



IPv6 过渡技术规模部署实践

郭泓伟, 张国贤

(中国电信股份有限公司无锡分公司 无锡 214001)

摘要:为应对 IPv4 地址枯竭,无锡城域网在规模部署双栈的基础上引入 NAT444、DS-Lite 等 IPv6 过渡技术,通过网络改造、溯源机制创新和流程优化实现 10 万双栈用户规模迁移。描述了 IPv6 的过渡背景,对 IPv6 过渡技术进行了概述,介绍了过渡技术规模部署的方案和情况。

关键词:下一代互联网;IPv6 过渡技术;NAT444;DS-Lite

doi: 10.3969/j.issn.1000-0801.2013.09.036

Scale Deployment Practice on IPv6 Transition Mechanism

Guo Hongwei, Zhang Guoxian

(Wuxi Branch of China Telecom Co., Ltd., Wuxi 214001, China)

Abstract: According to the situation with the exhaustion of IPv4 addresses, Wuxi Telecom has carried on scale deployment of IPv6 transition mechanism(NAT444、DS-Lite) on metropolitan network 100 000 broadband users have been migrated. The background of IPv6 transition mechanism was described, IPv6 transition technology was summarized, and scale deployment scheme and situation of transition technology were introduced.

Key words: next generation internet, IPv6 transition mechanism, NAT444, DS-Lite

1 IPv6 过渡背景

随着移动互联网、物联网、云计算、三网融合等互联网新业务的快速发展,IPv4 的地址空间已无法满足业务规模增长需求,仅以无锡城域网固网宽带业务为例,预计 2014 年底 IPv4 地址缺口将达到 24.3 万个。IPv6 技术由于具有海量的地址空间、简洁的数据分组头、“即插即用”的地址自动配置方式、更高的安全性和移动性支持等特点,将成为替代 IPv4 技术的必然选择。然而 IPv6 在设计时没有平滑过渡方面的考虑,IPv4 与 IPv6 网络之间在编址、报文转发、服务应用等方面均互不兼容、无法直接通信,现网大量应用系统在 IPv6 网络中无法使用。由于 IPv4 应用和业务的广泛性,特别是相关网络和应用的长尾效应,使得 IPv4 和 IPv6 网络将在较长时间内共存。IPv4 向 IPv6 演进包括双栈、隧道、协议转换等多种技术方案,综合考量用户感

知、部署成本、流量模型、技术成熟度等因素,双栈技术是现阶段运营商级网络向 IPv6 演进的优选方案。双栈技术方案为保证用户能够继续沿用原有的 IPv4 业务必然引起 IPv4 地址的持续消耗,不能解决 IPv4 地址短缺问题,城域网在规模部署双栈技术的基础上引入 IPv6 过渡技术,是运营商应对 IPv4 地址枯竭、向 IPv6 网络演进的必由之路。

2 IPv6 过渡技术概述

IPv6 过渡技术是在网络层面部署 IPv4 和 IPv6 双协议栈的基础上,业务层面引入 IPv6 地址和私网 IPv4 地址,通过 NAT(network address translation,网络地址转换)技术实现公网 IPv4 地址的复用,提高公网 IPv4 地址的利用率。目前无锡城域网固网宽带业务占用 IPv4 地址总量达到 85%,规模引入过渡技术能够有效缓解 IPv4 地址短缺问题。

现阶段可规模商用的运营商级过渡技术方案分为

NAT444 和 DS-Lite 两类。NAT444 方案在 IPv4 业务实现上采用 NAT 技术实现公网地址之间的映射转换,业务控制层设备为业务终端同时分配 IPv6 地址和私网 IPv4 地址;DS-Lite 方案在 IPv4 业务实现上采用 IPv4 over IPv6 隧道叠加 NAT 技术实现 IPv6 网络承载 IPv4 业务,业务控制层设备只为业务终端分配 IPv6 地址。

