的IPv6运营经验。

5.2 接入网的IPv6过渡场景

接入网一方面向下为用户侧设备提供接入，另一方面向上接入Internet。在过渡时期，接入网需要为各种终端用户设备和服务器设备提供IPv4和IPv6接入服务。接入网的基本场景如图2所示。接入网络的用户可以是直接以与接入网相同地址协议簇形式接入，也可能是用户侧设备（CPE）接入网络并且在其后连接了不同地址协议簇的主机；在接入网络另一侧的接入网边界路由器（BR）接入Internet。若接入网为双栈，则可以利用已有的IPv4和IPv6同时为用户提供双栈的接入环境。但是采用双栈开销巨大，而且并没有解决IPv4地址紧缺的问题，因此在过渡时期通过单栈的接入网结合过渡技术为用户提供IPv4和IPv6服务。

由于IPv4地址的耗竭以及中国国家的大力推动，IPv6接入网将逐渐建成。在此IPv6接入网中的用户会有与IPv4或IPv6 Internet交互的需求。用户设备可以直接以IPv6接入IPv6接入网，此时可以正常的与IPv6 Internet通信；但是访问IPv4 Internet时，出现异构网络直连的过渡场景。IPv6应用访问IPv4 Internet场景如图3所示。此时需要IVI或NAT64等翻译技术来实现。

IPv6接入网中的用户也可能使用IPv4应用或者一部分IPv4用户通过CPE接入IPv6网络。若用户访问IPv4 Internet则会出现如图4所示的IPv4-over-IPv6过渡场景。此时可以借助Public 4over6、Lightweight 4over6等隧道解决方案，用户设备（CPE或直连设备）作为隧道的发起点（TI），接入网另一侧的BR设备作为隧道的汇聚点（TC）。用户的IPv4分组在TI上根据具体方案进行封装，TI再将封装好的分组通过IPv6接口发往TC，TC收到此分组后进行解封装发往IPv4 Internet。另外，该场景下可以使用MAP-E无状态隧道技术或者是两次翻译方案。

若IPv6资源的丰富，而IPv6接入网中某些终端的上层应用由于技术等原因仍然使用的是IPv4，则会产生IPv4应用与IPv6 Internet进行通信的场景，如图5所示。在这种场景下需要进行用户侧的翻译。根据翻译点的不同所使用的翻译技术也有所不同。如果是直连设备，则翻译点就在直连设备上。此时可以使用主机侧翻译技术[24]；如果是CPE设备所连接的IPv4主机上的IPv4应用希望与IPv6 Internet通信，那么翻译点就在

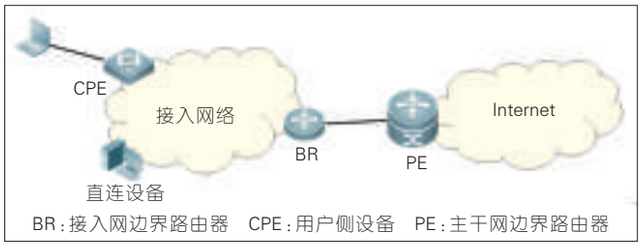
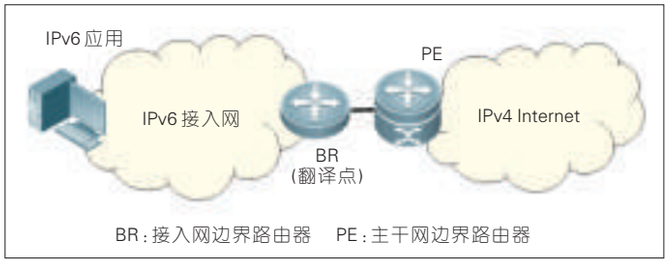
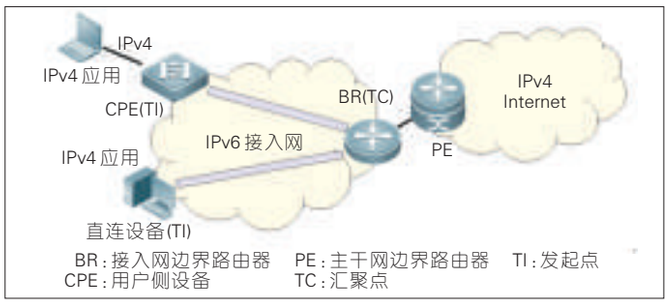


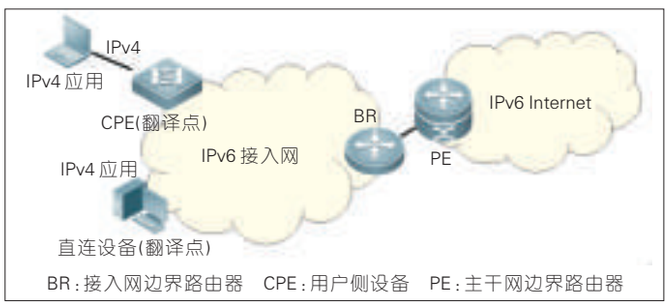
图2▶ 接入网基本过渡场景



◀图3
IPv6应用访问IPv4 Internet场景



◀图4
IPv4应用或终端穿越IPv6接入网与IPv4 Internet互通场景



◀图5
IPv4应用与IPv6 Internet互通场景

CPE上。

现有的已建设的接入网一般是IPv4单栈，出现的过渡场景与IPv6接入网的恰好相反。在此IPv4接入网中的用户会有与IPv6和IPv4 Internet交互的需求。用户设备可以直接以IPv4接入IPv4接入网，此时可以正常地与IPv4 Internet通信，但是访问IPv6 Internet时，会出现异构网络直连的过渡场景，此时可以采用IVI等翻译技术来实现。IPv4接入网中的用户也可能使用IPv6应用或者一部分用户使用IPv6通过CPE接入IPv4网络，此时用户若访问IPv4 Internet则会出现异构网络直连的过渡场景，可以借助IVI等翻译技术来实现，但是并不提倡这种在全网都是IPv4的情况下只将终端或应用升级成IPv6的场景；若用户访问IPv6 Internet则会出现IPv4-over-IPv6过渡场景，此时可以借助6rd等隧道过渡技术来实现互访。

5.3 互联网内容提供商过渡场景

对于互联网内容提供商(ICP)而言需要为Internet提供服务，并且由于ICP的服务系统构架升级相对普通用户要困难得多，需要对其进行单独讨论。首先单栈接入网络应当为ICP提供双栈传输能力，使得ICP既可以提供IPv4的服务又可以提供IPv6的服务。单栈接入网使用隧道技术为ICP提供双栈环境如图6所示。假设ICP接入的是IPvX单栈网络，则其可以直接使用与接入网地址协议簇相同的IPvX Internet提供服务，而使用隧道技术来为IPvY Internet提供服务。

另一方面，考虑到ICP服务系统构架由于技术升级等原因可能是单栈的，则可以使用翻译技术来为异构的网络提供服务。IPv4 ICP为IPv6 Internet提供服务如图7所示。现有的ICP实际上大多数只支持提供IPv4的服务，随着IPv6规模的增大，这些ICP希望在没有升级之前为IPv6 Internet提供服务。在这种过渡场景下可以通过在ICP侧与IPv6网络的接入处部署NAT64有状态的翻译设备实现互通。与此相反的是ICP升级自己的服务，最终会出现只能提供IPv6服务的系统架构。此时IPv4 Internet可能还有一定的规模，则IPv6 ICP需要为IPv4 Internet提供服务。在这种需求场景下，可以通过在ICP侧与IPv4网络的接入处部署一个无状态的翻译设备，IVI可以很好地适应ICP在这种场景下的需求。

6 结束语

由IPv4向IPv6平滑过渡是中国发展下一代互联网的重要步骤之一，也是当前国际互联网发展的大势所趋。本文综合介绍了在当前IPv6过渡的环境下，全球的主流过渡技术，

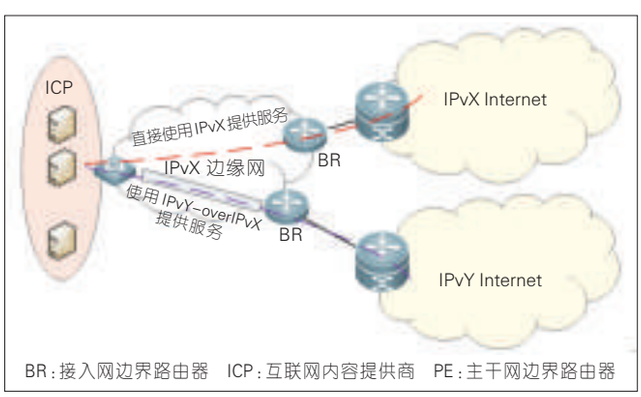
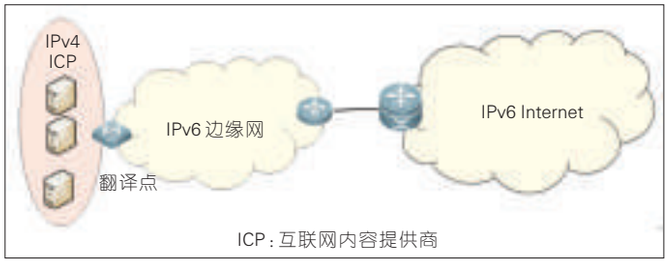


图6▶
单栈接入网使用隧道技术为ICP提供双栈环境



◀图7
IPv4 ICP 为 IPv6 Internet 提供服务

并对各个典型过渡场景进行了分析。针对不同的过渡场景，运营商和互联网内容提供商可以选择不同的过渡技术实现 IPv4 与 IPv6 的互通。通过各方的共同努力以推动下一代互联网的发展。

参考文献

[1] CARPENTER B, MOORE K. Connection of IPv6 domains via IPv4 Clouds [S]. RFC3056. 2001.
[2] CLERCQ J D, OOMS D, PREVOST S, et al. Connecting IPv6 is lands over IPv4 MPLS using IPv6 provider edge routers (6PE) [S]. RFC4798. 2007.
[3] TEMPLIN F, GLEESON T, THALER D. Intra-site automatic tunnel addressing protocol (ISATAP) [S]. RFC5214. 2008.
[4] HUITEMA C. Teredo: Tunneling IPv6 over UDP through network address translations (NATs) [S]. RFC4380. 2006.
[5] NORDMARK E. Stateless IP/ICMP translation algorithm (SIIT) [S]. RFC2765. 2000.
[6] AOUN C, DAVIES E. Reasons to move the network address translator – protocol translator (NAT-PT) to historic status [S]. RFC4966. 2007.
[7] BAGNULO M, MATTHEWS P, VAN BEIJNUM I. Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers [S]. RFC6146. 2011.
[8] LI X, BAO C, CHEN M, et al. The CERNET IVI translation design and deployment for the IPv4/IPv6 coexistence and transition [S]. RFC6219. 2011.
[9] WU J, CUI Y, LI X, et al. The transition to IPv6, Part I: 4over6 for the China Education and Research Network [J]. IEEE Internet Computing, 2006,10(3):80-85.
[10] CUI Y, WU J, LI X, et al. The transition to IPv6, Part II: The softwire mesh framework solution [J]. IEEE Internet Computing, 2006, 10(5):76-80.
[11] WU J, CUI Y, METZ C, et al. Softwire mesh framework [S]. RFC5565. 2009.
[12] DESPRES R. IPv6 rapid deployment on IPv4 infrastructures (6rd) [S]. RFC5569. 2010.
[13] TOWNSLEY M, TROAN O. IPv6 rapid deployment on IPv4 infrastructures (6rd) [S]. RFC5969, 2010.
[14] WU P, CUI Y, WU J, et al. Transition from IPv4 to IPv6: A state-of-the-art survey [J]. IEEE Communications Surveys and Tutorials, to be published.
[15] DURAND A, DROMS R, WOODYATT J, et al. Dual-stack lite broad-band deployments following IPv4 exhaustion [S]. RFC6333. 2011.
[16] CUI Y, WU J, WU P, et al. Public IPv4 over access IPv6 network [R]. draft-cui-softwire-host-4over6-05. 2011.
[17] CUI Y, WU J, WU P. DHCPv4 over IPv6 transport [R]. draft-ietf-dhc-dhcpv4-over-ipv6-05. 2012.
[18] CUI Y, DONG J, WU P, et al. Tunnel-based IPv6 transition [J]. IEEE Internet Computing, to be published.
[19] CUI Y, WU J, WU P, et al. Lightweight 4over6 in access network [R]. IETF draft. 2011.
[20] DESPRES R, PENNO R, LEE Y, et al. IPv4 residual deployment via IPv6 -- A unified stateless solution (4rd) [R]. IETF draft. 2012.
[21] TROAN O, DEC W, LI X, et al. Mapping of address and port (MAP) [R]. IETF draft, 2012.
[22] TROAN O, MATSUSHIMA S, MURAKAMI T, et al. Mapping of address and port(MAP) [R]. IETF draft. 2012.
[23] 崔勇, 董江, 徐明伟, 等. IPv6过渡场景分析 [R]. 中国通信行业研究报告, 2012B7. 2012.
[24] TSUCHIYA K, HIGUCHI H, ATARASHI Y. Dual stack hosts using the bump-in-the-stack technique (BIS) [S]. RFC2767. 2000.

作者简介

孙静文，北京邮电大学网络技术研究院宽带网研究中心在读硕士研究生；研究方向为 IPv4/IPv6 过渡技术、下一代互联网；已发表学术论文1篇。

孙琪，北京邮电大学网络技术研究院宽带网研究中心在读硕士研究生；研究方向为 IPv4/IPv6 过渡技术、下一代互联网；已发表学术论文1篇。

吴鹏，清华大学计算机科学与技术系在读博士研究生；研究方向为 IPv4/IPv6 过渡、下一代互联网、网络体系结构；已发表学术论文5篇。

综合信息

中兴通讯创新方案落地欧洲 Mtel IPTV 多屏融合平台商用

【本刊讯】2013年3月6日消息，2013年2月中兴通讯宣布其为保加利亚运营商 Mtel 建设的 IPTV 多屏融合平台正式商用。该 IPTV 平台将为保加利亚用户提供 TV、PC、Tablet 和手机的多屏终端 TV 业务，Mtel 将通过固网/移动宽带为用户提供更有竞争力的视频互动体验，从而提升宽带价值。

Mtel 是保加利亚最大的运营商，是 TAG 奥地利电信集团的全资子公司。Mtel 于2011年9月与中兴通讯携手建设实验局，12月签署为期5年的合作协议。按照协议，中兴通讯将为 Mtel 提供包含中间件、DRM、CDN 等端到端的定制化解决方案，分阶段完成业务部署。

凭借全球成熟的商用经验，中兴通讯将协助 Mtel 大力发展宽带用户，提升用户体验，为用户提供“任何时间、任何网络、任何终端的任何多媒体服务”。中兴通讯通过研究保加利亚用户的使用习惯推出 IPTV 多屏融合业务，使用户不仅可以“躺在沙发上看电视”，还可以通过网络在书房、办公室、上班路上等场合随时随地享受视频带来的乐趣。

DOI: 10.3969/j.issn.1009-6868.2013.02.002
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130228.1703.003.html

统一的IPv4/IPv6翻译与封装过渡技术——IVI/MAP-T/MAP-E

IVI/MAP-T/MAP-E: Unified IPv4/IPv6 Stateless Translation and Encapsulation Technologies

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0007-005

包丛笑/BAO Congxiao
李星/LI Xing
（清华大学 网络科学和网络空间研究院，北京 100084）
(Network Science and Cyberspace Research Center, Tsinghua University, Beijing 100084, China)

摘要： 采用无状态翻译技术和双重翻译/封装技术IVI/MAP-T/MAP-E定义IPv4/IPv6地址的映射算法和协议翻译/封装算法，在保持互联网地址端对端的特性的基础上，使IPv4和IPv6互联互通，支持由IPv4和IPv6双向发起的访问。无状态翻译技术和双重翻译/封装技术有着不可替代的综合优势，是IPv4/IPv6过渡的必经之路。

关键词： 无状态翻译；地址映射；协议翻译

Abstract: Stateless translation and stateless double translation/encapsulation technologies (IVI/MAP-T/MAP-E) define the address mapping and protocol translation/encapsulation algorithms between IPv4 and IPv6. IVI/MAP-T/MAP-E technologies maintain end-to-end address transparency between IPv4 and IPv6 and support communication between IPv4-only and IPv6-only end systems. IPv4/IPv6 translation is a critical technique for IPv4/IPv6 transition.

Key words: stateless translation; address mapping; protocol translation

基于IPv4协议的互联网是世界上最重要的信息基础设施，但是只有2^{32}个地址空间的IPv4地址已经分配完毕。为了解决IPv4地址耗尽问题，目前有两种技术路线：采用IPv4地址翻译技术（NAT44）或升级到IPv6。NAT44是IPv4公有地址与IPv4私有地址之间的有状态翻译技术，已经非常成熟，但由于其破坏了端到端特性，只能支持单向发起的通信。多年以来，NAT44在用户接入端被广泛使用，但在核心网络上使用时，需要在NAT44翻译器上维护大量状态。为了长远地解决IP地址的问题，建设下一代IPv6网络，发展IPv6信息资源，发展IPv6用户势在必行。10多年前，业内已经认识到IPv4地址枯竭问题，发明了下一代互联网协议IPv6，该协议具有2^{128}个地址空间，从根本上解决了地址耗尽的问题。因特网工程任务组（IETF）最早推荐从IPv4向IPv6过渡采用双栈技术和隧道技术，全世界很多运营商在不同规模上进行了IPv6的试验，若干信息提供商也提供了IPv6的服务[1]。但截至2012年，全世界IPv6网的流量平均不到IPv4的1%。实践表明，升级到双栈不仅没有给运营上带来直接的收益，反而影响了用户的体验。这就是为什么双栈和隧道技术应用10多年，却没有推动完成IPv4互联网向IPv6互联网过渡的原因。从根本上看，网络的价值在于其用户数。对于新建的IPv6网络，其用户数不可能与IPv4互联网上的用户数可比，如果IPv6的用户不能与IPv4的用户互联互通，则IPv6网络没有任何存在的价值。因此过渡的核心问题是新建IPv6网络必须与IPv4互联网互联互通。两种不同协议之间的互联互通，只能通过翻译技术解决，但是由于IETF在设计IPv6协议时，没有充分意识到与IPv4协议兼容的重要性，具有很高的技术难度。随着纯IPv6网络建设案例的增多和研究的深入，IETF在IPv4/IPv6翻译技术，特别是无状态翻译技术取得了突破性进展，形成了系列RFC标准和工作组草案，为IPv4到IPv6过渡提供了新的技术方案。

收稿日期：2012-12-23
网络出版时间：2013-02-28

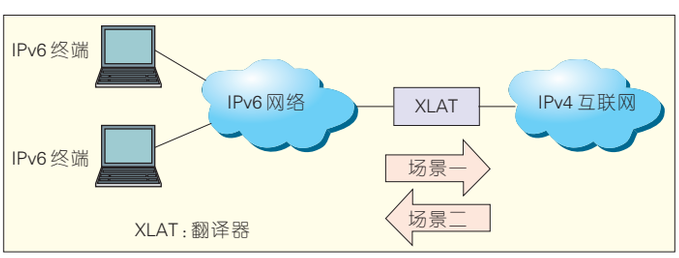
1 无状态IPv4/IPv6翻译技术

互联网的基本特性为“无连接”的体系结构，路由器不需要维护状态，IPv4/IPv6翻译器本身也是一个路

由器，因此无状态的IPv4/IPv6翻译器对于运营商来讲更具有价值。同时，无状态IPv4/IPv6翻译（IVI）技术具有可扩展性、可管理性、安全性好的特点，并支持双向发起的通信。IVI的名称借用了罗马数字的表示方法。在罗马数字中IV表示4，VI表示6，IVI表示IPv4和IPv6的互联互通。

1.1 IPv4/IPv6翻译技术的应用场景

由于IPv4的地址空间为2^{32}，IPv6的地址空间为2^{128}，极其悬殊，因此不加限制条件的IPv4/IPv6翻译器在理论上讲是不可行的。在IETF标准RFC6144中定义了IPv4/IPv6翻译的8个应用场景。翻译器的两边一边是IPv4另一边是IPv6。其变化之一在于哪一边是自己控制的网络，哪一边是互联网。其变化之二在于哪一边发起通信。无状态IPv4/IPv6翻译技术的应用场景如图1所示[2]。场景一为IPv6网络上计算机发起对IPv4互联网上计算机的访问，场景二为IPv4互联网上计算机发起对IPv6网络上计算机的访问。



◀图1 无状态IPv4/IPv6翻译技术的应用场景

场景一和场景二意味着新建纯IPv6网络，通过翻译器XLAT为IPv6用户提供对IPv4互联网的通信。无状态IPv4/IPv6翻译技术可以适应于场景一和场景二，有状态IPv4/IPv6翻译技术只能适应于场景一。IPv4/IPv6翻译技术的核心是需要解决地址映射和协议翻译问题。

1.2 地址映射和域名翻译

由于IPv4和IPv6的地址空间差距巨大，用IPv6表示IPv4是毫无问题的，可以通过无状态的映射方法，映射后的IPv6地址称为转换地址。用IPv4表示IPv6是难点，可以动态维护映射表，作有状态的地址映射，或在IPv6地址中选择一个子空间通过无状态的方法与IPv4地址映射。映射后的IPv6地址称为可译地址。所有这些地址映射算法均在IETF标准RFC6052中定义。嵌入了IPv4地址的IPv6地址格式如图2所示[3]。

其中Prefix是IPv6网络前缀，根据不同的前缀长度，嵌入IPv4地址，并且在64至71位间保持为全0，以便与IPv6地址结构中的u-bit兼容。Suffix为后缀，在基本的地址映射中为全0，预留给传输层端口的编码，以便把一个IPv4地址映射为若干IPv6地址，达到高效地、无状态地复用稀缺的公有IPv4地址资源的目的。此外，RFC6052要求转换地址和可译地址使用同样的前缀（Prefix），从而可以自动获得最优路由。

当纯IPv6计算机发起对IPv4互联网的访问时，必须获得相应的转换地址，这个工作由域名翻译器DNS64根据上述映射算法完成，由RFC6147定义[4]。DNS64是同时接入IPv4网络和IPv6网络的域名服务器（DNS），能够把A记录动态翻译成AAAA记录。具体步骤为纯IPv6计算机通过DNS64查询所需域名的AAAA记录，如AAAA记录存在，则DNS64直接返回AAAA记录给纯IPv6计算机；如AAAA记录不存在，则DNS64查询域名对应的A记录，并根据RFC6052定义的映射算法，生成AAAA记录返回给纯IPv6计算机。无状态的翻译器支持IPv4互联网发起的通信，在这种情况下，需要静态配置的DNS46，当IPv4互联网上的用户发起对IPv6计算机的访问时，DNS46则返回对应IPv6计算机AAAA对应的A记录[5]。

1.3 协议翻译

两种不同协议栈之间的互联互通必须解决的第二个问题是协议翻译，所庆幸的是IPv4和IPv6之间协议的差距并不很大，是可译的。具体协议翻译算法由RFC6145定义，包括版本号映射，IPv4的服务类型与IPv6的流量等级映射，IPv4的总长与IPv6的载荷长度映射，IPv4的存活期与IPv6的转发计数映射，IPv4的传输层协议与IPv6的下一个头映射，IPv4地址和IPv6地址映射等[6]。IPv4与IPv6协议翻译的最大难点是分片处理，因为IPv4可以支持路由器分片或端系统分片，但IPv6仅支持端系统分片。对于IPv4已经分片的分组，IPv6必须增加分片扩展头，以便端系统重组。此外，IPv4和IPv6网络所能支持的最大传输单元的大小（MTU）是不同的，同时由于IPv4的基本头为20个字节，而IPv6的基本头为40个字节，因此在从IPv4到IPv6的翻译过程中必然遇到MTU超出的问题。此外，IPv4的传输控制协议（ICMP）和IPv6的传输

PL	0	32	40	48	56	64	72	80	88	96	104	112	120
32	Prefix	v4 (32)				u	Suffix	zero					
40	Prefix	v4 (24)				u	(8)	Suffix	zero				
48	Prefix	v4 (16)				u	v4 (16)	Suffix	zero				
56	Prefix	(8)				u	v4 (24)	Suffix	zero				
64	Prefix					u	v4 (32)	Suffix	zero				

Prefix：前缀　Suffix：后缀

▲图2 嵌入了IPv4地址的IPv6地址格式

控制协议(ICMPv6)也有很多不同,需要分别处理。

值得指出的是,当IPv6网络中的路由器(通常并不使用可译地址)返回ICMPv6分组时,翻译器无法找到对应的IPv4地址,造成翻译后的ICMP分组的源地址不可溯源。IETF最新发布的RFC6791定义了对这个问题的有效处理方法[7]。

RFC6145也是有状态IPv4/IPv6翻译器所依据的协议翻译算法。有状态IPv4/IPv6翻译器中的状态维护技术由RFC6146定义,主要规定了IPv6地址和端口到IPv4地址和端口的动态映射表的生成、维护和销毁算法[8]。

2 无状态双重翻译/封装技术(MAP系列)

IPv4/IPv6翻译技术可以使IPv4和IPv6互联互通,但仍有3个问题需要解决。第一个问题是由于IPv4地址耗尽问题,无状态IPv4/IPv6翻译必须能够复用公有IPv4地址以高效使用IPv4地址资源;第二个问题是目前有的应用程序并没有IPv6的版本(如Skype),也有的应用程序嵌入了地址的信息(如Ftp);第三个问题是对于终端用户往往需要分配一个64位的前缀,而不是单个的IPv6地址。无状态双重翻译/封装系列技术MAP-T/MAP-E可以解决这些问题。MAP是Mapping Address and Port的缩写,意指无状态地对地址和端口进行复用,与双重翻译(MAP-T)或封装(MAP-E)技术组合,可以解决无状态复用公有IPv4地址的问题和上述的应用程序问题,同时可以支持前缀分配。MAP-T/MAP-E目前是IETF的工作组文档[9-10],其他相关的工作组文档还有DHCPv6扩展[11]和部署考虑[12]。

2.1 双重翻译模式的应用场景

无状态双重IPv4/IPv6翻译模式(MAP-T)的应用场景如图3所示。

图3中,核心翻译器称为BR,对于IPv6为路由器,对于IPv4为可以复用IPv4地址的IPv4/IPv6翻译器。第二次翻译在家庭网关CE上进行,CE对于IPv6为路由器,对于IPv4为IPv4/IPv6翻译器,并且根据下述端口映射算法对于传输层的端口进行映射。IPv6接入网部署认证和IPv6前缀分配设备AAA数据库和DHCPv6服务器。IPv6接入网上可以部署使用可译地址的纯IPv6服务器,通过CE或BR的一次翻译能够对于用户的纯IPv4终端和IPv4互联网上的用户提供服务。IPv6接入网的用户设备接到家庭网关上,可以是纯IPv4终端设备,它通过CE一次翻译访问网内纯IPv6服务器的信息资源,同时它通过CE和BR双重翻译访问IPv4互联网上的信息资源,并与其他用户互访。该用户设备还可以是IPv4/IPv6双栈终端设备,直接访问网内和IPv6互联网上的信息资源,并通过CE和BR双重翻译访问IPv4互联网上的信息资源,也能与其他用户互访。该用户设备也可以是纯IPv6终端设备,直接访问网内和IPv6互联网上的信息资源,通过BR一次翻译访问IPv4互联网上的信息资源,并与其他用户互访。

2.2 封装模式的应用场景

封装模式(MAP-E)的应用场景如图4所示。

图4中,核心封装/解封装器称为BR,对于IPv6为IPv6路由器,对于IPv4为可以复用IPv4地址的IPv4 over IPv6封装/解封装器。家庭网关CE对于IPv6为路由器,对于IPv4为IPv4 over IPv6封装/解封装器,并且根据下述端口映射算法对于传输层的端口进行映射。IPv6接入网部署认证和IPv6前缀分配设备AAA数据库和DHCPv6服务器。IPv6接入网的用户设备接到家庭网关上,可以是纯IPv4终端设备,它通过CE和BR封装/解封装访问IPv4互联网上的信息资源,并与其他用户互访;该用户设备还可以是IPv4/IPv6双栈终端设备,直接访问网内和IPv6互联网上的信息资源,并过CE和BR封装/解封装访问IPv4互联网上的信息资源,也能与其他用户互访。值得指出的是,封装模式MAP-E无法支持在IPv6接入网内部署纯IPv6服务器,也不支持可以与IPv4互联网互联互通的纯IPv6终端设备。

2.3 端口映射

MAP的核心技术之一是无状态地址和端口映射算法,其思想是利用16位的传输层(传输控制协议(TCP)、数据报协议(UDP))端口对于IPv4地址进行扩展。不复用IPv4地址时,一个终端设备可用的并发TCP或UDP的端口数为65 536;如复用比为16,则一个终端可用的并发TCP或UDP的端口数为4 096;如复用比为128,则一个终端可用的并发TCP或UDP的端口数为512。根据统计,一个普通终端的并发TCP或UDP的端口数为有限的,因此可以利用无

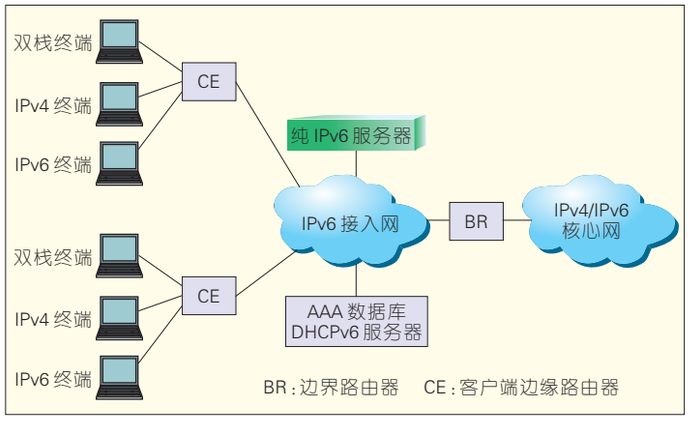
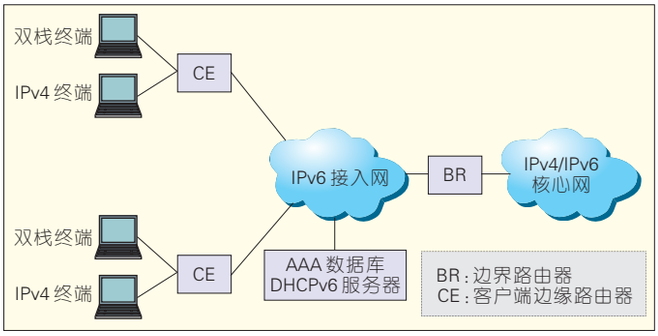


图3 无状态双重IPv4/IPv6翻译模式的应用场景



◀图4 MAP-E的应用场景

状态地址和端口映射算法高效率地复用公有IPv4地址资源。在使用无状态地址和端口映射算法时，需要给每一个终端定义一个端口标识集（PSID），端口标识集和可用端口的映射关系由扩展的模算法来决定。

扩展的模算法的定义为：

（1）给定PSID，该端系统可以使用的传输层端口P为：$P=R\times M\times j+M\times K+i$，其中$R$为复用比，$M$为连续端口数，$i$和$j$为整数变量。

（2）给定传输层端口P，该端系统的PSID为：$P=\mathrm{floor}(P/M)\%R$，其中floor为只舍不入的取整算法，%为常规定义的模运算符。

扩展的模算法是一个适应性很广的算法，即可以使持有不同PSID终端所使用的传输层端口在整个端口空间均匀分布，也可以按块分布，还可以制订每一块包含的连续端口数量。此外，扩展的模算法还可以支持类似与无分类地址域间路由（CIDR）类似的地址聚类，即对应于特定的PSID，可以定义PSID长度，对于可用端口进行聚类使用。

通过扩展的模算法，在给定复用比、连续端口数量和端口聚类长度的条件下，可以通过PSID的值计算出特定终端可以使用的所有的TCP或UDP端口；也可以对于任意给定端口计算出对应的PSID，实现端系统的无状态公有IPv4地址复用，因而可以极大地减小管理开销，并极大地提高安全性和可溯源性。由于ICMP和ICMPv6没有源和目标端口的域，只有标识域（ID），因此要对标识域ID作扩展的模算法映射。

2.4 地址格式

MAP的地址格式是RFC6052的扩展，如图5所示。

与RFC6052的区别主要有：

（1）MAP的地址格式是RFC6052当Prefix长度为64的一个特例，其Prefix里包含IPv6 Prefix、EA-bits（由IPv4子网标识和PSID组成，用于唯一标识不同的用户）和Subnet-id（用于标识一个用户使用的大于等于/64的IPv6子网）。

（2）在MAP中Suffix不为0，而是嵌入了PSID。

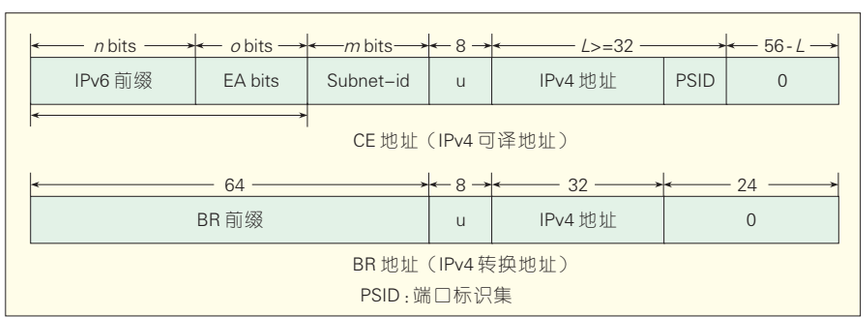
（3）MAP对于转换地址和可译地址使用不同的Prefix，以解决为终端用户分配前缀，而不是响应单个地址的要求。

利用EA-bits可以为每一个家庭网关CE分配唯一的Prefix，不使用EA-bits，而为每个CE分配不同的Prefix也可以达到同样的目的。采用EA-bits的好处是可以进行地址聚类，可扩展性好；不采用EA-bits的好处是IPv6前缀与IPv4地址独立。这两种方法各有优缺点，可以根据不同需要进行选择。

2.5 统一双重翻译和封装模式的机制

无状态双重IPv4/IPv6翻译可以支持纯IPv4应用程序（如Skype），同时对于嵌入IP地址的应用程序（如Ftp）也不需要IPv4/IPv6之间的应用层网关（ALG），此外双重翻译不需要DNS64和DNS46。无状态双重翻译可以看成是具有头压缩功能的、无状态IPv4 over IPv6的封装技术。无状态双重翻译技术（MAP-T）和无状态封装技术（MAP-E）采用同样扩展的模算法和同样的地址格式（在封装模式下BR地址可以蜕化为单个地址），因此具有众多的相似性，唯一的不同是数据流处理模式。在双重翻译模式（MAP-T）下数据流的处理依据为翻译，由RFC6145定义，在封装模式（MAP-E）下数据流的处理为数据封装，由RFC2473定义[13]。

MAP-T模式的优点是可以蜕化为一次翻译，有利于过渡到纯IPv6网络，但仍然保持与IPv4互联网的互联互通。同时，在IPv6接入网内的IPv6数据报文没有封装的数据结构，可以使用IPv6路由器上的所有网络层和传输层的管理和控制功能，而MAP-E必须对于数据报文进行解封装，才能进行管理和控制。MAP-E模式的优点是可以完全保持IPv4报文承载的所有信息，同时不需要对传输层的校验和进行修改。由于RFC2473定义的封装模式与传输层的TCP、UDP均由IPv6头结构的下一个头定义，因



▲图5 MAP地址格式

此，只有从IPv4到IPv6的处理需要定义采用翻译模式还是封装模式，从IPv6到IPv4的处理可以根据下一个头自动适应性完成翻译或封装模式的选择。因此，MAP-T和MAP-E可以根据需求灵活配置，其分析参见MAP测试文档[14]。

2.6 统一无状态/用户状态/有状态

无状态是指IPv4/IPv6地址和传输层端口之间的映射关系完全由算法决定，设备不需要维护映射状态表。有状态是指IPv4/IPv6地址和传输层端口之间的映射关系根据会话的5元组动态生成，设备需要维护动态生成的映射状态表。用户状态是指IPv4/IPv6地址和传输层端口之间的映射关系对于各个用户定义，设备只需要维护用户映射状态表。无状态翻译技术不仅可以与无状态封装技术统一起来，也可以与有状态的翻译技术NAT64和有状态的隧道技术Dual-stack Lite[15]统一起来。因此MAP-T/MAP-E家庭网关CE，不经任何修改就可以与有状态翻译器NAT64或Dual-stack Lite的AFTR完成有状态双重翻译或有状态隧道的功能。由于无状态和有状态是两个极端的情况，MAP-T/MAP-E家庭网关CE也可以不经任何修改支持任何用户状态的场景。

3 过渡路线图

虽然IPv4地址已经分配完毕，但全世界IPv6的普及率仍然非常低。为了保证全球互联网的健康和可持续发展，必须制订正确的过渡路线图。10年前IETF制订的“以双栈为主，辅之以隧道，在没有其他选择时用翻译”的策略值得反思，理由为：

（1）这一政策在过去10余年里并没有完成从IPv4到IPv6的过渡。

（2）对于中国这样的国家，已经没有更多的IPv4公有地址实施双栈，而通过NAT44利用私有地址实施双栈并不能鼓励IPv6的过渡。

随着无状态翻译技术（IVI）和无状态双重翻译技术（MAP）的成熟，我们建议应建设纯IPv6网络，实施“以翻译技术为主，辅之以封装，在没有其他选择时用双栈”的策略。具体技术方案为：

（1）新建纯IPv6网络，当通信的对端也为IPv6是，采用IPv6通信。

（2）当通信的对端为IPv4是，优先采用一次无状态IPv4/IPv6翻译技术进行通信。

（3）当应用程序不支持IPv6，或应用程序嵌入IPv4地址时，采用无状态双重IPv4/IPv6翻译技术进行通信。

（4）当需要保持IPv4报文所有的信息，或处理传输层加密的报文，采用封装技术进行通信。

（5）在过渡的中后期，双重翻译将无缝地退化成一次翻译，最终关闭一次翻译器，进入纯IPv6时代。

采用以上建议的过渡线路图，可以使我们自己的网络率先过渡到IPv6，并高效地利用公有IPv4地址资源与IPv4互联网互联互通，从而在IPv4到IPv6的过渡过程中保持主动。这一技术方案完全符合中国发展下一代互联网的路线图和时间表，即“在2011—2015年的过渡阶段，政府引导全社会向IPv6过渡，IPv4与IPv6共存，新建网络必须为IPv6并实现与IPv4的互通”。目前无状态IPv4/IPv6翻译技术（IVI）已经发布5个IETF的RFC标准，MAP技术已经形成4个IETF工作组草案。IVI技术已有思科、中兴通讯、华为等设备厂家的产品支持，并在CNGi-CERNET2上正常运行2年以上。MAP技术已有思科等设备厂家的产品正式发布，得到意大利电信、日本软银、德国电信、美国Charter等多家国际运营商的支持和关注，产业链正在逐步形成。作为唯一能够使IPv4和IPv6互联互通的无状态翻译技术和双重翻译技术IVI/MAP，预计在近几年会得到大发展。

参考文献

[1] NORDMARK E, GILLIGAN R. Basic transition mechanisms for IPv6 hosts and routers [S]. RFC 4213. 2005.
[2] BAKER F, LI X, BAO C, et al. Framework for IPv4/IPv6 Translation [S]. RF C6144. 2011.
[3] BAO C, HUITEMA C, BAGNULO M, et al. IPv6 addressing of IPv4/IPv6 translators [S]. RFC 6052. 2010.
[4] BAGNULO M, SULLIVAN A, MATTHEWS P, et al. DNS64: DNS extensions for network address translation from IPv6 clients to IPv4 servers [S]. RFC 6147. 2011.
[5] LI X, BAO C, CHEN M, et al. The CERNET IVI translation design and deployment for the IPv4/IPv6 coexistence and transition [S]. RFC 6219. 2011.
[6] LI X, BAO C, BAKER F. IP/ICMP translation algorithm [S]. RFC 6145. 2011
[7] LI X, BAO C, WING D, et al. Stateless source address mapping for ICMPv6 packets [S]. RFC 6791. 2012.
[8] BAGNULO M, MATTHEWS P, VAN BEIJNUM I. Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers [S]. RFC 6146. 2011.
[9] LI X, BAO C, DEC W, et al. Mapping of address and port using translation (MAP-T) [R]. draft-ietf-softwire-map-t-00. 2012.
[10] TROAN O, DEC W, LI X, et al. Mapping of address and port with encapsulation (MAP) [R]. draft-ietf-softwire-map-02. 2012.
[11] MRUGALSKI T, TROAN O, BAO C, et al. DHCPv6 options for mapping of address and port [R]. draft-ietf-softwire-map-dhcp-01. 2012.
[12] SUN Q, CHEN M, CHEN G, et al. Mapping of address and port (MAP) -- Deployment considerations [R]. draft-ietf-softwire-map-deployment-00. 2012.
[13] CONTA A, DEERING S. Generic packet tunneling in IPv6 specification [S]. RFC 2473. 1998.
[14] LI X, BAO C, HAN G, et al. MAP Testing Results [R]. draft-xli-softwire-map-testing-00. 2012.
[15] DURAND A, DROMS R, WOODYATT J, et al. Dual-stack lite broadband deployments following IPv4 exhaustion [S] .RFC 6333. 2011.

