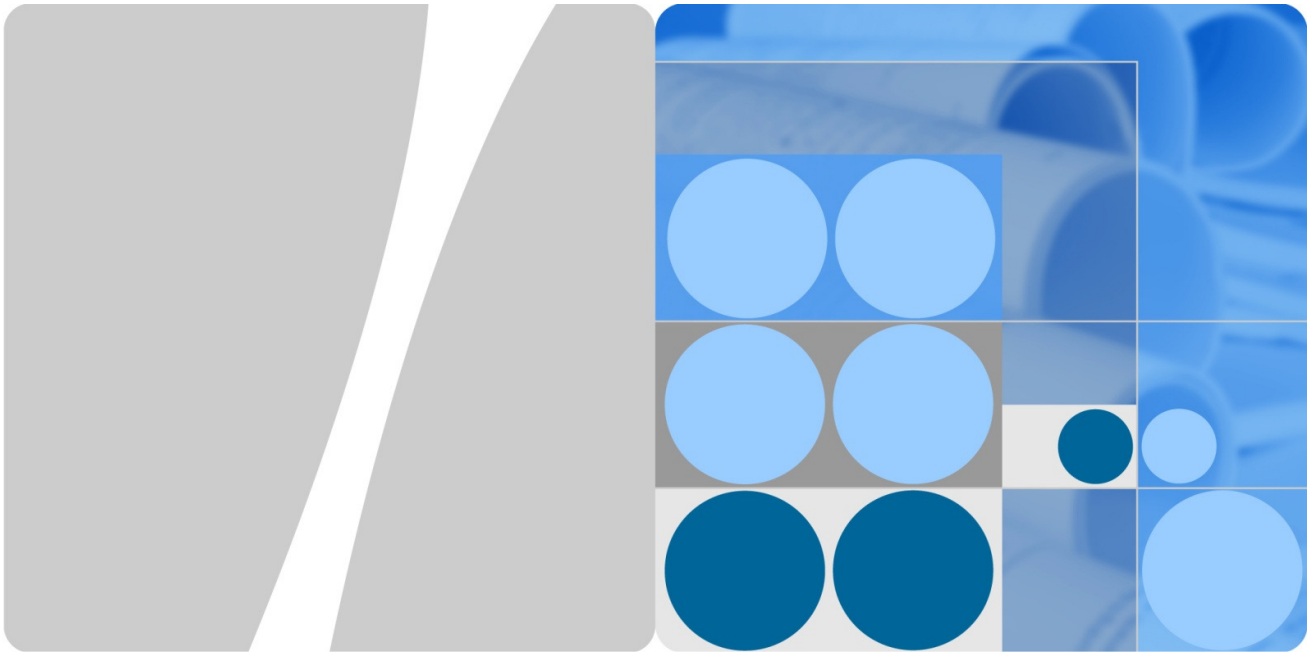

IPv6 网络端到端演进技术白皮书

文档版本　01

发布日期　2011-09-30

华为技术有限公司

商标声明


和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： http://www.huawei.com

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目 录

附图目录

附表目录

IPv6 网络端到端演进技术白皮书

关键词： IPv6, Dual Stack,双栈, CGN, NAT, NAT444, NAT44, NAT64, DS-Lite, BNG/CGN融合

摘　要：

随着 IPv4 地址短缺问题日益加剧，从 IPv4 网络向 IPv6 迁移已经成为运营商当前的工作重点。虽然 IPv6 是解决地址短缺的终极方案，但由于业务延续性的需要，IPv4 网络将继续保留相当长时间。在网络向 IPv6 演进的过程中，部署 NAT 将无可回避，从而形成 IPv4 私网地址、IPv4 公网地址和 IPv6 地址共存的复杂网络。

本文详细分析了这种复杂性产生的根源，描述了网络演进中的关键技术，给出了如何顺利进行 IPv6 早期部署、应对 IPv4 地址短缺、平滑过渡到 IPv6 网络的建议，并简要分析了主流运营商的 IPv6 演进路线选择及其考虑因素。

1 前言

1.1 IP 网络面临的两大问题

IP 网络的发展是由业务和应用驱动的，IP 协议要求每个网络终端都具有唯一的可寻址的 IP 地址。IANA 已于 2011 年 2 月将最后的 5 个 A 类 IPv4 地址段一次性地分配完毕，以后获取可用的 IPv4 地址将非常困难。现阶段 IP 网络面临两大问题：

新业务需要大量 IP 地址

移动终端和物联网等海量终端应用是 IP 地址需求的最大来源。现有的 IPv4 地址 IPv4 只能提供 43 亿地址，其中可用的约 30 亿，地球上平均每人分配不到一个地址，再加上历史原因造成的地址分配不均衡，部分运营商地址短缺的问题很严重，在开展业务时只能使用私网地址（RFC1918）。

业界已经提出了 500 亿连接的概念：全球家庭宽带网络约 20 亿终端，移动网络约 50 亿终端，估计到 2020 年全世界将有 500 亿个物联网终端通过网络互相连接。有激进观点甚至认为 500 亿的估计太保守。现有 IPv4 网络绝对不能满足 500 亿终端的要求。

IPv6 与 IPv4 完全不兼容

解决地址短缺问题是 IPv6 发展的最大动力。IPv6 的核心协议标准自 2008 年以来并没有太多变化，可以满足 IPv6 组网要求，固定和移动宽带应用类标准也已基本完善。目前最大的问题在于 IPv6 与 IPv4 完全不兼容，现有的绝大多数应用程序在只有 IPv6 协议栈的计算机上甚至无法运行。程序初始化时要调用 socket，IPv6 与 IPv4 的调用参数不同，如果程序没有考虑到 IPv6 参数调用，则会启动失败。如要在 IPv6 协议中运行现有的应用，则需要进行源代码修改、重新编译和安装。

1.2 应对方案

解决地址短缺问题有两个途径，IPv6 和 NAT，但这两个途径都有各自缺陷。

NAT 方案

NAT 方案通过引入私网 IPv4 地址和地址转换技术解决地址短缺问题，本质上是一种多用户共享公网地址的方案。该方案基本上不存在应用互通问题，已被广泛商用部署，但是不能支持大量 IP 地址（每网络约增加一个 A 类地址）。