2.1 NAT444

NAT444 方案中业务控制层设备为宽带用户同时分配 IPv6 地址和私网 IPv4 地址,统一在运营商侧部署 NAT 设备——CGN(carrier-grade NAT)。私网 IPv4 业务通过 CGN 访问公网 IPv4 地址资源,CGN 实现公网 IPv4 地址之间的映射转换;IPv6 业务直接访问 IPv6 地址资源无需经过 CGN。桥接型 CPE 用户在 CGN 上实现一次 NAT;路由型 CPE 用户在 CPE 上实现第一次 NAT、在 CGN 上实现第二次 NAT。NAT444 业务实现模型如图 1 所示。

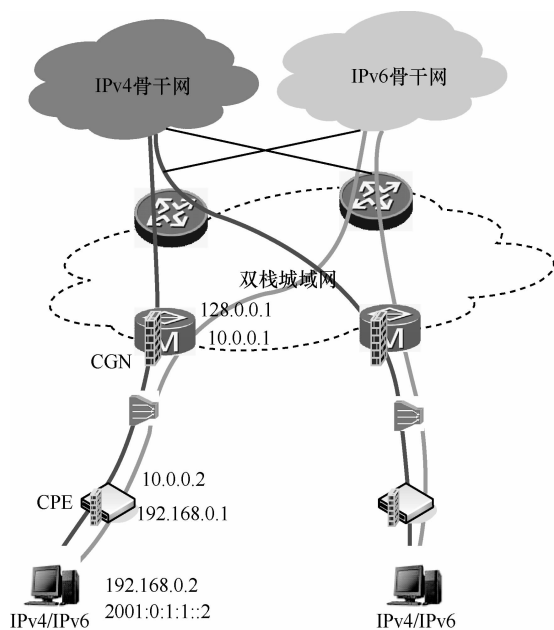


图 1 NAT444 业务实现模型

2.2 DS-Lite

DS-Lite 方案中,业务控制层设备为宽带用户只分配 IPv6 地址,具备 DS-Lite 功能的路由型 CPE 为用户终端分配私网 IPv4 地址和 IPv6 地址,CPE 与 AFTR(address family transition router) 建立基于 IPv6 的 DS-Lite 隧道,CPE 将用户私网 IPv4 流量通过 DS-Lite 隧道转发到 AFTR 设备,AFTR 设备负责终结隧道,完成 IPv4 报文的 NAT。DS-Lite 业务实现模型如图 2 所示。

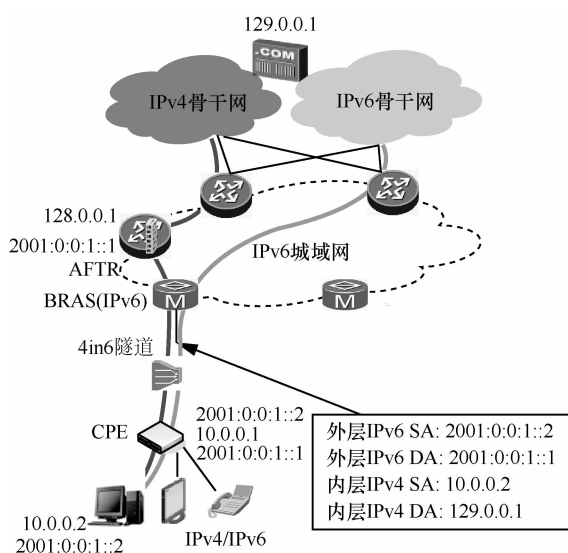


图 2 DS-Lite 业务实现模型

3 过渡技术规模部署

作为中国电信集团公司下一代互联网现场试验本地网之一,无锡城域网在公网双栈规模部署的基础上,从 2011 年开始以过渡技术为重点对固网宽带拨号业务进行规模改造,从网络、业务、运营支撑、终端等层面对标现有城域网的网络运营和业务提供体系,通过网络和支撑系统改造,全面支持 IPv6 和 NAT444、DS-Lite 过渡技术的端到端网络及业务部署。结合两种过渡技术的特点,对批量改造用户采用 NAT444 方案、对新装和纤改用户采用 DS-Lite 方案,进行过渡技术改造,截至 2012 年底,无锡城域网具备了 20 万 IPv6 用户的网络覆盖能力,基于过渡技术部署完成了 10 万双栈用户的改造。