作者简介

包丛笑，清华大学网络科学与网络空间研究院副研究员；研究方向为计算机网络和多媒体通信系统；5篇IETF无状态IPv4/IPv6翻译技术RFC文档核心作者。

李星，清华大学电子工程系、网络科学与网络空间研究院教授、博导；主要研究方向为信号与信息处理、计算机网络和多媒体通信系统；8篇IETF RFC文档核心作者。

DOI: 10.3969/j.issn.1009-6868.2013.02.003
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130204.1018.002.html

PNAT——面向移动互联网的双重翻译过渡技术

PNAT: Double-Translation Transition Technology for Mobile Internet

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0012-004

陈刚/CHEN Gang
邓辉/DENG Hui
段晓东/DUAN Xiaodong
（中国移动研究院，北京 100053）
(China Mobile Research Institute, Beijing 100053, China)

摘要： 基于中国移动提出的双重翻译过渡技术——PNAT，对解析层面的3种双重翻译技术：BIH、464xlat、4rd的特点以及各技术间的共存互补关系进行了深入探讨。双重翻译技术能使得现阶段网络中基于IPv4的用户和应用无需修改就能运行在IPv6网络里，从而提高内容提供商、业务提供商对IPv6的支持力度；双重翻译技术能使得业务流量被主动引导到IPv6网中，从而促进网络运营设施在短时间过渡到IPv6。

关键词： IPv6过渡技术；翻译；BIH技术；464xlat技术

Abstract: In this paper, we propose a double-translation transition technology called PNAT. We describe three kinds of double-translation technologies at the analytic level— BIH, 464xlat and 4rd— and the way in which they complement each other. Double-translation allows numerous IPv4 applications to be supported in IPv6 networks without any modifications being needed. It allows content and service providers to be supported on IPv6 networks. Traffic can be steered to IPv6, which allows prompt transitioning from IPv4 to IPv6.

Keywords: IPv6 transition technology; translation; BIH technology; 464xlat technology

随着IPv6运营实践日益广泛，取得的经验证明IPv6新技术部署需要对现有IPv4应用以及用户有良好的后向兼容性，否则IPv6部署将会受到过渡初期业务、应用的桎梏。比如目前主流应用，Skype等即时通信软件都不支持IPv6；APP Store上已有超过50万余种的应用，升级这些应用来支持IPv6更是步履维艰。面对现状，一种应用/用户无感知、透明的IPv6过渡技术可从根本上克服IPv6应用的问题，从而吸引产业链等众多环节来加速IPv6部署进程。两次网络地址翻译（PNAT）技术正是在这种背景下由中国移动提出的旨在进行IPv6快速部署的一种过渡技术。

1 IPv6推动难点和PNAT过渡理念

IPv6技术最大优势是用来解决网络运营商所面临的IPv4地址紧缺的问题。除此之外，IPv6并没有比IPv4带来更明显的优势。虽然IPv4的设计基于20世纪70年代中期的技术水平，但目前互联网取得的巨大成功已经证明了IPv4的设计是非常成功的。IPv6的设计一方面最大限度地保留了IPv4的诸多优点，另一方面试图在服务质量、安全和移动性上进行突破。但实践证明IPv6相比IPv4并没有本质提升。缺乏杀手级应用的IPv6并不能带来新的商业模式，因此IPv6部署缺乏强有力的推动力。

此外，发展IPv6涉及到产业链的多个环节，改造难度大。由于IPv4地址和IPv6地址格式有很大的不同，端到端系统（包括终端、接入网、承载网、业务网、应用、网管等）都需要做出调整和升级。这其中涉及到通信产业链的各个层面、各个环节，是一个系统级的统筹工程。终端和业务应用始终是IPv6推动过程中的薄弱环节。终端系统和芯片发展的滞后使IPv6无法真正落实到实际用户；大量业务不支持IPv6又导致网络过渡过程因缺乏内容源而滞缓。在这种环境下要发展IPv6、加速IPv6过渡进程需要在IPv6过渡理念上进行突破。

网络问题终究要靠网络运营商自身来解决。发挥网络运营商对网络部署把控的天然优势，利用网络新技术可主动加速IPv6过渡进程。采

收稿日期：2012-12-28
网络出版时间：2013-02-04
基金项目：国家科技重大专项（2011ZX03002-001-01）

用接入IPv6化的思路，能够对产业链其他环节起到带动、示范作用，最终实现IPv6在各个层面的突破和成熟。随着运营商定制手机、家庭网关等用户侧设备的增加，将为PNAT过渡策略营造了良好的发展环境。

具体而言，采用PNAT的过渡理念包括两个方面：其一，使得现阶段网络中IPv4应用程序和用户透明运行在IPv6网络里，降低推动内容提供商、业务提供商支持IPv6的阻力；其二，将业务流量引导到IPv6网中，这种流量引导会促进网络运营在短时间过渡到IPv6。

基于以上考虑，PNAT提出了用户侧翻译联合网络侧翻译的双重翻译架构和部署理念。首先，采用翻译技术保证了数据传输的效率，相比隧道过渡技术，不引入额外的传输负荷，适合移动网应用和部署；其次，用户侧翻译使得用户终端可以在IPv6单栈接入情况下，兼容所有IPv4业务，从而使IPv6用户得到在IPv4网络中同样丰富的业务选择和体验；最后，网络侧NAT64翻译又使得IPv4业务源在IPv6过渡早期能够快速提供IPv6接入服务，实现了IPv4、IPv6的自由互通。

2 基于双重翻译的技术方案

PNAT是联合用户侧翻译和网络侧翻译的“两次翻译”方案，也是中国移动在IETF首次提出的一种IPv6过渡理念[1]。为了推广该项技术，PNAT力争走自主知识产权和国际标准化并重的道路。在过去的3年中，PNAT技术在IETF和3GPP开展了一系列标准化工作，目前已取得一定进展，形成了较为完善的方案体系。

2.1 BIH方案

BIH[2]是一种基于主机IPv4/IPv6有状态翻译技术，这种机制可以支持IPv4的应用程序和IPv6业务间的互通，从而在IPv6的环境中全面实现对IPv4应用兼容。BIH功能模块如图1所示。

BIH主机翻译模块包括两种模式：Socket应用编程接口（API）翻译和包头翻译。Socket翻译主要是在Socket API层面对IPv4和IPv6的Socket进行翻译。包头翻译则是在IPv4传输控制协议/网间协议（TCP/IP）协议栈生成IPv4报文后再对包头进行翻译。这两种工作模式的区别主要在于进行主机翻译的位置不同。Socket翻译在协议栈与应用程序的接口层面进行翻译，BIH执行在API接口层；在包头翻译方式中，BIH执行在网络层时，IPv4包被拦截并转换成IPv6，并使用RFC6145[3]定义的算法来完成IP转换。在具体实施中可根据不同的功能要求，分别选择包头和Socket翻译方式。

除了翻译功能外，BIH还具备域名服务器（DNS）解析扩展功能，负责将应用程序发出的DNS解析请求转换成为A/AAAA两个请求，并根据网络DNS回复类型来构建主机映射表项，配合翻译功能最终完成IPv4-IPv6翻译。

BIH实现了移动终端上有状态的NAT46翻译，目前在IETF进行了标准化，已正式发布为RFC6535。当网络不进行DNSsec部署时，终端可以通过DNS64获得远端IPv4业务的合成IPv6回复，联合网络侧有状态NAT64/DNS64执行两次有状态翻译功能。

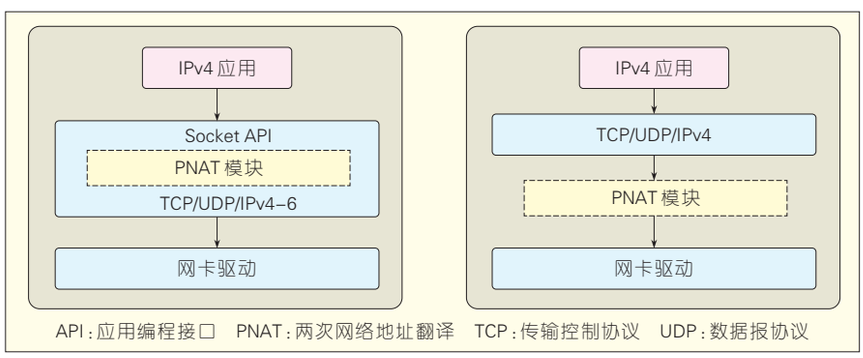
2.2 464xlat方案

464xlat[4]方案实现了IPv6网络中中国移动产品——MIFI终端的接入，即移动终端作为无线路由器来支持多个IPv4用户接入，将IPv4数据包翻译成为IPv6在网络上发送。464xlat的部署场景如图2所示。

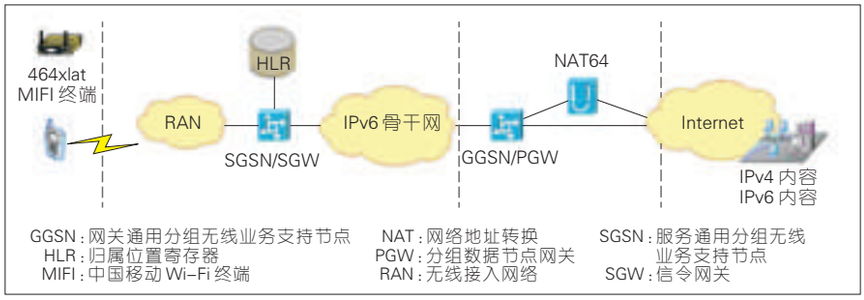
464xlat分为了用户侧翻译（CLAT）和网络侧翻译（PLAT）两部分。CLAT部署基于RFC6145的无状态NAT46翻译过程，PLAT采用标准的有状态NAT64[5]技术。

464xlat的CLAT功能可以部署于MIFI，也可以安装于移动终端充当无线路由器功能。在功能部署上，464xlat能够实现众多IPv4终端的接入，无感知地将IPv4用户数据迁移到IPv6网络中。为了实现以上通信过程，在464xlat CLAT包括无线路由网关功能、地址翻译器、NAT44、DNS代理等。

464xlat具体实施过程中无线路由网关通过Wi-Fi信号向下为多个终端提供接入服务，完成IPv4私有地址的分配和上传数据的接收处理。所连接终端将以CLAT地址为默认网关和DNS服务器。CLAT上的DNS代理功能将截获IPv4主机发出的DNS解析请求，并实现该解析请求在IPv6网中DNS服务器的查询，从而能够避免对DNS数据包的两次翻译。464xlat中由于业务仍处于IPv4，因此相应的IPv4 DNS回复将通过CLAT DNS代理后返回各IPv4主机。IPv4用户获得DNS回复后将发起IPv4业务数据请求，464xlat CLAT将对接收的IPv4数据包执行无状态NAT46翻译，完成业



▲图1 BIH功能模块



▲图2 464xlat的部署场景

务流在IPv6网络中的传递，配合网络侧NAT64功能来实现和远端IPv4业务的互通。

464xlat也可以联合BIH来完成IPv4用户向IPv6业务访问，通过IPv6接入网络为IPv4用户提供完整的IPv4/IPv6业务访问能力。

2.3 4rd方案

4rd[6]技术采用和464xlat相同的两次翻译网络架构，力求保证在两次翻译过程中信息零损失。联合网络无状态NAT64翻译处理，4rd实现了两次无状态的IPv4/IPv6翻译。

4rd借鉴了A+P的思想，在多个用户间共享一个IPv4地址，不同用户仅使用规定的端口号来进行通信。这种地址共享技术既能够满足NAT64无状态翻译中IPv6嵌入公有IPv4地址的要求，又能够在公有地址紧缺条件下，保证两次无状态翻译的进行。4rd具备如下特点：

（1）网络侧无状态翻译

4rd中用户侧的无状态NAT46翻译需要根据网络预先下发配置参数以及代理的IPv6前缀来实施无状态算法，通过计算得到共享IPv4地址信息。数据转发时，翻译模块也需要根据共享IPv4地址和端口信息来还原对应的IPv6信息，从而无需在用户侧构建翻译映射表即可完成IPv4到IPv6的转换。这种两次无状态翻译技术能够有助于克服网络侧单点失效问题，并能够降低冗余备份、Log处理等功能实施的复杂性。

（2）两次翻译过程中信息无丢失

为了保证对IPv4数据包中DF标志位、生存时间（TTL）、服务类型（TOS）等字段的无损传输，4rd提出了扩展IPv6分片头ID字段思想，能够保证经过两次翻译处理后，IPv4数据包头中各字段信息的无损还原。

（3）网络翻译功能简化

4rd提出设置校验和中立比特位来避免网络侧网关翻译过程中校验和的重新计算。这样考虑不仅降低了网关处理负荷，而且有利于网关对新型传输协议的扩展支持。此外，4rd也提出设置*V*比特的标识位，使得数据包更易于被网络管理设备识别和监管。

2.4 3种双重翻译技术的互补与共存

根据网络设备能力以及部署要求，以上3种双重翻译解决方案可以共存互补，实现IPv4用户无感知的IPv6过渡体验。

基于双重翻译的方案体系如图3所示。双重翻译架构下，网络侧部署NAT64网关，根据翻译设备的不同部署形式，终端侧可以采用相应翻译方案来进行配合。具体来讲，将出现如下3种情况。

（1）网络侧同时部署有状态NAT64和DNS64设备时，经过DNS64处理，IPv4通信对端将通过IPv6域名形式来表示。终端侧采用BIH翻译功能来实现IPv4/IPv6互通。

（2）网络侧仅部署有状态NAT64设备时，IPv4通信对端仍呈现IPv4域名形式。则翻译方案采用464xlat，在终端侧执行无状态的IPv4-IPv6翻译过程。

（3）网络侧仅部署无状态NAT64设备时，则采用4rd两次翻译方案，实现内嵌IPv4共有地址在多个终端间的共享复用。

各种技术在相应场景下的功能实施应用范围如表1所示。

3 PNAT产品化和开源开放

经过多年的技术发展和积累，PNAT不仅在标准化方面取得了收获，而且也成功实现了从理念创新向产品落地的稳步发展[7-17]。BIH模块功能已被成功移植到移动终端、家庭网关、桌面操作系统软件等三大类、多平台的各项产品中，并逐步走向现网试点和应用。

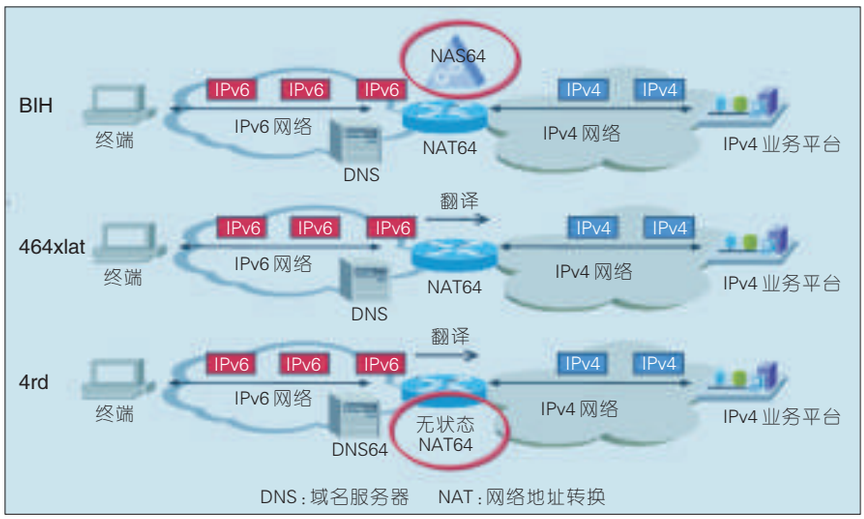
T-mobile USA公司也针对464xlat技术推出了手机终端软件插件，实施了3G现网测试及试点，避免了双栈3G终端在发起双倍PDP激活时，增加网络侧设备投入的问题，取得了良好的效果。

PNAT产品化也在走开源、开放的道路，BIH源代码于2011年正式宣布开源，并贡献给了国际开源社区——Linux Foundation，开源渠道的拓展将更加扩大PNAT/BIH产品应用的受众面，代码成熟度也将在无数开源开发者共同努力下进一步提升。

下一步PNAT会将继续推进各类原型的产品化，真正地进入市场应用。同时还将拓展PNAT产品使用的新场景，比如LTE阶段，用户就可以利用高速无线上行速率方便使用MIFI应用，利用一个LTE的IPv6终端，带动用户多个终端设备过渡到IPv6时代。

4 PNAT技术展望

PNAT克服了IPv4业务升级改造难、产业推动阻力大、网络运营复杂等问题，结合了移动互联网大力发展的趋势和移动用户永远在线的需求，开辟了一条移动网络向IPv6过渡的创新之路，并以此来推动IPv6技术在



▲图3 基于双重翻译的方案体系

▼表1 不同双重翻译技术的应用范围

网络侧设备能力	业务能力	用户侧翻译方式
有状态 NAT64+DNS64	IPv4	BIH
	IPv6	BIH
有状态 NAT64	IPv4	464xlat
无状态 NAT64	IPv4	4rd

全世界范围内的部署和应用。PNAT技术推广将致力以下几个方面。

（1）引导流量向IPv6网络迁移，从而提前IPv6网络部署、加速IPv6过渡。随着PNAT部署扩大，网络中IPv6流量将得以增长，IPv6影响将得以扩大。这种积极因素将带动整条IPv6产业链的发展和成熟。

（2）实现真正IPv6网络的部署。在IPv6单栈网络环境下，将支持众多IPv4应用，做到IPv6过渡对IPv4应用的无感知。

（3）借鉴移动网络2G到3G成功演进经验，确保网络平滑过渡。IPv6过渡应该学习移动网络从2G到3G的平滑升级以及2G/3G共存理念。从2G到3G升级看IPv4到Pv6演进，网络将是IPv4和IPv6两种技术长期共存的网络，因此应在发展IPv6的同时兼容IPv4，以保障和提高用户在IPv6网络中的业务体验。

（4）PNAT也是TD-SCDMA 3G网络部署IPv6单栈的一个很好的选择。目前双栈部署实践中，终端需要同时支持两套协议栈（也就是IPv4和IPv6）处理，这在终端实现复杂度和能耗方面往往不是最优的，而且网络侧也需要相应提供两套运维系统，因此IPv6单栈在这种考虑下是比较好的部署选择。PNAT能很好满足IPv6单栈网络中兼容众多IPv4应用的需求，是一个合适的选择。

参考文献

[1] HUANG B, DENG H. Prefix NAT: Host based IPv6 translation [S]. draft-huang-pnat-host-ipv6-00. 2009.
[2] HUANG B, DENG H, SAVOLAINEN T. Dual-stack hosts using "bump-in-the-host" (BIH) [S]. RFC 6535. 2012.
[3] LI X, BAKER F, BAO C. IP/ICMP translation algorithm [S]. RFC 6145. 2011.
[4] MAWATARI M, KAWASHIMA M, BYRNE C. 464XLAT: Combination of stateful and stateless translation [R]. draft-ietf-v6ops-464xlat-08. 2012.
[5] MATTHEWS P, BEIJNUM I, BAGNULO M. Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers [S]. RFC 6146. 2011.
[6] DESPRES R, PENNO R, LEE Y, et al. IPv4 residual deployment via IPv6 -- A stateless solution (4rd) [R]. draft-ietf-softwire-4rd-03. 2012.
[7] ARKKO J, BAKER F. Guidelines for using IPv6 transition mechanisms during IPv6 deployment [S]. RFC 6180. 2011.
[8] DURAND A, DROMS R, LEE Y, et al. Dual-stack lite broadband deployments following IPv4 exhaustion [S]. RFC 6333. 2011.
[9] DESPRES R. IPv6 rapid deployment on IPv4 infrastructures (6rd) [S]. RFC 5569. 2010.
[10] TSIRTSIS G. Network address translation -- Protocol translation (NAT-PT) [S]. RFC 2766. 2000.
[11] MATTHEWS P, SULLIVAN A, BEIJNUM I, et al. DNS64: DNS extensions for network address translation from IPv6 clients to IPv4 servers [S]. RFC 6147. 2011.
[12] IANA Considerations and IETF protocol usage for IEEE 802 parameters [S]. RFC 5342. 2008.
[13] BOUCADAIR M, MATSUSHIMA S, LEE Y, et al. Motivations for Carrier-side stateless IPv4 over IPv6 migration solutions [R]. [2012-12-22]. draft-ietf-softwire-stateless-4v6-motivation-04. 2012.
[14] BUSH R. The address plus port (A+P) approach to the IPv4 address shortage [S]. RFC 6346. 2011.
[15] BAKER F, BLACK D, NICHOLS K, et al. Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers [S]. RFC 2474. 1998.
[16] RAMAKRISHNAN K, FLOYD S, BLACK D. The addition of explicit congestion notification (ECN) to IP [S]. RFC 3168. 2001.
[17] BAO C, HUITEMA C, BAGNULO, et al. IPv6 addressing of IPv4/IPv6 translators [S]. RFC 6052. 2010.

作者简介

陈刚，北京邮电大学博士毕业；工作于中国移动研究院，从事IPv6、移动网络管理、IP承载网络技术等方面的研发工作；已发表学术论文20余篇；作为3GPP的报告人主持发布3GPP规范10篇。

邓辉，清华大学博士毕业；工作于中国移动研究院，从事移动互联网基础设施演进相关的研发工作；已发表学术论文40余篇。

段晓东，中国移动研究院网络技术研究所所长，从事IP承载网络、下一代互联网、数据业务管理及移动互联网的研发工作。

DOI: 10.3969/j.issn.1009-6868.2013.02.004
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130131.1747.001.html

面向NAT用户的IPv6过渡技术——Silkroad

Silkroad: An IPv6 Transition Technology for NAT Users

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0016-005

刘敏/LIU Min
（中国科学院计算技术研究所，北京 100190）
(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)

摘要： 传统的隧道技术不能满足隐藏在网络地址转换（NAT）域内的IPv6节点的通信要求。针对此问题，提出了一种新的面向NAT用户的IPv6过渡技术Silkroad，它可以使位于一个或多个IPv4 NAT内的节点获得IPv6的连通性。Silkroad支持所有类型的NAT用户和IPv6网络进行互联，且不需要升级现有的NAT设备；此外，Silkroad不需要任何特殊的IPv6地址前缀，能够为NAT用户分配固定的IPv6地址，并具有更高的安全性。

关键词： IPv6协议；过渡技术；网络地址转换技术；隧道技术

Abstract: Traditional tunneling methods are not suitable for IPv6 nodes, which are hidden in the network address translator (NAT). In this paper, we propose a new IPv6 transition mechanism, called Silkroad, for new NAT users. Silkroad allows nodes in one or several IPv4 NATs to use IPv6 connectivity. It also allows all types of NAT users to connect with IPv6 networks without updating NAT devices. In addition, Silkroad does not need a special IPv6 addressing prefix and can assign stable IPv6 addresses to users. It is also highly secure.

Keywords: IPv6 protocol; transition technology; NAT technology; tunnel technology

由于整个网络从IPv4升级到IPv6需要一个漫长的过程，所以IPv6网络在其发展的初始阶段使用的主要是现有IPv4网络的传输设施。鉴于此，IPv6过渡的关键点在于如何兼容现有IPv4网络下众多的主机和路由器设备。在兼容IPv4网络的同时部署IPv6网络，将会使整个网络的过渡平滑而自然。在这种愿景的驱动下，出现了很多IPv4承载IPv6（IPv6 over IPv4）的隧道技术，它们都是将IPv6数据置于IPv4包的载荷部分进行数据的收发，负载的类型为41。然而，这些方法都不适用于位于网络地址转换（NAT）域内的IPv6设备，其主要原因是NAT一般都会对入口的包进行过滤，拒绝包括负载类型41在内的很多种数据包。即便允许它们通过，由于NAT内所使用的都是私有IPv4地址，也无法使用IPv6到IPv4（6to4）及很多其他的隧道机制。

NAT是解决IPv4地址不足的一种临时性技术[1]。由于IP地址资源短缺现象很严重，所以NAT技术使用得非常普遍。在NAT内，用户使用私有地址进行通信，在NAT外，则被转换为一个或若干个公有地址。由于私有地址并不是全局唯一的，所以外部互联网的节点不能够主动访问NAT内的节点，这严重破坏了端到端的连通性，因此它们更迫切需要过渡到IPv6。目前有4种NAT类型[2]，分别是完全锥型NAT、受限锥型NAT、端口受限型NAT和对称型NAT。一般NAT不允许任意负载类型的数据包通过，这致使很多IPv6的隧道技术无法直接应用于NAT内的IPv6设备。因此，提供面向NAT用户的IPv6过渡技术已经成为了一个非常急迫的任务。

一种解决上述问题的方案是由微软提出的Teredo隧道技术[3]，也是目前除了本文所提的Silkroad隧道技术之外的唯一解决方案。它采用IPv6在UDP中（IPv6-in-UDP）的封装方式，将IPv6的数据包置于私有IPv4数据包的数据报协议（UDP）载荷中，而不需要升级现有的NAT设备。然而，Teredo隧道技术却存在着很多不足。首先，Teredo隧道技术不能为用户分配固定不变的IPv6地址，其地址内嵌了相关的隧道参数，即用户的私

收稿日期：2012-12-18
网络出版时间：2013-01-31
基金项目：国家高技术研究发展（"863"）计划（2006AA01Z257）；国家自然科学基金项目（60602058、60572120）；国家科技重大专项（2009ZX03003-002-02）

有地址和UDP源端口经过NAT转换后的外部地址和外部端口。由于外部地址和外部端口在每次初始化时的随机性，用户在使用Teredo隧道技术进行通信时往往会被NAT转换成不同的外部地址或外部端口，从而获得和上次不同的IPv6地址。这种IPv6地址经常变化的情况破坏了NAT用户端到端的连通性，丧失了IPv6本来的优势。其次，Teredo隧道技术不支持对称型NAT用户。由于对称型NAT根据数据包目的地址来转换它的私有地址和UDP源端口，如果目的地址不同，那么转换后的外部地址和外部端口也就不同。Teredo协议本身采用的是客户端-客户端隧道模式，通信对端就是隧道端点，这样随着通信目的地的变化隧道参数也在不断变化，这违背了Teredo中IPv6地址配置后隧道参数不变的假定。最后，Teredo的机制本身存在着严重的安全漏洞，不能有效地阻止非法用户利用Teredo中继器和IPv6网络互联。由于Teredo中继器的无状态特性，任何没有经过身份验证的非法用户只需要构造出符合Teredo格式的源IPv6地址就可以轻松享用中继器提供的数据转发服务。非法用户通过这种方式获得IPv6连接后，便可以对IPv6网络上的节点实施攻击。

1 Silkroad协议设计

同Teredo相比，本文提出的Silkroad协议可以为NAT用户提供不变的IPv6地址，并可适用于对称型NAT中，使系统在安全方面得到了很好的保障。

1.1 Silkroad体系结构

Silkroad典型的体系结构如图1所示，其实现机制主要包括3个实体部分：Silkroad客户端（SC），Silkroad接入路由器（SAR），Silkroad导航器（SN）。其中SC为位于NAT之后想要使用IPv6网络的终端用户，它是一个双栈节点，向应用程序提供透明的IPv6连接。SAR为SC提供身份认证及中继数据的服务，它同时与IPv4与IPv6网络相连，也是一个双栈节点。SN位于IPv4网络中，拥有SC的注册信息和SAR的部署情况，它的主要功能是为SC选择一个就近的SAR进行接入，并对各个SAR的运行状况进行监视。

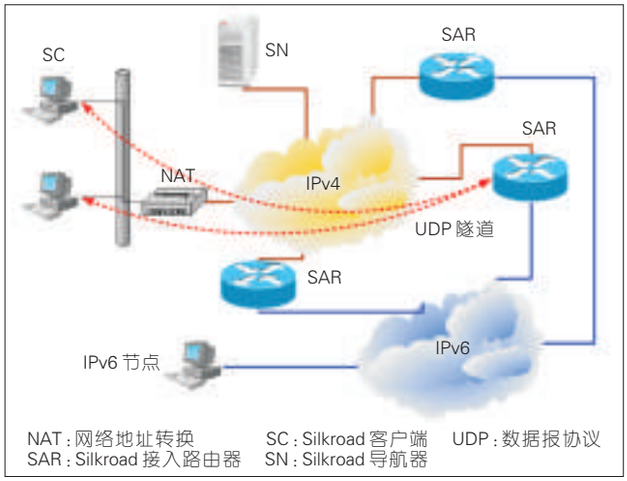


图1 Silkroad体系结构

1.2 Silkroad操作规程

1.2.1 SAR的确定

在Silkroad协议操作的开始阶段，SC选择适当的SAR来为自己提供Silkroad隧道服务。换句话说，SC必须知道它的SAR的IPv4地址。SC获取SAR网络信息的方式有多种，例如SC可以从系统管理员处获得SAR的信息，也可以用域名服务器（DNS）去查找SAR，还可以用预先配置或设定的IPv4任播地址查找SAR。当然，因特网业务提供商（ISP）也可通过其他方式把SAR的地址或域名通告给SC。无论如何，SAR的信息一旦确定，将自动保存到SC的配置文件中。如果SC没有明确地改变它的配置，此SAR将成为SC的默认SAR。

1.2.2 IPv6地址的获取

为了获得IPv6地址，SC通过UDP隧道向SAR发送一个常规的路由请求，而SAR则通过UDP隧道向SC回复一个常规的路由通告，其中包含所请求的IPv6地址。SAR返回的响应报文还包含一个控制选项，可用于存放SAR收到路由请求时所看到的SC地址信息（经NAT映射后的IPv4地址和端口）。事实上，当一个地址请求报文到达SAR时，可能经过了多级NAT设备，此时SAR看到的地址请求报文的源地址和端口应该是最靠近SAR的那个NAT所映射的地址和端口。SAR将会把该地址和端口写入响应报文的控制选项字段，和IPv6的地址响应信息一起返回给SC。

由SAR分配给SC的IPv6地址永久分配给SC，这样即使SC的IPv4地址是动态分配的，也可保持一个固定的IPv6地址。而SC也可以要求分配临时的IPv6地址，这样的地址会有一个生存期，超过生存期后地址将被收回，除非SC提交延长地址生存期的请求。一旦一个SAR给SC分配了IPv6地址，这个SAR就会在自己的映射表里面维护一个地址映射，将SC的IPv6地址同SC的IPv4映射地址及映射端口绑定在一起，该绑定可以随着SC的IPv4地址的改变而相应动态更新。

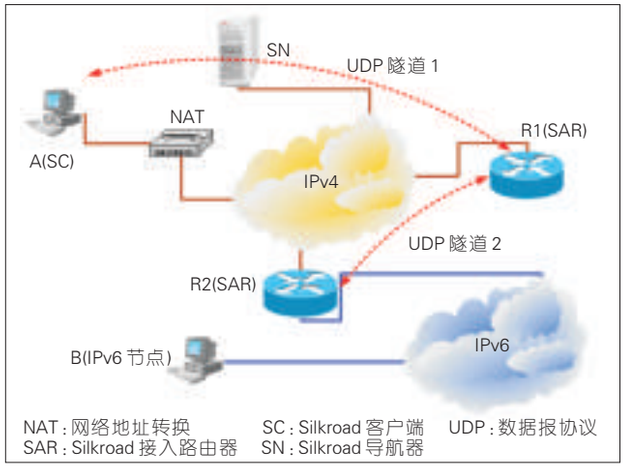
1.2.3 NAT类型的确定

当SC收到地址响应后，将把控制选项字段中的IP地址和端口与自己发送路由请求时绑定的本地IP地

址和端口进行比较。如果不匹配，则SC位于NAT域内；否则SC不需要使用Silkroad服务。为了获得更高的传输效率，SC可以选择在SAR的帮助下确定它所使用的NAT类型。为此，SC发送“测试请求”，而SAR返回“测试应答”，基本过程可参照文献[2]中的描述。一旦确定，NAT类型将被SC保存。

1.2.4 从SC到普通IPv6节点的报文传输

图2显示了一个SC和一个普通的IPv6节点之间的报文传输，其中A是NAT域中的SC，R1是A所选择的SAR，B是一个普通IPv6节点，R2是靠近B的SAR。A若要向B发送IPv6包，A首先要把这个IPv6包封装在IPv4的UDP数据包中，并把它发送给R1。此报文称作Silkroad包，这种包的格式如图3所示。



◀图2 从SC到普通IPv6节点的报文传输

▲图3 Silkroad包

NAT会接收此数据包，用映射地址和映射端口替换UDP的源地址和源端口，然后把包转发出去。此包通过IPv4的UDP隧道1后被R1接收，R1查看IPv6目的地址并在当前的通信对等节点列表中检查这个IPv6地址对应的表项是否存在以及该表项是否有效。如果这个IPv6地址对应的表项不存在或无效，R1将登录它的SN去查找是否B是由另一SAR提供服务的SC。此查找过程与DNS的处理过程类似。SN将在自己的管理域中查找SAR的信息，如果没找到对应表项，它将查询上层SN。SN中的默认路由是直接接入IPv6网络的SAR，这就意味着当R1找不到可以转发数据包到目的IPv6地址的SAR时，会直接将该数据包转发到IPv6网络上，之后数据包便开始遵循IPv6的路由规则进行常规的数据转发。