目前支持 NAT 的应用程序多数是 Client-Server 结构，由 Client 端主动发起业务连接。NAT 完全支持这种基于单向简单连接的应用。如果应用的连接关系比较复杂，则需要 NAT 支持该应用的 ALG。P2P 应用则需要应用程序与 NAT 协作（例如家庭网关支持的 UPnP），由于性能和安全问题这种协作在运营商 NAT 环境下难以被支持。

最重要的是，对于大型网络私网地址也仍然是有限的，NAT 只能缓解地址的短缺问题，最终仍然需向 IPv6 迁移。

未来网络单纯依赖 NAT 来解决地址短缺是不可行的。

IPv6 方案

IPv6 可提供大量的 IPv6 地址，从而彻底解决地址短缺的问题。但是如果部署一个全新的纯 IPv6 网络，现有的绝大多数基于 IPv4 的应用将无法在此 IPv6 网络上运行。运行在纯 IPv6 网络上的应用与 IPv4 互通在现阶段相当困难。

现阶段单靠 NAT64 设备进行网络层翻译处理在应用部署上仍然存在一些问题。随着时间推移，越来越多的 IPv6 程序将会采用类似 STUN/ICE 之类的技术在应用层面来处理与 IPv4 网络互通的问题，由应用程序自身来感知 NAT64 的存在并获得相关地址转换信息，以进行相应的互通处理。

现有网络依赖纯 IPv6 来解决地址短缺是不可行的。

综合上面的分析，运营商网络同时支持 IPv4 和 IPv6 双协议栈在相当长的时间内是必然的，而且若干年内必然是 IPv4 公网、IPv6 网络、IPv4 私网并存，与现有 IPv4 网络相比复杂度将大大增加，对网络建设、运维将产生重要影响。

2 部署 NAT 对现有网络和业务的影响

2.1 NAT 对网络设备的影响

部署 NAT 并没有改变 IPv4 基础协议，现有的网络设备都不需因此做调整，部分应用系统需针对私网地址作配置修改。

需要在网络中增加地址转换功能，可以通过新建独立的 NAT 设备或在现有设备上增加支持 NAT 功能的模块实现。

2.2 NAT 对网络性能的影响

NAT 不仅仅增加了处理时延，还增加了网络和路由的复杂性，而且 NAT 本身是流量的汇聚点，每 session 的备份在运营商环境中难以实施，一旦发生故障可能需要终端用户操作干预以重新建立 session，从而降低了网络的可靠性。

NAT 还对网络的安全性能产生不利影响：某些应用如 IPSec 和 DNSsec 很难穿透 NAT；由于 NAT 的有状态特性，更加容易被作为 DOS 的攻击目标；NAT 上集成的 ALG 特性也往往是 NAT 的安全薄弱环节。这些缺陷降低了网络的安全性和业务提供能力。

2.3 NAT 对网络维护的影响

故障定位更加困难

IP 网络的故障定位本身就相对复杂一些，引入 NAT 后，原有的 Ping/Traceroute 对 NAT 不再有效，网络故障定位更加复杂，使得业务中断时间变长。

某些业务在 NAT 环境下可能不能正常应用。对从公网 IPv4 地址切换到私网 IPv4 地址的用户来说，可能出现原有的应用在 IP 地址切换后不能正常工作的情况。

部署 NAT 导致的法规遵从

大多数国家存在因特网溯源的法规要求，部署 NAT 需配置日志服务器，以记录用户的网络访问记录。如果对每个用户的每个会话都记录，将可能导致每 NAT 数十 MB/s

的日志流量，在这样大的日志流量下，需要高性能大存储的日志服务器，此外由于性能问题，日志记录丢失的可能性很大。

增加的 NAT 设备、故障定位难度增加、用户申告以及溯源的法规遵从导致网络维护难度和工作量增加。

2.4 NAT 对业务和应用的影响

除了增加维护难度以及降低网络的可用性、安全性和性能外，NAT 技术还对业务和应用产生不良影响：

降低用户业务体验

增加的处理时延降低应用性能；需要额外的信令交互穿越 NAT，导致应用的响应速度降低；一些应用将不同终端通过共享 IP 的访问当作攻击风险，要求输入额外的验证字符串导致用户操作复杂；可能由于某用户发送恶意邮件导致其他共享同一 IP 地址的用户也被列入黑名单；采用地址做定位的业务也将不再有效；降低 P2P 性能甚至导致某些功能不可用等等。

影响运营商业务

如果 NAT 部署于用户与 DPI 之间，DPI 功能将失效。

某些运营商增值业务需要基于 IP 地址获取用户信息，NAT 会对这类应用产生影响，因为现在地址为多用户共享；对 IP 语音通讯类的应用往往需要增加应用层网关。

增加应用层的成本

主要是增加了应用开发的成本，以及应用的运营维护成本。NAT 有多种不同的行为模式，应用程序要探测行为模式以及进行相应的处理，导致软件开发相当复杂；网站需要额外记录端口信息；很多应用程序都需要特定的中间服务器处理私网地址用户的相互访问。

阻塞某些应用

IPSec 和 DNSsec 很难穿透 NAT，基于此类技术的应用将无法工作；而且各厂商的 NAT 实现差异较大，同一厂商不同平台的 NAT 实现也往往存在差异，可能出现某些应用在某些型号 NAT 设备下能够运行而在另外某些型号 NAT 设备下不能运行的情况。

如果 BNG 为 CPE 分配私网 IPv4 地址，CPE 又为终端分配私网 IPv4 地址，业务将在 CPE 和 NAT 上进行两次地址转换。两次地址转换对多数应用没有影响，但是也存在少数应用在这种情况下不能正常工作。

尽管有这些不利影响，NAT 仍然被广泛商用部署。主要原因地址短缺已经大范围存在，NAT 是唯一可保护现有 IPv4 投资的技术。支持 NAT 已成为新开发应用的必选功能，且 NAT 相关标准不断发展完善，新开发的应用可以利用这些技术，通过增加应用层的复杂度规避网络层的限制。