3.1 NAT444 和 DS-Lite 方案现网部署

双栈技术是互联网向 IPv6 演进的基础,其 IPv4 的业务实现仍然采用公网 IPv4 地址,为解决 IPv4 地址短缺的难题,必须引入过渡技术,实现公网 IPv4 地址的复用。由于城域网部署 NAT444 和 DS-Lite 方案将在现网引入私网 IPv4 地址,现网规模部署需重点关注以下因素。

- CGN/AFTR 设备公网地址端口分配方式分为动态分配和预分配两类,预分配端口方式由于有效降低了业务溯源的信息量,成为运营商规模部署过渡技术的基础。
- 过渡期间,一台 BRAS (broadband remote access server, 宽带接入服务器) 承载用户 IPv6、IPv4 公网并存,为保障用户在业务迁移改造过程中无感



知,需不改变用户使用习惯,即用户仍然通过原有账号拨号上网,通过 AAA 服务器控制 BRAS 为用户分配的地址属性。

- 业务溯源的实现是过渡技术方案规模部署的关键因素。传统 NAT 的溯源采用 Syslog 方式,CGN 将地址映射信息(私网地址和映射公网地址+端口段)通过 Syslog 协议发送日志服务器。根据运营业务溯源的要求,日志的输出要保证实时性,但 Syslog 协议不是一个高效率的协议,难以保证实时性。同时城域网部署 Syslog 方式需新增日志服务器,增加了业务溯源实现的复杂度;AAA 服务器除了提供查询界面,还需要支持和日志服务器之间的公私网映射关系表查询接口,业务溯源时需完成 2 次查询。一是通过公网地址+端口查询 AAA 服务器用户信息;二是 AAA 服务器向日志服务器查询地址映射信息。因此 Syslog 方式不适用于城域网规模部署 NAT444 和 DS-Lite,RADIUS 动态上报和静态算法是 2 种运营商级的溯源方式。
- CGN/AFTR 网关是 NAT444 和 DS-Lite 方案的关键设备,基于 CGN/AFTR 网关部署的网络层次分为集中式和分布式,在网络部署和业务溯源机制等方面差别较大。

3.1.1 端口分配方式

NAT 技术的应用将出现多个用户共享一个公有 IPv4 地址的情况。为维护每个用户的公平性,必须对用户使用的公有地址端口数量进行合理分配,规定每个用户使用端口的数量和范围,确保用户之间不会互相影响,同时实现溯源。目前 CGN 设备有如下两种端口分配方式。

- 动态为每一个穿过设备的连接分配一个端口,并记录此映射关系的信息。用户访问互联网的每个连接均会产生 1 条日志记录。动态的端口分配方式由于 CGN 会生成大量的日志信息,并基于溯源的需求日志信息上传时日志服务器将占用大量网络资源,因此在运营商规模部署中不适用。
- 基于用户预分配端口块,在用户的第一个连接穿过 CGN 设备时,为该用户预分配一可用端口段,产生 1 条日志记录,后续其他连接均使用该端口段,不产生新的日志记录。该方式有效降低了日志信息量,适用于运营商规模部署。然而当用户耗尽已分配的端口块后,新的连接无法建立,这将影响

用户的业务体验。为减少对用户业务的影响,结合网络安全因素(防止少量用户占用大量 CGN 设备资源),需根据现网业务使用情况确定合理的私网和公网映射参数。无锡城域网部署 NAT444 和 DS-Lite 方案时,基于现网用户业务端口占用情况(现网业务端口使用情况分析详见第 3.2 节),确定以下端口分配模型:每个私网地址预分配 4 096 个端口、公网地址的前 4 096 个端口保留(规避常用服务端口范围),私网与公网之间的映射比为 12:1~15:1。