图2中，确定了R2的地址后，R1就会将IPv6包和控制选项字段一起通过隧道发送给R2。R1和R2之间的传输可以遵循不同的方式，它们可以用另一条UDP隧道传输IPv6包，如图2中的UDP隧道2。这种方法的优点是转发之前R1的处理比较简单，只需要替换掉UDP的源地址和源端口以及目的地址和目的端口即可。然而，UDP封装将带来一定的开销。Silkroad还可以像传统的隧道机制那样，使用IPv6在IPv4中（IPv6-in-IPv4）的封装方式，把IPv6包作为IPv4包的载荷在SAR之间发送。但是这种方法不支持控制选项字段，而且R1在转发之前必须拆包并重新封装。ISP可以指定SAR之间的传输方式，默认的处理是当包中有控制选项字段时，SAR通过UDP隧道转发包；否则，使用传统隧道转发包。事实上，控制选项字段一般只出现于一次会话的前几个报文中。在Silkroad隧道机制中，SAR需要监听UDP和IP两种封装，但是无论在SAR之间采用哪种传输方法，SC只需要支持UDP封装。为简单起见，在下面的描述中，我们假设SAR之间都使用UDP隧道转发报文。当R2接收到数据包后，如果发现B是它所负责的SC，那么R2就会把该数据包通过隧道转发给B，否则，R2将会从UDP载荷中解封装出IPv6数据包再发往IPv6网络上给B。注意R1和R2可能是同一个SAR，此时，UDP隧道2将不存在。

1.2.5 从普通IPv6节点到SC的报文传输

图2中，B是一个普通的IPv6节点，A是一个SC，R1是A所对应的SAR。当B想要向A发送IPv6包时，B只需要遵循IPv6规则，由于A的地址空间属于R1的管辖区域，此数据包可以正常转发到R1上。当R1接收到该数据包后，发现A是它所负责的SC，然后就通过UDP隧道1转发该数据包给A。

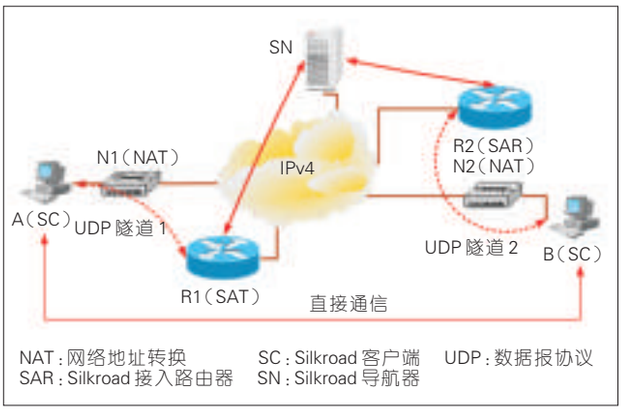
2 路由优化

为了达到更高的传输效率，Silkroad机制在某些情况下可以进行路由优化。

2.1 两个SC间的路由优化

两个SC间的路由优化如图4所示。假设A和B是两个SC，分别位于NAT为N1、N2的后面。R1、R2分别是A和B所选择的SAR，它们也可能是同一个SAR。事实上，A和B都能识别Silkroad报文。因此，A和B可以直接互通，条件是N1和N2不能同为对称型NAT。当然，在直接通信前，A和B需要借助R1、R2交换必要的参数信息。

在A与B通信时，如果A想进行



◀图4
两个SC间的路由优化

路由优化，A首先应在R1的协助下确定N1的NAT类型。然后，A将自己的映射地址和映射端口以及NAT类型封装到Silkroad报文的控制选项字段内发给B。如果N1和N2都是完全锥型NAT，则B在响应报文的控制选项字段内嵌入自己的映射地址和映射端口，通过UDP封装直接返回给A，而后A、B即可通过UDP隧道直接通信。如果N1和N2都是对称NAT，则A、B无法在无SAR协助的情况下进行直接通信。在其他的情况下（受限锥型NAT/端口受限锥型NAT/对称型NAT），则A、B需要发送更多的测试报文，在非完全锥型NAT设备上建立对应的映射关系，而后才能直接通信。在这种情况下，建议用户只在需要传输大量数据时进行路由优化，如果交互信息量比较少，则不需要发送更多的测试报文来进行路由优化。

2.2 同一链路上两个SC间的路由优化

同一链路上两个SC间的路由优化如图5所示。假设A和B为同一链路上的两个SC，均通过名为N1的NAT接入Internet，默认SAR为R1。这里我们不考虑N1的类型，只介绍A、B间如何进行路由优化，以避免所有的报文都经过R1、N1转发。

我们假设A和B已经分别从R1获得了它们的IPv6地址。当然，如果A和B都用相同的IPv4映射地址，那么它们可能位于同一个NAT域中。然而，即使IPv4映射地址不同，它们也可能位于同一个NAT域中。另一方面，即使位于同一个NAT域内，相互之间也不一定能直接通信。所以，必须采取一些措施来确定它们之间是否是直接可达的。有许多方法可以达到这一目的，这里只是给出一种可行方案。如果A想要判断它和B是否在同一链路上，A可以把自己的私有地址放在数据包的控制选项字段中，通过N1和R1发送给B。当B接收到此包时，同时向R1以及A的私有地址返回报文，内含自己的私有地址。如果A能够收到这两个报文，则A可以确认它和B在同一链路上，可以直接互通。后继报文将利用A、B的私有地址封装，只在本地链路上传播，而不必再经过R1和N1转发。

3 安全考虑

通常，Silkroad服务被局限在一个ISP的范围里，ISP应该在它的边界上实行IPv4入口过滤，阻止SAR的地址在ISP外部被用作源地址。ISP也应该对它的用户实行IPv4入口过滤，尤其是SC，使得它们不能伪造包的IPv4源地址。

在SN和SAR的交互中，可采用安全的简单网络管理协议（SNMP）。此外，若想更简单地在远程终端上执行Shell命令（RSH），则可以采用标准的IP安全协议（IPsec）。需要指出的是SC和SAR间通信过程的安全问题。由于两者间的隧道不是面向连接的，因此如果SC单方面强行释放了隧道，SAR会继续转发隧道数据，导致数据包被错误地发到原来的IPv4地址。Silkroad机制采用了保持活动（keep-alive）机制来解决这个问题。当检测到隧道失效后，SAR立即停止向离线用户转发数据包。另一方面，SC必须确保它使用的Silkroad隧道是有效的，这可以通过检查是否定期收到来自SAR的数据包来进行验证。

NAT使SC失去了端到端的连接，但却可以获得更高的安全性，因为NAT可以被看作是一种防火墙。然而，在Silkroad的帮助下，SC在IPv6网络中变成可达的节点，这就使其成为攻击者的潜在目标。为了确保通信的安全性，SC可使用IP安全服务，如使用全局IPv6连接的认证头（AH）或封装安全载荷（ESP）。SC和SAR之间认证过程的实现将有效地阻止中间人攻击，如果认证过程经过对称或非对称加密系统加密，中间人攻击就很难实现。

4 Silkroad原型系统

Silkroad机制已经在重庆的IPv6

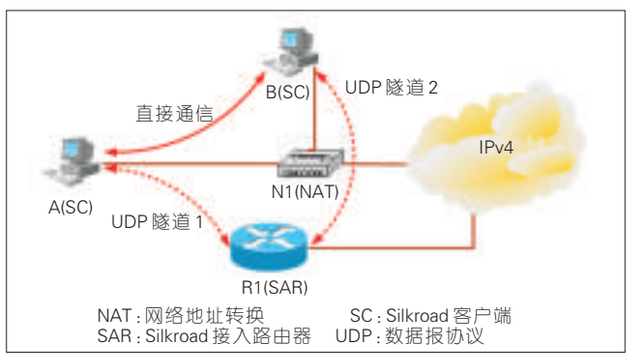
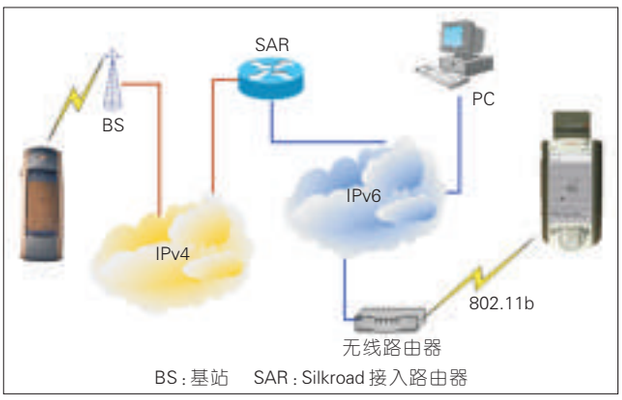


图5▶
同一链路上两个SC间的路由优化

城域网示范项目中得到了实现验证，并已应用于中国联通的中国下一代互联网（CNGI）试验网。我们基于Silkroad实现了手机与手机或手机与个人数字助理（PDA）及PC用户间的IPv6通信。图6展示了Silkroad的原型系统，演示的内容是手机用户通过Silkroad接入IPv6网络，然后和IPv6网络中的PDA及PC用户进行基于IPv6协议的网络五子棋游戏。



◀图6
Silkroad原型系统

当两个基于IP的手机进行端到端通信时，至少有一方需要具有公有IP地址，但目前移动运营商网络不支持IPv6，同时由于缺少IPv4公有地址，它们只能向手机用户分配私有IPv4地址。这样，手机用户本身就存在于NAT之中，即便支持IPv6也无法进行端到端的通信。通过本文的Silkroad系统，手机用户不仅可以和手机用户通信，还可以和PDA及PC用户进行通信，任何一方只要输入对方的IPv6地址就可以开始这个通信过程。可见，Silkroad在运营商网络由IPv4向IPv6的过渡过程中能起到重要的推动作用。

5 结束语

本文提供了一种可以使IPv4中的NAT用户获得IPv6网络通信的协议：Silkroad，它相比Teredo而言支持所有类型的NAT用户，能为用户分配固定不变的IPv6地址，并且具有更高的安全性。我们将Silkroad部署在示范性网络平台上，实现了手机与手机及手机与PDA或PC间的基于IPv6的端到端通信。Silkroad为NAT用户提供了接入IPv6网络的有效方法，在IPv4向IPv6的过渡过程中将会起到重要的推动作用。目前，Silkroad已经成为了中国通信行业标准[4]。其作为中科院知识创新工程重要方向项目“IPv6网络关键技术研究和城域示范系统”的重要成果，通过了由中科院组织的科技成果鉴定。鉴定专家委员会认为Silkroad技术具有明显的技术创新和实用价值，具有国际领先水平。

参考文献

[1] SRISURESH P, HOLDREGE M. IP network address translator (NAT) terminology and considerations [S]. RFC2663. 1999.
[2] ROSENBERG J, MAHY R, MATTHEWS P, et al. Session traversal utilities for NAT (STUN) [S]. RFC5389. 2008.
[3] HUITEMA C. Teredo: Tunneling IPv6 over UDP through (network address translations) NATs [S]. RFC4380. 2006.
[4] YD/T 1635-2007. IPv6网络技术要求--面向网络地址翻译（NAT）用户的IPv6隧道技术 [S]. 2007.

作者简介

刘敏，中国科学院计算技术研究所研究员、博士生导师，中国通信标准化协会理事，中国计算机学会服务计算专业委员会委员，中国计算机学会高级会员，CCF YOCSEF委员；主要从事移动互联网、未来网络技术的相关研究；已承担基金项目12项；已发表论文50余篇；已获得专利授权19项、专利受理12项、软件登记13项。

综合信息

中兴通讯荣获首届GTI Award融合方案大奖

【本刊讯】2013年2月26日，GTI在巴塞罗那举办了GTI之夜，活动齐聚了全球30多家顶尖的TDD运营商和数十家产业链伙伴，并颁发了首届GTI Award，表彰为TD-LTE推广做出卓越贡献的厂商。中兴通讯凭借领先的技术实力和良好的实践表现，获得了FDD/TDD融合解决方案创新大奖。

早在2011年，中兴通讯就同瑞典和黄Hi3G建设了全球首个，也是最大的FDD LTE/TD-LTE双模商用网络，成为全球双模网络的示范。该网络由中兴通讯独家承建，全面应用了FDD LTE/TD-LTE融合解决方案，基于相同软硬件平台研发，具备完全平滑的演进和升级优势。

中兴通讯高级副总裁曾学忠在领奖时表示：“4G技术中，TDD已经被越来越多的运营商所采用。融合组网的解决方案，使得一些FDD运营商在遇到‘瓶颈’时有一个新的解决思路。中兴通讯很高兴能为GTI和TDD产业链的繁荣做出贡献。”

2012年3月，Hi3G与中兴通讯签署了扩容合同；同年12月18日，中兴通讯同中国移动香港公司合作建设的双模网络宣布正式商用，又开创了新的里程碑。

DOI: 10.3969/j.issn.1009-6868.2013.02.005
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130207.0955.001.html

下一代互联网4over6软线隧道过渡技术

4over6 Softwire Tunneling Transition Technology for Next-Generation Internet

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0021-004

崔勇/CUI Yong
陈煜驰/CHEN Yuchi
（清华大学 下一代互联网实验室，北京 100084）
(Next Generation Internet Lab, Tsinghua University, Beijing 100084, China)

摘要： 针对下一代互联网IPv6过渡关键技术问题，从隧道过渡角度展开研究，提出了面向全网IPv6过渡的4over6软线隧道技术方案。技术方案针对全网IPv6过渡需求，设计了主干网及接入网分别适用的4over6软线隧道技术，形成了统一的隧道过渡技术体系，通过主干网异构路由隔离混传、接入网自适应异构动态接入以及透明化自动隧道等机制实现下一代互联网IPv6平滑过渡。

关键词： 下一代互联网；软线；隧道；IPv6过渡

Abstract: This paper focuses on issues with key techniques in IPv6 transition for next-generation Internet. In this paper, we propose a 4over6 softwire tunneling mechanism. We propose different 4over6 softwire tunneling technologies for the backbone and access networks. These technologies are used to create a unified tunneling transition system and satisfy the requirements for IPv6 transition across the whole network. Smooth IPv6 transition can be achieved using mechanisms such as heterogeneous routing separation and transport in the backbone, dynamic self-adaptation and heterogeneous access in the access network, and transparent automatic tunneling.

Keywords: next-generation Internet; softwire; tunneling; IPv6 transition

基于IPv4协议的互联网经过20多年的飞速发展，在全球范围内已经取得了巨大的成功。但是，随着互联网规模的持续增长和新需求、新业务的发展，基于IPv4协议的网络环境已经难以应对，网络规模的急剧扩张突显了IPv4的一系列严重问题，包括地址资源紧张、路由可扩展性不足、网络地址转换（NAT）技术[1]带来的端到端特性破坏等。

IPv6协议[2]是下一代互联网中替代IPv4的网络基础协议。由于IPv6与IPv4并不兼容，互联网在从IPv4向IPv6过渡过程中面临着异构路由、可扩展性、状态维护等问题。此外，由于大多数网络资源、服务和应用仍基于IPv4实现，终端用户对IPv4仍然存在依赖性，因特网业务提供商（ISP）对于IPv6升级存在一定惰性，以及网络设备尤其是接入网设备对IPv6的支持不足等因素，可以预见IPv4与IPv6将在较长一段时期中共存。而在二者共存期间，如何较快较好地实现全网向IPv6过渡、尤其是接入网的过渡，已经成为制约下一代互联网发展的关键。

针对IPv6过渡问题，目前国际上已提出了许多过渡技术方案[3]，这些方案按技术思想分类，大体上可分为隧道技术和翻译技术[4-5]。本文从隧道过渡思想出发，提出了面向全网的IPv6隧道过渡框架体系，以及主干网和接入网过渡各自适用的4over6软线隧道技术。目前4over6软线隧道过渡方案及相关技术已在国际互联网标准化组织IETF形成了RFC4925、RFC5565及RFC5747等3项国际标准，并在中国通信标准化协会形成了2项中国通信行业标准。

1 全网IPv6隧道过渡框架与策略

面向全网的IPv6过渡技术方案需要满足以下需求[6]：

（1）保持端到端特性，这是实现终端节点与互联网间无障碍双向互联互通需要满足的基本要求。

（2）保持异构网络可扩展性，即IPv4、IPv6网络各自路由的独立性和可扩展性，这是在过渡时期在对既有IPv4网络进行最小化改动的情况下平滑实现IPv6过渡的重要保证。

收稿日期：2013-01-18
网络出版时间：2013-02-07
基金项目：国家“973”计划项目（2009CB320501）

（3）对上层移动应用呈现透明性，这是在IPv6过渡时期维护既有的IPv4应用，保证用户良好网络体验的关键保障。

全网IPv6隧道过渡问题[7]按照网络层次不同可划分为主干网IPv6过渡及接入网IPv6过渡，其隧道过渡框架如图1所示。主干网向上连接ISP网络，同级连接其他的主干网，向下连接接入网。接入网向上连接主干网，向下为各种不同类型的用户提供接入。在接入网与主干网相邻一侧有若干边界路由器（BR）与边缘路由器（PE）相连接。主干网隧道过渡要求PE升级为双栈，PE之间使用隧道为与主干网地址簇异构的分组提供透明传输。接入网隧道过渡要求BR以及用户侧边缘设备（CE）维护双栈，CE为与接入网不同地址簇的分组建立隧道发往BR，这样CE与BR间使用隧道为与接入网地址簇异构的分组提供透明传输[8]。

全网隧道过渡路线是一个IPv4到IPv6 over IPv4到双栈到IPv4 over IPv6到IPv6的过程。过渡初期，一些IPv6接入网孤岛出现，这些网络需要互访以及访问IPv6互联网。此时由于IPv4传输仍是主流，大量主干网核心路由器（P）仍是IPv4单栈，将这些P路由器都升级为双栈比较困难。因此应优先将少数PE路由器升级为双栈，并通过6over4隧道为IPv6孤岛提供跨越IPv4传输服务。主干网P路由器会在运行IPv4基础上启动IPv6协议栈，即升级为双栈。随着双栈P路由器规模增大，相应运维开销也逐渐增大，而IPv4传输需求越来越少，因此P路由器会逐渐关闭IPv4协议栈而只运行IPv6单栈；PE路由器则仍维持双栈，通过4over6隧道为IPv4网络提供IPv6网络跨越传输服务；最终IPv6在全网中完全部署，网络全面实现向IPv6的过渡。

2 IPv6主干网4over6软线隧道技术原理与应用

主干网边缘处的PE与上级、接入网、同级主干网之间运行外部边界网关协议（EBGP），互相传递域间路由信息，而主干网内侧的骨干路由器（PR）间则通过内部边界网关协议（IBGP）进行路由信息交互。主干网的主要任务是提供IPv4和IPv6的接入和传输服务，在IPv6单栈主干网中需要部署实现主干网4over6软线隧道技术[9]。

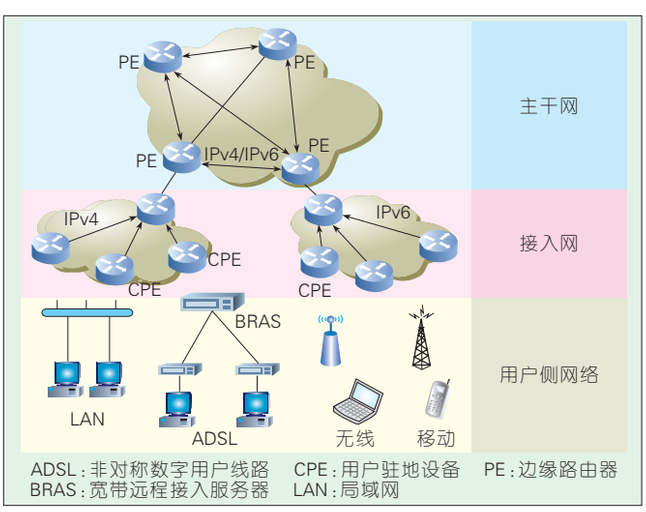
4over6软线隧道技术通过自动隧道机制来建立4over6软线隧道。4over6软线隧道技术对IPv6主干网的核心路由协议——多协议边界网关协议（MP-BGP）进行4over6扩展[10]。该扩展使MP-BGP报文携带IPv4目的网络的信息和隧道端点信息并发送到IPv6主干网的另一端，通过IPv4/IPv6路由异构隔离混传使PE间能建立无状态的4over6软线隧道。PE与CE之间可以通过域内或域间IPv4路由协议来交互IPv4路由，也可以由CE配置缺省路由到PE。PE需要完成分组的封装、解封装及转发处理。在入口PE通过IPv4路由找到合适的出口PE后，入口PE需要采用IPv4-in-IPv6封装机制来封装并转发原始IPv4分组。出口PE在IPv6主干网侧收到IPv4-in-IPv6封装分组后，对分组进行解封装，并转发到相应的IPv4目的网络。

目前主干网4over6软线隧道技术已在IETF形成了2项国际标准RFC5565和RFC5747，并已在当前世界上最大的IPv6单栈主干网——CNGI-CERNET2上完成了部署，形成了有百所高校校园网参加的4over6软线隧道过渡试商用系统。主干网4over6软线隧道技术在CNGI-CERNET2部署如图2所示。主干网4over6软线隧道技术结合各大高校的实际网络运营情况进行部署，能够很好地与校园网已有的路由相互隔离开，不对现有的校园网产生影响，同时又能很好地满足部署的需求。部署与试验结果表明，主干网4over6软线隧道技术对主干网改动小，设备配置简单，可扩展性高，维护负担小，保护了原有网络设备投资，能够支持WEB、P2P等各种IPv4典型应用，具有良好可运营性。

3 IPv6接入网4over6软线隧道技术原理

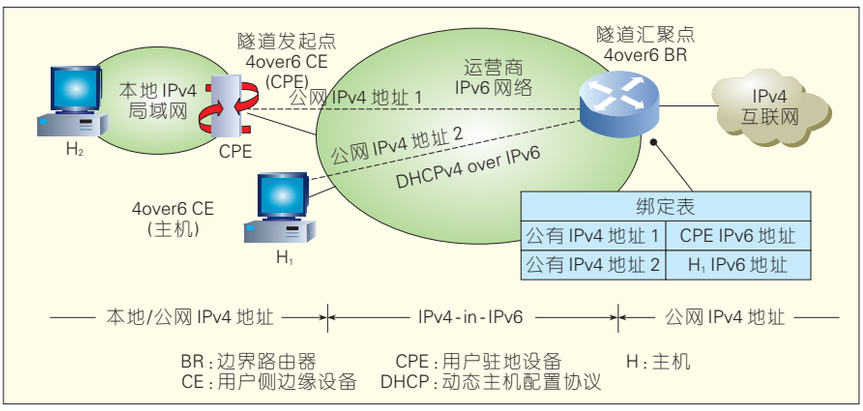
接入网的主要任务是为用户边缘设备提供接入主干网或互联网的服务。为达到这一目的，在IPv6单栈主干网中需要部署实现接入网4over6软线隧道技术[11-12]。

接入网4over6软线隧道体系结构如图3所示。作为4over6隧道发起点的CE既可以是终端主机，也可以是用户驻地设备（CPE），如家庭网关等，统称为4over6 CE。作为4over6软



◀图1 全网IPv6隧道过渡框架示意图

◀图2 主干网4over6软线隧道技术在CNGI-CERNET2部署示意



▲图3 接入网4over6软线隧道体系结构

线隧道汇聚点的BR由ISP维护，称为4over6 BR，部署在ISP IPv6网络（IPv6接入网）与IPv4互联网（或主干网）交界处，所有出入ISP网络的IPv4报文均要经过4over6 BR。

控制层面上，IPv6网络中的4over6 CE通过自适应异构动态接入机制[13]接入IPv4互联网。

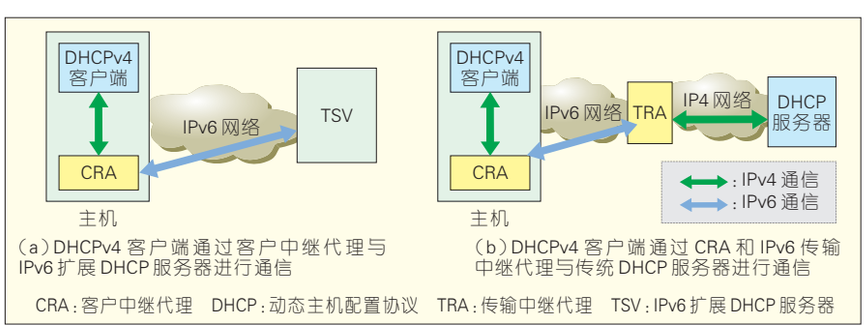
为4over6 CE分配IPv4地址的DHCP服务器由ISP维护，既可与4over6 BR联合部署，也可作为独立DHCPv4服务器与4over6 BR部署在同一IPv4网络中。由于4over6 CE与4over6 BR间通过纯IPv6接入网连接，传统的动态主机配置协议（DHCP）不能在IPv6网络中正常运行，因此接入网4over6软线隧道技术对传统DHCP进行了扩展，采用了基于IPv6传输的DHCPv4 over IPv6扩展机制来进行IPv4地址分配。DHCPv4 over IPv6机制包含两种场景，如图4所示。图4（a）所示为DHCPv4客户端通过客户中继代理（CRA）与IPv6扩展DHCP服务器（TSV）进行通信，而图4（b）所示为DHCPv4客户端通过CRA和IPv6传输中继代理（TRA），与传统DHCP服务器进行通信。

CRA与DHCPv4客户端处于同一主机中，负责对DHCPv4客户端与TSV/TRA之间的DHCPv4信息进行中继。CRA与DHCPv4客户端通过IPv4进行通信，与TSV/TRA则通过UDPv6报文进行通信。TRA部署在IPv6网络与IPv4网络之间，负责对CRA与纯IPv4的传统DHCPv4服务器之间的DHCPv4消息进行中继。TSV是支持IPv6报文传输的DHCP服务器。它部署在IPv6网络中，与CRA或TRA进行IPv6报文承载的DHCP数据传输，从而完成与DHCP客户端的交互，并为客户端主机分配IPv4地址。通过该机制，DHCP客户端无需进行任何改动，即能获取TSV分配的IPv4地址及其他资源。

数据层面上，在4over6软线隧道两端，4over6 CE与4over6 BR分别完成IPv4-in-IPv6的封装和解封装工作。

当4over6 CE需要向4over6 BR发起通信时，4over6 CE首先根据本机IPv6地址以及事先得知的4over6 BR端IPv6地址，对IPv4报文进行IPv6报文头的封装，原本的IPv4报文成为了IPv6报文的负载。之后4over6 CE将该报文发送至IPv6网络，报文经由IPv6网络被递交给4over6 BR。4over6 BR在收到该报文后，首先对该IPv6报文进行解封装，将其还原为IPv4报文，之后4over6 BR将报文交由IPv4网络进行后续转发过程。

当4over6 BR需要向4over6 CE发起通信时，4over6 BR首先检查在本机维护的4over6 CE端IPv6地址与IPv4地址映射表，通过匹配IPv4报文目的地址与映射记录，找出对应的IPv6地



▲图4 DHCPv4 over IPv6机制应用场景

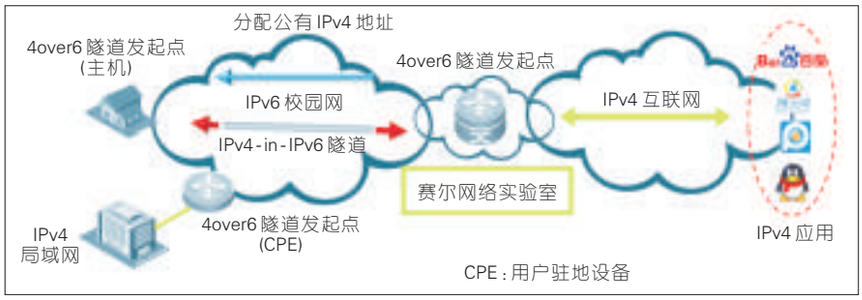
址作为IPv6头的目的地址，并以4over6 BR本机的IPv6地址为源地址，对IPv4报文进行IPv6报文头的封装，IPv4头及以内的内容成为了IPv6报文的负载。之后4over6 BR将该报文发送至IPv6网络，报文经由IPv6网络被递交给4over6 CE。4over6 CE在收到该报文后，解封装其IPv6报文头，将报文还原为IPv4报文，并转交给上层应用。

4over6 BR通过轻量级的用户级状态维护来记录4over6 CE的IPv4-IPv6地址对应信息。4over6 BR并不关心4over6 CE建立的会话数，仅维护4over6 CE的IPv6地址与IPv4地址的映射，因此这是一个每用户有状态、每流无状态映射。用户级状态维护机制的运用极大地减少了4over6 BR的运维开销，减轻了ISP网络负担，提升了终端用户的网络体验。

4 IPv6接入网4over6软线隧道技术系统研制与部署试验

清华大学完成了基于Linux、Windows7及Android系统的4over6软线隧道软件系统实现。软件系统源代码于2012年在开源社区Github上开源，开源项目将为扩大技术方案影响、吸引同行共同参与完善技术方案做出重要贡献。

目前接入网4over6软线隧道过渡方案及相关技术已在IETF softwire及dhc工作组形成了多项标准草案，引起了业界的广泛关注。在推动技术方案标准化过程中，清华大学也带动中国多家设备厂商进行了4over6软线隧道技术的设备实现。基于多种软硬件平台的技术实现形成了接入网4over6软线隧道技术原型系统体系。在此基础上，清华大学在IPv6校园网进行了接入网4over6软线隧道技术试验部署，如图5所示。此外，中国电信在湖南省网部署了面向宽带网的IPv4与IPv6双栈环境，完成了轻量级4over6软线隧道技术LAFT6的现网试验部署[14]。各项实验与试验部署结果显示，接入网4over6软线隧道技术实现简单，易于部署，运维开销小，扩展性强，能对既有IPv4应用和服务提供良好支持。



▲图5 清华大学IPv6校园网试验部署

5 结束语

本文针对下一代互联网IPv6过渡关键技术问题，从隧道过渡角度展开研究，提出了面向全网IPv6过渡的4over6软线隧道技术方案。技术方案针对全网IPv6过渡需求，提出了主干网及接入网分别适用的4over6软线隧道技术，形成了统一的隧道过渡技术体系，通过主干网异构路由隔离混传、接入网自适应异构动态接入以及透明化自动隧道等机制实现全网IPv6平滑过渡。目前4over6软线隧道技术方案已形成3项IETF国际标准及2项国家通信行业标准。基于技术方案的研究和标准化，4over6软线隧道技术已完成多项系统研制、应用与部署。各项应用、实验与部署结果表明，4over6软线隧道技术具有实现简单、易于配置与部署、运维开销小、对IPv4应用及服务能提供良好支持等优势，为推动下一代互联网IPv6过渡提供了重要解决方案。

参考文献

[1] SRISURESH P, EGEVANG K. Traditional IP network address translator (traditional NAT) [S]. RFC3022. 2001.
[2] DEERING S, HINDEN R. Internet protocol version 6 (IPv6) specification [S]. RFC2460. 1998.
[3] WU P, CUI Y, WU J, et al. Transition from IPv4 to IPv6: A state-of-the-art survey [J]. IEEE Communications Surveys and Tutorials, to be published.
[4] WU P, CUI Y, XU M, et al. Flexible integration of tunneling and translation for IPv6 transition [J]. Springer Networking Science, 2012, 1(1/2/3/4):23-33.
[5] WU P, CUI Y, XU M, et al. PET: Prefixing, encapsulation and translation for IPv4-IPv6 coexistence [C]//Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM'10), Dec 6-10,2010, Miami, FL, USA. Piscataway, NJ,USA:IEEE, 2010:5p.
[6] CUI Y, DONG J, WU P, et al. Tunnel-based IPv6 transition [J]. IEEE Internet Computing, to be published.
[7] LI X, DAWKINS S, WARD D, et al. Softwire problem statement [S]. RFC4925. 2007.
[8] CUI Y, WU P, XU M, et al. 4over6: Network layer virtualization for IPv4/IPv6 coexistence [J]. IEEE Network, 2012, 26(5):44-48.
[9] WU J, CUI Y, METZ C, et al. Softwire mesh framework [S]. RFC5565. 2009.
[10] WU J, CUI Y, LI X, et al. 4over6 transit solution using IP encapsulation and MP-BGP extensions [S]. RFC5747. 2010.
[11] CUI Y, WU J, WU P, et al. Public IPv4 over access IPv6 network [R]. IETF draft. 2012.
[12] CUI Y, SUN Q, BOUCADAIR M, et al. Lightweight 4over6: An extension to the DS-lite architecture [R]. IETF draft. 2012.
[13] CUI Y, WU P, WU J, et al. DHCPv4 over IPv6 transport [R]. IETF draft. 2012.
[14] SUN Q, XIE C, LEE Y, et al. Deployment considerations for lightweight 4over6 [R]. IETF draft. 2012.

作者简介

崔勇，清华大学计算机科学与技术系教授，中国通信标准化协会理事，国际互联网标准化组织IETF IPv6过渡Softwire工作组主席；研究领域为网络体系结构、IPv6与下一代互联网、泛在网络和云计算等；已主持基金项目12项；已发表学术论文100余篇，出版专著2部。

陈煜驰，清华大学下一代互联网实验室在读硕士研究生；研究领域为下一代互联网IPv6过渡技术领域。

DOI: 10.3969/j.issn.1009-6868.2013.02.006
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130228.1703.001.html

域名系统对IPv6协议支持分析

Making DNS Compatible with IPv6 Protocol

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0025-004

冷峰/LENG Feng
王伟/WANG Wei
张跃冬/ZHANG Yuedong
（中国互联网络信息中心，北京 100190）
(China Internet Network Information Center, Beijing 100190, China)

摘要： IPv6协议作为下一代互联网的网络层协议，具有诸多新特性与功能，而域名系统（DNS）作为重要的互联网基础架构，需要全面支持IPv6新特性。从域名系统的资源记录及协议扩展、IPv4到IPv6的过渡、域名解析各环节对IPv6的准备、IPv6为域名系统带来的新机遇等几方面对IPv6环境下域名系统的域名服务进行了分析和研究。

关键词： 域名系统；IPv6协议；递归解析服务器；权威解析服务器

Abstract: IPv6 is the next-generation network layer protocol and has many new features and functions. The domain name system is an important part of the Internet infrastructure and must fully support the new features of IPv6. In this paper, we analyze DNS resource record and protocol extension, transition from IPv4 to IPv6, the preparation of DNS for IPv6, and opportunities that IPv6 brings to DNS.