3 部署 IPv6 对现有网络和业务的影响

当前的相关产业链中，路由器、交换机、终端操作系统对 IPv6 协议的支持较完善，CPE、接入网、SPOP、网管、应用业务系统的支持相对较弱。

目前部分电信类业务如 TV/Voice 迁移到 IPv6 还比较困难（不仅仅依赖于承载网络改造完成，还需要更换终端和业务服务系统），成本和成熟度不佳，现阶段进行规模商业部署应慎重。

3.1 IPv6 对网络设备的影响

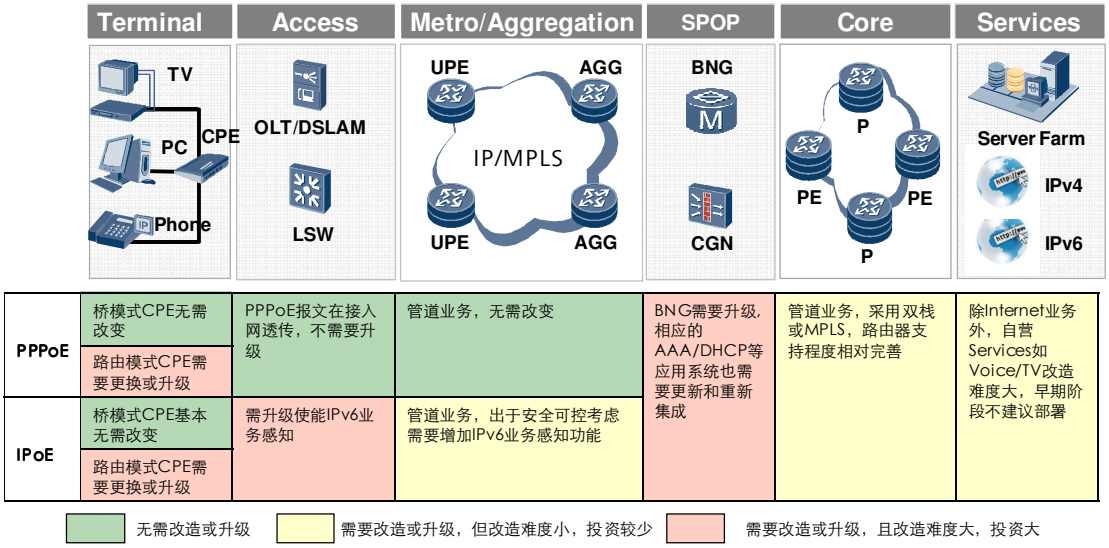

图1 部署 IPv6 对网络设备的影响

IP 骨干和城域汇聚

骨干网和汇聚对 IPv6 业务主要是管道承载，路由器支持 IPv6 相对成熟，通过升级或使能 IPv6 特性即可。

城域汇聚/交换机作为二层透传的管道。通过 PPPoE 承载的 IPv6 业务可以透传，但通过 IPoE 承载 Native IPv6 时，有些协议报文将作为未知组播处理，带来安全风险，需要类似 DHCPv6 Snooping/MLD Snooping/ND Snooping 之类的增强特性，从而导致升级要求。

SPOP

SPOP 设备主要包括 BRAS/SR/AAA/DHCP/DNS，一般需要升级。引入 IPv6 对计费、在线查询、Radius 会增加新的接口或协议，需要 BRAS 厂商、计费业务厂商重新进行开发和集成测试。

接入

DSLAM/OLT/ONU/MxU/交换机作为二层透传的管道。通过 PPPoE 承载的 IPv6 业务可以透传，但是通过 IPoE 承载 Native IPv6，有些协议报文将作为未知组播处理，带来安全风险，因此需要类似 DHCPv6 Snooping/MLD Snooping/ND Snooping 之类的增强特性，从而导致升级要求。

接入层设备的海量部署特征使得升级或更换成本相对较高。

CPE

桥接型 CPE 可以支持 IPv6 透传；路由型 CPE 往往需要更换以支持 IPv6。

在 IPv6 的早期部署阶段和规模商用阶段,采用的解决方案可能存在很大的差异。比如，早期试验阶段为快速部署 IPv6 可能采用 6RD 方案，商用阶段采用双栈方案，地址不足可能转向 DS-Lite 方案，这就需要 CPE 能够通过远程修改配置或远程软件升级灵活支持不同的技术方案。

终端的海量部署特征导致升级或更换成本很高。

网管和 OSS

网管和 OSS 系统要升级以处理 IPv6 相关 MIB 和用户管理接口，一般需进行软件升级和集成验证。

其他

DNS 支持双栈比较容易，只需要更改配置，老版本 DNS 可能需要软件升级。

不同厂商的防火墙、负载分担设备、DPI 对 IPv6 支持程度有较大差异，目前比较少见成熟商用系统。这些设备在数据中心内被广泛应用。

3.2 IPv6 对网络性能的影响

在路由器上开启 IPv6 双栈对高性能路由器的影响非常小。可能降低某些设备的可用资源，比如 BRAS 在线用户数，AAA 系统处理能力等，某些负载较重的场景下需要仔细评估。

部署 IPv6 一般对网络端到端时延抖动丢包率没有影响，在现有的 IPv6 路由数目下对故障收敛性能影响很小。

3.3 IPv6 对网络维护的影响

开启 IPv6 会增加网络的维护工作量和技能要求，但 IPv6 对网络维护的冲击比较小，具备维护 IPv4 能力的工程师可以在较短时间内掌握 IPv6。

早期部署阶段需要进行所有系统的集成、验证、部署，后续存在新业务开发、业务发放等其他维护工作，这是新技术引入的一般规律。

3.4 IPv6 对业务和应用的影响

在 IP 网络上部署 IPv6 对现有网络业务和应用基本没有影响，用户可以额外获得访问 IPv6 资源的能力。部署 IPv6 往往需要调整 DNS 等业务系统，不正确的配置或有缺陷的软件将影响用户体验。