3.1.2 AAA 服务器控制用户属性

BRAS 将用户名、密码等信息上送到 AAA 服务器进行接入认证,AAA 服务器通过 RADIUS 认证回应 access-accept 报文并下发 RADIUS 用户扩展属性值(域名),双栈 BRAS 根据用户域名将绑定到不同属性地址池(IPv6 地址、公网/私网 IPv4 地址)。BRAS 为用户分配 IPv4 地址(公网或私网)和 IPv6 地址。私网地址转换由 CGN 完成。AAA 服务器根据用户属性向 BRAS 下发用户地址类型域,实现 BRAS 灵活分配 IP 地址,PPPoE 上网用户拨号无需添加域名,不改变现有习惯。AAA 服务器可针对不同厂商 BRAS 设备通过不同 RADIUS 属性下发用户地址类型域。

3.1.3 分布式部署方案

分布式部署方案中 CGN 和 BRAS 集成,采用 RADIUS 动态上报方式溯源。

在 BRAS 上配置 2 块具备 NAT 功能的 CGN 单板,CGN 单板之间负载均衡、互为主备。BRAS 采用 20 Gbit/s 以上平台设备,CGN 采用 20 Gbit/s 吞吐量、600 万并发 session 的板卡。每 BRAS 上 1 块 CGN 可以满足用户的 NAT 要求,考虑可靠性,每个 BRAS 插 2 块 CGN 单板实现流量的主备备份和负载分担。正常情况下,IPv4 私网流量在 2 块 CGN 单板上负载分担,1 块 CGN 板卡出现故障,流量将快速切换到另 1 块 CGN 板卡上。CGN 板卡之间的备份方式分为温备份和热备份,在温备份方式下,CGN 板卡故障切换时用户 session 中断,在热备份方式下,2 块 CGN 板卡实时同步用户 session,因此 CGN 板卡故障切换时用户 session 仍然保持。由于热备份方式需要占用双倍的板卡资源,通过业务测试分析,温备份已能够满足高速上网业务需求。

引入私网双栈技术导致公私网一对多的网络地址端口转换(network address port translation,NAPT),解决用户溯源问题是过渡技术规模应用的关键因素。分布式部署方

案采用“RADIUS 动态上报溯源方案”。

用户上线时,当 BRAS 收到 AAA 服务器发来的 access-accept 报文后,为终端用户分配 IPv4 私网地址,并预留针对此地址做 NAT 后的源地址及源端口块范围。BRAS 通过扩展 RADIUS 属性,在 accounting-request/update(计费开始、更新)消息中上报用户地址信息,AAA 服务器获得用户地址(私网 IPv4、IPv6)、转换后的公有地址、端口范围等信息通过私有属性封装在 accounting-request 报文中上送 AAA 服务器,AAA 服务器同步扩展支持私有属性,维持与用户信息的对应关系实现“RADIUS 动态上报溯源”方案。

3.1.4 集中式部署方案

集中式部署方案采用独立 CGN 设备旁挂核心路由器(core router,CR),采用静态算法方式溯源。

新建 4 台独立 CGN 设备旁挂 CR 部署在城域网出口,CGN 与 CR 全双归互联。升级现网 BRAS 支持双栈,AAA 服务器通过 RADIUS 认证回应 access-accept 报文下发 RADIUS 用户属性(域名),BRAS 为用户分配 IPv4 地址(公网或私网)和 IPv6 地址。CR 部署策略路由,将私网流量牵引到 CGN 设备,私网地址转换由 CGN 设备完成。采用链路聚合将多条上行和下行链路捆绑,通过逻辑链路区分 CGN 上行和下行链路,保证上下行链路状态联动,提高可靠性。由于同一私网用户通过不同 CGN 进行 NAT 转换可能对信令流和媒体流分离的应用如 RTSP、PPTP、FTP 等产生影响,因此 CR 策略路由部署需确保在 CGN 设备无障碍的情况下,同一私网用户通过相同 CGN 进行 NAT。