Keywords: domain name system; IPv6 protocol; recursive DNS server; authoritative DNS server

域名系统（DNS）的主要功能是通过域名和IP地址之间的相互对应关系，来精确定位网络资源，即根据域名查询IP地址，反之亦然。DNS作为重要的互联网基础资源[1]，链接底层物理设施和上层业务应用，由于任何通信链接都需要依赖IP地址和域名完成标识、定位和路由，包括超文本传输协议（HTTP），电子邮件等众多上层应用都是建立在DNS体系基础之上的。

IPv6是互联网工程任务组（IETF）设计的用于替代现行版本IP协议（IPv4）的下一代网络层协议。与IPv4相比，IPv6具备多项优势。首先，它具备更大的地址空间[2]；其次，IPv6的地址结构和地址分配采用严格的层次结构；再次，IPv6协议增加了增强的组播支持和对流的控制，更支持多媒体应用；另外，IPv6协议具备更高的安全性，用户可以对网络层数据进行加密和校验；最后，IPv6是一个可扩充协议。

同为互联网基础资源，DNS对IPv4支持已经经过了长期实践的检验，那么在IPv6取代IPv4成为主流IP协议的过程中，在DNS支持方面存在如下的问题：

（1）IPv6的新特性和DNS的支持密不可分，那DNS是否已经具备了全面支持IPv6的能力？

（2）在IPv6漫长的过渡期内，IPv4和IPv6必将共存，这对DNS技术会造成什么样的影响？

（3）域名解析各环节（权威解析服务机构、递归解析服务机构、客户端）是否已经做好了IPv6和DNS技术升级的准备？

（4）IPv6会为DNS技术带来什么新机遇？

本文针对上述问题对IPv6时代的DNS技术进行了分析研究，并得出结论。

1 DNS对IPv6的扩展

1.1 资源记录

针对IPv6协议的新特性，域名系统的资源记录在IPv4基础上进行了扩展，为互联网全面支持IPv6协议打下了基础。域名解析可分为正向解析和反向解析两类，下面对它们进行分别讨论。

1.1.1 正向IPv6域名解析

IPv6域名相对于IPv4，将正向解析记录扩展为“AAAA”，作为IPv4至IPv6地址的过渡手段；同时增加了“A6”类型资源记录，该记录可完全支持IPv6协议特性。

“AAAA”记录被用来解析完全限定域名到IPv6地址，它是128位的IPv6地址，相较于IPv4地址扩大了4倍，但由于此类型资源记录是将IPv4“A”记录的简单扩展，因此不能完成

收稿日期：2012-12-26
网络出版时间：2013-02-28
基金项目：下一代互联网可信域名服务系统产业化项目（CNGI-09-03-04）

对IPv6的层次性等特性的支持。

“A6”记录格式是在RFC2874基础上提出的，它由前缀长度、IPv6地址和域名前缀组成。“A6”记录是把一个IPv6地址与多个“A6”记录建立联系，每个“A6”记录都只包含了IPv6地址的一部分，结合后拼装成一个完整的IPv6地址。

与“AAAA”记录相比，“A6”记录支持更多IPv6的新特性，如地址汇聚、地址更改等。它的特性包括：“A6”记录允许一个IPv6地址分散于多个记录中，或允许在不同的区域；域名系统记录即委派地址被“DNAME”记录所取代。

1.1.2 反向IPv6域名解析

与IPv6的两种正向域名解析对应，反向IPv6域名解析可由两种形式表示：

（1）与“AAAA”记录对应的是“IP6.INT”形式的反向IPv6域名解析记录，它以“.”分隔，以半字节16进制数字格式来表示，低位地址在前，高位地址在后。

（2）与“A6”记录对应的是“IP6.ARPA”形式的反向IPv6域名解析记录，它以“[”开头，16进制地址居中，地址后添加“]”，是一种二进制的串格式。

但是，由于一次完整的地址解析需要按照地址的分配层次关系到不同的域名系统服务器进行查询，并且所有的查询都成功才能得到完整的解析结果，与原有IPv4协议下的域名系统解析相比，无疑会延长解析时间，增加出错几率。因此，IPv6协议在技术方面需进一步改进DNS地址链功能，找寻提高IPv6域名解析速度的办法[3]。

1.2 协议扩展

由于IPv6协议的新特性，其IP地址长度由32位扩展至128位，扩大了4倍[4-5]，导致DNS的查询包的大小增长明显，所以域名服务协议为解决此问题进行了协议扩展。目前已经制订的国际标准有：

（1）DNS Extensions to Support IP Version 6 (RFC3596)[6]。

（2）DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers (RFC6147)[7]。

（3）Common Misbehavior Against DNS Queries for IPv6 Addresses (RFC4074)[8]。

正在制订的国际标准有：

（1）IPv6 AAAA DNS Whitelisting Implications (draft-ietf-v6ops-v6-aaaa-whitelisting-implications-03)。

（2）EDNS0 Option for Indicating AAAA Record Synthesis and Format (draft-korhonen-edns0-synthesis-flag-02)[9]。

（3）Some Considerations on the Load-Balancer for NAT64 draft-wang-behave-nat64-load-balancer-0。

1.2.1 DNS扩展机制

RFC1035指出，对于DNS服务器，递归解析时采用用户数据报协议（UDP）。由于IPv4协议中，UDP协议传输的数据包如果超过512字节，将被截断并分片传输，此时DNS协议中Flags字段Truncated位被置位。即当DNS响应数据大于512字节的时候，数据只返回512字节，剩余的数据将被丢弃。此时递归解析服务器通常会使用传输控制协议（TCP）重发原来的查询请求，允许返回的响应超过512个字节。在这种情况下，会对用户体验造成影响。

DNS扩展机制（EDNS0）允许递归解析服务器公布其支持UDP数据包的最大分片，更便于传输大于512字节的数据包。

权威解析服务器接收到DNS查询请求时，它对来自查询请求中的可选扩展资源记录（OPT RR）字段中的数据报协议的分片大小进行标识，以根据递归服务器指定的最大数据报协议的数据包大小发送回复。

（1）EDNS0查询

递归解析服务器发送查询请求之前，它检查其高速缓存，确定目标权威解析服务器是否支持EDNS0。如果该权威DNS服务器支持EDNS0，递归解析服务器会将OPT RR附加到它发送的查询的附加部分。若反之，递归解析服务器将不会发送OPT资源记录。OPT RR记录内容只与UDP传输层消息关联，其将递归解析服务器的数据报协议的分片大小存储在其字段中，列出递归解析服务器可在其网络中传输的最大数据报协议的分片字节数。

（2）EDNS0响应

DNS权威解析服务器接收包含公布最大数据报协议的分片包大小的OPT RR的查询时，它将基于此OPT记录指定的限制来数据报协议的响应包的大小。如果DNS权威解析服务器接收到不包含OPT RR的查询，它会假定递归解析服务器不支持EDNS0。在这种情况下，DNS权威解析服务器将其数据报协议的响应包限制为最大512字节。

1.2.2 域名系统安全扩展

DNSSEC作为DNS的安全扩展，其协议标准和技术实现也已经成熟，并且兼顾了DNSSEC和DNS服务的兼容性。因此，从理论上讲，DNSSEC服务同样支持IPv6的资源表示和解析方式，并且能在IPv6网络上稳定可靠地运行。就像在IPv4网络和IPv6网络中DNS应保持统一体系，在IPv6环境下DNSSEC也应同时遵循这一建议。IPv6环境下，DNSSEC在密钥分发流程、数据完整性和原始性认证机制上与IPv4网络中的DNSSEC并无多大差异，区别仅仅是在于返回资源的记录类型由A记录扩展至“AAAA”记录[10]。

目前，DNSSEC在部署过程中面临着密钥管理软件功能需更加完善等问题，而IPv6环境下DNSSEC的部

署还面临着IPv4向IPv6网络过渡中由协议转换所带来的加密算法、资源记录报文、相关标准等问题。当然，由于IPv6网络本身从技术标准、软硬件、基础设施等方面，都是一个在运行中不断完善的过程，因此也需要对其在实际应用和运行过程中发现、跟踪、调整和改进。而DNSSEC在密钥管理、业务衔接方面还需积累实施部署和运行的经验。因此在IPv6环境下，DNSSEC的部署也将会是一个从小范围试点到大范围推广的逐步推进过程。

2 过渡技术

在未来很长一段时间内，IPv6将与IPv4网络共存。由于IPv6协议的新特性，在DNS实现细节中将与IPv4网络有所区别（如资源记录），在由IPv4向IPv6过渡的过程中，两类网络间的DNS互通同样需要过渡技术来实现[11]。

2.1 DNS64

DNS64（DNS6-4）技术是实现IPv6网络与IPv4网络互访的技术手段之一。其主要思想是在DNS解析过程中将DNS查询信息中的A记录（IPv4地址）合成到“AAAA”记录（IPv6地址）中，并返回合成的“AAAA”记录至IPv6侧用户。

它的具体实现是：当DNS64接收到来自IPv6客户端的“AAAA”记录DNS查询后，它首先尝试查找该“AAAA”记录的解析结果。如果没有目标节点的可用“AAAA”记录，DNS64执行一个A记录的查询。对返回的每个A记录，DNS64合成一个“AAAA”记录。合成“AAAA”记录的对应域名与原始A记录的相同，但是原始A记录中的IPv4地址被替换为一个IPv6地址。这个IPv6地址由IPv4地址与DNS64的附加参数经过计算得到。合成的“AAAA”记录会返回给IPv6客户端，客户端将会使用IPv4接收端的IPv6地址建立一个IPv6通信。这个报文会经通过路由上的某台IPv6/IPv4转换设备送往IPv4网络。

DNS64功能可以在域名系统体系中的如下3个位置中的任意位置部署实现：某个权威域的权威解析服务器；递归解析服务器；客户端主机，与本地解析器结合。这3种部署各有优缺点，具体如何实现需根据实践结果调整。

2.2 双栈技术结合快速回退算法

支持双协议栈的IPv6节点间互通时使用IPv6协议栈，IPv4节点间互通时使用IPv4协议栈。当IPv6节点访问IPv4节点时，该IPv6节点将申请临时IPv4地址并得到网管路由器的隧道终端（TEP）IPv6地址。此时，IPv6节点将原始IP数据库包封装成一个4over6的IP数据包（包头为IPv6报文头，目的地址为TEP地址，内容为IPv4包），4over6包经过IPv6网传到网关路由器，网关路由器将其IPv6头去掉，将IPv4包通过IPv4网络送往IPv4节点。

网关路由器需要将IPv6地址与IPv4临时地址对应关系进行存储，用于上述过程反向通信。

IPv4向IPv6过渡时期，域名系统所采用的双栈技术是在IPv6与IPv4网络双栈技术基础上，在域名系统中同时存储A记录及“AAAA”记录，由于节点既可以处理IPv4协议，也可以处理IPv6协议，无论域名系统服务器回答“A”记录还是“AAAA”记录，都可以进行通信。双栈技术无疑增加了域名系统服务器缓存的记录数量，并要即时进行地址转换及映射工作，由此，域名系统服务器的负担将大大增加。

双栈技术的使用带来了其他问题，当终端同时获得AAAA和A记录时，会优先尝试AAAA记录，如果访问AAAA记录出现超时，才会使用A记录，在AAAA记录有问题的情况下会导致用户体验受损。快速回退算法（Happy Eyeball）的出现弥补了这一缺陷，该算法要求终端同时尝试AAAA记录和A记录，然后使用最先返回的结果，此举可以显著增强用户体验。

2.3 DNS-ALG结合NAT-PT技术

NAT-PT是一种纯IPv4应用和纯IPv6应用之间的互通方式。通过NAT-PT，可以实现IPv4主机对IPv6主机的访问，反向访问也同样可以实现。关于IP地址转换以及协议转换的工作都由加载在边界网关路由器上的NAT-PT设备实现。翻译网关要向IPv6域中发布一个路由前缀，凡是具有该前缀的IPv6包都被送往网关路由器。

网关路由器为了支持NAT-PT功能，需要建立IPv4地址池，预留一部分全球IPv4地址，IPv6向IPv4域中转发报文时临时分配给IPv6主机，并保存该IPv4与IPv6地址映射关系。但是，由于协议对地址格式的定义不同，直接通过地址访问难以实现，只有通过域名才能实现访问，所以网关路由器还需要DNS-ALG的支持，才能在IPv4网络节点访问IPv6网络节点的过程中发挥作用。

DNS-ALG是在NAT-PT基础上封装的一个应用层网关，主要职责为将IPv4与IPv6网络间域名查询请求所返回资源记录进行转换，即A记录与“AAAA”记录的转换工作。

3 域名解析各环节对IPv6的准备

域名解析包括权威解析服务系统、递归解析服务系统和客户端。权威解析服务系统又包括根服务器、顶级域名解析和二级及以下权威域名解析3个部分。各环节目前对IPv6的支持程度不一。

3.1 根服务器

全球13个根服务器中，目前还有3个根服务器未支持IPv6，分别为C

根、E根、G根。而其他10个根均支持IPv6解析。

3.2 顶级域名

全球权威服务器里面有43%支持IPv6解析。中国国家顶级域.CN从2008年初开始提供IPv6解析，目前每天IPv6解析量大概在1 000万～2 000万次之间，约占总解析量的1%。

3.3 二级及以下权威域名

二级及以下权威域名服务系统的基础建设普遍比较薄弱，运维能力也参差不齐。为了抽样了解二级及以下权威域名服务系统对IPv6的支持，我们选择在中国境内使用最广泛的“.CN”、“.COM”和“.NET”顶级域下的122 055 373个二级及以下域名作为检测对象，所涉及的权威服务器数量3 628 091台。结果表明，权威服务器的IPv6解析支持率仅为0.77%，这是权威解析系统中对IPv6支持最为薄弱之处。

3.4 递归服务器

递归服务器对于EDNS0的支持已经非常广泛，对于递归服务器而言，只要保证定期升级解析软件，并增加冗余带宽来应对IPv6查询带来的更大流量，即可保证对IPv6新特性的支持。

3.5 客户端

目前，主流终端操作系统自身配备的解析器均可以支持IPv6解析，不过对于Happy Eyeball等过渡时期的新技术也需要通过开发升级来更好地进行支持。

4 DNS的新机遇

在DNS为IPv6提供支持的同时，IPv6同样为DNS带来了新的机会。

在20世纪80年代DNS协议设计初期，为保证域名系统在当时网络条件下的服务可达性，设计者推荐使用面向非连接的UDP协议，并将报文长度限制在最小传输单元512字节内，以确保报文在极端恶劣的网络环境下也能尽量迅速准确地传递。而512字节的报文长度限制是目前仅有13个根服务器的原因。在IPv6环境下，UDP最小传输单位增加至1 280字节，报文空间的扩展使得DNS协议可以增加更多属性，并使得根服务器的增加成为可能。

5 结束语

DNS已经具备了全面支持IPv6的能力，虽然某些特性的支持程度需要继续提高。在IPv6演进过程中，多项有关资源记录、协议扩展、IPv6过渡方面的域名系统标准已经或正在被制订和应用，域名系统正在全面支持IPv6应用。另外，域名系统对于如何解决在IPv6情况下降低解析时间和出错机会、如何更好的支持IPv6即插即用特性、如何解决DNS服务器的自动发现机制等方面会继续努力和提高。

在IPv6漫长的过渡期内，IPv4和IPv6两类网络间并存和互通所需要的过渡技术也已经成形，正准备接受实践的检验。DNS64技术可实现两类网络间的互访，而双栈技术和Happy Eyeball的结合解决了两类网络间并存的问题。

域名解析各环节正在进行IPv6和DNS升级的准备。就权威和递归解析服务环节而言，需要定期升级解析软件以保证对IPv6新特性的更好支持，需要新增IPv6接入，并增加冗余带宽来应对IPv6查询带来的更大流量；对客户端而言，各种终端操作系统的解析器也需要通过开发升级来实现对新技术的支持。

在DNS为IPv6提供支持的同时，IPv6同样为DNS带来了新的机遇。在IPv6环境下，UDP最小传输单位增加至1 280字节，报文空间的扩展使得DNS协议可以增加更多属性，并使得根服务器的增加成为可能。

随着IPv6的应用日益稳定增长，作为重要的互联网基础资源，域名系统将会与IPv6的新特性结合更加紧密，更好的适应互联网发展，提升用户体验。

参考文献

[1] 李丹, 吴建平, 崔勇, 等. 互联网名字空间结构及其解析服务研究 [J]. 软件学报, 2005,16(8):1445–1455.
[2] 何宝宏. IPv6面临的挑战与机遇 [J]. 电信网技术, 2011(5):35–38.
[3] 张鸿. 剖析IPv6时代的域名系统 [N]. 计算机世界, 2002.
[4] CRAWFORD M. DNS extensions to support IPv6 address aggregation and renumbering [S]. RFC2874. 2000.
[5] BUSH R. Representing Internet protocol version 6 (IPv6) addresses in the domain name system (DNS) [S]. RFC3363. 2002.
[6] THOMSON S. DNS extensions to support IP version 6 [S]. RFC3596. 2003.
[7] BAGNULO M. DNS extensions for network address translation from IPv6 clients to IPv4 servers [S]. RFC6147. 2011.
[8] MORISHITA Y. Common misbehavior against DNS queries for IPv6 addresses [S]. RFC4074. 2005.
[9] VIXIE P. Extension mechanisms for DNS (EDNS0) [S]. RFC2671. 1999.
[10] GUDMUNDSSON O. DNSSEC and IPv6 A6 aware server/resolver message size requirements [S]. RFC3226. 2001.
[11] 张鸿, 钱华林. IPv6对域名系统的需求及其解决方法的研究 [J]. 微电子学与计算机, 2003(1):35–38.

作者简介

冷峰，北京邮电大学硕士毕业；中国互联网络信息中心工程部副主任，负责域名服务平台规划工作；主持和参与基金项目5项。

王伟，南开大学博士毕业；中国互联网络信息中心副总工，负责国家域名系统技术工作；已发表学术论文8篇，申请专利2项，参与制订CCSA标准9项、IETF标准1项。

张跃冬，中国科学院计算技术研究所博士毕业；中国互联网络信息中心运行管理中心副主任，专注于网络服务平台架构、域名系统安全、域名运行监测分析等方面的工作。

DOI: 10.3969/j.issn.1009-6868.2013.02.007
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130228.1703.002.html

ICP 的 IPv6 演进

IPv6 Transition of ICP Networks

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0029-004

侯金刚/HOU Jin'gang
（腾讯科技有限公司 网络平台部，北京 100080）
(Department of Infrastructure Services, Tencent, Beijing 100080, China)

摘要： 探讨了因特网业务提供商（ICP）如何综合应用双栈、应用网关、地址翻译等技术推动 IPv6 的演进。基于不同型态、不同质量的应用将导致 IPv6 改造采用的技术方案、实施的复杂性不同，提出了准备-试点-梳理-实施 4 步走的路线图，以帮助控制 IPv6 演进的风险，使演进过程有序进行。

关键词： IPv6 协议；互联网业务提供商；双栈；应用网关

Abstract: In this paper, we discuss how Internet content providers use technologies such as dual stack, application gateway, and network address translation to improve IPv6 migration. The range of applications and application qualities may complicate IPv6 migration solutions and their implementations. In this paper, we propose a solution that has four steps: preparation, trial, brush, and implementation. This solution helps control risks during IPv6 migration and can be used for smooth IPv6 transition.

Key words: IPv6 protocol; ICP; dual stack; application gateway

1 IPv6 是下一代互联网的核心

互联网行业在全球以及中国一直处于高速发展的阶段，也面临着一系列新的挑战和机会。归结起来，互联网应用有三大的发展趋势：移动化、社区化、个性化。简单总结为 SoLoMo（Social + Localize + Mobility），也即随着 3G/4G 移动网络的发展，互联网应用基于的终端从 PC 越来越多地变成手机及各种 Pad 设备，而且个性化以及基于当时用户场景和位置信息的应用需求越来越多。用户习惯的改变，以及应用模式的革新对底层的传统的支撑网络提出了新的要求：

- 网络需要更扁平，更普遍可达，以支持远比今天的 PC 终端数量更大的设备的接入。
- 网络支持通过不同终端进行互联网访问的用户的身份和位置信息的识别与区分。
- 网络更灵活的支持用户间的 SIP 类、P2P 类互访，而不单单是用户到服务器的单一访问模式。

然而，以 IPv4 作为地址标识的当今的互联网存在着很多的局限，越来越阻碍着新一代互联网应用的部署：

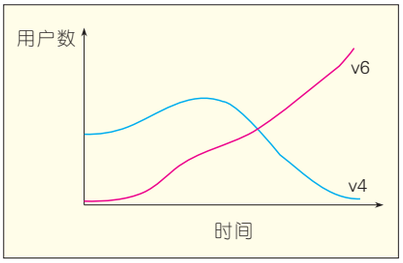
- IPv4 的公网地址已经分配殆尽，无法进一步提供给各互联网应用的服务，更不用说呈几何级数增长的终端设备；2011 年 2 月 3 日，ICANN 宣布全球 IPv4 地址分配殆尽，接着 APNIC 也在 4 月 15 日宣布可正常分配的 IPv4 地址块也没有了。
- 传统的 IPv4 网络在不同的区域，用户接入、移动终端等地方采用了不能全局路由的私网地址，并辅以网络地址转换（NAT）、代理服务器（Proxy）等转换技术，造成用户信息的不完整，带来如性能差而难以定位等问题。
- IPv4 网络没有很好地提供端到端服务质量的保障，对于互联网应用中采用类似 P2P 机制的存在诸多限制，影响用户访问体验。

而下一代地址编码 IPv6 则采用 128 比特编码，比采用 32 比特编码的 IPv4 地址空间大的多，理论上可以支撑地球上所有对象的地址标识需要，因此完全可以做到接入互联网的任何对象的全局唯一地址，无论是 PC 终端、移动终端，甚至是物联网下的任何物体；采用 IPv6 对互联网下的对象进行全局唯一标识，也就为互联网应用识别用户身份、提供端到端的服务质量保证创造了条件。因此，理论上 IPv6 比 IPv4 可以更好的支持新型互联网应用的持续不断发展，为用户体验的改善提供更好的支撑[1-6]。

2 IPv6 演进中存在着诸多挑战和问题

IPv4 和 IPv6 用户数发展预测如图 1 所示。IPv4 往 IPv6 演进是大势所趋，然而这个过程不会一蹴而就，需要解决诸多技术难题以及非技术因

收稿日期：2012-12-24
网络出版时间：2013-02-28
基金项目：国家发改委（发改办高技[2012]291 号）



▲图1 IPv4和IPv6用户数发展预测

素，而且这个演进过程必将是漫长的和复杂的。

首先，用户端可能只支持IPv4、IPv6或二者都支持，网络也是IPv4的网络与IPv6的网络并存，这必然带来互联网的应用为了兼容性的考虑，不得不同时支持IPv4与IPv6。

第二，Web 2.0时代中互联网应用并不都是用户端同服务器端的通信，也存在着大量的用户端间的直接的访问，如文件传输、内容下载、语音与视频通信等等；在纯的IPv4网络或者纯的IPv6网络里，用户端间直接建立网络连接进行通信并不存在问题，但IPv4与IPv6的用户如何通信就是一个非常值得探讨的事情，这是现实业务改造中面临的诸多技术难题之一。

第三，也是最主要的一个问题，很多的互联网应用都是历经多年、多人开发完成的，其间很少考虑到IPv6，所采用的程序并不都用域名或其他的名字服务，支持的地址标识字段也仅仅适合IPv4而不适合IPv6，如果把这些应用的代码重新梳理一遍并改造使之支持IPv6，跟重写所带来的工作量也差不了多少。腾讯几年前为了让QQ支持的注册用户数突破21亿，即把其中的32位整数改为仅支持正数，就修改了500多个功能模块，130多万行代码，涉及到2 000多名程序员，耗时15个多月[2]。而如果让QQ支持IPv6，把其中的IP地址字段进行修改，同时相关模块支持IPv6地址的管理，这个工作量将比前者更大，复杂性也更高！

第四，互联网应用依赖的基础设施，服务器、操作系统、网络设备、防火墙、四层交换机、域名服务器（DNS）等过去只是为了支持IPv4而用，如果要把它们全部升级改造成支持IPv6将也是一项耗时巨长、投资巨大的工作。

第五，除了编码外，IP地址在互联网应用中也被用来识别用户的位置信息，以便向用户可以更精准地推送相关内容，譬如一个北京的用户登陆到互联网上，应用可以给他推送北京的天气信息，而其中主要的判断依据就是这个用户的IP地址信息；过去多年，互联网应用根据IPv4的地址分配信息以及用户的动态访问信息，建立起一个相对比较完备、比较准确的地址库，用来判断用户的位置信息，以便进行内容的推动、网络就近性访问的调度。然而，IPv6的使用量还很少，分配规则也不清晰，必将造成依赖IP地址作为位置标识的很多互联网应用演进到IPv6后造成很多功能的失效。

3 IPv6演进中关键技术探讨

因特网业务提供商（ICP）向IPv6演进涉及的技术和方案多而复杂，包括业务应用、代理接口、客户端、基础网络等多个层面，单单基础网络上就很难有一种理想的技术适合各种场景，满足各种不同类型的互联网应用的IPv6改造。因此，人们制订如下标准或原则来判断不同的技术对于应用IPv6改造的契合性，作为技术选择的依据：

（1）透明性。对于已有的互联网应用无需改动或者仅需少量的外挂式代码的变化就可以支持IPv6，而无需全部的代码遍历并修改。

（2）兼容性。增加新的技术模块或组件在支持IPv6的同时继续支持IPv4。

（3）渐进性。IPv6的功能可以逐步添加而无需只有整个应用全部的代码改造完成才可以使用。

这3个原则简称TCG，作为判定互联网应用IPv6演进使用的技术手段的规则。

对于互联网应用的IPv6演进，本文从ICP的角度出发，假设的场景是：支持IPv6协议的终端用户通过电信运营商的IPv6网络来访问互联网公司自己的数据中心内的应用；考虑到兼容性，同时有IPv4的终端用户也来访问这些应用；至于终端用户和电信运营商的网络如何支持IPv6或兼容IPv4，则不在本文的讨论范围。

根据以上假设，下面介绍下在ICP的IPv6演进中通常使用的3项技术方案：

（1）双栈技术

双栈技术采用应用从前端接入到中间逻辑处理以及后端数据管理，乃至应用本身的运营支撑模块全部改成IPv4和IPv6支持，整个应用程序中凡是涉及到使用网络地址编码的部分，之前只是支持IPv4的格式，现在需要增加关于IPv6的支持，对于其中依赖IP地址作为位置信息标识的还需要单独建立IPv6的地址库与位置的映射关系。同时，该应用路径上所依赖的基础架构包括DNS、全局负载均衡设备（GSLB）、路由器、交换机、四层交换机、防火墙、分布式拒绝服务攻击（DDoS）防护设备、后端的商业支持系统/运营支持系统（BSS/OSS）都需要全部改造使之既支持IPv4，又支持IPv6。这项技术对于互联网应用的改造非常彻底，但需要非常大的工作量，相当于原有代码的重写，对于老的、源代码缺乏维护的应用比较难操作。

（2）地址翻译

地址翻译技术在互联网应用的前端接入模块前部署一个NAT/PAT设备，把给应用分配的IPv6地址部署在这个设备上，并把IPv6地址与老的IPv4地址的对应关系也部署在这个设备上。当IPv6的用户访问互联网应用时，会经过这个翻译设备，把IPv6的源地址、目的地址转换成一对IPv4地址对，去访问真正的应用服务器，应用服务器返回的数据包经过翻

译设备时，再转换成真正的 IPv6 数据包，返回最终用户。这项技术无需应用进行修改，只是需要在 NAT/PAT 设备上维护 IPv6 与 IPv4 的映射关系，对于应用需要更多的地址信息的，在设备上提供映射关系的信息作为数据接口给到后端的应用使用。这项技术同时也可以应用在解决 IPv4 公网地址不足的场景，互联网应用使用的服务器全部配置私有 IPv4 地址，多个服务器对应一个公网地址，通过采用不同的传输控制协议（TCP）端口进行区分。

（3）应用网关

应用网关技术类似地址翻译技术，也是在互联网应用的前端接入模块前部署应用网关设备，前端支持用户的 IPv6 访问，后端的互联网应用保持 IPv4，应用网关设备在两边做代理，把 IPv6 的数据包与 IPv4 的数据包进行转换。所不同的是应用网关工作协议的第七层，把用户的 IPv6 连接终结在网关设备上，同时建立跟后端互联网应用服务器的新的 IPv4 连接，并在两个连接间起到代理的作用，其最大的价值在于可以从应用层识别应用协议，并加以处理，可以针对 IPv6 与 IPv4 的转换，提供更多应用层的处理，包括内容的过滤、安全控制以及应用状态的保持、转换等等。应用网关技术工作原理如图 2 所示。应用网关技术也可以应用在 IPv4 与 IPv6 的用户间进行 P2P 的通信。

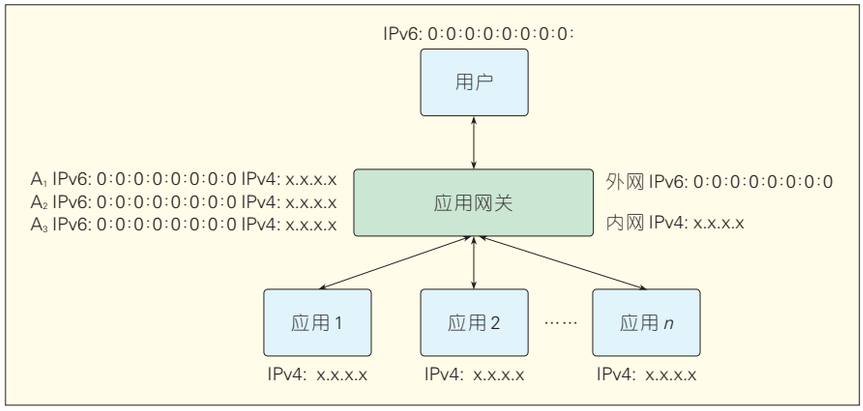
以上 3 种技术的 TCG 对比如表 1 所示。

对于 ICP 来讲，通常存在不同模式的应用，譬如纯粹的基于超文本链接标记语言（HTML）的 Web 应用，基于客户端的 C/S 结构的应用，核心模块由第 3 方开发、外围 BSS/OSS 自己开发的应用等等，不一而举，难以采用一种统一的技术或方案进行 IPv6 的演进。

一个技术框架方案来支持不同种应用的 IPv6 演进如图 3 所示。

图中，IX 为代表互联网应用所在的数据中心同外部运营商互联的网络路由器，A1、A2、A3、A4、A5、A6 为互联网应用。6 种互联网应用综合利用双栈技术、应用网关技术以及地址翻译技术来实现 IPv6 的演进，A1、A2、A3 无论是只支持 IPv4、IPv6 还是两者都支持，都需要通过一个应用网关被外面的用户通过 IPv4 网络或者 IPv6 网络所访问，应用网关需要支持 IPv4 间、IPv6 间以及 IPv4 与 IPv6 间的七层协议转换，可以应用在类似 B/S 结构的 Web 类应用，应用的服务器无需采用公网地址；A5、A6 两种应用，都已支持 IPv6，可以直接被外面的 IPv6 用户所访问，而 A4 老的 IPv4 应用需要经过一层地址翻译，才能被外面的 IPv6 所访问。

经过上述技术分析，结合诸如腾讯互联网公司的实际业务情况，可以看出，IPv6 改造采用的技术万不可一刀切，一定是由简入繁、多案并存、逐步过渡的方式。



▲图 2 应用网关技术

▼表 1 3 种技术的对比

技术	双栈技术	地址翻译	应用网关
透明性（T）	低	高	高
兼容性（C）	高	中	中
渐进性（G）	低	高	高

4 IPv6 演进策略

互联网应用型态的多样性，历经多年版本迭代的复杂性，决定了 IPv6 演进的长期性与困难。因此，可以如下的原则确定演进的策略与步骤：

（1）基础架构先行，应用后至；首先在网络、服务器、操作系统、DNS 等方面激活 IPv4、IPv6 双栈，所谓“先修路，后跑车”。

（2）应用先易后难，逐步推进；优先解决 Web 类的，逻辑简单的应用；

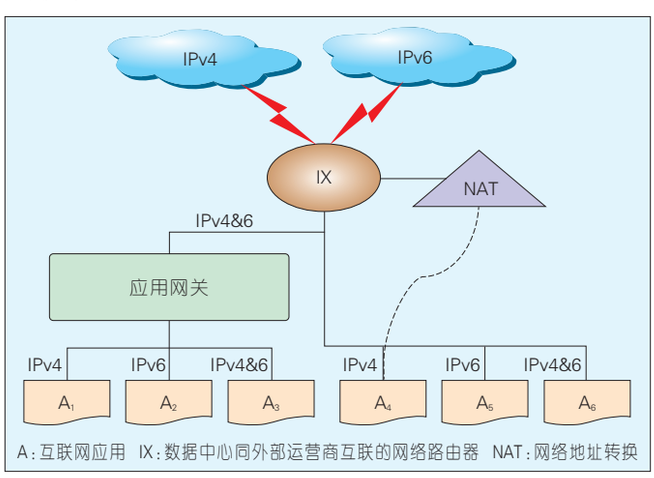


图 3▶ IPv6 演进方案

而后解决平台型的、C/S结构的应用；次之，解决代码很久的应用；最后，解决运营支撑类的应用。

（3）从试点开始，逐步推广。

下面提供一个4步走的路线图，供参考：

第一步，准备：评估各项IPv6演进技术的适配性、需要的环境，同步电信运营商网络的节奏。

第二部，试点：搭建一个IPv6演进的试点平台，让基础架构支持双栈，准备采用的各种IPv6演进的技术模块如双栈、地址翻译、应用网关均部署，选取一两个代表性的应用进行改造，并使之IPv6访问。

第三步，梳理：把待改造的互联网应用进行梳理，应用的型态、代码量、代码的质量、对于IP地址的依赖情况等一一梳理清楚，并同准备采用的演进技术进行对齐，确定各自的演进工作量、复杂度与风险。

第四步，实施：根据第三步的梳理结果和制订的演进方案，进行实施操作，并控制好演进过程中的风险。

就腾讯公司而言，从2010年开始正式启动IPv6的演进项目，成立了跨公司各个团队的联合项目组，包括因特网数据中心（IDC）资源、业务应用、运营支撑、安全支持、政府关系多个部门在内的专家一道来共同解决IPv6的演进，主要的进展：

- 构建了一个包含200台服务器规模的IPv6试验平台，这个平台从网络接入、路由器、防火墙、四层交换模块（LVS）、网络交换、服务器、操作系统等都端到端的支持双栈协议，并接入了CERNET的IPv6网络。
- 在这个平台上通过应用网关的方式接入了www.qq.com及www.pengyou.com两个公司的主应用，并且修改了公司的域名服务器，使得这两个域名自然的既支持IPv4的访问，又支持IPv6的访问。
- 截至2012年底，试验平台上的IPv6用户数达200多万户，流量达到300 Mb/s。

下一步的计划：

- 按照中国国家发改委的IPv6演进项目的要求以及中国电信、联通、移动三大运营商的IPv6改造节奏，在深圳、上海、天津3地接入三大IPv6线路，支持应用的双栈连接。
- 在2013年，完成5个主要网站类应用www.qq.com、www.pengyou.com、www.qzone.com、open.qq.com及t.qq.com的IPv6应用支持改造，主要采用地址翻译技术、应用网关技术，以及在部分模块尝试端到端双栈技术。
- 同步完成关于IPv6接入的DDoS防护、安全管理模块的改造支持等。
- 2014年后，启动客户端应用，包括QQ、微信等应用的IPv6迁移改造工作。
- 2015年后，启动腾讯公司平台上合作的第三方应用的IPv6迁移改造工作。

5 结束语

IPv4地址分配殆尽，向IPv6演进是唯一可行之路，在此基础上人们将推动下一代互联网朝着普遍可达、端到端服务质量可期的目标走。双栈、地址翻译、应用网关等都能帮助IPv6的演进，然而，不同型态的应用，不同质量的应用，会使得IPv6的改造采用的技术方案、实施的复杂性不同。准备–试点–梳理–实施，4步走的路线图可以帮助控制IPv6演进的风险，使得整个演进过程有序进行。

参考文献

[1] 雒江涛, 李晴川. 基于云存储的分组域监测系统 [J]. 重庆邮电大学学报（自然科学版）, 2012,24(6):675–681.
[2] World Ipv6 Launch [EB/OL]. [2011–03–25]. http://www.worldipv6launch.org. 2011.
[3] 李维贤. IP城域网通过BGP协议接入客户网络 [J]. 数字通信,2012,(3):69–71.
[4] 腾讯公司内部项目资料 [EB/OL]. [2011–09–21]. http://www.tencent.com. 2011.
[5] 廉志玲, 张代忠, 张小菊. 一种基于查表的解速度模糊算法 [J]. 雷达科学与技术, 2011,9(4):49–52.
[6] 王淑惠, 谭清中, 唐彦, 肖亮. IPv6是物联网最佳的寻址技术 [J]. 数字通信, 2011,(3):28–31.

作者简介

侯金刚，北京大学计算机科学硕士毕业；腾讯公司网络平台部副总经理、总架构师，负责公司IDC资源的规划与分布策略、网络架构的研发与日常运营工作；已发表学术论文3篇。

综合信息

中兴通讯发布创新的全系列LTE小基站回传解决方案

【本刊讯】2013年2月4日，中兴通讯近期发布了其创新的全系列LTE小基站回传解决方案。这一解决方案不仅满足了目前日益增长的移动宽带需求，也为后续小基站的大规模部署奠定了坚实的基础。

智能终端的广泛使用给人们的日常生活带来了极大的便利，但同时也给运营商带来了很多新的问题。为了更好地提升用户数据业务的体验，网络模型将逐步从宏站为主转变为宏、微并重的覆盖模式。小基站将在今后网络建设中扮演越来越重要的角色。

中兴通讯的全系列LTE小基站回传方案可以满足上述要求。其最大亮点在于，依托丰富的产品线资源，有机整合了多种传输技术，比如分组承载网、固网宽带（如PON/xDSL/Cable等）、微波传输、以及最新的Wi–Fi和TD–LTE无线回传技术。在实际组网中可以根据实际情况任意组合，满足各种场景的要求。

DOI: 10.3969/j.issn.1009-6868.2013.02.008
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130228.1703.004.html

IPv6-only 网络互通技术

Preliminary Research on IPv6-only Network Interoperability Technology

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0033-003

刘笑寒/LIU Xiaohan[1]
闫双舰/YAN Shuangjian[1]
孙玉勤/SUN Yuqin[2]
（1. 中国科学院计算机网络信息中心，北京 100190；
2. A10 网络有限公司，北京 100738）
(1. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China;
2. A10 Networks Inc., Beijing 100738, China)

摘要： 对 IPv6-only 的网络互通技术从过渡技术、访问场景、应用需求 3 方面进行分析，对网络构建从应用、数据、状态、操作系统 4 个维度进行探讨。对 IPv6-only 网络访问 IPv4 internet 这一典型场景，采用 NAT64/DNS64 翻译技术进行系统构建，从操作系统、浏览器、网络应用（网页、即时通信、流媒体）等方面进行测试，并基于测试结果对 IPv6-only 的实际部署进行讨论。

关键词： 下一代互联网；纯 IPv6；网络互通技术；NAT64/DNS64 翻译技术

Abstract: In this paper, we analyze IPv6-only interoperability in terms of transition technologies, visiting scenarios, and application requirements. We also analyze network construction in terms of applications, data, state, and OS. We propose a practical system for accessing IPv4 internet from an IPv6-only network. This system uses NAT64/DNS64 translation technology. We test the OS, web browser, and Internet applications such as web sites, instant messaging, and streaming in this system. We use these test results as the basis for a discussion on the practical deployment of IPv6-only network.