如果用户开启 IPv6，但没有到 IPv6 互联网的可靠连接，则可能产生业务质量下降，这是由于某些应用或操作系统优选 IPv6 造成的。一个例子是 Windows Vista（SP1 之前版本），操作系统只要存在双栈则优选 IPv6，从 DNS 获得 IPv4 以及 IPv6 地址后先采用 IPv6 发起业务连接请求，如果不成功则等待几十秒超时后才切换到 IPv4 连接。如果用户启用双栈又不能访问 IPv6 因特网，表现出的行为之一就是网页打开速度特别慢。

尽管 IPv6 被作为寄予厚望的 IPv4 替代者，但发展一直相对缓慢，只有在部分竞争激烈和政府强制的区域发展较快，最主要的原因是不能与现有 IPv4 兼容，导致切换的成本和风险较高。**但 IPv6 作为唯一的最终地址短缺解决方案，虽然发展速度可能会受市场影响，但不会改变向 IPv6 转移这一大方向。**

4 网络演进的关键技术

4.1 过渡技术概述

如上分析，支持 IPv4 公网、IPv4 私网和 IPv6 共存，以及平滑过渡是网络演进的关键技术,支持 IPv4 私网和公网共存的技术是 NAT,支持 IPv4 向 IPv6 过渡的技术(过渡技术必然支持 IPv4/IPv6 共存）有很多，发展至今被公认为适合运营商部署的只有双栈、DS-Lite 以及演进后期的 NAT64（也是 NAT 技术）。6RD 技术曾被寄予厚望，但试验后发现不适合运营商环境，主要原因是：

- 缺乏产业链支持
- 完全依赖现有 IPv4 网络，后续向 IPV6 演进需要二次网改
- IPv6 通过隧道封装在 IPv4 网络上，通过独立的 6RD 网关解封装，无法管理到用户和业务，QoS 和流量管理难以实施，导致网络不可管理
- 如为桥接式 CPE 需 PC/终端安装特定驱动程序，导致低成本的桥接式 CPE 很难广泛实施；而路由型 CPE 需要支持 6RD，与其它方案相比又无成本优势

双栈方案和 DS-Lite 方案都是基于双栈的，区别在于 DS-Lite 允许运营商只部署 IPv6 单栈网络，但是参与通讯的终端必须同时为 IPv4 或者同时为 IPv6。只有 NAT64 才支持 IPv6 终端访问 IPv4 网络服务。这三种过渡技术的对比如表 1 所示：

表1 三种过渡技术的对比

	双栈	DS-Lite	NAT64
终端/主机	IPv4/IPv6 单栈或双栈	IPv4/IPv6 单栈或双栈	IPv6 单栈
CPE	支持桥接型 CPE；路由型 CPE 需要升级；最终网络路由型 CPE 占绝对主流	只支持路由型 CPE，CPE 需支持 DS-Lite 功能	不涉及
接入网	PPPoE 不涉及； IPoE 接入需要接入网支持 IPv6 特性	PPPoE 不涉及； IPoE 接入需要接入网支持 IPv6 特性	不涉及
SPOP	双栈	IPv6 单栈，一般集成 DS-Lite 转换网关功能	不涉及

	双栈	DS-Lite	NAT64
骨干网	双栈	双栈或 IPv6	不涉及
特殊网关设备	无	DS-Lite 转换网关（AFTR）。但单独部署将降低网络的可管理性	NAT64 网关
产业成熟度	好	一般，对应用层无要求，且无需 NAT444 两次转换	差，应用层支持少

4.2 双栈（Dual Stack）

IPv4/IPv6 网络共存决定了双栈（Dual Stack）是向 IPv6 演进的基础， IP 骨干网和终端无法回避双栈，在绝大多数场景下，BNG 也要支持双栈，可以说， BNG 和 CPE 是向 IPv6 演进的关键环节。

IP 骨干网

IP 骨干网支持双栈，主要有两种技术选择，Native 双栈和 MPLS 6PE/6vPE。

- Native 双栈是在现有的 IPv4 路由器和链路上，再使能 IPv6 和路由协议，IPv6 报文与 IPv4 报文一样，直接封装在链路层上转发；
- MPLS 6PE/6vPE 则是保持现有的 MPLS 网络不变，在 PE 上将 IPv6 报文封装到 MPLS 中进行转发。

两种方式都有广泛的实际部署应用。

表2 Native 双栈与 MPLS 6PE/6vPE 比较

	Native 双栈	MPLS 6PE/6vPE
实施成本	所有设备使能 IPv6，实施成本稍高	PE 使能 IPv6，P 不必使能 IPv6，实施成本较低
协议部署	所有路由器运行 IPv6 IGP/BGP	MPLS 核心不变，PE 之间部署 MP-BGP
可扩展性	没有限制	没有限制
维护成本	需在所有节点维护新引入的 IPv6 协议和路由	IPv6 作为 MPLS 新业务，对现有维护冲击较小，维护范围限制在 PE
业务支持	单播/组播	组播不成熟，可支持 VPN 应用

有观点认为，双栈对设备资源和维护的要求高，不是合适的解决方案。但对骨干路由器而言，本身就是高性能设备，且路由器专用的转发和查找芯片技术仍然在快速发展，未来并不存在资源瓶颈；至于维护，网络发展的历史已经充分说明，维护能力和资源应当匹配新业务新技术的发展；所以骨干网的演进路线很可能就是双栈，直到最后 IPv4 网络关闭。

BNG

IP 边缘节点统称为 BNG（Broadband Network Gateway），BRAS 和 SR 都属于 BNG 的表现形式，BNG 主要用于用户和业务控制，包括分配地址。

但 BNG 支持双栈不仅仅是为用户同时分配 IPv4 地址和 IPv6 地址，在会话管理、计费、QoS、安全等方面都需要修改以适应 IPv6 业务。由于 BNG 往往与网管、AAA、DHCP、Policy 等业务系统有接口，因此 BNG 支持双栈还需要这些周边系统配合，在导入 IPv6 之前应该进行详细的集成测试。