与分布式部署方案不同,由于 CGN 不与 AAA 服务器交互,集中式部署方案采用静态算法方式溯源。AAA 服务

器和 CGN 采用一致的私网地址和转换后的公网地址+端口段的映射算法,综合网管(IPOSS)通过 SNMP 向 AAA 服务器和 CGN 统一下发映射算法的参数。

3.1.5 分布式和集中式 NAT444 部署场景

分布式和集中式 NAT444 部署场景分析如表 1 所示。

3.2 城域网规模部署情况

3.2.1 批量迁移

无锡城域网 5 台业务量密集的 BRAS 通过各配置 2 块 CGN 板卡部署分布式方案,IT/IP 支撑系统在新装、移机、批量割接、私改公等流程完成适配过渡技术(包括 NAT444 和 DS-Lite)改造,目前已规模迁移 10 万双栈用户。迁移双栈用户以批量割接为主,由于采用 AAA 服务器控制用户地址属性,因此用户迁移时只需在 IT/IP 支撑系统批量更改用户地址属性,对在线用户无影响,用户下线后再次拨号,将获取私网地址,整个迁移过程用户无感知。

3.2.2 障碍分析

根据现网规模部署过渡技术后的用户障碍情况分析,固网拨号业务引入私网 IPv4 地址对主流应用,包括即时通信、网页浏览、在线游戏、下载、视频、E-mail、网上购物、网上银行、证券交易、用户自建 VPN/VPDN 等均不会产生影响,目前障碍发生率约为 0.3%,产生影响的业务包括以下内容:

- 需要公网对拨号终端的访问,如自建服务器、远程监控、远程终端管理等业务;
- 用户的特殊应用。

城域网在过渡技术流程设计时,已考虑私网引入对部分业务的影响,制定过渡技术(NAT444 和 DS-Lite)改公网双栈的回退流程:将 IT/IP 相关支撑系统内用户地

表 1 分布式和集中式 NAT444 部署场景分析

部署方式	集中式部署	分布式部署
总体成本	部署节点少,相对集中,但需要新部署 CGN 设备	仅占用 BRAS 槽位,不需要单独提供接口、电源,容易部署
用户管理	CGN 部署位置较高,无法获得用户信息,用户策略控制和溯源较复杂	CGN 和 BRAS 集成,溯源方案简单,便于实现基于用户的精确化策略控制
业务流量	同城本地化流量将绕到城域网核心路由器和 CGN 处理;CR 需要部署策略路由,配置复杂;私网地址需要在全网宣告	流量模型无变化;私网地址无需全网宣告
可靠性	CGN 需要维护大量 session,单点故障影响范围大,对 CGN 设备可靠性要求高,组网复杂	CGN 需要维护 session 较少,单点故障影响范围小
应用场景	适用于私网用户业务量小、分散分布、BRAS 不支持 CGN 插卡的区域	适用于私网用户业务量大、密集分布的区域



址属性改为公网双栈,用户下线后再次拨号,将获取公网 IPv4 地址。

3.2.3 私网用户统计与分析

无锡电信城域网已规模部署 10 万双栈用户,业务类型为 PPPoE 拨号上网。通过对并发用户数、并发 session 数、端口占用数实时监控分析,可了解用户的业务特征、CGN 板卡资源占用情况,适时调整 NAT 映射相关参数(分配端口块、可用 session 数等)和 CGN 承载用户数,避免私网引入对用户业务的影响。

以下选择梁溪路局点 Me60 进行用户统计分析,目前梁溪路 Me60 承载私网双栈用户数 2 万,每个用户(PPPoE session)预分配 4 096 个端口,已空出公网 IPv4 地址 42 个 CDMA 网段(约 1.05 万个地址)。