Keywords: next-generation Internet; IPv6-only; network interoperability technology; NAT64/DNS64 translation technology

1 IPv6-only 网络环境背景分析

2012 年 6 月的 IPv6 日，世界范围内全面启动 IPv6。在中国，移动、联通和电信大范围开展现网 IPv6 试点。在中国下一代互联网发展建设峰会上，与会专家明确了重视 Android 系统的 IPv6 二次开发，并主要采用双栈技术（RFC4213）[1]的方针。

随着 IPv6 研究工作和实际部署的不断开展，网络迟早会进入纯 IPv6（IPv6-only）的时代。现在人们提出了很多 IPv6 的过渡计划，大多都遵循 RFC5211[2]中提出的 3 个渐进阶段：准备阶段、过渡阶段、后过渡阶段。

收稿日期：2012-12-04
网络出版时间：2013-02-28

目前很多的资源还在 IPv4 环境中，因此 IPv6 和 IPv4 网络之间的互通尤为重要。在 IPv6-only 环境下，现有的硬件设备和软件可以应用到什么程度，访问网络会出现什么问题，以及如何解决这些问题并没有明确的官方资料。

在研究领域，爱立信的 J.Arkko 等于 2010 年开始构建 IPv6-only 办公环境，并于 2012 年 4 月确定 RFC6586[3]。WIDE Project 从 2011 年 9 月开始在 WIDE Camp 构建 IPv6-only 实验环境，并提出草案[4]。

2012 年 11 月中国下一代互联网示范工程（CNGI）专家委员会主办的 IPv6 过渡技术国际测试大会，对业内支持 IPv6 过渡技术的产品进行了摸底测试，推进了设备的产业化进程。

本文对 IPv6-only 的网络互通技术从过渡技术、访问场景、应用需求 3 方面进行分析，对其中涉及的 IPv6-only network 访问 IPv4 Internet 场景，采用 NAT64/DNS64[5-7]翻译技术进行系统构建和应用测试分析，并对实际 IPv6-only 环境部署遇到的问题和网络设计进行讨论。

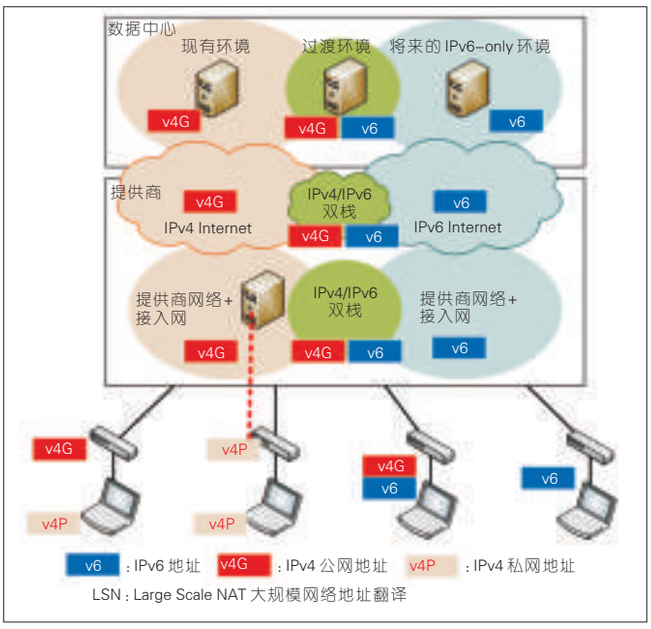
2 IPv6-only 网络互通技术分析

2.1 过渡技术分析

数据中心、网络提供商的网络环境以及网络地址的渐进情况如图 1 所示。在现网 IPv4 环境下，普遍采用网络地址翻译（NAT）技术解决地址缺乏问题。随着 IPv4 地址分配完毕和 IPv6 的发展，过渡环境中普遍采用双栈。新出厂的终端将支持 IPv6，用户可以在数据中心、网络提供商提供的 IPv6-only 网络环境下工作。

IPv4 到 IPv6 的过渡技术[8]分为双栈、隧道和翻译 3 种。3 种过渡技术各有其优缺点，应针对不同场景配置不同的过渡技术。

从 IPv4 到 IPv6 的过渡形态来看，



◀图 1
网络地址渐进情况

对于 IPv6-only 环境的新用户，资源访问应包括 4 个类型：本地 IPv6 访问、一次翻译、双重翻译、隧道封装。

当用户访问服务器时，首先采用本地模式，即 IPv6 to IPv6。若主机和服务器在不同环境下，主机尝试一次翻译，即 IPv6 to IPv4。若一次翻译访问失败（应用层网关（ALG）发生问题等原因），主机尝试双重翻译，即 IPv6 to IPv4 to IPv6。若一些应用（如 IPSec）在双重翻译模式下无法工作，则采用隧道封装进行访问，即 IPv4 over IPv6。

2.2 访问场景分析

本文涉及的 IPv6-only 环境的新用户，应用程序可能是 IPv4 或 IPv6 的，在边缘网络访问 IPv4 Internet 或 IPv6 Internet。

对于 IPv4 Internet 的访问可分为[9] 3 种场景：

（1）使用 IPv4 访问 IPv4 Internet。

（2）使用 IPv6 访问 IPv4 Internet。

（3）使用 IPv6 访问 IPv6 Internet。

场景（1）对应于现在的过渡阶段，大量资源都在 IPv4 Internet 上，很多应用还只支持 IPv4。可采用 IPv4 over IPv6 隧道技术，如 Public 4over6、Lightweight 4over6、DS-Lite 技术等。用户设备作为隧道发起点，网络出口的路由器设备作为隧道汇聚点。也可采用双重翻译技术，如 MAP-T、464XLAT 技术。

场景（2）对应于的用户设备支持 IPv6 单栈。可采用一次翻译技术，如有状态的 NAT64/DNS64、无状态的 IVI 技术等。网络边缘路由器支持双栈。

对于 IPv6 Internet 的访问，场景（3）即全网已升级为 IPv6。可采用 IPv6 单栈进行本地访问。若有用户设备的上层应用仍使用 IPv4，则需要进行用户侧的翻译。

2.3 应用需求分析

IPv6-only 边缘网络中用户的应用需求可分为：

（1）IPv4 环境下流行应用。

（2）IPv4/IPv6 可以互通解决的各种应用。

（3）IPv6 环境下的杀手级应用。

应用（1）包括目前流行的操作系统、网络应用、网络安全等。其中网络应用包括客户/服务器（C/S）架构和对等（P2P）架构的应用。

应用（2）指针对现有特定社会需求或科学研究而开展的需要 IPv4/IPv6 互通解决的应用，如异构或同构网络间大数据传输等科研应用。

目前运营商的地址短缺问题也可采用大规模网络地址翻译（LSN）或者地址+端口（A+P）等方式解决，此类应用也列于应用（2）。

应用（3）指通过 IPv6 能够提供更好解决方案的应用，如网络安全等。

3 IPv6-only 网络构建及应用测试

3.1 网络构建分析

IPv6-only 网络构建可通过 4 个维度进行分析：

（1）应用：可分为普通应用程序和应用层网关（如嵌入地址的网络内容）。

（2）访问：可分为本地、一次翻译、双重翻译、隧道封装。

（3）状态：分有状态和无状态。

（4）操作系统：可从 IPv6 地址角度和域名解析角度进行分析。RFC4291[10] 和 RFC5156[11] 中列举了各种用途的 IPv6 地址，如单播地址、链路本地地址、站点本地地址、6to4、Teredo 等各种特殊的地址。

根据以上分析，我们选取 IPv6-only 边缘网络环境中，新用户使用 IPv6 访问 IPv4 Internet 这个典型场景进行第一步的网络环境构建和互通技术分析。

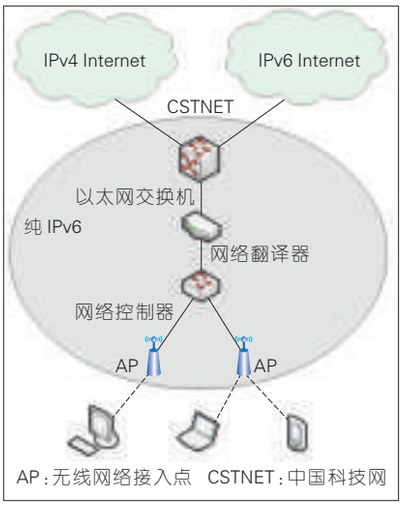
3.2 网络构建

我们于 2012 年 5 月初在中国科学院计算机网络信息中心 4 号楼办公区构建 IPv6-only 实验网络，针对 IPv6-only 访问 IPv4 Internet 场景，采用 NAT64/DNS64 技术。

系统构建如图 2 所示。网络翻译器下行连接网络控制器、无线网络接入点（AP），对用户终端提供网络连接；上行连接三层以太网交换机，接入中国科技网（CSTNET）。

本实验环境采用 A10 Networks 的 AX2600-GCF 网关实现 NAT64/DNS64；

用户终端关闭 IPv4，只打开 IPv6；采用无状态地址分配[12]，NAT64 转换前缀为 2001:cc0:2003:2::/64；DNS64 服务器地址为 2001:cc0:2003:2::243；网络配置时间为 2012 年 5—7 月。10 名研发人员使用 PC 机、手机、平板电脑等终端设备在此网络环境下测试。



▲图 2 网络系统配置

3.3 网络应用测试

网络应用测试主要测试 IPv4 环境中流行的应用，包括操作系统和网络应用。网络应用包括客户/服务器架构和对等（P2P）架构的应用。

3.3.1 操作系统

测试了 Windows7、WindowsXP、Ubuntu10.04、Fedora16、Android4.0、Mac10.6 等操作系统。

Windows7、Ubuntu10.04、Mac10.6、Android4.0 支持 IPv6-only。其中，采用 Android4.0 的用户终端需进入开发模式配置静态 IPv6 域名服务器（DNS）。Mac10.6 系统用于邻居发现的地址（默认 IPv6 地址）和用于接送发送标识的地址需指定为同一地址。

WindowsXP 不支持 IPv6-only。其原因是 WindowsXP 没有安装 IPv6 的部分支持协议如 DHCPv6 等。

Fedora16 无法支持 IPv6-only。在实验中配置了 IPv4 后可使用 IPv6。Fedora 的网络管理器的 IPv6 基于 IPv4 连接。2012 年 5 月更新的 Fedora17 可支持 IPv6-only。

3.3.2 网络应用——网页

在 IPv6-only 环境下，我们测试了 IE9、Firefox、Chrome3 款浏览器，均可访问 IPv4（只有 A 记录）网站，其中 Chrome 须在配置页面打开 IPv6 功能。

目前的网站分为仅有 A 记录（指向 IPv4 地址的记录）的网站（Yahoo 中国、新浪等），以及同时具有 AAAA 记录（指向 IPv6 地址的记录）和 A 记录的网站（Yahoo.com、Google.com 等）。

访问网站时，根据 NAT64/DNS64 工作机制，DNS64 首先查询 AAAA 记录，当存在 AAAA 记录时，客户端直接通过路由器进行 IPv6 访问；当 AAAA 记录不存在时，查询 A 记录，DNS64 将网站地址转换为 IPv6 地址发送给客户端，客户端通过 NAT64 翻译发给路由器后，访问网站。

我们进行了将 DNS64 设置为仅查询 A 记录的实验，顺利访问 Yahoo 中国、新浪等 IPv4 网站，验证了 NAT64/DNS64 机制的有效性。

3.3.3 网络应用——视频

我们访问了新浪、搜狐、优酷、土豆、酷六等网站的视频，其中仅新浪的大部分视频可顺利访问。经抓包分析，新浪的大部分视频采用按照域名对视频文件进行查询的方式，而其他网站视频采用的是按照 IPv4 地址进行查询。NAT64/DNS64 无法解决 IPv4 地址嵌入的情况。

3.3.4 网络应用——客户端应用程序

我们测试了 QQ、MSN、Renren、Skype 等即时通信客户端程序，均无法通信，即利用数据报协议（UDP）的客户端通信基于 IPv4，无法在 IPv6-only 环境下工作。

根据 RFC6120[13] 中对可扩展通信和表示协议（XMPP）的定义，服务器的标识等应采用域名。因此相应地，Web QQ、Web MSN、Web Renren、Web Weibo 等基于 XMPP 协议或超文本传输协议（HTTP）的通信可以正常工作，除非嵌入 IPv4 地址。

3.4 测试结论与讨论

从测试结果来看，构建 IPv6-only 网络环境是可行的，满足办公的基本需求。目前很多硬件软件都支持 IPv6，但并不能完全支持 IPv6-only，推荐采用最新的操作系统，如 Window7、Mac10.6、Fedora17、Android4.0、Ubuntu10.04 以上版本。

本次实验采用网络层的翻译机制 NAT64，对于嵌入 IPv4 地址的应用程序，需要设置应用层网关配合 NAT64 工作，或采用应用层的翻译机制。以文件传输协议（FTP）为例，工作方式分为主动模式和被动模式，有 IPv4 环境中的 FTP 协议[14] 和 IPv6 环境中的 FTP 协议[15]。在本次实验场景的翻译过程中需要在应用层对嵌入 IP 地址的 FTP 指令及相应连接参数做逐条翻译，如 EPSV 到 PASV，EPRT 到 PORT 等[16]。

IPv6-only 的实际部署，如对于不能访问的采用地址进行索引的视频的处理、只支持 IPv4 的网络程序等，需要网络内容服务商（ICP）的支持。现阶段仍推荐采用双栈。

实验采用手工配置 DNS 和无状态自动分配地址，需进一步测试基于路由器公告（RA）的 DNS 自动发现和 DHCPv6 功能。

4 总结及未来研究计划

本文对 IPv6-only 网络互通技术从过渡技术、访问场景、应用需求 3 方面进行分析。对 IPv6-only 网络中的用户访问 IPv6 Internet 的场景采用 NAT64/DNS64 进行系统构建，在基本应用测试的基础上进行了分析。

近期讨论的实验包括 IPv6-only 环境中 IPSec 等安全性测试和 FTP 等需配置应用层网关的应用测试等。

➡下转第 43 页

DOI: 10.3969/j.issn.1009-6868.2013.02.009
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130220.1136.001.html

IPv6 测试研究与实践

IPv6 Testing: Research and Practice

中图分类号：TN915.04 文献标志码：A 文章编号：1009-6868 (2013) 02-0036-004

宋林健/SONG Linjian
李震/LI Zhen
（北京天地互连信息技术有限公司，北京，100028）
(Beijing Institute of Internet, Beijing,100028, China)

摘要： 从 IPv6 协议测试范围和测试方法两个方面对 IPv6 协议测试进行了研究分析，并从设备、网络和网站系统等 3 个方面探讨论了 IPv6 测试的规范和体系。基于近年 IPv6 测试领域相关成果，介绍了 IPv6 论坛的测试体系 IPv6 Ready 和 IPv6 Enabled，以及全球在 IPv6 测试方面的实践和探索。

关键词： IPv6 测试；互操作性；一致性；测试范围；测试方法

Abstract: In this paper, we discuss the scope and methods of IPv6 protocol testing. We discuss IPv6 testing framework in terms of devices, network, and website testing. We also discuss international studies on IPv6 Ready and IPv6 Enabled and give an overview of global IPv6 testing practice.

Keywords: IPv6 testing; interoperability; conformance; scope of testing; method of testing

2011 年 2 月 3 日，互联网地址编码分配机构 IANA 的报告宣称全球最后 IPv4 地址块已经分配完毕[1]，正式宣告世界范围内 IPv4 地址已经枯竭。IPv4 地址的缺乏制约和影响了许多发展中国家互联网的发展速度，采用 IPv6 协议提升移动互联网的承载能力、可扩展性和可管控性是当前互联网领域的主要发展方向。

IPv6 在将近 15 年的发展和部署过程中，与 IPv6 相关的协议内容涵盖了 IPv6 基本协议、IPv6 组播、IPv6 移动协议、安全与管理协议、物联网协议、以及 IPv6 过渡协议等多个部分。IPv6 测试也从最初的 IPv6 基本协议测试，发展到如今对包含多种形式的 IPv6 终端、IPv6 路由交换设备、网络过渡设备，支撑系统、安全管理系统、互联网业务应用等多方面为整体的协议、设备、网络和网站测试体系研究。IPv6 测试研究成为新一轮全球互联网研究的热点，受到业界广泛认可，并作为验证 IPv6 互联网领域相关设备、物联网设备、服务/内容提供商（SP/CP）、运营商对 IPv6 相关协议及标准的符合程度，检验 IPv6 部署方案是否具备商用水平的关键手段。

以 IPv6 为标志的下一代互联网测试研究一直受到各国政府的大力支持。早在 2005 年欧盟启动 Go4it 计划，建设欧盟的下一代互联网测试平台[2]。2008 年 6 月，美国国防部宣布启动 IPv6 产品认证计划，政府各部门 IT 网络采购需全面满足 USGv6 文件和 IPv6 Ready 测试的设备通过认证确保相关产品能够满足国防部的任务要求。日本政府采用了全球 IPv6 论坛 IPv6 Ready 标识认证计划。日本的主要信息终端厂商如索尼、东芝、日立、松下等厂商的产品都支持 IPv6，在 Phase I 中的日本企业超过 30%，Phase II 中占 50%。

本文将从 IPv6 协议测试入手，讨论测试的范围和方法。然后介绍 IPv6 测试的规范和体系，其中包含设备、网络和网站系统等各个方面。最后文章将介绍近些年 IPv6 测试领域相关的实践和成果，其中包括 IPv6 论坛的测试体系 IPv6 Ready 和 IPv6 Enabled，以及在全球领域 IPv6 测试方面的实践和探索。

1 IPv6 协议测试研究

随着互联网飞速发展，网络规模越来越大，再加上 IPv6 的广泛部署和应用，稳定高效的 IPv6 协议将成为网络运行安全和稳定的保障。本节将从 IPv6 协议测试范围和测试方法两个方面来对 IPv6 协议测试进行研究分析。

1.1 IPv6 协议测试范围

因特网工程任务组（IETF）是国际上推动 IPv6 标准化的主要力量，自 1994 年以来已经制订了 1 000 多项有关 IPv6 的标准，相关内容涵盖命名、寻址、路由、演进机制、安全、管理等领域[3]。如今 IPv6 的核心标准已经完成包括 IPv6 基本协议、地址结构、数

收稿日期：2012-12-27
网络出版时间：2013-02-20
基金项目：国家科技重大专项（2010ZX03004-002-03）

据包头、邻居发现机制、地址分配机制等。目前的主要工作集中在IPv4向IPv6过渡、基础网络对业务网络的支撑和适应，以及基础网络自身的管理、完善和演进上。虽然IPv6相关协议很多，从国际IPv6论坛Ready Logo认证[4]的测试标准来看IPv6协议测试集中在寻址、路由、移动、安全、管理等核心协议上。IPv6 Ready Logo涉及的IPv6协议和设备如表1所示。

▼表1 IPv6 Ready Logo涉及的IPv6协议和设备

IPv6协议类型	IPv6相关标准	测试设备
IPv6核心协议	RFC1981、RFC4861、RFC4862、RFC2460、RFC5095、RFC4291、RFC4443、RFC2474、RFC3168	路由器、主机
地址分配协议DHCPv6	RFC3315、RFC3646、RFC3736	客户端、服务器、Realy代理
安全协议IPsec	RFC2404、RFC2410、RFC2451、RFC3602、RFC3566、RFC3686、RFC4301、RFC4303、RFC4305、RFC4312	终端节点、安全网关
安全协议IKEv2	RFC4306、RFC4307、RFC4718	终端节点、安全网关
组播协议MLDv2	RFC3810、RFC4604	路由器、监听节点
管理协议SNMPv6	RFC3416、RFC3418、RFC2578、RFC2579、RFC2580	代理、管理节点
设备类型协议CPE	RFC6204、RFC5942、RFC4294	用户驻地设备
移动协议MIPv6	RFC3775、RFC3776、RFC4877	CN、本地代理、MN
移动协议NEMO	RFC3963、RFC3775、RFC3776、RFC4877	本地代理、移动路由器
传输协议SIP	RFC3261、RFC3264、RFC4566、RFC2617、RFC3665	UA、终端节点、B2BUA、Proxy、注册器
移动通信协议IMS	TS24.229、TS.33.203 RFC3261、RFC3265、RFC3320、RFC3327、RFC3455、RFC3485、RFC3608、RFC3680、RFC4320、RFC4566、RFC4896、RFC5049、RFC3310、RFC3329、RFC2617、RFC3665	终端

CN：通信对端　MN：移动节点　UA：用户代理　B2BUA：背对背的用户代理

1.2 IPv6协议测试方法

协议测试传统上来说分为一致性测试、互操作性测试（又称互通性测试）和性能测试3部分，在IPv6协议测试领域亦如此。就IPv6协议测试而言，全球主要的研究工作集中在IPv6协议的一致性测试方面。对IPv6协议的一致性测试研究又主要集中在协议形式化描述、测试结构、抽样测试集和测试例生成方法等几个方面。比如对移动IPv6的协议形式化描述和一致性测试方法[5]，RIPng协议的一致性研究[6]等。也有对自动化测试方法进行的研究，如对邻居发现协议的自动化测试[7]。

互操作性测试会根据具体协议和测试目的的不同采用不同的方法和手段，一般没有固定的标准和模式。欧洲电信标准化协会（ETSI）曾经出版过一个IPv6测试框架[8]和一个互操作性测试方法[9]，定义了IPv6协议互操作性的基本概念、单元、流程和形式化描述方法，为IPv6测试的互操作性测试提供一个样例和模板。在性能测试方面，IETF有一个专门的工作组BMWG在做这方面的标准，比如三层设备的转发性能测试标准RFC2544。其中关于IPv6性能测试在RFC5180中有相关的描述和规定。

国际上在IPv6测试研究方面比较领先的有日本TAHI项目组、欧洲电信标准化协会（ETSI）、美国New Hampshire大学的IOL实验室。中国在IPv6协议测试领域研究的工作主要来自中科院计算所的李忠诚教授、清华大学的尹霞教授和北京邮电大学马严教授的研究室。

2 IPv6评测规范和体系研究

制订IPv6支持度评价指标体系以及相应的测试规范，是衡量设备、网络、网站系统升级改造是否支撑IPv6的大规模推广部署，设备是否能够满足演进过程中的各种组网需求的关键。

2.1 IPv6设备测试

对设备的测试要求如同对协议测试要求，也分功能性测试（包含一致性与互通性）与性能测试，具体测试方法的研究在前面已有具体描述，本节主要作为IPv6评测规范体系的一部分，列举设备测试所涉及的协议和测试方法。

IPv6设备测试主要包括：制订IPv6网络设备（包括终端设备、接入设备、交换设备及路由设备）、IPv4与IPv6网络互通设备的测试指标和测试规范，以指导对设备的IPv6功能进行验证，检查设备是否达到相关RFC要求的功能，及对设备的IPv6相关性能进行测试。设备的IPv6功能和性能测试如表2所示。

2.2 IPv6网络测试

IPv6网络测试主要包含基础网络和业务支撑系统两部分。目前，IETF IPPM工作组定义了网络通用测度框架，涵盖网络性能、网络可用性、网络扩展性和网络可靠性等，主要测试指标包括：吞吐量、丢包率、延迟、网络有效性、服务有效性、链路利用率、路径容量和可用带宽、资源利用率以及多业务承载能力和服务质量等。根据上述指标的测量，从功能方面要求网络具备IPv6地址的分配能力，能够实现IPv4/IPv6混合流量或纯IPv6流量的转发，能够承载多业务及路由学习；从性能方面，主要是测评网络IPv6及IPv4/IPv6混合流量模型的吞吐量、丢包率、往返时延以及与过渡技术方案相关的性能指标，包括隧道性能和翻译性能；从网络扩展性方面，主要是测评网络路由规模（包括IPv4/IPv6路由）及路由收敛能力。

IPv6环境下的业务支撑系统要求兼容现有IPv4用户使用习惯，在IPv6网络环境中，提供之前IPv4环境下的所有业务支撑和网络管理功能，

▼表2 设备所具备IPv6功能

测试项目	测试内容	测试类别
IPv6核心协议	IPv6基本协议、ICMPv6协议、DHCPv6协议、邻居发现协议、路径最大传输单元发现协议、地址结构及类型协议、地址配置协议等	功能测试
IPv6路由协议	包括BGPv4+、OSPFv3、RIPng等协议	功能测试
IPv6过渡协议	DS-Lite、NAT64、6rd、IVI、4over6等协议	功能测试
IPv6组播协议	MLDv2、IGMPv3、PIM-SM、PIM-DM等协议	功能测试
IPv6网管协议	SNMPv3、SNMP over IPv6等协议	功能测试
IPv6安全协议	IPsec、真实源地址认证协议等	功能测试
IPv6基本转发性能	RFC2544中规定的吞吐量、丢包率、延迟、背靠背帧	性能测试
路由协议性能	BGP4+、OSPFv3、RIPng的路由表大小、路由收敛性能	性能测试
组播协议性能	MLDv2、IGMPv3、PIM-SM、PIM-DM等协议支持的组播组数量、组播数据复制转发能力	性能测试

相关业务支撑系统具有单独IPv6域名，实现IPv6的所有业务属性。业务支撑系统主要包含数据库系统、数据存储服务系统、客户关系管理（CRM）、安全防范系统（SPS）、传输服务访问点（TSAP）、认证计费授权系统、网关系统等几部分。分别在功能性，可靠性和兼容性等方面有要求，比如域名服务器（DNS）要求能够支持AAAA格式的查询报文和响应报文，通过递归查询的方式对用户的AAAA查询予以响应；能够支持AAAA格式的域名记录，对用户IPv6网站实现IPv6域名的本地授权解析；支持IPv6单栈环境下响应用户发起的纯IPv6的DNS查询。

2.3 网站评测

网站IPv6支持度的主要测试指标如表3所示。

3 IPv6测试实践与成果

3.1 IPv6 Ready和Enabled

从IPv6测试的实践来说，国际IPv6论坛一直处于领导地位，已经形成以全球IPv6 Ready和IPv6 Enabled为核心的网络协议测试和网站评测体系。IPv6 Ready Logo测试认证包括一致性测试和互通性测试，其认证主要目的是保障IPv6设备和应用软件互联互通能力，为全球IPv6商业部署树立信心，IPv6 Ready Logo既包括核心网协议认证，也包含CPE路由器认证等，同时也提供DHCPv6、IPsec、SIP、MLDv2、IKEv2、MIPv6、NEMO等扩展协议认证。为了适应IPv6网络大规模部署，IPv6论坛正筹划对重要的IPv6过渡协议，如4over6、NAT64、6rd等国际标准在IPv6 Ready的框架下进行开发测试。

IPv6 Enabled Logo认证是由中国团队（天地互连）主导开发，主要目标是鼓励更多的应用和服务支持IPv6，并且能提供持续的基于IPv6的服务。IPv6 Enabled Logo参与2011 IPv6 Day活动，为网站提供IPv6支持度测试服务，第2年继续与ISOC合作，为参加World IPv6 Launch的网站提供IPv6测试服务。IPv6 Enabled Logo的中国团队同时还积极参与国际项目的合作，协同共建IPv6 Observatory平台[10]，对全球IPv6网站和网络的发展，以及IPv6培训教育情况进行检测和监控。

根据IPv6论坛显示的数据，截至2012年11月，全球已发放1320个IPv6 Ready Logo，其中发放IPv6 Ready第一阶段认证Logo共有480个，第二阶段认证共840个，并且中国厂家的设备测试量占到全球设备测试量的20%。在IPv6 Enabled认证方面，截至2012年12月19日，全球共有1 895个网站获得全球IPv6 Enabled WWW Logo，其中中国有499个，排名全球第一，是第二名获得数量的近3倍。全球获得IPv6可用认证（IPv6 Enabled）的网络服务提供商（ISP）共有131家，中国共有4家ISP获得认证。

3.2 IPv6互联通测试活动

除了对IPv6协议本身的测试，不

▼表3 网站IPv6评测指标

测试指标	说明
FQDN域名解析	检查网站FQDN域名是否能支持AAAA解析，并能返回解析出来的IPv6地址
主页连接测试	通过测试工具发起HTTP、HTTPS、FTP、POP3、SMTP等连接至网站域名对应的IPv6地址，检查网站是否能够返回正确的响应
链接测试	按照主页连接检测方法检查站内链接是支持IPv6连接并能正确响应
接口测试	网站IPv6支持度评测需要考察相关外部通信是否支持IPv6连接等
网站关键网元设备IPv6/IPv4协议栈测试	此项考核指标主要测试各关键网元设备（Web服务器、缓存服务器、数据库服务器、CDN服务器、Squid服务器、负载均衡设备、DNS服务器、防火墙、IPS/IDS设备）IPv6协议栈，主要检查其对IPv6基本协议（RFC2460、RFC4861、RFC4862、RFC4443、RFC1981）的支持程度
网站关键网元Web核心软件IPv6/IPv4测试	此项考核指标主要测试网站中的Web服务软件、DNS系统软件、数据库系统软件、Squid软件、负载均衡系统、CDN服务器软件以及网站的各种应用包括视频、语音、文件共享等对IPv4/IPv6双栈的支持
连接响应速度测试	测试网站IPv6/IPv4连接的响应速度
负载测试	为了测量Web系统在IPv6/IPv4连接上的性能，以保证Web系统在需求范围内能正常工作
网站上传下载速度测试	在本地和远程通过测试工具进行IPv4/IPv6上传和下载文件多次测试，并分别详细记录基于IPv4和IPv6的上传和下载的测试过程和网络速度
网站支持IPv6的连续性	对网站进行IPv6支持度周期性维护性测试，即进行定期（例如每一周）的上述功能测试，并记录和统计所以测试结果，输出测试报告

CDN：内容分发网络　FTP：文件传输协议　IPS/IDS：入侵防御系统/入侵检测系统
DNS：域名系统　HTTP：超文本传输协议　POP3：邮件协议第三版本
FQDN：完全合格域名/全称域名　HTTPS：安全超文本传输协议　SMTP：简单邮件传输协议

少研究者把目光投向了基于应用和用户体验的活动，使用用户真实的日常应用来发现IPv6世界中还存在的问题。诺基亚公司在2010年针对用户体验，搭建一个纯IPv6环境下的网络来对IPv6网络，终端和应用进行研究[11]。该研究发现基于Web的大多数应用都能够正常工作，然而针对一些上层包含IPv4地址的应用，部分防火墙包分片功能，陈旧的通信软件还未能很好支持IPv6。日本WIDE联盟在2011~2012年成功举办了2次IPv6体验大会[12]。与诺基亚公司不同，日本WIDE联盟的IPv6体验实验基于IPv6过渡技术，通过还原已知问题的手段，透过用户面对面调研来反映用户体验。同时对一些特定热点问题如路径最大传输单元（PMTU）和MTU不匹配问题，IPv4 Fallback问题等进行了研究和实验。

面对IPv6过渡期间的各种技术和方案，为了更好地了解技术的成熟度和产业接受程度，2012年在北京成功举办了全球首届IPv6过渡技术测试大会，对DS-Lite、NAT6、4over6、IVI在内的数10种过渡技术进行互操作性测试。参会的单位大家39家，共61款设备进行了测试。本次IPv6过渡技术测试大会主要从3个方面来进行测试：过渡协议功能性测试、多厂家的互通性测试和端到端业务承载能力测试。本次过渡测试给设备厂商开发IPv6过渡产品很大的驱动力，从协议测试上看仅部分标准成熟的过渡技术具备试商用条件，而大部分协议还在标准化和产业化的初步阶段。从厂商设备研发情况看，仍然有很多参测设备为原型产品，还不能满足商用的需求。本次测试大会主要针对基本功能和互通性测试，测试中发现不同厂商对协议某些地方的理解存在一定的偏差，比如对隧道大包的处理，比如对一些过渡协议中的自动配置协议的支持还不够好，设备的互通性尚存在一些问题。

4 结束语

由于IPv6过渡进程的深入，IPv6测试不仅仅只停留在协议本身的测试。大量的不同厂家的设备、复杂的网络以及内容丰富网站应用给IPv6测试带来了巨大的挑战。本文总结了以往在IPv6协议测试研究领域相关的工作，从设备测试、网络测试和网站测试3个方面介绍了IPv6评测规范和体系方面的研究成果，还介绍了IPv6论坛的测试体系IPv6 Ready和IPv6 Enabled，以及全球IPv6测试领域的实践和探索，为IPv6测试领域关键问题和发展趋势的研究打下基础。

参考文献

[1] SMITH, LUCIE, LIPNER. Free pool of IPv4 address space depleted [R]. Number Resource Organization. Retrieved 3 February 2011.
[2] Go4it [EB/OL]. [2012-09-15]. http:// www.go4-it.org. 2012.
[3] 陈运清, 王茜, 胡捷, 等. 构建运营级IPv6网络 [M]. 北京:电子工业出版社, 2012.
[4] IPv6 Ready Logo [EB/OL]. [2012-06-25]. http://www.ipv6ready.org. 2012.
[5] 张玉军, 李忠诚. 移动IPv6测试中的层次化协议描述和测试生成方法 [J]. 电子学报, 2004, 32(F12):30-34.
[6] 喻星, 尹霞, 王之梁, 等. IPv6中RIPng协议的一致性测试研究与实践 [J]. 计算机工程, 2006,32(20):117-119.
[7] WANG Z, YIN X, WANG H, et al. Automatic testing of neighbor discovery protocol based on FSM and TTCN [C]//Proceedings of the 2004 Joint Conference of the 10th Asia-Pacific Conference on Communications, and the 5th International Symposium on Multi-Dimensional Mobile Communications(APCC/MDMC'04):Vol 2, Aug 29-Sep 1,2004,Beijing,China. Piscataway, NJ,USA: IEEE, 2004:805-810.
[8] ETSI TS 102 351. Methods for testing and specification (MTS);Internet protocol testing (IPT); IPv6 testing: Methodology and framework [S]. 2012.
[9] ETSI TS 102 237-1. Telecommunications and Internet protocol harmonization over networks (TIPHON) release 4; interoperability test methods and approaches; Part 1: Generic approach to interoperability testing [S]. 2012.
[10] ARKKO J, KERANEN A. Experiences from an IPv6-Only Network [S]. RFC6586. 2012.
[11] HAZEYAMA H, HIROMI R, ISHIHARA T. Experiences from IPv6-Only networks with transition technologies in the WIDE camp spring 2012 [R]. IETF draft. 2012.

作者简介

宋林健，清华大学博士毕业；北京天地互连信息技术有限公司高级研究员，主要从事下一代互联网体系结构方面的研究，负责IPv6过渡及IPv6测试相关技术的研究和标准制订；已参与基金项目3项，发表学术论文10篇。

李震，西安电子科技大学硕士毕业；北京天地互连信息技术有限公司下一代互联网测试中心主任；主要从事下一代互联网IPv6标准和测试方面的研究，负责相关测试标准和技术的跟踪、研究和标准制订；已参与基金项目3项，主持完成3项IPv6测试国际标准，发表学术论文3篇。

综合信息

中兴通讯再破TD-LTE技术难关 全球首次实现跨频段载波聚合（CA）

【本刊讯】2013年2月17日消息，中兴通讯在北京首度实现F+D跨频段载波聚合，现场测试速率分别达到430 Mbit/s（4×4MIMO模式）、223 Mbit/s（2×2MIMO模式）。中国移动多个省份的领导及专家现场见证、分享了这一突破性的成果。本次技术演示是继1月份中国移动TD-LTE广州扩大规模试验网环境下实现D频段载波聚合223 Mbit/s峰值后的再一次技术难关突破。

载波聚合（CA）是LTE-A最核心的技术，它可以把相同频段或者不同频段下的两个以上的载波合并为一个信道，成倍地提高TD-LTE小区的峰值速率。

DOI: 10.3969/j.issn.1009-6868.2013.02.011
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130301.1044.001.html

计算通信：超量信息无线传输的深度探索

Computing Communications: Wireless Transmission of Mass Information

中图分类号：TN929.5 文献标志码：A 文章编号：1009-6868 (2013) 02-0040-04

摘要：计算通信是融合计算和通信这两个信息科技的新兴学科。它通过信息传递和信息处理的深度耦合，在网络中高效地传递函数流，突破超量信息无线传输的“瓶颈”。计算通信基础理论、网络计算增益析取机理、及计算通信普适性是计算通信未来的主要研究方向。

关键词：计算通信；超量信息；无线传输

Abstract: Computing communications is a new paradigm that combines computing and communications. Through tight integration of information transmission and processing, computing communications delivers a function flow (rather than a traditional information flow) in the network and can fundamentally solve challenges in the wireless transmission of mass information. Promising research areas in computing communications include computing communication theory, exploitation of network computing gain, and ubiquitous computing communication.