IPv6 地址分配方式与 IPv4 有较大的差异，由于 IPv6 地址空间巨大，推荐为用户分配一个合适的网段前缀（参考 RFC3633），这也是路由型 CPE 将占主流的原因。出于管理需要，建议为 CPE WAN 接口分配一个单独的/128 主机地址。

CPE

CPE 终端是网络演进部署的关键。无论采用 PPPoE 还是 IPoE 架构，长期来看，路由型双栈 CPE 将是主流。

CPE 采用 DHCPv6 Prefix Delegation 从 WAN 接口向 BNG 获取 IPv6 前缀，并将其子网化，分配给 LAN 接口。在 LAN 侧，用户终端可以通过 DHCPv6 或 SLAAC 方式获得地址，CPE 应该同时支持这两种模式，因为不同用户终端的能力可能存在相当大的差异。CPE 应具备升级到支持 IPv6 组播的能力。目前对于家庭网络的 IPv6 组播架构还不是很完善，CPE 应保持硬件 Ready 状态，使得通过软件升级即可支持 IPv6 组播业务。未来高清电视、互联网视频对带宽的要求很高，CPE 应该对任意比例的 IPv4、IPv6 混合业务具备高速转发能力。

CPE 需要完全保留原来 IPv4 的功能，如 DHCP、NAT、端口映射、UPnP 等。随着 NAT 技术的发展，CPE 应具备可升级能力支持新协议部署。

CPE 应当支持通过远程软件升级或修改配置即可支持不同的 IPv6 业务部署方案。向 IPv6 网络演进中存在调整部署方案的需求，比如早期部署 6RD 或双栈，后期由于 IPv4 地址不足需要切换到 DS-Lite，需要无需入户就可支持演进技术方案的迁移，以降低部署成本。

4.3 NAT

在网络演进的过程中，NAT 无处不在，以各种改头换面的方式体现出来：在网络演进初期阶段，NAT44 或 NAT444（两次地址转换，一次在 CPE，一次在 CGN）是解决 IPv4 地址短缺的必由之路，在网络演进中后期阶段，NAT64 是解决 IPv6 访问 IPv4 的关键技术，此外业界还提出了很多避免 NAT444 转换的技术。

各种不同的 NAT 技术以及其变种（如 L2 Aware NAT、DS-Lite），在运营商环境下部署都需要考虑以下因素：性能容量足够支持规划用户数、高可靠以保持用户业务感知、可管理、方便溯源、对应用程序友好。

CGN（Carrier Grade NAT）或 LSN（Large Scale NAT）都是 NAT 在运营商应用场景下的名称。由于没有什么标准来定义 CGN 是否达到运营商级别，所以提出 LSN，在本文中我们还是使用 CGN，它是大多数所熟悉的术语。

CGN 技术需求

- **性能容量：**运营商环境下，每 NAT 支持的用户数非常多，可能达到十万级别的用户数，每用户的平均流量可能在几百 kbits/s 左右，NAT 设备需要 100G 级别的转发能力。实验表明，Web2.0 网页的点击会生成数十个 TCP 连接，P2P 应用会生成超过 100 个会话，每用户预留 1000 个端口配额可满足一般用户需求，NAT 设备需要具备每秒新建百万会话、维持千万活动会话的能力；
- **可靠性：**通过部署 NAT 设备冗余和设备内板级备份来提高 NAT 网络的可靠性，当主用设备故障后能够自动切换到备用设备上。与一般业务不同，NAT 会话是有状态的，会话生成和老化非常快，CGN 上可能达到每秒百万会话的状态变化，备份会话需要进行协议交互，这种情况下备份几乎不可能做到可靠；而且一般认为绝大部分 NAT 会话的生存时间都极短，备份价值非常小。所以目前被

广泛接受的是只对生存时间较长的会话进行备份；

- **用户管理：**CGN 应用环境下需要进行端口配额管理，避免少数用户滥用导致其他共享 IP 地址的用户无法正常使用网络。此外，一般还需要保持外部地址的惟一性和端口的奇偶性，规避某些特殊端口（例如可能被判别为病毒的端口）等，需要 CGN 具备可管理的特性；
- **地址溯源：**部署 NAT 还必须考虑溯源问题，溯源同时需要应用层面支持，比如网站的访问日志不能仅仅记录 IP 地址，还需要记录端口信息。按照普通实现按 session 记录日志，在 CGN 环境下，日志流量可能高达数十 MBytes/s，需要极高性能的日志处理和存储系统，提高了运维成本。CGN 可通过支持端口预分配技术，一次为用户预留数百上千个端口，降低 NAT 日志规模至千分之一或更小。
- **应用程序友好：**

  NAT 根据实现有多种运行模式

  - Full cone NAT
  - Address-Restricted cone NAT
  - Port-Restricted cone NAT
  - Symmetric NAT

其中，Full cone NAT 的限制最少，对应用程序最友好，相对地安全风险较高，安全和方便往往对立。在运营商应用环境中，建议使用 Full cone NAT 模式，原因是：现有的 IP 网络也是用户自己负责安全，限制过多将增加投诉和报障，从而增加运营商成本；过于严格的限制会增加应用程序的开发难度和部署运行成本。

- **按需部署**

CGN 通过配置或加载不同软件支持 NAT44、NAT64、DS-Lite，在演进过程的不同阶段通过使能不同特性满足业务需求，无需更换硬件以保护投资。

CGN 在网络中的部署

依赖于 NAT 在网络中的部署位置，可以有集中式和分布式两种部署方式。

- 集中式部署中 NAT 为全网私网用户服务，具体表现是在 CR（核心路由器）上增加 NAT 功能模块，或在出口路由器旁挂 NAT 设备。由于在 CR 上增加 NAT 功能模块不符合 CR 简单稳定可靠的网络设计原则，一般不建议；
- 分布式部署则是 NAT 只服务特定区域的私网用户，具体表现是在 BNG 上增加 NAT 功能模块，或在 BNG 旁挂 NAT 设备。