(1)并发用户数量变化趋势

并发用户数量变化趋势如图 3 所示。

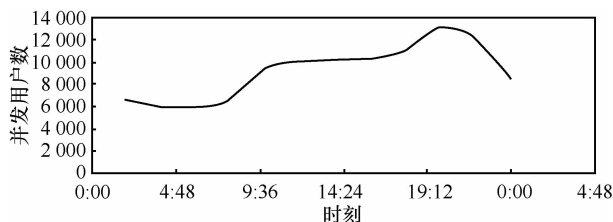


图 3 并发用户数变化趋势

(2)用户总 session 数变化趋势

用户总 session 数变化趋势如图 4 所示。

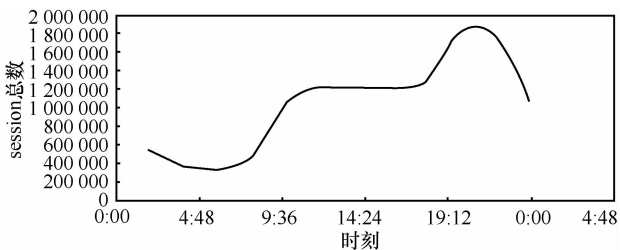


图 4 并发 session 数变化趋势

(3)平均每个用户 session 数变化趋势

平均每个用户 session 数变化趋势如图 5 所示。

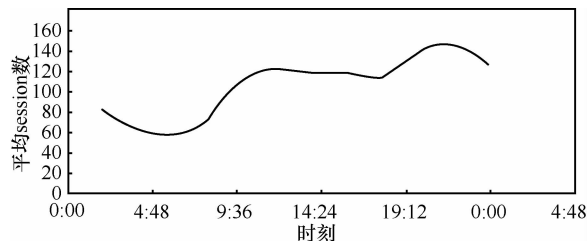


图 5 用户平均 session 数变化趋势

(4)session 数用户分布统计

session 数用户分布统计如图 6 所示。

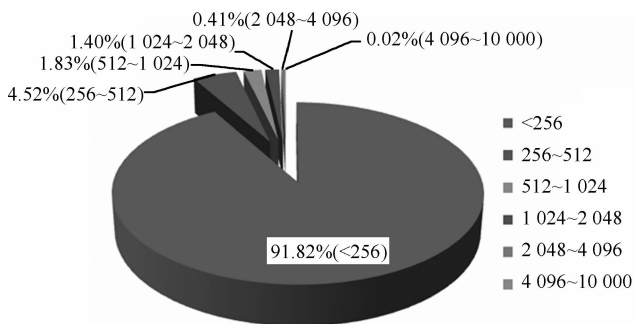


图 6 session 数用户分布统计

(5)实用端口数用户分布统计

实用端口数用户分布统计如图 7 所示。

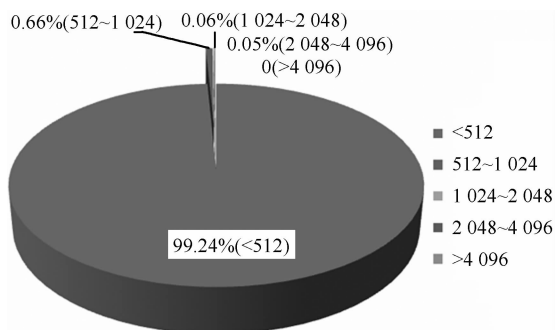


图 7 实用端口数用户分布统计

对图 7 分析可知。

- 私网用户 session 总数与并发用户数的变化趋势相同,随着并发用户数的增加,CGN 处理的 session 总数也相应增长,session 总数在晚忙时达到峰值;目前峰值 session 数约为 140 万,远低于单块 CGN 板卡 600 万的性能指标。
- 平均每个用户 session 数与并发用户数的变化趋势基本一致,随着并发用户数的增加,平均每个用户使用的 session 数也相应增长。
- session 数小于 256 的用户占 91.82%、达到 4 096 的用户占 0.02%,在 BRAS 将每个用户 session 数上限配置为 4 096,在保障 CGN 设备安全的前提下,能确保绝大部分用户使用私网地址后业务体验不降低。
- 实际端口使用数小于 512 的用户占 99.24%,BRAS 为每个用户预留 4 096 个端口能够满足私网用户的业务需求。

(下转第 190 页)