Key words: computing communications; mass information; wireless transmission

王新兵/WANG Xinbing
陶梅霞/TAO Meixia
刘辉/LIU Hui

（上海交通大学电子工程系 无线通信技术研究所，上海 200240）
(Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai200240, China)

- 未来无线通信网络发展的关键是指数增长的网络计算能力
- 计算通信的本质是通过信息传递和信息处理的深度耦合，在网络中高效地传递函数流
- 虚拟化网络和复杂业务是研究计算通信的主要挑战

通信网络深刻地改变着人们的生产、生活和学习方式，成为支撑现代社会经济发展、社会进步和科技创新的最重要的基础设施。大容量、异构性、智能化代表未来网络的发展趋势。在全球化日益加深的当今世界，信息的作用已经不仅仅是经济发展的支撑，更是决定经济发展成败的引擎。目前，下一代通信网络相关技术已成为各国竞争的焦点和制高点，引领了信息产业革命的第三次浪潮，将成为未来社会经济发展、社会进步和科技创新的最重要的基础设施，也关系到未来国家物理基础设施的安全利用。挑战支持大容量媒体数据传输网络的重大基础科学问题，发展中国具有自主产权的异构网络技术，推动中国在该领域的跨越式发展，具有十分重要的国际战略意义。

移动互联网作为人们交流及获取信息的一种高效便利的渠道，在中国向信息化社会转型的过程中起着倍增器作用。目前国际信息化总的趋势是从固定宽带转向移动宽带。到2015年之前，全球移动宽带用户数将达到38亿。中国目前无线通信用户数将达到6.5亿，在软、硬件领域都可望成为全球最大的产品供应国。无线通信已成为中国率先实现高新技术崛起的支柱行业之一。然而，随着信息向网络化、泛在化、融合化的方向发展，数字媒体内容呈现爆炸式增长的态势。移动通信的发展因此面临着巨大的技术挑战。最近的研究报告（Cisco）显示，未来5年全球移动数据量将增长66倍，其中2/3来自于无线视频业务。因此，移动用户将经常遭遇连接阻塞，下载时间过长，播放滞缓等不良体验。根据全球移动通讯系统协会（GSMA）的报告显示，移动宽带连接速度最慢的国家是印度和中国，而韩国、澳大利亚、新西兰、其余亚太地区和西方国家较快。作为中国战略性新兴产业，移动通信网络的重要作用日益凸显。如何全面提升移动网络的综合效能将直接

收稿日期：2013-01-23
网络出版时间：2013-03-01

影响到国民经济建设、国防现代化建设、社会和谐发展。

国际标准化组织、通信企业及科研院所将下一代无线通信系统的技术目标定位在100 MHz带宽内提供高达1 Gbit/s峰值数率。然而从现实情况来看，在过去40年里，蜂窝网的容量增长大约为平均每5年2倍。其中较大的贡献来自于网络规模的扩大及频谱的增加。近十多年里开发的移动通信的主流技术已在3G、4G中得到较充分地应用。在信源方面，多媒体压缩能力的增长大约为每5年2～3倍。其中正在开发的最新标准H.265只有20%的压缩能力来自“纯算法”；绝大部分提高来自“视觉体验”。显而易见，移动用户对无线多媒体跳跃式的需求增长是传统无线网络无法独立满足的。如何从根本上解决超量信息有效传输的是网络社会中信息科学目前所面临的重大科学问题。

1 计算通信概念的提出

世界各国高度重视数字媒体通信技术的研发、通信网络基础设施建设、无线通信标准制订。进入新世纪以来，为突破数字通信的技术瓶颈，各国纷纷推出数字媒体相关的研究计划。由于频谱资源日趋匮乏，开发新一代颠覆性无线通信技术，谋求移动通信的主导地位，已经成为国际信息科学技术前沿研究的汇聚点。

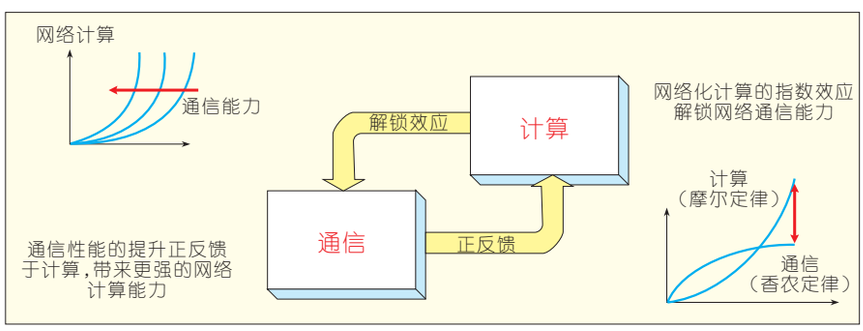
作为数字化革命的两个方面，计算和通信在主要目标上有所不同。网络计算主要侧重于在消耗最少资源条件下如何通过计算迅速地、可靠地解决问题；而无线通信则主要侧重于在消耗最少资源条件下如何快速地、可靠地传递信息。在过去的60年间，计算与通信两学科得到了不可思议的发展。一方面，图灵和诺依曼利用许多晶体管搭建了现代电子计算机的雏形，奠定了现代计算体系。而今，电子数字计算机成为人类社会的基石，而且已经演进到多处理器网络架构。在可以预见的未来，网络及移动节点的计算能力仍将遵循摩尔定律，继续其快速增长的势头。另一方面，1948年香农的革命性论文开始了现代数字通信理论的新纪元[1]。文章揭示了无线传输能力在根本上受到频谱资源和能量等因素制约，其承载能力成对数级别增长。从物理层来挖掘无线资源，虽然可以获取频谱利用率及功率利用率的提升，但这种提升是有限的，除了会带来设备复杂度及建网成本的增加，更有一种不可忽视的后果：能耗的巨大增长。

如何突破根据传统香农信息论的现有无线网络瓶颈束缚，取得跨越式的发展，已经成为无线网络进一步发展亟待解决的问题。信息论领域权威学者美国工程院院士斯坦福大学教授T. Cover曾指出：“通信是计算约束的，而计算是通信约束的”。由此可见，解锁无线网络潜能的关键正在于通信与计算的有效结合。如图1所示，传统通信的容量由香农公式所示呈现log函数增长趋势；而根据摩尔定律，计算能力却是呈现指数特性增长。特别是在大规模网络中，网络中的移动终端拥有越来越强大的计算能力，而且智能终端通过网络可以完成更加强大协同计算。将这种丰富的计算能力注入移动通信网中，可以解锁网络通信能力的巨大潜能。可以预见未来移动通信的一个动力源泉是强大的网络计算能力。

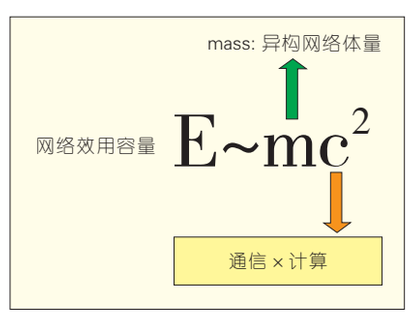
纵观无线通信发展历史，网络通信效能的本质提升必须建立在传输理念及网络架构有根本性变革的基础上。计算通信将融合计算和通信这两个信息科技的关键领域，对摩尔定律驱动下的“计算通信理论”这一正在崛起的学科作全面性的研究工作，以谋求超量信息无线传输的本质性突破。其核心思想是围绕终端的信息体验（QoE），通过信息传递和信息处理的深度耦合，同时运用计算和通信资源在网络高效地传递函数流。如图2所示，计算通信系统中的网络效用容量由3种关键因素决定：异构网络总体量、无线传输能力、网络计算能力。虽然无线传输能力的提升已遭遇香农瓶颈效应，网络体量的扩大也附带着巨大的成本，但网络的效用容量仍可通过丰富的计算能力取得跳跃式的增长。网络的效用容量和计算能力之间存在着某种互换关系，好比物体质量与极限速度的之间的对应关系。

2 计算通信的潜力

计算通信作为引领下一代无线通信网络发展的潜在突破性理论，可



▲图1 计算通信的解锁效用



▲图2 计算通信能力与网络效用容量的关系

以追溯到30年前的计算机网络时代。众所周知，路由是计算机网络（比如Internet）中最基本的数据传输装置。字节（比如数据包）可以通过网络的中间节点成功地从源节点传输到目的节点。从这个角度看，2个节点之间的数据包交换就实现网络间的通信，犹如交通道路上运行的汽车。然而当网络（移动）节点具备计算能力时，该经典的通信模型得到了突破。早在1984年，Gamal[2]第一次提出了在一个n个节点的完全连接双向图中计算恒等和奇偶函数的最小传输次数问题。随后，信息论领域大师，美国工程院院士MIT大学教授Gallager在文献[3]中设计了一种计算恒等和奇偶函数的协议，并指出该协议需要最多$O(n\log n\log n)$量级的传输次数，量化了计算带来的多节点无线网络中的巨大增益。20年之后，Goyal, Kindler以及Saks[4]证明了计算恒等函数至少需要该数量级的传输次数，因此Gallager的协议是复杂度最优的计算恒等函数算法。

Kumar等的工作[5-7]将计算通信理论运用到无线自组织网络和传感器网络中，开启了这个领域。随后的深入研究表明，在以"效用函数"为目标的通信系统中，整个网络可被视为一个面向应用的分布式计算系统，其计算能力直接决定了系统的效用容量。2000年，Ahlswede等人在文章[8]中指出信息流跟道路上的交通是大不一样的。他们指出，在计算通信网络中，对数据流可以进行代数编码，并再把编码后的信息传输到下一节点。最近十年，网络编码作为计算通信的一种特殊形式，迅速激起了研究的热潮，并且逐步被人们在理论上证明了基于网络计算的传输优势。最近，网络编码逐步被应用到实际系统中。Microsoft、HP、Intel、Cisco、Samsung等研究机构/公司都在实际的有线网络和无线网络中开展网络编码应用工作。

尽管计算通信的潜力初现端倪，国际上到目前为止在移动计算通信领域的研究仍是相对零散的，如文献[9-11]。Dougherty等的研究表明，现有的网络计算技术还无法提供无线网络应能达到的计算增益[9]。网络计算运用于无线多址接入的研究也刚起步[11]。由于移动网络的复杂性及超量信息的多元性，学术界还缺乏完整的理论体系及系统级突破性技术，能将计算通信的潜能转化为可观的网络增益。这对中国来说是个巨大的机遇。

3 计算通信的研究方向

针对超量信息无线传输的容量瓶颈、虚拟化网络中计算通信资源的动态多元性、及新兴数据业务的复杂移动性等挑战问题，计算通信的研究将涉及跨信息领域的交叉学科，主要研究点可以概括为3个方面。

3.1 计算通信基础理论

在摩尔定律的驱动下，无线网络和移动终端的计算能力都在继续以指数性速度增长，为获取有效网络容量增益提供了必要条件。建立"计算通信"的抽象数学模型，并形成普适的理论分析体系，是开启计算通信理论潜能的关键。由此产生的共性关键科学问题就是通过数学语言精准地描述异构网络及移动节点的计算能力，量化群体信息的传输有效性，并借此揭示网络计算能力对整体效用容量的影响。计算通信基础理论的研究涵盖3个主要部分。第一，对无线异构网络"效用容量"的数学定义及建模。在计算通信机制下优化的最终目标函数是群体用户的有效体验即所谓的"效用容量"，只有在"效用容量"框架下研究计算与通信的融合才有真正的意义。第二，基于计算能力的"相对信息理论"。若要在有限的通信能力下提高网络效用容量，就必须考虑相对网络计算能力的有效信息传输能力，探索计算能力与效用容量的转换规律。在此基础上，进一步对异构网络的效用容量进行优化，进而精确地分析各种计算参数对整体性能的影响。第三，在虚拟资源模型化的基础上，通过网络计算将模型化的多元虚拟资源有效的分配给各个网络结点，以最大化虚拟网络资源的利用率，达到包括效率和可靠性在内的最佳网络通信性能。

3.2 网络计算增益析取机理

多种并存的异构网及移动节点所提供的多元化的通信和计算资源通过网络化和虚拟化的方法，在逻辑上形成统一个完整统一的虚拟重构网。如何基于虚拟化的通信资源和网络化的计算资源，利用虚拟重构网的多元协同计算通信理论与机制，以有效获取网络通信中的最大计算增益，是计算通信涵盖的第二大研究点。这就要求首先对这些虚拟通信资源建立适当的描述模型，该模型应该能用统一的方法刻画不同性质的虚拟通信资源，并为多元虚拟资源的协同分配提供基础。其次，协同通信机制能够将所需的计算任务进行有效的分割，尽可能减少每个子计算任务之间的耦合，通过不同网络结点的分布式并行计算来获得协同通信机制要求的虚拟资源分配方案。可见，协同通信机制设计和网络化计算是相互耦合、相互制约、相互促进的整体。因此，这两个部分的联合设计和联合优化将成为本研究点的关键。

3.3 计算通信普适性

适用于大规模网络的计算通信理论必须具备可扩展性、可实现性、及可适用性。因此，需要研究如何充分发挥多元计算能力，建立可扩展的异构计算方法，以及研究可实现的复杂业务的适配机制。具体来说，就是研究大型泛在网络中业务与网络之间，节点与网络之间，以及节点与节点之间的协同计算、通信、及分享机制，将超量信息高效地传递到众多动态节点，从而保证用户体验到优质的

移动服务。

4 结束语

无线网络能力本质的提升必须建立在传输理念及网络架构有根本性变革的基础上。虽然现有无线网络容量的提升已遭遇香农瓶颈效应，但网络及移动节点的计算能力却仍呈现指数特性增长。“计算通信理论”是有着巨大潜力的新兴学科，将融合“计算”和“通信”这两个信息科技的关键领域，运用丰富多元的移动网络“计算资源”来解锁无线传输现阶段的瓶颈，释放巨大潜能，以获取无线网络有效容量的跳跃性提升。计算通信的研究应瞄准网络社会中信息科学面临的重大科学问题，研究体现无线网络异构性、适应移动节点多元性及超量信息复杂性的突破性理论论架构与关键技术，是未来无线通信网络的发展方向之一。

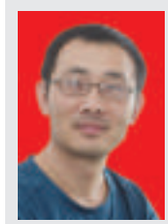
参考文献

[1] SHANNON C E. A mathematical theory of communication[J]. Bell System Technical Journal, 1948,27:379–423,623–656.
[2] GAMAL A .Reliable communication of highly distributed information[M]. GOPINATH T M, THOMAS B C.Open Problems in Communication and Computation, Berlin, Germany: Springer–Verlag, 1987: 60–62.
[3] GALLAGER R G.Finding parity in a simple broadcast network[J]. IEEE Transactions on Information Theory, 1988,34(2): 176–180.
[4] GOYAL N, KINDLER G, SAKS M. Lower bounds for the noisy broadcast problem[J]. SIAM Journal on Computing, 2008,37(6): 1806–1841.
[5] GUPTA P, KUMAR P R. The capacity of wireless networks[J]. IEEE Transactions on Information Theory, 2000,46(2): 388–404.
[6] GIRIDHAR A, KUMAR P R. Toward a theory of in–network computation in wireless sensor networks[J]. IEEE Communications Magazine, 2006, 44(4): 98–107.
[7] GIRIDHAR A, KUMAR P R. Computing and communicating functions over sensor networks[J]. IEEE Journal on Selected Areas in Communications,2005,23(4):755–764.
[8] AHLSWEDE R, CAI N; LI S Y R, et al. Network information flow[J]. IEEE Transactions on Information Theory, 2000,46(4):1204–1216.
[9] DOUGHERTY R, FREILING C, ZEGER K. Unachievability of network coding capacity[J]. IEEE Transactions on Information Theory, 2006,52(6): 2365–2372.
[10] AYASO O, SHAH D, DAHLEH M A. Information theoretic bounds for distributed computation over networks of point–to–point channels[J]. IEEE Transactions on Information Theory, 2010, 56 (12): 6020–6036.
[11] NAZER B,GASTPAR M. Computation over multiple–access channels[J]. IEEE Transactions on Information Theory, 2007, 53(10): 3498–3516.

作者简介

王新兵，上海交通大学教授、博士生导师、电子工程系副主任；主要研究移动多播、移动网络拓扑、认知无线电频谱分配；担任《IEEE TMC》等国际期刊编委；获IEEE通信学会亚太区杰出青年研究者奖等奖项；近五年发表20余篇文章，申请发明专利27项。

陶梅霞，上海交通大学教授、博士生导师、无线通信技术研究所副所长；主要研究方向为物理层网络编码、无线资源分配、MIMO技术等；担任《IEEE TCOM》等国际期刊编委；获IEEE通信学会亚太区杰出青年研究者奖等奖项；在IEEE重要期刊上发表论文40余篇，持美国专利2项。



刘辉，上海交通大学致远讲席教授、IEEE院士、国家“千人计划”成员、电子信息与电气工程学院副院长；主要研究方向是宽带移动通信、无线接入、移动多媒体广播；获中国国家专利金奖等奖项；出版2部英文专著，发表45篇论文，并拥有67项中外专利。

上接第35页

未来我们将着重于IPv6–only环境下双重翻译和隧道技术的应用及基于科学问题的网络互通研究。

IPv6–only是IPv4/IPv6过渡的完成阶段，面对目前因特网工程任务组（IETF）中提出的面对各种场景的过渡技术，IPv6–only如何在应用、访问、状态、操作系统等4个维度上同现有网络环境实现更好的互通是非常重要的研究领域，将促进IPv6的进一步发展和推广。

参考文献

[1] NORDMARK E, GILLIGAN R. Basic transition mechanisms for IPv6 hosts and routers [S]. RFC4213. 2005.
[2] CURRAN J. An Internet transition plan [S]. RFC5211. 2008.
[3] ARKKO J, KERANEN A. Experiences from an IPv6–only network [S]. RFC6586. 2012.
[4] HAZEYAMA H, HIROMI R, ISHIHARA T, et al. Experiences from IPv6–only networks with transition technologies in the WIDE camp autumn 2012 [R]. draft–hazeyama–widecamp–ipv6–only–experience–02. 2012.
[5] BAGNULO M, MATTHEWS P, VAN BEIJNUM I. Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers [S]. RFC6146. 2011.
[6] BAGNULO M, SULLIVAN A, MATTHEWS P, et al. DNS64: DNS extensions for network address translation [S]. RFC6147. 2011.
[7] IPv4 Exhaustion and IPv6 transition [J]. Nikkei Network, 2011, 10(7):30–35.
[8] AMOSS J J, MINOLI D. Handbook of IPv4 to IPv6 transition: Methodologies for institutional and corporate networks [M]. New York, NY,USA: Auerbach Publications, 2007.
[9] 崔勇. IPv6过渡场景分析 [R]. 北京: 中国通信标准化协会, 2012.
[10] HINDEN R S, DEERING S. IP version 6 addressing architecture [S]. RFC4291. 2006.
[11] BLANCHET M. Special–use IPv6 addresses [S]. RFC5156. 2008.
[12] THOMSON S, NARTEN T, JINMEI T. IPv6 stateless address autoconfiguration [R]. RFC4862. 2007.
[13] SAINT–ANDRE P. Extensible messaging and presence protocol (XMPP): Core [S]. RFC6120. 2011.
[14] POSTEL J, REYNOLDS J. File Transfer Protocol [S]. RFC0959. 1985.
[15] ALLMAN M, OSTERMANN S, METS C. FTP Extensions for IPv6 and NATs [S]. RFC2428. 1998.
[16] BEIJNUM VAN I. An FTP Application Layer Gateway (ALG) for IPv6–to–IPv4 Translation [S]. RFC6384. 2011.

作者简介

刘笑寒，新潟大学计算机系博士毕业；中国科学院计算机网络信息中心副研究员，庆应义塾大学SFC研究所客座研究员；研究方向为IPv6、无线传感网、无线网状网、可视光通信；已发表SCI/EI检索论文8篇。

闫双舰，郑州大学计算机科学与技术系学士毕业，中国科学院计算机网络信息中心在读研究生；研究方向为IPv6、无线传感网、数据挖掘、海量数据处理。

孙玉勤，上海交通大学信息与控制工程系学士毕业；美国A10网络公司中国区技术经理；研究方向为IPv6和应用交付。

DOI: 10.3969/j.issn.1009-6868.2013.02.012
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130302.1647.001.html

无线通信系统的环境适应性

Environmental Adaptability of Wireless Communication Systems

郭丹旦/GUO Dandan
唐雄/TANG Xiong
崔卓/CUI Zhuo
（中兴通讯股份有限公司，广东 深圳 518057）
(ZTE Corporation, Shenzhen 518057, China)

摘要：无线通信基站系统在现场应用中会受到散热、灰尘、油烟、湿度、腐蚀等环境因素的影响，而运营商降低建设成本和维护成本的考虑使得设备工作环境更加恶化。从实践角度出发，基于实际部署经验，系统性地分析了上述因素对设备可靠性的影响，提出综合设备设计和应用环节统筹考虑以减轻影响的思路，并针对设备设计和部署给出了具体建议。

关键词：无线通信；环境适应性；热设计；腐蚀；灰尘

Abstract: Environmental factors such as radiation, dust, lampblack, and corrosion affect wireless base station systems. With decreased capex and opex, there is a risk that the environment in which equipment operates deteriorates. In this paper, we draw on actual experiences in the field to discuss the effect of environmental factors on base station reliability. We suggest that environmental factors should be taken into consideration when designing network equipment, and we make specific recommendations about how equipment should be designed and deployed.

Keywords: wireless communications; environmental adaptability; thermal design; corrosion; dust

中图分类号：TN929.5 文献标志码：A 文章编号：1009-6868 (2013) 02-0044-05

运营商降低TCO，主要包含降低建设成本（CAPEX）和维护成本（OPEX）两方面。为了降低CAPEX，通信设备尤其是接入网设备的工作环境，从机房、空调方舱等被改变为包括室外柜、弱电井、楼道、车库、铁皮柜等各种复杂环境，同时对通信设备提出了体积小、重量轻、安装成本低的要求，还要维持设备本身的低成本；为了降低OPEX，设备要尽量降低功耗、尽可能地被室外化、工作环境相对过去高温化，以节省电费，同时还要求设备尽量减少维护，以降低维护人工的成本。接入网电信设备从精密、娇贵的设备，已经变成无所不在的IT化设备，甚至工作环境更为恶劣。简易机房土坯墙面脱落，设备工作环境的恶劣情况如图1所示。

收稿日期：2012-12-29
网络出版时间：2013-03-02

在这个背景之下，电信设备近年来面临各种新的环境适应性问题，设备设计的一些思路也需要有发生转变，从遵从标准到理解标准，满足应用场景，以适应不同的应用需求。

对设备在不同环境下工作以及设计约束的研究，传统上属于可靠性方面。但是，传统的可靠性方法，并没有针对实际工程应用给出足够的指导意见。设备设计制造需要对环境适应力的深入理解和把握，以及从系统多个维度上进行平衡。

1 研发过程的可靠性方法

传统的可靠性增长方法，主要从可靠性预计和可靠性指标分配开始，通过预算、设计、控制等过程，保证产品的质量[1]。

可靠性预计基于对器件失效模型的认识，通过概率与数理统计方法，首先对系统建立可靠性数学模型，然后评估其平均无故障时间（MTBF）、平均恢复前时间（MTTR）等指标。参照的方法很多来源于美国军用手册MIL-HDBK-217。但是，这些方法在实际应用中有非常大的局限性，存在估算数据不准确，参考意义不足等诸多问题[2-3]。

有一些人认为，这些方法估计不准确的主要原因是在于手册制订时间过早，电子工业经过多年的发展，手册不符合实际的情况，有一些通过经验进行修正尽量使得估计准确。但是仍然有很多因素使得估算不能准确进行：一个主要因素是失效模型受到诸多外界因素的影响，故障往往都不是模型中考虑的，多为过应力使用、非设计场景的恶劣环境等，无法在模型里考虑充分；另外一个主要因素是系统模型异常复杂，实际的分析

▲图1 设备工作环境恶劣情况

不可行。

以目前通信基站系统来看，一个系统中存在若干个单板，有一些故障模式并不能完全用串联或者并联描述，存在关联性。每个单板中可能存在超过100~200种、数千个物料，存在多种失效模式。失效模式、失效模式对应的器件范围，甚至一种失效模式在不同的情况下对系统的影响均不相同。这样，使得可靠性预计更加困难，难以有效实施。而且，随着IT化和商用货架产品（COTS）的广泛使用，可靠性分析对设计的指导作用更加有限。从设备设计和生产实际的经验认为，主要的设备故障和异常往往来自于环境等外部不可控因素影响，而不是部件自身老化或者偶发失效，这也使得可靠性预计的准确度大大降低。

有一些更激进的观点认为，可靠性预计已经变成数字游戏，对产品的质量和设计没有指导意义。能够保证设计系统可靠的是可靠性增长试验和失效物理（可靠性筛选和监控属于控制范畴，本文不讨论）。

失效物理通过研究某一种因素对部件的影响，分析部件会在何时、何处、何种原因，发生何种类型的失效。通过研究扩散、相变、腐蚀、应力、静电泄放等物理化学过程对器件的影响，来分析器件可能产生的问题。失效物理的分析为进一步的改进和增加可靠性提供了很好的基础（分析的例子如图2所示）。

目前可靠性增长试验中，应用较为广泛的包括四角测试、强加速寿命试验（HALT）等，还包括盐雾、灰尘、湿尘、振动、渗漏等。这些试验有意无意地采用失效物理分析的一些因素，例如HALT实际上考虑的是高低温应力、高低温循环带来的应力、叠加强振动等，来寻找系统设计的薄弱环节；盐雾考虑的是腐蚀的问题；振动主要评估结构在应力下的表现。因为设备寿命远远长于试验所能够接受的周期，在可靠性试验设计中，常常也采用加速的方法，加大应力（电压、温度、湿度、温变速率等），增加样本数量，来评估系统实际工作中的寿命。图3就是室内无线基站设备的加强灰尘试验，评估极端灰尘环境对设备连接可靠性和散热的影响。这些试验对改进设计、提高实际

▲图2 陶瓷电容横贯裂纹、芯片ESD损毁形貌

应用的可靠性起到了很大的作用，也是设计中保证设计指标的必要手段。但是为了控制分析的复杂度，试验剖面设计一般只针对某一类应力、机理或者失效模式，和现场应用的复杂环境有所区别。

系统设计上，要综合考虑可靠性方法、电路设计方法、结构设计、环境设计方法等，结合降低CAPEX和OPEX的要求，确保设备的可用性。

2 设备常面临的环境问题

设备工作的环境情况非常复杂。北大西洋公约组织将全球的气候根据温度和湿度作了划分，作为设计指导的依据[4]。温度和湿度对设备存在一定的影响，但是设备的工作和

▲图3 中兴通讯无线接入网设备的加强应力灰尘试验

更多的环境因素、人为操作因素相关联。美国军用标准MIL-HDBK-338B第7部分，对环境因素和对设备的影响进行了一些描述，但是也没有给出设计指导意见[5]。

因为环境对设备的影响相互关联，很难独立的进行分析。本文尽量将关联的因素进行归类，分析对设备带来的影响以及设计应用中需要进行的考虑。

2.1 散热及相关

温度对设备有很多方面的影响，与散热相关的设计是设备最关注的方向之一，并且和包括灰尘等方向相关联。

从可靠性角度来看，温度影响着器件内粒子的扩散速度，过高的温度会加速迁移的速率，最终导致器件的失效。同时，温度还会加速腐蚀的进行。温度的昼夜、季节变化导致设备各个部件的热胀冷缩。热胀冷缩率的不同，对器件封装、组装等各个环节产生循环的应力。温度对设备寿命的影响在可靠性分析中已经有很多的研究，一般认为，温度每升高10度，设备的寿命缩短为原来一半。

从可靠性预计角度来看，为了延长设备寿命，应该使设备保持较低的温升。实际上，为了满足日益增长的处理复杂度需要，设备的集成度持续提升，设备比以往要耗散更多的电力。要把这些热散出去，需要增加设备的体积，或者增加设备的风流量，增加辅助的散热设施。这些措施的采用，直接抬高了CAPEX；降低设备温度，还意味着风扇/空调的转速更高，作为运动部件的风扇，比电子零部件更容易失效，这也就意味着降低温度实际使得设备更容易失效；更大的风流量，也意味着防尘网需要更频繁的清洗，增加了人力维护成本；更大的风流量，还意味着更多的耗电、更大的噪音。作为设计折中，系统设计中，比较倾向于让器件的工作温度在保证降额的情况下，贴近高温区，减少散热带来的电费增加以及风扇磨损、噪音等相关问题。不但如此，系统设计中让设备工作温度靠近高温区，还可以降低设备内湿度。但是，贴近高温区，也可能使半导体器

件漏电导致设备消耗更多能量，需要平衡各因素进行考虑。

一般设备设计上已经考虑使得零部件在设计寿命内工作处于浴盆曲线的底部，在设计范围内工作时温度并不是设备发生故障的主要原因。设备损坏主要的原因常常来自于一些不可控制因素，使得工作环境超过设计能够支持的极限温度。从功能性能角度来看，温度会影响数字逻辑器件的工作频率，使得设计裕度被打破而导致功能异常；有的器件，例如恒温晶体振荡器（OCXO）内部具有加热恒温槽，在外部温度低于内部温度下才能起到恒温的作用，当温度范围超出标准时，时钟保持性能可能受到严重影响。这些都还是可恢复的异常。在更恶劣的情况下，芯片的漏电流随着温度升高以指数方式增大，在额定温度点附近功耗随温度快速增加，反过来带动芯片的温度进一步增加而导致热失效；高温下焊点、机械结构可能由于蠕变而失去强度；PCB板可能发生碳化、分层等失效[6]。

设备设计上更关注的是如何使得这些外部的异常更不容易发生，异常发生的时候系统如何自动保护，同时兼顾越来越精细的OPEX优化考虑。第一，对工作的环境提出了更明确的要求，根据实际的气候、业务模型等条件，把设备的工作环境作为一个系统来进行指标分配设计及综合成本评估，而不是只关注设备本身；第二，当服务质量许可的情况下，当出现异常状况时，系统通过自动降低负荷，甚至局部断电的方式进行自我保护，在异常解除时恢复工作，增强实际的可用性和可靠性。

随着多年的扩容、2G向3G换代以及多网多制式的共存，单个站点的容量密度也远高于过去。没有重新设计的机房或方舱，可能对设备的工作环境带来较大的影响。文献[7]给出了机房换气设计的要求。如果机房达不到要求，则会导致设备过温。在实际应用中，因为空调设备被盗、损坏，通风装置损坏或者过滤网被堵塞等情况经常出现。有的站点建设时间较早，容量很低，经过长期运行通风设施存在问题。存在问题的设施如图4所示。当进行更换扩容后，这样的站点经常频繁出现高温告警。

设备散热设计通过仿真、测试验证的方法在行业内已经广泛使用。通信设备内部相对环境均存在一定的温升，基本不存在除设备本身发热之外的其他热负荷，所产生的热量基本属于显热，主要通过强制对流进行散热，这样使得我们在考虑散热的计算时候相对较为简单。在保持对流空气温升一定的情况下，单位功耗需要的空气流量是一定的。这个是设备设计的物理限制，无法突破。散热需要的空气体积可用如下公式计算：$L = Qs/(Cp \times \rho \times \Delta T)$，其中$L$为空气的体积，$Qs$为显热，$Cp$为空气的平均比定压热容，在设备的工作范围内，可以认为是一个常量，ρ为空气的比重。假设设备设计最高工作环境温度为55℃，允许出风口空气温度为65℃，温升10℃，系统热负荷为700 W，则系统一个小时需要的风量约为200 m^2。

这只是一个指导性的结果，不能替代系统内部的热设计，但是综合考虑设备风速、通风口面积等设计，如果不能达到这样的风量，则只能降低设备的热负荷，增加允许的系统温升，或者采用其他补充的散热方式来满足散热要求。

在高海拔区域，因为气压下降，风冷的效果会受到进一步的影响。但是高原地区一般也不会出现高温等环境，设备的散热环境并不会将最差因素叠加。在整个热系统设计中，

▲图4 机房空调被盗损、封闭空间密集安装等，带来设备散热问题

可以根据实际情况，做出成本优化又能保证可靠性的设计。

综上所述，散热的设计不仅仅涉及到器件的可靠性，而是要考虑整个系统的工作模式以及降低CAPEX、OPEX的需求，结合设备外运行环境如噪声要求等协同解决，在各种相互矛盾的限定因素中平衡优化。在空间受限、集成密度高、空间局限的情况下，噪声和风扇的耗电相对就会较大，如果希望低噪音，就需要加大通风口的面积，降低风速，或者控制设备内处理功耗，降低集成密度。对于环境温度很高，甚至考虑采用压缩机等制冷设备散热，但是也可能会带来更高的能耗和噪音，同时压缩机热端同样也需要考虑如何散热。

2.2 灰尘、油烟

灰尘、油烟对设备最主要的影响体现在散热上。灰尘会堵塞防尘网或空气过滤设备，附着在散热器上的灰尘，还会直接影响器件散热。灰尘还会带来其他一些影响，例如在连接器上堆积的灰尘，可能影响到新插入组件的连接可靠性。一般情况下，连接器设计的滑动行程和摩擦力已经考虑了插入过程推开灰尘，但是偶发的大颗粒灰尘堆积存在隐患。中兴通讯软基站设备采用了连接器保护一体式假单板设计，经过多种试验分析，能够有效防止这样的问题发生，同时还兼顾平衡风阻的作用。

根据中国大气监测的情况，很多城市地区总悬浮颗粒物（TSP）平均保持在2级，即0.2 mg/m^3的水平上[8-10]。新闻报道显示在2006年4月沙尘暴天气下，北京TSP达到0.35 mg/m^3，峰值达到2 mg/m^3。按照多地区平均水平0.2 mg/m^3来算，根据上面散热能力的计算，700 W系统散热每小时气流中所包含的颗粒物约为0.04 g，每年通过系统冷却系统的悬浮颗粒物约为350 g。如果对空气过滤，这些灰尘会使得系统的维护周期大大缩短，维护工作量以及OPEX上升。实际

上，如果不过滤，大部分颗粒物会直接穿过系统，只有一小部分会在系统内，主要在气流受到阻碍的区域堆积，例如连接器。大量的分析认为，这样的颗粒灰尘对系统可靠性的影响并不大，系统防尘设计上，应该让这样的灰尘无障碍的穿过系统。

在多个现场采集的灰尘分析中，很多灰尘呈现絮状、纤维状，来源可能是植物（如杨树、柳树飘絮）、摩擦脱落的衣物纤维、植物焚烧产生的飘浮物等，这些纤维状灰尘吸附在空气过滤设备上，积累后就会增加系统风阻，影响系统散热，同时，随着空气过滤系统的网孔堵塞，更小颗粒的灰尘会被过滤，系统堵塞速度变快。因为气流摩擦产生静电的关系，絮状及颗粒状灰尘也会吸附在包括风扇扇叶、单板上，部分影响到系统的散热（如图5、图6所示）。在这种环境条件下，要通过过滤、隔离等手段，尽量避免灰尘进入设备，在防护设施上，也需要考虑增加通风面积、定期除尘等方式，保证整个环境的散热通畅。

灰尘中存在的盐类会吸收空气中的水分引起腐蚀。如果现场存在高湿度、甚至油烟，如老式建筑居民楼道、地下车库、农村民房等，灰尘会更容易黏附在设备上。在个别地方，甚至出现过设备防水百叶窗开口以及通风孔全部被灰尘及油烟混和物堵塞的现象。在这种场合下，往往结合腐蚀危害，需要整体考虑，采用例如热交换器柜等防护设备，在存在难以清理的油烟等环境条件下，应尽量避免安装设备，如不得不安装，应尽量采用密闭型的自然散热设备。

工业上对灰尘的处理有很多经验，包括惯性除尘、喷淋、过滤、静电吸附等多种方式得到应用[11]。对于通信设备，灰尘没有工业环境恶劣，而能够提供的动力、空间都非常有限；设备分散安装在各个站点，维护周期长甚至希望能够免维护，同时不允许出现高噪音、强烈振动。可以采用的主要就是惯性除尘、过滤等方法，减少一部分进入设备的灰尘。

▲图5 某地恶劣环境工作2个月的接入网设备防尘网被沙尘堵塞

▲图6 潮湿环境下防尘网灰尘结块黏附

从上面的分析可以看出，对于系统的防尘设计，也需要结合实际环境因素以及降低CAPEX和OPEX的需求。对很多室内应用，可以允许灰尘直接进入和穿过设备，减少维护开销；对于部分恶劣环境，考虑增加过滤装置，但需要考虑装置的容尘能力以及维护开销；对于部分运营商愿意进行设备维护，不希望灰尘进入设备的，可以使用防尘网；对于存在腐蚀性物质等的环境，要考虑采用内外环境隔离的设备。

2.3 湿度和腐蚀

从功能和性能角度，湿度和温度一起影响到空气和板材的介电常数，有可能减少高速设计的裕度，引发设备误码率增加等异常。从设备可靠性来看，湿度会加速腐蚀的发生，使得灰尘、有害气体等对设备的损害加剧。对于部分工艺不良的器件，空气中的水汽可能带来破坏性的后果。例如当半导体芯片钝化层不良时，在潮湿空气中可能发生内部分层，通过非偏置的高度加速应力测试（uHAST）试验可以识别此类工艺缺陷；密封不良的电阻器可能因为空气中含硫，发生硫化而损坏，需要通过选型规避。