两种部署方式各有优劣，如表 3 所示：

表3 集中式 NAT 和分布式 NAT 比较

	集中式 NAT	分布式 NAT
建设成本	由于设备数目更少，且集中管理，每用户建设成本相对低。	每用户建设成本相对高。
网络路由	路由部署比较复杂，流量迂回一般难以避免；由于 NAT 改变原来的流量模式可能导致网络调整。	网络结构变化小，只需要在 SPOP 调整路由，SPOP 之上无需变化。
维护和故障处理	每 NAT 的作用范围大，出现故障影响面广，故障定位相对困难；但是另外一方面，由于 NAT 设备少，集中维护，可快速处理故障。	每 NAT 的作用范围小，相对容易定位故障； NAT 多点部署导致维护稍复杂。
性能需求	对 NAT 设备的转发性能、session 表容量和 session 创建速率要求很高。	对 NAT 设备的性能要求一般，甚至可以集成到 BNG 中。
Log	需要较高性能和可靠性的 Log 服务器。	集成到 BNG 的 NAT Log 功能可由查询 AAA 端口分配记录实现，无需独立 Log 服务器。

NAT64

NAT64 也是一种地址转换技术，用于 IPv6 终端访问 IPv4 服务，使得 IPv6 主机能够部分兼容原有 IPv4 业务。IETF 不准备开发 IPv4 终端访问 IPv6 业务的标准，主要考虑只需新技术向下兼容，以及通过此策略推进 IPv6 发展。

NAT64 与 NAT44 的区别主要是地址转换基于 IPv6 源地址和源端口，而不是 IPv4 源地址和源端口，NAT64 需要与 DNS64 配合，可以在现有的 DNS 上增加 DNS64 功能。NAT64 替代了原来的 NAT-PT（RFC2766，NAT-PT 可支持 IPv4 终端访问 IPv6 服务），NAT-PT 被废弃的原因在 RFC4966 中说明了,NAT-PT 的实现需要集成 DNS-ALG 的功能，而在 NAT 设备中集成应用层协议，在运营商环境部署会带来一系列的问题难以克服：

- 如何控制终端在访问 IPv4 时选择 NAT-PT 设备作为 DNS；
- 负载分担和可靠性；
- 应用层攻击，以及性能问题（NAT 的 CPU、内存资源远不如服务器）；

与 NAT44 类似，NAT64 主要适合 Client-Server 模式的应用，且数据报文的内容中不携带地址和端口信息。引入 NAT64 是希望端到端部署纯 IPv6 网络从而保持网络的简单性，但是目前应用层还没有准备好。

由于现阶段宽带互联网应用部署纯 IPv6 的基础不成熟，因此 NAT64 在目前固网中的应用将相当有限，在移动互联网的应用要依赖运营商推动产业链。由于移动互联网开放程度不如宽带互联网，推动应用开发商和服务提供商的难度要小。

NAT64 可能以集中式部署为主，因为网络演进到 IPv6 占主流时，BNG 一般不会保留 IPv4 协议栈，而且 IPv4 业务量将比较少，分布式部署在成本上不合理。

4.4 BNG/CGN 融合

CGN 需要做到 NAT 资源的可管可控，以及基于用户的精细化策略控制，实现地址和 NAT 资源的电信化运营分配。传统 NAT 的管理体系架构一般以基于源地址网段/domain 等做到对用户组实施策略控制，而 BNG 的管理体系架构一般以用户帐号（单一用户）或控制域（用户组）进行管理；将两者融合以实现 BNG+CGN 的协同，可在目前的 BNG 用户管理架构下，以最好的兼容性实现用户 CGN 控制策略，包括 NAT 端口分配策略/端口段范围、NAT 会话数、ALG 策略等。

CGN/BNG 融合架构，可以提供可运营的 NAT 管理策略控制能力，实现用户地址和端口资源的有序化管理和分配，提高资源的利用效率，无需部署单独的 NAT Log，有

效降低运维成本；并可基于用户组策略进行差异化的 NAT 资源和 ALG 能力分配，提供电信级可管理的 NAT 业务。

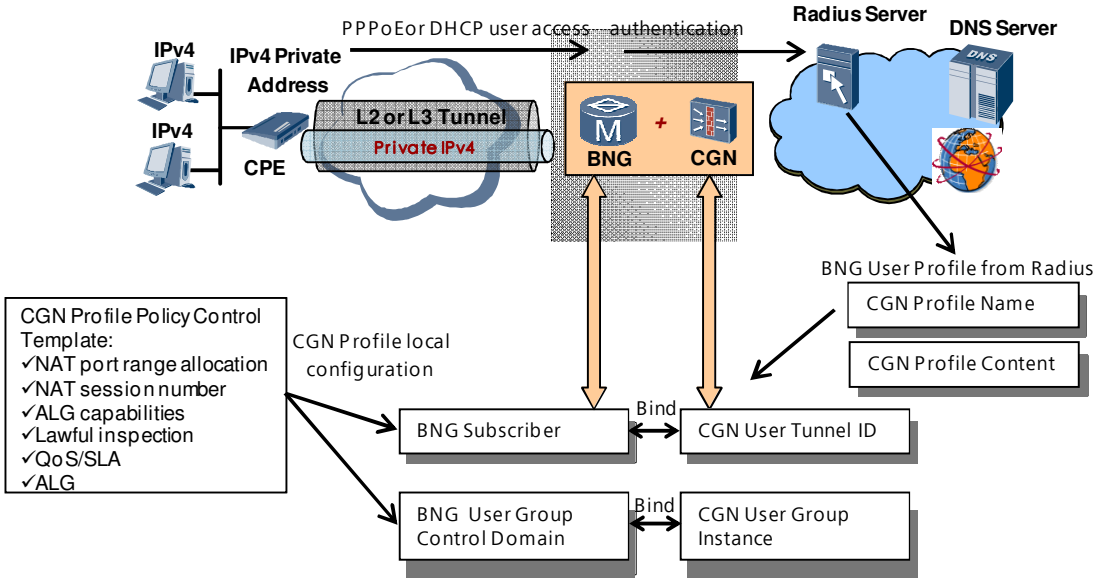

图2 BNG/CGN 融合架构

BNG/CGN 融合设备可支持 DS-Lite 或 L2 Aware NAT，在这个场景下，CPE 无需使能 NAT，BNG/CGN 融合设备基于源 IP 地址、源端口信息，加上 CPE 二层信息如 PPPoE Session ID 或 VLAN/QinQ，或三层信息如 IPv6 地址，进行 NAT 转换，与 NAT444 需要两次 NAT 相比可减少一次 NAT，对应用程序更友好。

4.5 DS-Lite

DS-Lite 是“轻量级的双栈”，相对 Dual Stack 需要全局支持双栈而言是一种局部双栈。它在 IPv6 网络中部署 IPv4-in-IPv6 隧道用于传输 IPv4 业务，而 IPv6 业务则直接通过 IPv6 网络传输。除此之外，DS-Lite 的另外一个关键技术是终端/主机采用自行分配的 IPv4 私网地址，运营商不管理用户的 IPv4 地址。用户的 IPv4 业务在 DS-Lite 转换网关上基于 IPv6 地址+IPv4 私网地址+IPv4 端口进行地址转换，映射到公网地址和端口。

用于封装、解封装的网络设备是部署于家庭的 CPE 和部署于网络的 DS-Lite 网关，它们各自有专有名称 B4 和 AFTR。

DS-Lite 比 Dual Stack 私网地址部署方案来说有如下优势：

- 无需两次 NAT。
- 无需管理用户侧私网 IPv4 地址。
- 原理上说网络侧只需要支持 IPv6，具有简单性。

部署 DS-Lite 可促进 IPv6 网络发展，因为隧道承载部分可以全部 IPv6 化。但是参考 6RD 的部署经验，在网络中单独部署 DS-Lite 网关（AFTR），与 6RD 网关将不会有大的差异，网络和流量依然会难以管理。在运营商环境下，用户管理和业务处理具有强耦合性，DS-Lite 集成到 BNG 中是更合理的部署方式，不然也可选择在 DS-Lite 网关与 BNG 间增加接口，这需要标准化组织支持。

5 网络演进部署建议

运营商需要根据自己面临的问题，采用不同的演进部署方案。

对于面临竞争压力和政府政策部署 IPv6 的运营商，可选择双栈部署方案；对地址短缺很严重的运营商，可考虑部署 NAT 方案或部署双栈+NAT 方案，而对于完全新建的网络后续可考虑部署 DS-lite（还要考虑相应的配套系统）。

当主流应用、终端都支持 IPv6 后，可考虑逐步关闭 IPv4 网络/协议栈，通过部署 NAT64 来解决少量的 IPv4 内容访问需求。

5.1 部署双栈，构建网络向 IPv6 迁移的基石

IPv6 部署应遵循“先核心、后边缘”的部署原则。首先骨干网/IGW 支持双栈，根据现有网络部署技术和运维需求可选择 Native 双栈和 MPLS 6PE/6vPE 两种路线，如果现网为 MPLS 网络可选择 6PE/6vPE 路线，升级 PE 路由器支持双栈，无需对现有骨干 P 路由器做更改。

早期部署的业务需求主要来自于 VPN 专线。只需要升级或新建双栈 PE 路由器，可满足要求支持 IPv6 的专线客户需求。

SPOP 是 IPv6 公众宽带业务部署的核心和难点。由于 PPPoE 业务无需改造接入网，HSI 早期部署建议首先从 PPPoE 宽带用户开始。通过升级或新建 BRAS 以及认证计费系统支持个人用户双栈接入。对于桥接式的 CPE 不需要更换，对路由型 CPE 需要升级为支持双栈，CPE 应可以通过软件升级支持后续演进。

最初引进 IPv6 阶段，也可以新建或升级少数 BRAS 双栈节点，其他区域 PPPoE 用户以 L2TP 接入双栈 BRAS。

对 VAS（Value Added Service）系统的升级改造应同时进行，以增加 IPv6 业务流量。

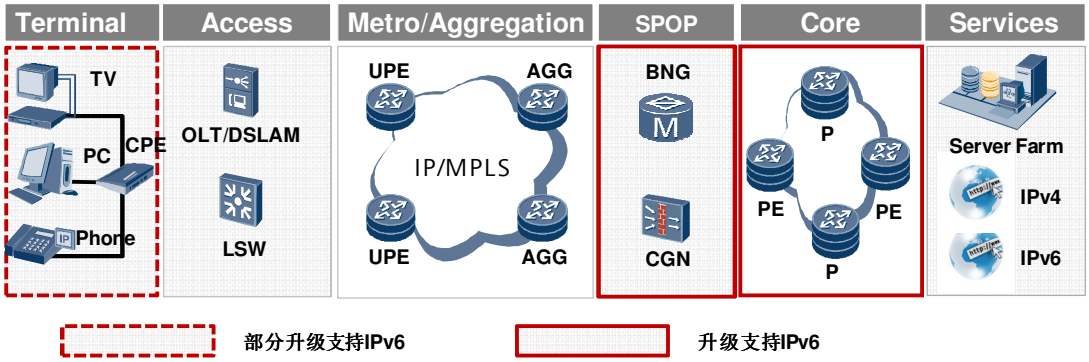

图3 早期部署阶段，升级骨干网，改造 SPOP 和 CPE

5.2 部署 CGN，解决地址短缺，为业务平滑迁移提供保障

目前 IPv4 业务仍然占据主流，而 IPv4 地址短缺成为多数运营商需要面对的问题，NAT 是解决地址短缺的必然选择。一般地，引入 NAT 会导致网络流量模型变化，使得流量管理/用户业务识别的难度增加，甚至使得某些业务受损。

BNG/CGN 融合方案，通过 BNG 集成 DS-Lite 或 L2 Aware NAT，业务只需一次 NAT，而且这种分布式 NAT 结构降低了对设备的性能需求，容易隔离故障和定位，运营商也无需管理终端的 IPv4 地址、无需部署 NAT Log，来提高用户业务感知。缺点在于：更换终端的成本较高（现有 CPE 终究要更换以支持 IPv6）；依赖 BRAS 厂商的支持；依赖于使用 IP 地址识别用户的系统如 DPI/流量管理等都将不能正常工作。