一般认为，金属在洁净大气中，在相对湿度小于60%~70%的干燥大气中，发生腐蚀的速度非常慢，当相对湿度大于60%~70%时，腐蚀速度大大加快。如果空气中存在H_2S、SO_2等气体时，腐蚀速度也会加快。临界湿度随着空气成分、金属成分不同而有差异，积尘中的粒子也会增加吸附而导致腐蚀速度增加。但是总体上可以认为，通过控制使相对湿度小于60%，可以防止大部分大气腐蚀的发生[12]。

控制湿度的一个重要手段就是控制温度。设备中空气被加热时主要是显热增加，饱和水气压增大，绝对水气压并没有变化，导致空气的相对湿度降低。假设设备入口处的空气湿度接近饱和，设备内空气温升达到10度左右，即可以使得空气相对湿度降低到60%以下，避免腐蚀的发生。这个方法存在局限性：第一，为了保证低湿度，对空气进行加热，使得设备工作的温度升高，对于需要高温运行的设备，相当于恶化了设备的工作温度环境；第二，空气在设备中是逐渐被加热的，在进风口附近，湿度较高，灰尘堆积也较多，更容易发生腐蚀。

因为上述原因，一般认为，大部分单板上因为自热温升，可以认为正常环境腐蚀发生很慢。通过湿尘试验也可以对设备的自身对灰尘、湿度、盐分的抵抗能力进行预评估。对于腐蚀更容易发生的位置，如风扇、设备进风口等部件，要考虑防护。当环境十分恶劣时，需要考虑增加隔离等防护措施。对于维护要求低的部件，可考虑三防工艺，但是三防工艺会增加成本，加长加工周期，影响维修，喷涂还会增加对环境的污染。

根据经验，化工厂、港口、地下车库等，湿度很大，存在大量有害气体或盐雾，甚至机房环境在蓄电池使用不当的情况下，也会出现漏酸等情况，对机房内设备带来危害（如图7所示）。无防护的设备的腐蚀尤易发

生在通风部件，对腐蚀的部件进行成分分析显示，腐蚀物主要的成分为硫和氯（如图8所示）。

目前的基站设计电路组装密度很高，中兴通讯软基站室外产品一般都采用纯自然散热或者热交换器柜，避免直通风场景。虽然直通风或者透气过滤膜能够降低设备成本，但是灰尘堵塞（透气膜产品）、对进气预加热以降低湿度的设计降低了散热能力，加之有害气体、灰尘、盐雾直接进入设备等问题，降低了产品的适用范围。对于环境洁净、抗腐蚀能力强、成本敏感，或者内部有蓄电池的设备，可考虑使用直通风设备。

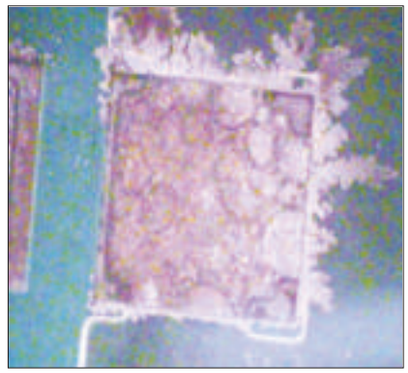
2.4 其他问题

设备及相关配件在长期工作中，还会面临各种其他问题，例如雷击、水浸（水滴）、日晒、人为因素等等。

在夏季，雷击和水浸的问题相对较多。光缆加强筋、动力线引雷，加上有的机房接地装置安装不良甚至被窃，导致设备接地不良，可能使设备造成雷击伤害。设备工作环境的控制是一方面，另一方面室内设备的一些防护指标，按照标准已经不能满足现场的实际要求，实际控制要在标准上有所提升，但是又不需要达到室外的防护标准，形成对实际工作环境理解后的设计要求。设备设计上也往往会根据使用的经验，做高于标准要求的设计。

有的机房存在渗漏滴水，在6—7月雨季，雨水从机房屋顶渗漏，进入室内设备并可能在设备内部形成局部积累，造成短路，设备设计上要尽

▲图7 蓄电池漏酸可能对机房内设备带来危害



▲图8 高湿硫化氢环境下的PCB腐蚀

量做到防滴，但是这样又可能增加设备尺寸、并对上下直通风道等热设计方式带来影响。

室外安装设备长期暴露在日光下，日晒会为带来设备额外的热负荷，需要使用遮阳棚/罩等方法避免日光直晒；日光照射导致的冷热循环盈利，以及紫外线对漆面分解，可能使得漆皮老化剥裂，塑胶套管脆化。

这些因素和设备相关，但很多已经不属于设备设计解决的范围，需要从多方面分析和解决。运营商和施工单位在进行机房设计建设的时候，需要考虑实际应用的不确定因素，控制施工质量，根据实际环境需要进行合理的辅料选材，为设备提供接近于设计指标的工作环境。综合设备异常损坏维修以及业务中断的成本，而不是一味的考虑材料成本，进行成本优化的工程设计与施工。

3 结束语

综上所述，通信设备的环境适应性设计，在可靠性物理的基础上，涉及到材料学、热力学、环境科学等多专业配合，需要结合很多实际工程环境和设备工作场景研究的经验，综合考虑运营商降低CAPEX和OPEX的需求，在一系列相互冲突的限定条件之中进行优化平衡。同时，通信设备环境适应性设计，在日益复杂的应用场景面前，无法只考虑设备自身，而是对环境也需要提出精细的要求，研究设备和环境的交互作用，通过设备制造商与运营商的共同配合与努力，确保整体方案的优化和可靠。

参考文献

[1] Extreme climatic conditions and derived conditions for use in defining design/test criteria for NATO forces materiel [R]. NATO STANAG 2985.
[2] 李永红, 徐明. MIL-HDBK-217可靠性预计方法存在的问题及其替代方法浅析 [J]. 航空标准化与质量, 2005(4):40-43.
[3] 曾纪科. 美军标的修订与塑封微电路在军用设备上的应用 [J]. 电子产品可靠性与环境试, 1998(4):45-48.
[4] CROWE D, FEINBERG A. Design for reliability [M]. Boca Raton, FL,USA:CRC Press, 2001.
[5] HNATEK E R. Practical reliability of electronic equipment and products [M]. New York, NY, USA: Marcel Dekker, 2003.
[6] MIL-HDBK-338B. Electronic reliability design handbook [S]. 2012.
[7] Lecture: Air movement and natural ventilation [EB/OL]. [2012-03-25]. http://www.arch.hku.hk/teaching/lectures/airvent/.
[8] GB 3095—1996. 环境空气质量标准 [S]. 1996.
[9] 甘肃省环境保护厅. 甘肃省环境质量概要（2010年度）[R]. 2010.
[10] 内蒙古自治区环境保护厅. 2006年内蒙古自治区环境质量状况公报 [R]. 2006.
[11] MODY V, JAKHETE R. Dust control handbook [M]. Park Ridge, NJ,USA:Noyes Data Corp, 1988.
[12] 陈鸿海. 金属腐蚀学 [M]. 北京: 北京理工大学出版社, 1995.

作者简介

郭丹旦，北京理工大学硕士毕业；中兴通讯股份有限公司系统架构师，从事无线通信软基站BBU硬件架构设计工作。

唐雄，中兴通讯股份有限公司BBU系统部部长；从事无线通信系统开发工作。

崔卓，吉林大学硕士毕业；中兴通讯股份有限公司系统架构师，从事无线通信软基站软件架构设计工作。

DOI: 10.3969/j.issn.1009-6868.2013.02.013
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130305.1103.001.html

GoTa 4G 宽带多媒体集群系统的研究

GoTa 4G Wideband Multimedia Trunk System

赵先明/ZHAO Xianming, 徐云翔/XU Yunxiang, 朱伏生/ZHU Fusheng
（中兴通讯股份有限公司，广东 深圳 518057）
(ZTE Corporation, Shenzhen 518057, China)

中图分类号：TN929.5 文献标志码：A 文章编号：1009-6868 (2013) 02-0049-05

摘要：在 LTE 系统中引入了群组预建立和邻区群组信息广播机制，使得集群终端在小区之间切换/移动无需和基站之间的信令交互而成为一种无缝切换过程，避免了切换过程中的信令交互和延迟；提出一种基于随机过程的群组呼叫建立过程的方法，使得话权申请的过程不依赖于信令链接，有效地提高了系统的性能。

关键词：集群；LTE 系统；随机过程；切换；开放式集群架构

Abstract: The group pre-establishment mechanism and neighbor group information broadcast mechanism have been added to LTE systems. In new LTE systems, handover of trunk terminals between cells does not require any signaling between base stations. Handover is seamless and without delay. In this paper, we propose group call establishment based on a random procedure. This ensures that the process of granting a call request is not based on signal connection. This solution improves system performance.

Key words: trunk; LTE; RACH; handover; GoTa

集群通信系统，按照实现的技术原理，可以分成两大类：（1）专业集群系统有北美数字集群标准 P25、陆上集群无线电（TETRA）和数字无线电标准（DMR）等。在这些集群系统中，空口协议栈和非接入层根据集群应用的特点和要求，有针对性地进行了设计。（2）基于个人无线通信系统的集群系统有基于码分多址（CDMA）的开放式集群架构（GoTa）系统和基于 LTE 的 GoTa 系统。这些无线集群技术的发展经历了从模拟到数字，从窄带数字集群到多媒体数字集群的过程。无线集群系统发展历程如图 1 所示。

窄带数字集群系统在语音集群业务的基础上可以提供有限的数字业务，以 TETRA 系统为例，一个载频的 TETRA 系统，具备 28.8 kb/s 的数据业务的承载能力，但是不支持视频，并且对增值业务的支持能力比较弱。

中兴通讯在 CDMA 系统的基础上，提出了 GoTa 集群通信系统，GoTa 系统除了能够提供传统的集群业务以外，还继承了 CDMA 系统的能力，如群发短信、高清视频、3.1~4.9 Mb/s 的数据承载能力等。

GoTa 4G 系统，是在 LTE 系统的基础上，引入 GoTa 集群核心网调度子系统（DSS）的多媒体宽带集群系统。由于 LTE 系统的控制面所面对的对象是单个终端，因此，所有的控制流程，如呼叫的建立、被叫过程和切换过程，首先需要保证用户和基站/核心网建立可保证安全的通信链路，然后再执行相应的操作，如切换、业务呼叫等。这就给 GoTa 4G 带来两个负面影响：

（1）信令风暴。集群应用是一发多收的通信模式，在任何时刻，只允许一个用户拥有话权，其他的群组用户只能接收话权用户的数据。当大量的用户从一个小区切换到另外一个小区的时候，或者话权用户释放话权，群组中所有的非话权用户申请抢

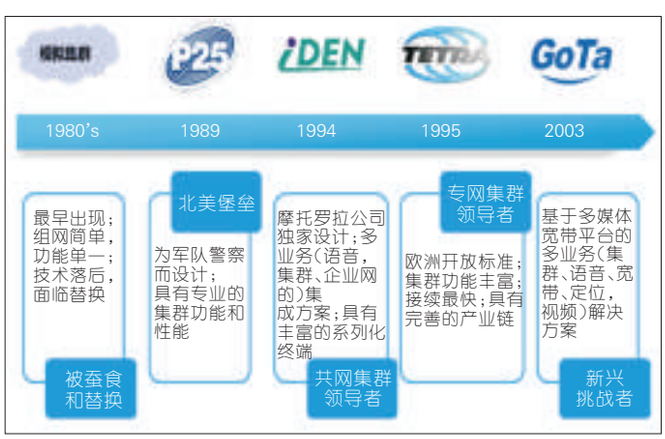


图 1 无线集群系统发展历程

收稿日期：2013-01-06
网络出版时间：2013-03-05
基金项目：国家科技重大专项（2013ZX03001-023）

占话权的时候，如果这些用户处于非链接态，就会造成空口控制面的信令风暴。

（2）呼叫/话权抢占延迟性能。对于处于下行链路空闲周期（IDLE）状态的终端，根据个人通信系统的协议规范，必须先和基站/核心网建立通信链路，然后建立业务承载，终端才具备发送业务数据能力。这样的处理流程很难满足呼叫和话权抢占的系统延迟需求[1-4]。

为解决这两个问题，本文提出了两个应对方案：群组预建立和广播机制，保证非话权用户的小区间的无缝切换；呼叫/话权抢占随机过程。

这两种解决方案有效地解决了GoTa 4G系统的性能问题，达到并且超过了现有集群系统的性能指标。

1 GoTa 4G系统简介

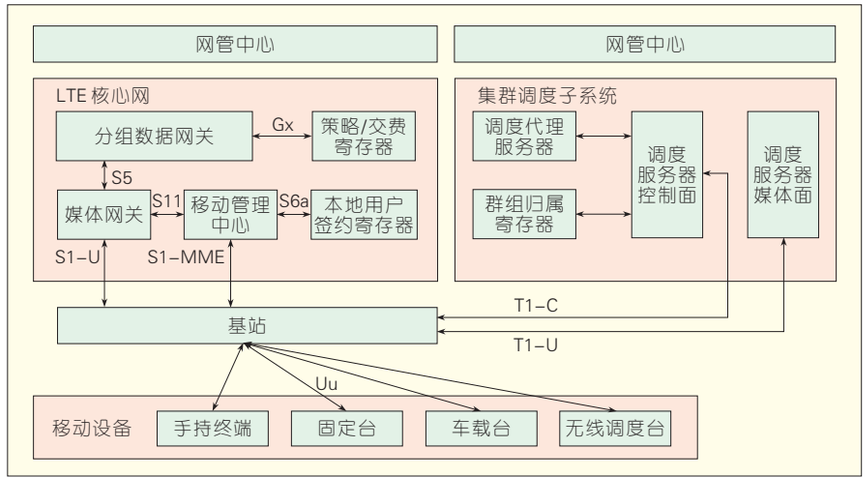
为达到数字集群系统的IP化、数据宽带化和业务多样化的演进目标，本文提出了基于LTE的GoTa系统演进系统（GoTa 4G）。系统充分地继承了GoTa的优点和功能，引入了LTE系统的宽带、扁平化系统和全IP特性，并且进一步提出了核心网分离的组网策略。系统架构如图2所示。

GoTa 4G系统整体网络架构为扁平化组网架构，由4部分组成：

- LTE无线接口E-UTRAN
- 演进的分组核心网（EPC）
- 集群核心网调度子系统（DSS）
- 集群终端

公网核心网LTE EPC为普通的符合标准协议规范的LTE核心网设备。在组网构成上，LTE EPC与集群的核心网隔离。当需要为终端提供普通的LTE业务服务时，需要布置LTE EPC相关设备。这种组网方式，遵循公网共用，专网隔离的总体原则。DSS系统功能与GoTa系统保持一致。

DSS系统主要目标客户是提供集群系统或业务的运营商。DSS系统由：集群调度服务器（PDS）、群组归属寄存器（PHR）和调度代理服务器（DAS）等组成。运营商通过配置PDS和PHR设备为集群用户提供集群调度服务，并提供集群业务的电信级鉴权、授权和计费服务。同时，运营商可以通过选配DAS设备，为集团用户提供调度台服务，为集团用户建立集群调度的虚拟专网。运营商还可以通过选配GoTa应用服务器（GAS）为客户提供行业定制业务。



▲图2 GoTa 4G系统架构

集群服务器（PDS）执行集群群组呼叫（PTT）呼叫处理，如鉴别PTT用户、建立PTT呼叫、判断PTT请求；接收上行链路来的PTT语音包，并分发到下行链路。对于空口，PDS负责动态管理群组临时无线身份标识（T-RNTI），保证在整网里，激活的群组的组标识（Group ID）与T-RNTI一一对应。PDS媒体面（M-PDS）和基站（eNB）存在之间的链路接口分成T1-C和T1-U接口，分别对应集群控制面接口和集群用户面接口。为应对集群应用场景的寻呼、信令连接、承载建立和用户鉴权等过程，在eNB和用户终端（UE）之间需要保证在无线链路链接（RRC）信令上携带的网络侧的信元要能够区分出网络侧的信元是路由到移动管理中心（MME）还是路由到PDS。

GoTa 4G系统为有效地兼容LTE系统，在不改变LTE基带规范的前提下，增加了若干条逻辑信道，以支持集群业务和控制功能。

上述设计保证了GoTa 4G系统可以支持大规模的公网组网（支持GoTa 4G业务和普通LTE数据业务）和大规模专网模式。

2 群组预建立和邻区群组信息广播机制

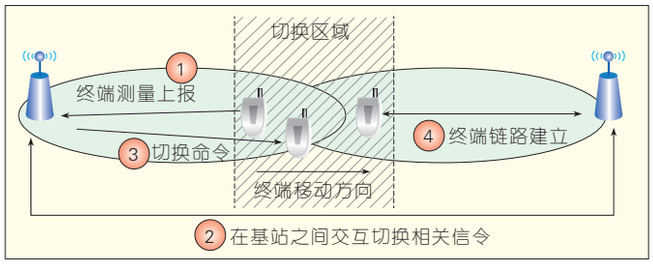
2.1 切换过程及在集群系统中存在的问题

LTE系统中，终端在两个小区之间的切换流程如图3所示。当终端移动进切换区时，向服务小区所属的基站发送测量上报，基站根据终端的测量上报，确定终端切换的目标小区，并和目标小区交互切换相关的信令。切换相关的信令交互完成后，基站向终端发送切换命令（HO Command），终端执行目标小区的接入过程，建立无线链路，完成一次小区间的切换过程。

在集群系统中，由于是一组用户在通话，因此，当组内用户在一个小区内发生群体切换时，就会造成：

（1）处于IDLE状态的用户首先切换到连接（CONNECTION）状态，然后执行向目标小区的切换。

（2）在一个极端的时间内，目标小区需要处理终端切入的信令处理



◀图3
个人无线通信系统小区间切换流程

峰值。

这就会造成两个小区的信令风暴，直接表现出的现象是系统发生拥塞或者切换失败导致的掉话。例如，一辆大巴车上，群组用户集中分布且都处于非话权态（接收数据），当大巴车穿越两个小区的交叠区时，可能会出现如下场景：

（1）所有的用户都进入链接态，然后执行切换流程。建立链接态需要5条信令，切换测量控制需要2条信令，切换过程需要3条信令（不包括两个小区间基站设备交互的信令），因此需要在极短的时间内，产生$n\times 13$条信令。n为用户的个数，当n超过基站的处理能力以后（在每个传输时间间隔（TTI）内，基站能够调度的用户数是有限的），或者产生掉话，或者切换延迟增加。

（2）LTE系统中切换性能与非竞争随机接入资源相关，非竞争随机接入的资源是有限的，当在某一个时刻大量的用户发起切换流程时，会导致切换接纳拥塞。

2.2 群组预建立和邻区群组信息广播机制描述

非话权态用户，接收下行群组公共业务信道，可以处于IDLE。对于处于非链接态的终端的小区间切换策略，执行小区重选过程，过程为：

（1）终端按照一定的规则周期性地检测服务小区和邻区的覆盖质量。

（2）当服务小区和邻区的覆盖质量满足一定条件时，终端自主地切换到目标小区，接收目标小区的广播消息而不用通知原服务小区。

群组非话权态用户，接收群组下行公共信道的数据，如果已经在邻区预建立了群组下行公共信道并且信道的相关配置信息能够被终端获取，那么终端也可以自主切换到目标小区接收群组下行数据，而不用和原服务小区进行信令交互。

具体的流程如下：

（1）当某群组内的用户在小区内分布时，如果处于工作状态（接收或者发送数据），那么则以一定的规则向基站上报状态。规则可以是基于覆盖的测量，也可以是周期的上报。

（2）基站接收到某群组的用户状态上报以后，认为这些用户具备向其邻区切换的潜在的可能性，因此，基站请求邻基站发送这个群组的下行数据，建立群组的下行承载。

（3）当邻基站完成群组的下行承载建立后，通知基站，并告知群组下行承载的相关参数配置。

（4）基站在小区内，向群组内用户广播邻基站的群组下行承载信息。

（5）当终端移动到两个小区间的切换区时，基于覆盖测量的结果，自行完成从服务小区向目标小区的切换，接收目标小区的下行群组数据，而无需原服务基站的参与和控制，完成无缝切换的流程（如图4所示）。

以话权用户和非话权用户在两个基站A/B间的切换流程为例，说明无缝切换的流程，如图5、图6所示。

话权UE切换流程如图5所示。集群群组呼叫建立后，此时UE A为群组A中话权用户，UE A由基站A下的小区移动到基站B下的小区。基站A为UE A切换前所在基站，基站B为UE A切换后所在基站。基站B下的小区之前已有非话权用户，并已建立群组A下行广播承载。

话权UE切换流程步骤如下：

步骤1，UE A为群组A中的用户，根据测量配置，向基站A上报测量配置。

步骤2，基站A进行切换判决。

步骤3–6，基站A通过控制面PDS，向基站B发送切换请求消息。基站A为话权用户UE A分配上下行承载资源，用于承载话权数据。资源配置完成后，基站B向PDS回切换请求响应（Handover Request Ack）消息。

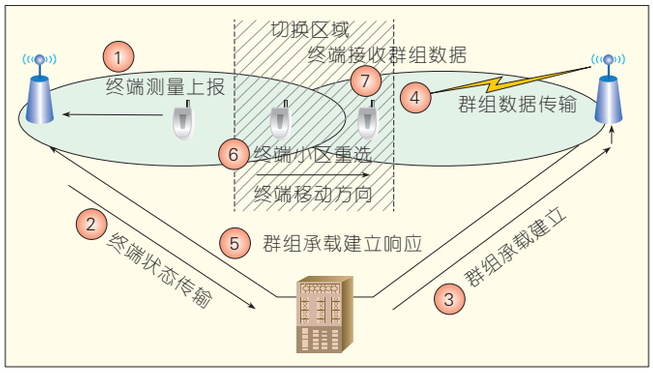
步骤7–8，基站A向UE A发送切换命令（Handover Command）。UE A收到后，切换到基站B下的小区。

步骤9，UE A向基站发送切换完成（Handover Complete）消息，进而发送给控制面PDS。

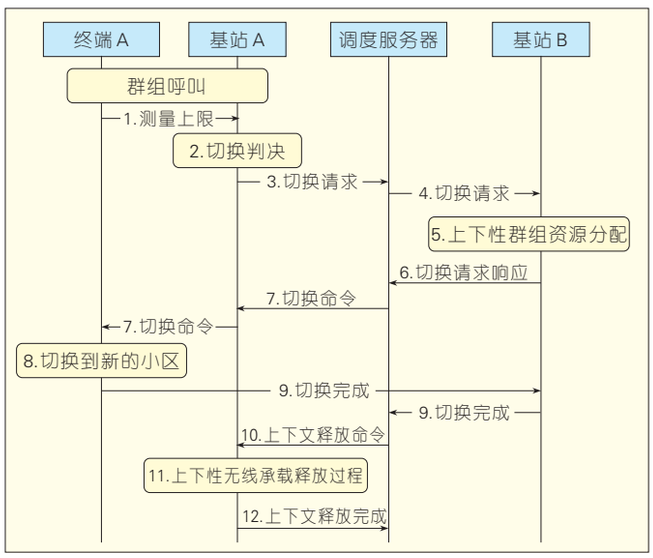
步骤10–12，控制面PDS向基站A发送初始上下文释放命令（Initial Context Release Command）消息，通知基站释放UE A上下文信息。基站通知UE A释放相关承载资源。空口释放完成后，基站向控制面PDS回初始上下文释放完成命令（Initial Context Release Complete）响应消息。

非话权空闲态UE切换流程如图6所示。集群群组呼叫建立后，此时UE A为群组A中非话权用户，正在基站A下接收集群下行广播数据，UE A由基站A下小区移动到基站B下小区。基站A为UE A切换前所在基站，基站B为UE A切换后所在基站。基站B下小区之前已有非话权用户，已建立群组A下行广播承载。

在组呼过程中，基站A将服务小区及相关邻区为群组所建立的信道资源，通过集群控制信道（TCCH）信道周期性发送给UE A。UE A处于空闲状态接收集群业务，在移动过程中，根据广播消息（SIB）内容，依照重选准则进行测量，确定是否进行小区重选过程。当满足重选准则时，UE A发现TCCH下发信道信息包含目标小区信道资源。此时，UE根据TCCH信道内容，自行选择到基站B下的目标



◀图4
基于群组预建立和广播机制的小区间无缝切换流程



◀图5
话权UE切换流程（下行广播资源已建立）

小区，在目标小区接收数据。

非话权空闲态UE切换流程的大致步骤如下：

（1）在组呼过程中，基站A将服务小区及相关邻区为群组所建立的信道资源，通过TCCH信道周期性广播发送。

（2）UE A根据测量信息，决定重选到基站B下的小区中。UE A读取TCCH信道广播信道资源，发现已有将要重选过去的目标小区信息。

（3）UE A自行重选至基站B下小区接收集群下行广播数据。

3 呼叫/话权抢占随机过程

3.1 基于LTE系统的集群系统呼叫/话权抢占过程中存在的问题

根据传统的个人无线通信系统协议规范，一个终端，在具备和系统交互信息的能力之前必须和系统之间建立无线链路。相应的，对于处于IDLE状态的非话权用户（听用户），当需要发起呼叫或者抢占话权的时候，首先需要和系统之间建立无线链路，之后发送呼叫或者抢占话权信令，完成抢占过程。

由于集群系统是一对多的通话模式，当话权用户释放话权的时候，会有多个用户同时发起话权申请流程，而这些发起话权申请流程的用户中只能有一个用户能够获得话权，相应地，其他的发起话权申请的用户所建立的无线链路，都属于无效的链路。这种工作模式会带来如下方面的影响：

（1）某个时刻多个用户发起话权申请触发的链路建立过程会带来信令风暴，造成系统拥塞。

（2）无线链路的建立过程，是一个比较耗时的过程，以LTE系统为例，建立RRC链路，所需要的时间在50 ms ~ 80 ms之间（不包括鉴权），而集群系统要求群组建立延迟小于300 ms，话权抢占延迟小于150 ms。

3.2 基于LTE系统的集群系统呼叫/话权抢占随机过程描述

在集群系统中，话权抢占或者呼叫流程，都是获取话权的过程，由于是多个用户同时抢占唯一的话权，因此，可以将此过程看成是一个随机的过程。

用户的呼叫/话权抢占信令在构建起和系统的通信链路之前，通过空口的随机接入过程的MSG3发送，基站将MSG3中所携带的呼叫/话权抢占信令转发给调度服务器，由调度服务器判决用户获得话权的能力。图7为话权申请流程示意图。图8为话权抢占流程示意图。

4 LTE宽带多媒体集群系统呼叫和切换流程性能分析

采用话权呼叫/抢占随机过程的性能延迟分析如表1所示。采用正常的话权呼叫/抢占过程的性能分析

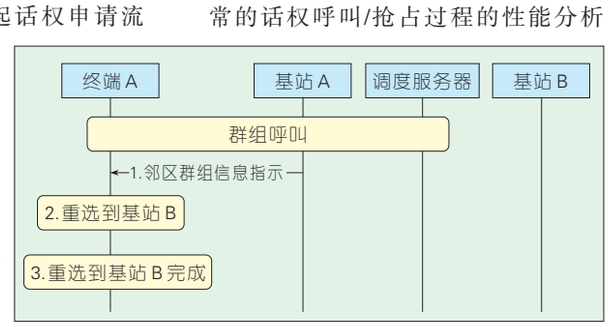
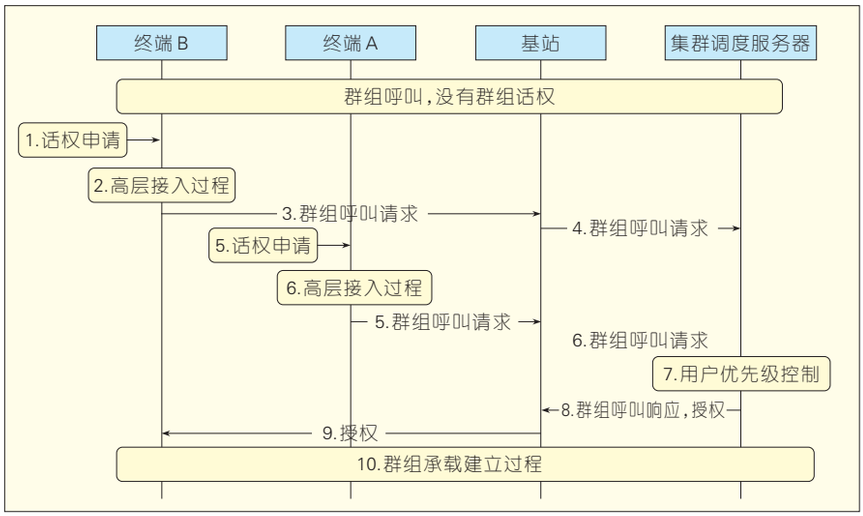
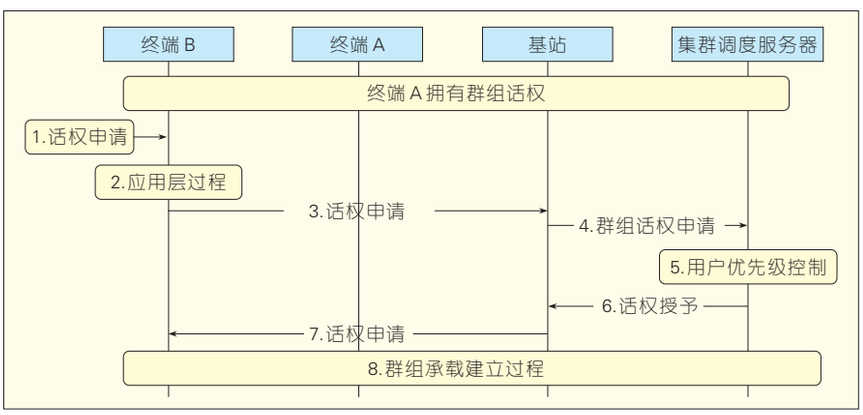


图6▶
非话权空闲态UE切换流程（下行广播资源已建立）



▲图 7 话权申请流程示意图



▲图 8 话权抢占流程示意图

如表 2 所示。从分析结果看，延迟可以控制在 110 ms。

由于邻区的群组下行承载已经预先建立并且在服务小区广播，因此，当非话权态用户从服务小区向目标小区切换过程中，无需和服务小区/目标小区执行切换相关的信令过程。随机话权抢占过程仅仅要求最终获得话权的用户和服务小区执行承载建立的信令交互过程，因此，优化了系统的信令负荷和性能指标。其中，基站设备间交互的信令减少 $n\times2$ 条，原服务小区空口信令减少 $n\times11$ 条（RRC 建立 5 条，默认承载建立 2 条，测量控制 2 条，测量上报 1 条，HO Command 1 条），目标小区空口信令减少 $n\times2$ 条，其中 n 为发起集体切换的集群用户数。

5 结束语

基于 LTE 系统的集群系统可实现多媒体宽带集群业务，但是直接照搬 LTE 的过程并不能满足集群业务的需求，并且由于集群业务的特殊性，会带来在 LTE 系统比较罕见或者概率较低的问题。

对现有的 LTE 系统的逻辑过程进行适当的改造，使其能够满足集群系统的性能需求，对实现多媒体宽带集群业务具有重要意义。

▼表 1 话权呼叫/抢占随机过程性能分析

步骤	时延/ms
主叫随机接入过程（Msg 3）	15
主叫上行承载建立过程	0
话权申请请求	40
通知被叫，话权用户变更	15
被叫侧应用层处理	40
合计	110

▼表 2 话权呼叫/抢占过程性能分析

步骤	时延/ms
RRC 链接建立（主叫随机接入过程）	70
默认承载建立	80
话权申请请求	40
通知被叫，话权用户变更	15
被叫侧应用层处理	40
合计	245

参考文献

[1] 3GPP TS36.321. Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification [S]. 2012.
[2] 3GPP TS36.211. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation [S]. 2012.
[3] 3GPP TR 36.213. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures [S]. 2012.
[4] 3GPP TR 36.331. Evolved Universal Terrestrial Radio Access (E-UTRA);Radio Resource Control (RRC) [S]. 2012.

作者简介

赵先明，哈尔滨工业大学通信与电子系统专业博士毕业；中兴通讯股份有限公司高级副总裁，哈尔滨工业大学兼职教授，中国科学技术大学博士生导师；长期从事移动通信系统的技术研究、应用开发与产业化工作；获国家科技进步一等奖、二等奖各 1 项；已发表学术论文 2 篇，出版专著 1 部。

徐云翔，中兴通讯无线经营部系统工程师；长期从事无线通信领域的研究和设计工作。

朱伏生，中兴通讯无线经营部政企网总工程师；长期从事无线通信产品，尤其是集群产品的研究和开发工作。

DOI: 10.3969/j.issn.1009-6868.2013.02.014
网络出版地址：http://www.cnki.net/kcms/detail/34.1228.TN.20130305.1713.002.html

交换机中同步以太网技术的研究及其实现

Synchronous Ethernet Technologies in Ethernet Switch

中图分类号：TN919.21 文献标志码：A 文章编号：1009-6868 (2013) 02-0054-04

曾富前/ZENG Fuqian
（中兴通讯股份有限公司 有线研究院，北京，100191）
(Wireline Product R&D Institute ZTE Corporation, Beijing, 100191, China)

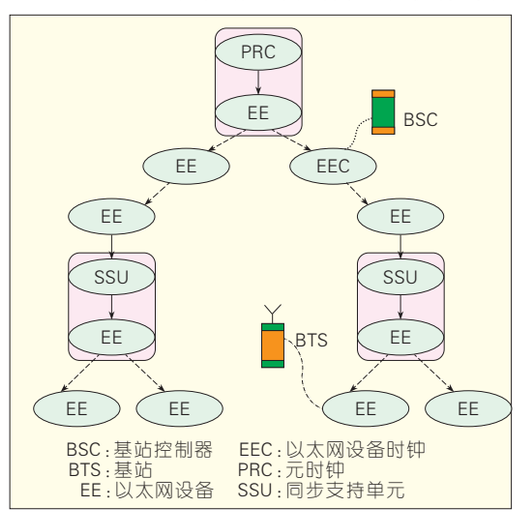
摘要：提出了一种支持同步以太网交换机工作在主从模式状态下同步时钟网络架构，及其在不同工作模式下的时钟同步机制；提出了交换机工作在从模式状态系统同步链路时钟的优先级选择策略，以及工作在主模式状态下交换机系统时钟指定选择原则。实践表明，这种同步以太网时钟架构和同步机制能很好地满足网络应用。

关键词：同步以太网；以太网同步状态消息；锁相环；漂移

Abstract: In this paper, we propose a synchronous clock network architecture for the Ethernet switch (ES) so that synchronous Ethernet (Sync-E) is achieved when the ES is in slave or master mode. We also propose different synchronous clock mechanisms for the ES in different modes. We introduce the priority selection strategy of synchronous clock link when the ES is in slave mode. The specific selection principle for the system clock when the ES is working in the master mode is also introduced. Finally, we suggest that this synchronous Ethernet clock architecture and mechanisms can meet needs of network applications.

Key words: synchronous Ethernet; Ethernet synchronization messaging channel; phase locked loop(PLL); wander

电信网络正在从基于时分的电路交换向基于分组的包交换的下一代网络发展，电信服务都可以统一到统一的包交换平台上来。以太网在网络汇聚中扮演越来越重要的作用，现在无线网络汇聚和骨干网络以及在城域网中都在应用以太网。不少接入网需要接入单元间时钟的同步，比如基站、无源光网络（PON）接入以太网，这就涉及到同步以太网的问题。接入单元有的需要频率同步，比如无线网络中GSM、UMTS/W-CDMA、LTE（FDD）只需要频率同步；有的还需要相位和频率都同步，比如CDMA2000、LTE（TDD）、TD-SCDMA、WiMAX等网络。同步以太网技术用于分布式网络的频率同步，但其不支持相位同步。使用同步以太网的网络，网络的各个节点均需支持同步以太网。传统的运营网络一般建有自己的SDH/PDH同步网络，以确保整个网络的频率同步。目前通过以太网来承载传统通信网络的频率同步和时间同步信息有着现实的迫切要求。图1是同步以太网在网络中的典型应用场景[1-2]。最上端是元时钟（PRC），每个同步网络单元都含有以太网设备时钟（EEC）。在正常情况下EEC是锁定某一输入业务端口的频率，业务输出端口频率和输入端口同步。这种链路拓扑结构能实现EEC至上而下的频率同步，最终和PRC时钟同步，这种同步方式类似SDH/SONET同步网。网络拓扑中的同步支持单元（SSU）作用是减少链路抖动和漂移的累计，另外也可以作为一个节点时钟给别的设备提供同步。在上游设备时钟出现问题，比如PRC时钟丢失时，自身可提供保持功



▲图1 同步以太网的典型应用拓扑

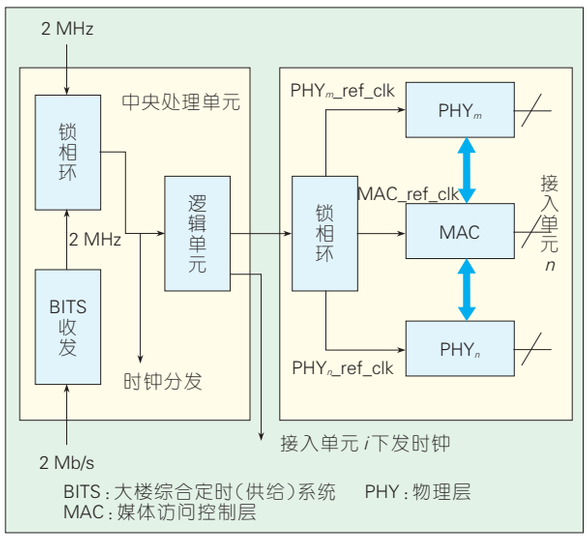
收稿日期：2012-12-27
网络出版时间：2013-03-05

能，直至上游设备时钟恢复。

1 以太网中的同步技术

1.1 同步以太网主设备

对于同步以太网主设备，其功能模块时钟拓扑如图2所示。设备一般是在设备外接标准的高精度时钟源，比如2 MHz时钟或者2 Mb/s数据或者是别的时钟源输入接口。2 MHz时钟通过锁相环输出，时钟的频率及其参数可以通过锁相环参数配置。2 Mb/s数据通过大楼综合定时（供给）系统（BITS）收发芯片接收链路码流提取2 MHz时钟，然后将提取的链路时钟2 MHz送给锁相环，再由中央处理单元中锁相环通过配置输出一个设置时钟。锁相环输出所需要时钟送给逻辑单元，逻辑单元分发出若干路下发时钟到各个接入单元，再由接入单元的锁相环输出到接入单元交换芯片媒体访问控制层（MAC）、物理层（PHY）作为其工作时钟，实现时钟源和接入单元交换芯片和PHY芯片的时钟同步。主设备通过接入单元的PHY链路数据输出。从设备接入单元根据主设备接入单元输出链路数据，通过提取链路时钟，然后根据从设备时钟同步原理而实现主交换设备和从交换设备的同步。以太网从设备不需要外接时钟源。



▲图2 同步以太网主设备功能模块时钟拓扑

1.2 同步以太网从设备

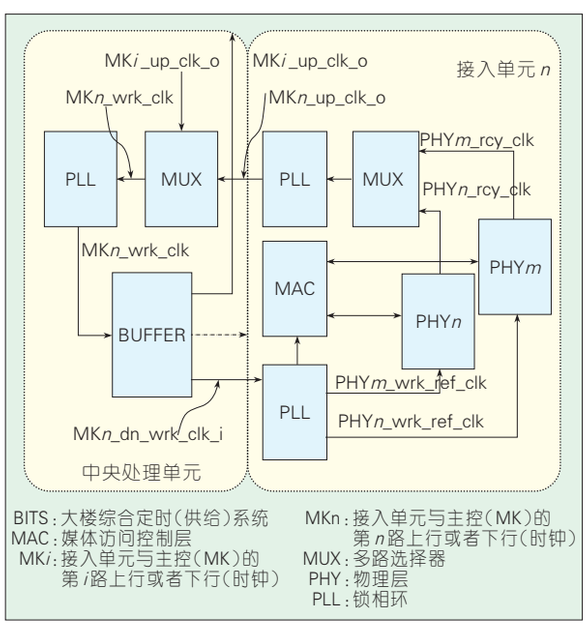
同步以太网从设备利用从主设备互联的数据流端口，通过PHY或者MAC芯片数据接收端集成的时钟数据恢复（CDR）来提取接收数据链路的时钟。这个时钟除了可以用作发送端时钟外，还可以用于芯片对接收数据的解串行化和解复用高速串行数据成低速并行数据，恢复功能可从发送方数据恢复时钟。图3为同步以太网从设备时钟拓扑框图。选择从哪路端口恢复时钟是可以配置选择的。根据电路所采用PHY的数量，恢复出来的每路高精度时钟，送给一个多路选择器或者逻辑，按照一定的规则选择一路恢复时钟送给一个锁相环（PLL）。PLL输出某一频率的高精度时钟送给中央处理单元。中央处理单元通过一个复用器按照一定的选择规则选择多个输入的接入单元时钟的一个，作为主控锁相环PLL的输入时钟源，然后中央处理单元PLL输出一高精度的时钟，通过时钟驱动缓存器分发给各个接入单元。各个接入单元根据中央单元输入的时钟，通过配置锁相环输出接入单元MAC和PHY所需要的频率，作为接入单元MAC和所有PHY的发送工作时钟，从而实现交换设备与主设备同步。一般接入设备只需MAC和PHY和主设备同步。

1.3 以太网同步状态消息ESMC的应用

一般来说，同步以太网既可以工作在普通模式，也可以工作在同步模式，默认是普通以太网方式。工作在同步模式时，接收端除了从输入信号提取输入信号频率并送给时钟系统进行处理进行时钟分发外，还获取并处理以太网同步消息信道（ESMC）状态消息，获取链路提供的时钟质量等级（QL）。

QL可以作为本地进行时钟切换依据，根据当前获取的链路的时钟质量等级，以判断系统是否切换到更高级别的时钟。在同步以太网模式，链路发送端锁定系统时钟用于数据发送，同时还产生ESMC消息发送链路时钟质量等级QL到下游环节，ESMC包含链路同步状态消息（SSM），用于表示EEC时钟等级，这个信道是按照组织特性慢协议（OSSP）协议提供。同步状态消息的数据格式如表1所示[3]。

表1中，目的地址是组播地址，低速协议类型及其协议子类型固定为88-09-0A[4-5]。在ESMC PDU中数据



▲图3 同步以太网从设备时钟拓扑

▼表1 ESMC PDU格式

8位字节序号	ESMC域
1—6	目的地址=01-80-C2-00-00-02
7—12	源地址
13—14	低速协议类型=88-09
15	低速协议子类型=0A
16—18	ITU-OUI=00-19-A7
19—20	ITU子类型
21	版本号(4比特) 事件标志(1比特) 保留(3比特)
22—24	保留的(3个八位字节)
25—1532	QL未来增强型类型长度值(4个八位字节)和填充数据
最后4个字节	FCS

ESMC:以太网同步消息信道
FCS:帧校验系列
OUI:IEEE注册机构分配的组织唯一标示符
QL:质量等级
TLV:类型长度值

和填充区域用4个字节来表示时钟的质量等级,其时钟质量等级采用类型长度值(TLV)来表示,其数据格式如表2所示。在同步以太网中,SSM用于表示上游设备给下游设备所同步跟踪的时钟源的质量。SSM包含在ESMC PDU中,周期性由上游设备发给下游设备作为下游设备选择时钟源的重要根据。下游设备时钟子系统根据端口收到的SSM级别,总是选择最高级别的时钟源,当高级级别时钟源发生故障后,再根据统计的当前的SSM等级,选择次之的时钟源。如果存在等级相同的时钟源就按照定义的优先级选择时钟源,如果所有链路时钟均不可用,选择本地时钟作为当前系统时钟。如果丢失所有的参考时钟源,锁相环PLL进入保持(holdover)状态;SSM消息需要沿着链路依次逐级传递,每级EEC都支持SSM消息的处理。时钟子系统通过这种方法选择系统源时钟,以使系统保持稳定工作。同步以太网ESMC中的SSM定义了2种时钟质量等级,如表3所示。

表3中,QL-EEC1主要适用于同步以太网设备时钟架构为2048 kb/s的体系,QL-EEC2主要面向同步以太网时钟架构为1544 kb/s的同步体系。

▼表2 QL TLV数据格式

比特数	含义
8比特	类型:0x01
16比特	长度:0x04
4比特	0(未使用)
4比特	同步状态消息编码

▼表3 同步以太网SSM消息

时钟	消息	同步状态消息编码
EEC1	QL-EEC1	1011
EEC2	QL-EEC2	1010

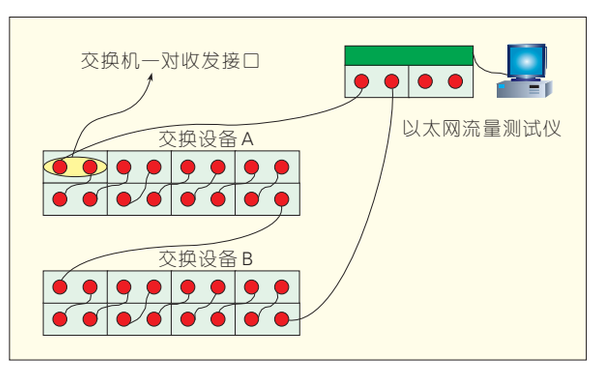
EEC:以太网设备时钟 QL:质量等级

2 同步以太网的调试方法

2.1 同步以太网基本调试方法

采用若干套交换设备分别级联组网也可以混合组网,这时两个节点间上下游构成主从设备。这里以同步以太网组网中两个上下游节点为例,上游设备配置为主设备、下游设备配置从设备,从设备端口提取上游设备链路时钟。一个同步以太网流量测试环境如图4所示。

图4中,以太网流量测试仪器的一对端口的收发端口与设备A的第一个以太网接口收发端口收发互连。图4中设备每个方格为一对以太网收发端口,上下一组以太网端口构成一对虚拟局域网(VLAN)。相邻一对VLAN通过网线或者光纤连接。设备A最后一个以太网的收发端口和设备B的收发端口收发互联。上下两对以太网接口构成一对VLAN,相邻VLAN通过网线或者光纤依次连接。设备B最后一对以太网的收发和以太网流量测试仪器的另一个收发端口互联。设备A配置成主设备,设备B为从设备。设备B从连接设备A的一个输出端口的接口提取链路时钟,通过设备B的时钟子系统的一系列变换形成设备B发送链路时钟。设备B通过读取PHY或者MAC的状态判断接入端口是否锁定上游设备链路时钟;另外,也通过以太网流量测试仪器发送以太网报文进入设备A,并通过一系列的设备端口连接和软件配置,使得输出设备A的报文进入设备B,然后最终回到以太网流量测试器的输入端口,同时在设备B到设备A方向通过以太网流量测试仪器对发反方向报文,这样可以同时验证两个方向在同步以太网模式下端口报文收发情况,报文可采用线速随机长度报文。通过这些方法可以观察同步以太网设备运行的同步稳定性。



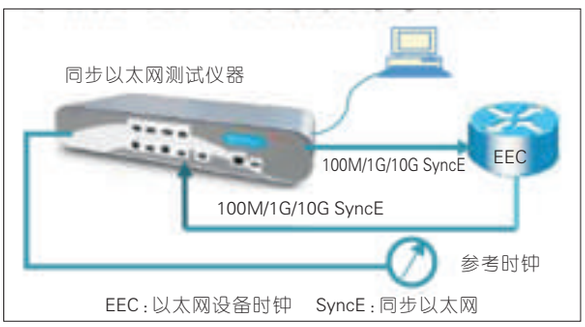
▲图4 同步以太网流量测试示例

2.2 同步以太网的时序特性

对于支持同步以太网的设备,其时序特性主要是满足ITU-T Recommendation G.8262/Y.1362标准中支持2048 kb/s体系EEC-option1选项和支持1544 kb/s体系的EEC-option2选项,符合业务端口输出漂移最大时间间隔误差(MTIE)和时间偏差(TDEV)的指标要求。对于输入端口需要满足输入端口漂移容限MTIE和TDEV模板的要求,端口还需满足抖动方面的标准指标要求[6-8]。满足这些指标要求便于实现上下游同步以太网设备单元间的互操作性,便于设备间的对接。对于同步以太网设备端口的漂移特性主要是以MTIE、TDEV来表征。MTIE反映信号的频

率偏差和相位的瞬态变化，TDEV 反映信号的频率成分。MTIE 是指在采样间隔 τ_0，总观测时间 $\tau=n\tau_0$，测量周期 $T=(N-1)\tau_0$，时间间隔误差（TIE）峰峰值的最大值，其中观测周期包含在测量周期中。

对于同步以太网输入输出漂移指标必须符合标准中的模板要求。图 5 所示测试环境可以测试同步以太网业务输出输入端口漂移最大时间间隔误差和时间偏差，可以用于业务端口的输入抖动容限和端口的抖动转移特性，还可以模拟 ESMC 状态测试时钟状态切换。不同的测试，仪器需要不同的配置设置。漂移测试需要外接一个高精度的稳定时钟源。



▲图 5 同步以太网部分指标测试环境

漂移指标主要包括漂移容限、漂移转移特性，漂移转移特性主要是考察设备对输入漂移的抑制能力。抖动测试主要是测试输出抖动和输入抖动容限，测试输出抖动是否满足标准要求[9]，测试输入端口容许的最大输入抖动，设备正常工作也不告警。

3 结束语

本文通过对同步以太网主从设备同步时钟链路部分的研究和设计，对同步以太网设备的基本功能测试，通过对同步以太网端口漂移、抖动指标等重要指标的测试，可以很好、可靠地实现同步以太网功能，满足同步以太网设备在网络中的应用。

参考文献

[1] 王文鼐, 王斌, 糜正琨. 分组通信网的同步与定时技术（2）[J]. 中兴通讯技术, 2011,17(2):56–60.
[2] 李勤. PTN 时钟同步技术及应用 [J]. 中兴通讯技术, 2010,16(3):26–30.
[3] ITU–T G.8264/Y.1364. Distribution of timing through packet networks [S]. 2008:3–17.
[4] KENNETH H. Tellabs synchronous Ethernet to transport frequency and phase/time [J]. IEEE Communication Magazine, 2012,180(8):152–160.
[5] ITU–T G.810. Definitions and terminology for synchronization networks [S]. 1996:16–19.
[6] ITU–T G.8261/Y.1361. Timing and synchronization aspects in Packet Networks [S]. 2006:16–21.
[7] ITU–T G.8262/Y.1362. Timing characteristics of synchronous Ethernet equipment slave clock (EEC) [S]. 2007:2–29.
[8] ITU–T .813. Timing characteristics of SDH equipment slave clocks (SEC) [S]. 2003.
[9] ITU–T O.172. Jitter and wander measuring equipment for digital systems which are based on the synchronous digital hierarchy (SDH) [S]. 2005:18–26.

作者简介

曾富前，电子科技大学硕士毕业，现在中兴通讯从事交换机数通产品的研发工作，主任工程师，曾从事 3G 无线 CDMA 基站产品、宽带无线接入产品硬件研发；已发表论文 1 篇，申请发明专利 2 项（获授权 1 项）。

综合信息

中兴通讯发布 UMTS 性能提升方案 频谱效率提升 40%

【本刊讯】2013 年 2 月 26 日消息，2013 年全球移动通信大会（MWC）期间，中兴通讯发布了创新 UMTS 性能提升解决方案，该方案可提升频谱效率 40%，帮助运营商大幅提升 UMTS 网络性能，以适应智能手机普及率提升等带来的移动数据业务流量爆炸式的增长。

UMTS 性能提升解决方案包含 En–Performance、Capacity SON、多用户检测和 CRAN RRU 4 个子方案。

En–Performance 子方案主要通过信令风暴抑制、控制面和用户面硬件资源共享、承载预配置、并行软切换等功能提升网络性能，可使 RNC 性能提升 30 ~ 100%，基站控制面性能提升 60%以上。

Capacity SON 子方案通过专利的功率控制技术，无需终端配合，可实现小区上行容量提升 35 ~ 50%；下行采用信道功率优化技术，可实现下行容量提升 10%，此外，基于负荷的无线参数自优化可改进用户体验。

多用户检测子方案采用干扰消除（IC）技术降低用户之间干扰，大幅提升系统容量，实验室测试表明，该技术带来的增益在 30 ~ 60%。目前，该技术已经在多个国家成功商用部署，整网性能提升在 10%以上。配合新一代 768CE 大容量基带板，该方案将提升网络容量。

CRAN RRU 子方案对多 RRU 组网技术进一步增强，满足室内、高速铁路覆盖等复杂场景的性能需求。发射功率从 80W 提升到 120W，并引入多天线联合检测等功能，提升多 RRU 场景下的功率和码资源利用率，下行容量增益提升 50%，大幅增强了单站点多 RRU 组网下的站点性能。

作为主流的无线网络端到端解决方案供应商，中兴通讯已为全球移动运营商部署超过 140 张 GSM/UMTS 无线网络。通过部署 UMTS 性能提升解决方案，中兴通讯将帮助移动运营商进一步提升网络质量，在激烈的市场竞争中处于领先之地。

DOI: 10.3969/j.issn.1009-6868.2013.02.015
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20130301.1044.001.html

大数据 2

韩晶/HAN Jing,宋美娜/SONG Meina
(北京邮电大学计算机学院,北京 100876)

[编者按]数据是与自然资源一样重要的战略资源,大数据技术就是从数量巨大、结构复杂、类型众多的数据中,快速获得有价值信息的能力,它已成为学术界、企业界甚至各国政府关注的热点。本讲座将分3期对大数据进行讨论:第1期介绍了大数据的提出、含义、特点,大数据和云计算的关系以及大数据典型应用;第2期介绍大数据获取、存贮、搜索、分享、分析、可视化等方面的关键技术,并对当前热点技术—可视化进行重点分析;第3期将探讨数据流挖掘等实时数据分析技术,介绍大数据中非结构化数据处理和挖掘技术,并给出大数据发展面临的挑战与应用前景。

中图分类号:TN91 文献标志码:A 文章编号:1009-6868 (2013) 02-0058-05

5 大数据生态系统

5.1 大数据生态系统

2011年6月,IBM架构师Stephen Watt在《Deriving new business insights with Big Data》文中对大数据生态系统进行了简单描述,提出大数据生态系统实际上就是数据的生命周期,即数据采集、存储、查找、分析和可视化的过程[1],见图1。

在这样的生态系统中,每个环节都存在着不同的商业需求,而需求的出现必然会导致创新的产生。所以,在每一个环节都有不少企业在深耕自己所在的领域,试图通过新技术和新方法来实现新的商业模式。

5.2 大数据生态图谱

随着大数据生态系统的逐步形成,很多人在尝试绘制和更新大数据生态系统图谱,希望通过对大数据领域的公司、技术、产品进行细分,及时了解到大数据生态系统全貌。在众多图谱当中,比较有代表性的是美国On Grid Ventures公司Matt Turck等人于2012年10月绘制更新的大数据生态图谱V2.0[2],如图2所示。

尽管各个图谱的分类方法、全面性、时效性、权威性各不相同,但我们仍可以观察到:

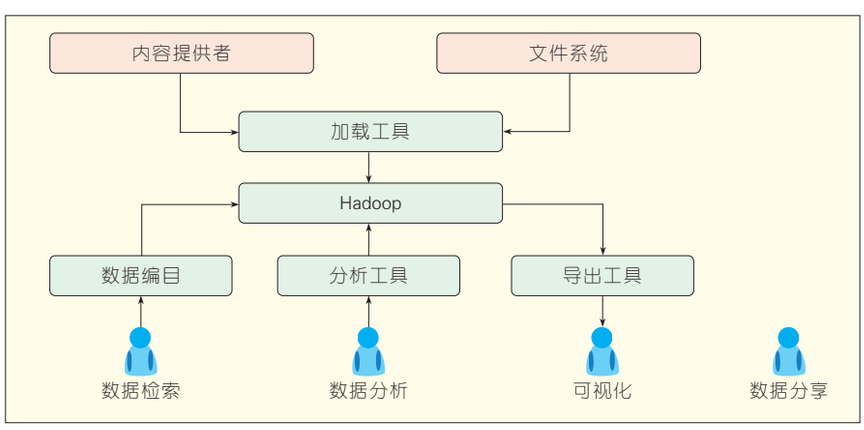
(1)大数据领域的企业主要集中在数据集市、数据存储(基础设施)、数据分析、数据应用4个层面,其中数据应用层面又包含数据服务、数据检索、商务智能,可视分析等。这正符合数据科学中对数据全生命周期管理的描述。此外,很多企业业务覆盖大数据多个层面,有的企业甚至已经建立了完整的大数据栈,成为"大数据应用服务提供商"。

(2)在大数据领域,活跃着的除了IBM、ORACLE等众多知名公司外,像Splunk、Tableau等专业大数据公司也及时跟上了大数据的浪潮,成功地获得了投资者和业界的关注。

(3)开源软件与大数据的结合迸发出惊人的颠覆性力量,更多厂商开始使用开源大数据工具,以支持其大数据业务。

大数据生态系统中覆盖大量的技术和产品,其中一些在大数据技术发展道路中起到了巨大的推动作用。

(1)Hadoop



▲图1 大数据生态系统

收稿日期:2013-01-05
网络出版时间:2013-03-01

▲图2 大数据生态系统

在大数据时代,Hadoop可以说是最耀眼的明星。凭借其开源和易用的特性,Hadoop不仅是大数据时代数据处理的首选,也是拥有海量数据处理需求的公司的标准配置。此外,许多商业创新也都围绕Hadoop展开的,并在大数据时代占据一席之地,如Cloudera推出的软件发布包可以帮助企业更方便地搭建以Hadoop为中心的数据管理平台;MapR则将Hadoop的速度改造为原来的3倍;海量数据管理软件商Platfora旨在提供一个更为友好且更具操作性的用户界面,它可以兼容包括Cloudera和MapR等多种Hadoop版本,能够大大降低使用Hadoop的门槛;而AsterData(已被TeraData收购)的核心技术SQL-to-MapReduce可将海量非结构化数据的处理技术和结构化数据的数据仓库技术结合,以弥补传统数据仓库的公司所欠缺的高速处理海量非结构化数据的能力。

(2)NoSQL

与Hadoop密切相关的NoSQL也一直是大数据领域的热点。NoSQL凭借其高性能和可扩展性等优势,成为关系数据库的强劲对手,在大数据时代占据一席之地。根据存储模型和特征,NoSQL大致可分为列存储、文档存储、key-value存储、图存储、对象存储、XML数据库等类型,虽然也存在个别数据库可被归为多种类别的现象,其中,HBase、MongoDB、Cassandra、CouchDB、Neo4j、HyperTable等NoSQL已被相当多的企业和开发人员所熟知。

(3)NewSQL

无论NoSQL是被解释为NoSQL,还是后来的Not Only SQL,其不支持结构化查询语言(SQL)语言的特性为开发人员带来诸多不便。因此,为了同时满足高性能和支持SQL两个方面,NewSQL被设计出来。NewSQL作为全新的关系数据库产品,或将关系模型的优势发挥到分布式体系结构中,或将关系数据库的性能提升到不必进行横向扩展的程度,这使得NoSQL面临前所未有的挑战。典型的NewSQL有VoltDB、Marklogic、Xeround、NuoDB等。

(4)Data Marketplace

除了解决大数据处理、存储问题之外,开放数据资源也在相当大程度上加速了大数据技术的发展。目前大部分的企业所面对的数据都是由内部系统或者交易记录日志之类的东西所产生的,然而如果能够获得企业自己无法获得,或者已经被处理过的外部数据,那么内外数据融合分析后产生的价值将不可估量。因此,能够下载或者访问数据集,自然而然也就成为了商业需求,甚至美国政府都推出了官方的数据集网站。

2009年5月,美国联邦政府正式启用了官方公共数据资源分享网站Data.gov,其数据内容涵盖了所有美国联邦政府行政部门在运营管理过程中采集、生产或转换而来的、有潜在价值的、可供再次开发利用的数据集。Data.gov鼓励个人开发者使用政府发布的数据集,开发出新颖的应用。值得一提的,该网站于近期正式对外发布了全新的"开源政府平台"(OGPL),该平台的代码将会对全球的开发者开放。

在中国,数据堂(datatang.com)是目前最为专业的科研数据共享服务平台,该平台致力于为全球科研机构、企业及个人提供科研数据支持,其数据内容主要是科研数据集,同时也提供浮动车历史数据、路况历史数据和车牌数据等,用户也可以上传发布自己的数据。通过该平台不仅使得中国的科研机构、企业、高校和个人之间可以充分共享数据,也促进各类科研数据价值的最大化。

在全球范围的大数据热潮中,对于大多数企业来说,大数据与自己有什么关系?如何快速直观地理解和发现大数据中的价值?没有足够"大"数据的情况下如何才能在大数据时代获益?虽然这些问题还没有完美的答案,但许多企业已经进行了积极的尝试,通过数据可视化尝到了大数据的甜头。

6 可视化和可视分析

在众多描述大数据的词语中,

"金矿"、"油田"等的描述最为常见，这意味着人们开始意识到大数据中蕴含着丰富的价值。然而，巨大的数量、数据的固有复杂性及未知的分析目标都放大了任务的难度。如果能够有一种简单的方式对数据规律进行直观展现，必将使大数据中的价值得到快速理解和发现，可视化就是这样的方式。

6.1 数据可视化、信息可视化和可视分析概述

可视化由来已久，1861年法国工程师Charles Joseph Minard绘制了《拿破仑征俄战役图》可以看作可视化领域的经典案例。到了18世纪后期数据图形学诞生，抽象信息的视觉表达手段一直被人们用来揭示数据及其他隐匿模式的奥秘。随着20世纪50年代计算机图形学的出现，信息技术加速了可视化的演变。时至今日，可视化已经发展为数据可视化、科学可视化、信息可视化、可视分析这几大方向。

数据可视化起源于20世纪50年代，其基本思想是将数据库中每个数据项作为可视化图形中单个元素，同时将数据的各个属性值以多维数据的形式表示，通过从不同维度观察数据而达到对数据深入洞察和分析的目的。

科学可视化是一个典型的交叉学科，源于1987年布鲁斯·麦考梅克等人编写的网络文件系统(NFS)报告《Visualization in Scientific Computing》(意为"科学计算之中的可视化")。科学可视化主要是将具有几何结构的三维数据转换为图像，应用领域涵盖科学和工程的多个方面。

信息可视化也是一个跨学科领域，出现于20世纪90年代，旨在为许多应用领域之中大规模非数值型信息资源的视觉呈现提供支持，这些信息资源可能是软件系统之中众多的文件、大规模并行程序的日志踪迹信息、网站内容等。与科学可视化相比，信息可视化侧重于异质数据集，如非结构化文本当中的点。

可视分析则起源于2005年，它是一门通过交互可视界面来分析、推理和决策的科学，通过将可视化和数据处理分析方法结合，提高可视化质量的同时也为用户提供更完整的大规模数据解决方案[3]。如今，针对可视分析的研究和应用逐步发展，已经覆盖科学数据、社交网络数据、电力等多个行业。

虽然在这几大方向之间的边界还未完全清晰，不过，其相互关系和区别可以总结如下：数据可视化外延不断扩大，可以认为数据可视化包含科学可视化、信息可视化和可视分析；科学可视化处理的是那些具有天然几何结构的数据；信息可视化处理的是异质的抽象的数据结构；可视分析则主要通过意会、推理、互动融合的方式来挖掘数据中的问题和原因。

可视化融合了问题的求解和艺术表现方式两个方面，允许我们同时通过理性和感官方式来感受数据，那么怎样才是成功的可视化？Noah Iliinsky在《数据可视化之美》一书中提到[4]，一个称得上"美"的可视化，必须具备新颖、充实、高效和美观4个关键要素。新颖性体现在必须从崭新的视角观察数据，传统可视化展现方式(如柱形图)虽易理解，但不够新奇有趣，是不足以激发读者新的理解的；充实性体现在可视化一定要为读者提供获取信息的途径，从而向读者传递信息甚至知识；高效性指成功的可视化须尽可能直截了当，而不允许展示太多与目标和主题无关的信息；美观是指合理的图形构建(坐标轴、布局、色彩、线条等)是实现可视化之美的必要因素。这四要素必须同时具备，否则不能对数据进行有意义地呈现。

6.2 可视化之美

美丽的可视化作品不同于传统的可视化，它们能够通过创造不同于惯例的图形构建方式，揭示数据显性和隐性的特征，使读者在对可视化效果感到惊喜的同时收获启示。通过以下的一些案例我们可以充分体会到这一点。

(1)电信数据可视化——《都市移动族》

当今城市被通讯数据所充斥，每个打电话发短信的人都生成特定时间地点的数据包，然而这些数据中有何规律？2008年，法国faberNovel公司对巴黎国际音乐节和新年夜产生的手机数据进行监测和可视化，帮助法国电信运营商Orange建立《都市移动族Urban Mobs》(图3)[5]。它不仅让我们发现城市活动中丰富的一面，同时也使电信运营商在流量分析、业务推荐等方面获得启示。

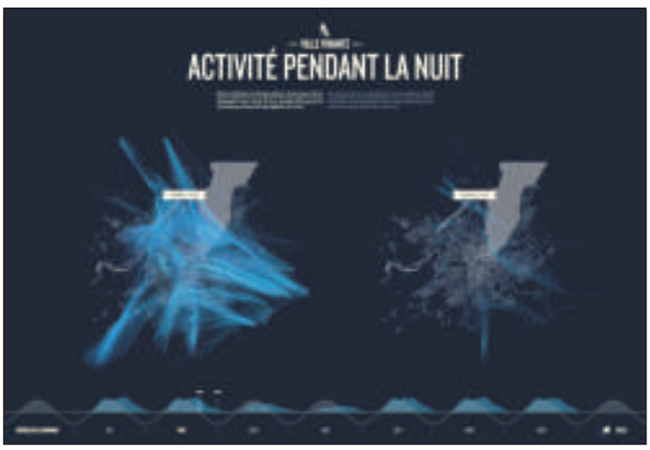
(2)电信数据可视化——《活力日内瓦》

手机可以看作是实时记录并上传用户地理位置信息的移动传感器，2011年，瑞士日内瓦市政府与Interactive Things公司合作，将市民每天在日内瓦市的行动轨迹的手机GPS数据进行记录，并制作城市生活(Ville Vivante)[6]动态显示瑞士电信每时每刻的数据流向。图4展示的是晚上六点到午夜之间人们移动的轨迹。这种融合基于位置的服务(LBS)和电信数据的可视化方式不仅使政府和公众对城市生活有了重新认识，同时也产生不可估量的经济政治效益。

(3)智慧城市——《实时新加坡》

现代城市中每天都在产生海量的数据，如何才能让政府和市民更快了解城市每时每刻的变化，帮助政府提高管理效能，为市民提供生活便利？2011年，美国麻省理工大学可感知实验室为新加坡建立了"LIVESingapore"实时新加坡平台[7](图5)，该平台能够为公众提供实时的城市活动及环境信息。其中，"实时通讯"显示新加坡语音通讯、短信及网络使用情况，"等时地图"实时呈

◀图3
都市移动族 Urban Mobs



◀图4
活力热内瓦

现新加坡居民交通耗时情况;“雨天打车”结合降雨监测和出租车数据进行可视化,从而在雨天智能调配出租车;“城市热岛”将新加坡区域温度与能源消耗的关系进行可视化。通过对城市生活、环境数据的可视化,可助力提高城市公共服务质量,改善市民生活,真正意义上实现智慧城市。

(4)北京大学 PKUVIS 微博可视分析工具

结合社会网络分析中的概念和可视化的呈现方法,佐以统计和智能数据挖掘的方法,可以为海量复杂社会网络提供快速、直观和智能的分析和呈现方法[8]。2012 年北京大学可视化与可视分析小组开发了支持可视化浏览和分析微博热点事件的“PKUVIS 微博可视分析工具”(图 6)[9]。该工具将一条条独立的微博连接起来,通过直观的视图清晰地呈现出一个事件中微博转发的过程,从而让用户能够迅速地发现事件中的关键人物、关键微博、重要观点,同时通过可视化的方式可以更好地分析新浪微博传播脉络以及事件的发生与发展的过程。

(5)电力大数据可视化

美国 Space-Time 是一家提供新一代地理空间和可视化解决方案的创业公司,2011 年,Space-Time 为美国加州独立系统运营商设计了一套可以实时监控电力传输系统能源基础设施的可视化软件 Space-Time Insight(图 7)[10],该可视化系统通过控制室中的一个 80 英寸的显示屏,在地图上实时展示长达 25 000 km 的输电线路状况,工作人员一旦发现一个地区出现了问题,就可以根据该地区问题的严重性和临近地区的反应来做出决策。不仅简化了日常运营复杂度,还在尽可能降低影响的情况下解决问题。这种大数据可视化实践对中国的电力大数据分析展示乃至整个能源相关行业都具有巨大的参考价值。

6.3 开源可视化工具

如果读者已经被以上可视化案例所吸引,并且愿意尝试将企业数据进行可视化,那么开源的数据可视化编程语言和环境将会是不错的选择。可视化领域中重要而常用的可视化编程语言和环境有 Processing、Processing.js、R、D3、Impure、ParaView、Circos 等,它们具备的一个共同特点就是为用户提供了常见的专业可视化模版或图形库,用户可以通过简单调用即可很快实现可视化效果,此外,由于软件的开源优势,专业用户可以根据其需求,对图形源代码进行定制修改。

在可见的未来,大数据可视化机

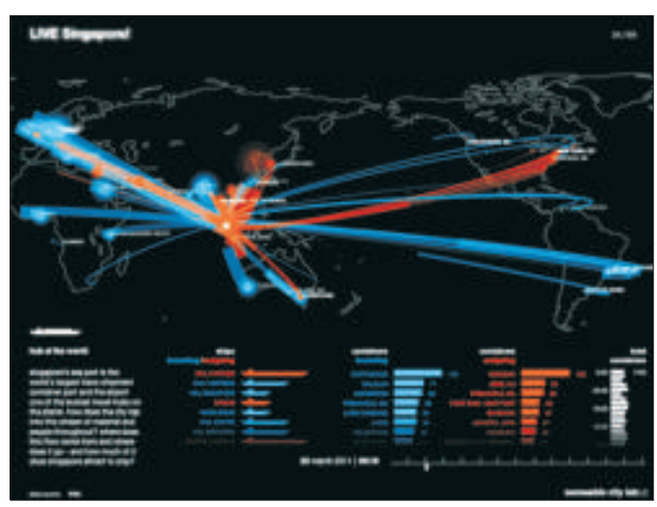


图5▶
实时新加坡

◀图6 北京大学PKUVIS微博可视分析工具微博可视化

◀图7 Space-Time Insight电力大数据可视化

遇挑战并存[11]，大数据可视化将越来越广泛地为各领域所使用，也将引发新一轮的投资热潮，而构建面向电子政务、电信、电力等特定行业大数据的可视分析工具是一个可以深入探索的重要发展方向。 （待续）

参考文献

[1] Deriving new business insights with big data [EB/OL]. http://www.ibm.com/developerworks/library/os-bigdata
[2] Big data landscape v2.0[EB/OL]. http://www.ongridventures.com/2012/10/23/the-big-data-landscape/
[3] 俞宏峰. 大规模科学可视化[J].中国计算机学会通讯,2012, 8(9): 29-37.
[4] STEELE J, ILIINSKY N. Beautiful visualization [M]. Sebastopol, CA,USA : O'Reilly Media, 2010.
[5] Urban Mobs[EB/OL]. http://www.urbanmobs.fr/en/
[6] Ville Vivante[EB/OL].http://www.villevivante.ch/
[7] LIVE Singapore[EB/OL].http://www.live-singapore.com.sg/
[8] 袁晓如,张昕,肖何等.可视化研究前沿及展望[J].科研信息化技术与应用,2011, 2(4):3-13.
[9] PKUVIS微博可视分析工具[EB/OL].http://vis.pku.edu.cn/weibova/weiboevents/
[10] Space-time insight[EB/OL]. http://www.spacetimeinsight.com/
[11] 黄伯仲,沈汉威,克里斯托弗·约翰逊等.超大规模数据可视分析十大挑战[J].中国计算机学会通讯,2012, 8(9): 38-43.

作者简介

韩晶，北京邮电大学计算机学院在读博士；主要研究方向为大数据管理；已发表论文10余篇，申请专利和软著9项。

宋美娜，北京邮电大学计算机学院教授；主要研究方向为分布式系统、服务计算、数据工程等；已发表论文50余篇，申请专利20余项。

综合信息

中兴通讯iCity智慧城市解决方案进入GSMA“全球移动大奖”短名单

【本刊讯】2013年2月21日消息，GSMA公布了第十八届“全球移动大奖”的入围短名单，中兴通讯的iCity智慧城市解决方案在600多个提案中脱颖而出，进入“无线智慧城市最佳创新”奖项的短名单，也是进入该短名单唯一的智慧城市综合解决方案。

GSMA“全球移动大奖”是移动行业领域的最高奖项，由160多名评委共同打分评定，包括行业分析师、媒体、学者以及14家运营商的技术执行官。这是中兴通讯面向政府和企业的解决方案第一次进入该类奖项短名单。

中兴通讯iCity智慧城市解决方案涵盖“维稳定、促增长、保民生”三大领域，包含电子政务、智慧交通、平安城市、智慧教育等12项重点应用，首次在业界提出包含“信息(Information)”、“智能(Intelligent)”、“创新(Innovation)”以及“我和城市(I with City)”的“4I智慧城市理念”，并引入云计算平台。中兴通讯此次能够进入短名单，是中兴通讯产品实力的最佳体现。中兴通讯希望能够创建更加智慧的城市，助力经济发展。