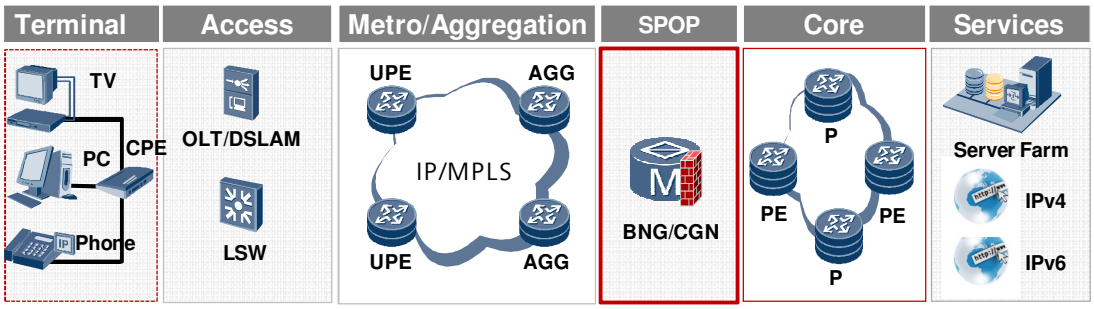

图4 部署 BNG/CGN 应对地址短缺

集中式 NAT 目前已经被广泛部署。在网络出口放置集中式 NAT，对私网地址用户的业务进行重定向到 NAT 设备上进行地址转换。独立式 NAT 部署的优点在于设备可获

得性好，原有网络和业务系统几乎不需要做更改；缺点在于集中式设备故障影响面广，且故障定位较困难；需要建设溯源系统；对 IPv6 部署没有明显促进。

在双栈网络中，解决地址短缺依然可以采用集中式 NAT 方式。

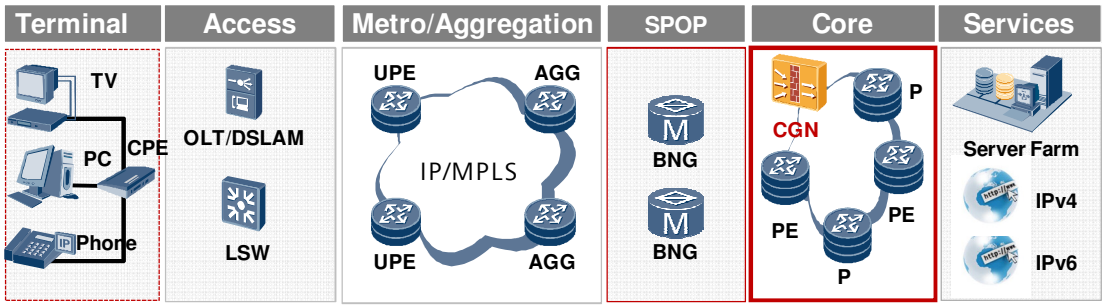

图5 部署集中式 NAT 应对地址短缺

业界已经达成共识，DS-Lite 也是同时解决地址与加速 IPv6 发展的一个选择，已成为 Native 双栈方案之外的唯一一个主流方案。其相关标准逐渐成熟，并被部分 Top 运营商选中作为演进路线，终端和设备供应商已有支持 DS-Lite 的商用产品发布。

两种主流的演进路线是可以兼容的：在初期部署阶段可采用 Native 双栈，同时进行 DS-Lite 试验（DS-Lite 只使用网络的 IPv6 协议栈），在条件具备的情况下可逐步切换到 DS-Lite 方案。

5.3 向纯 IPv6 网络过渡

向纯 IPv6 网络过渡主要解决两个问题：接入网支持 IPv6，以及部署 NAT64。

随着接入方式逐步向 IPoE 迁移，接入网支持 IPv6 业务感知的需求逐渐迫切。在 IPoE 接入场景下，IPv6 协议报文一般为特定组播报文，而且原有 DHCP option、DHCP Snooping、IP/MAC 绑定、组播 Snooping/proxy 等安全特性都需要针对 IPv6 进行开发。此外在 IPv6 协议下，ICMPv6 攻击成为接入层的新攻击点，Neighbor Discovery Snooping 成为接入层设备必备的安全特性。

随着 IPv6 网络部署的推进，对成本、功耗、海量且简单部署的需求将导致纯 IPv6 终端爆炸性增长。部分终端可能存在访问 IPv4 业务的需求，通过部署支持 NAT64 的电信级 NAT，可使得 IPv6 终端以可管理的方式访问 IPv4 网络和业务。

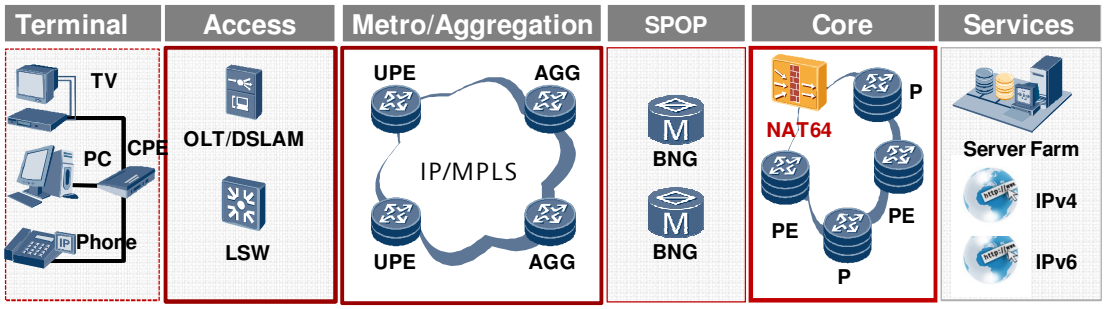

图6 升级接入网，部署 NAT64 和纯 IPv6

6 全球运营商 IPv6 演进方案选择

根据调查，全球大部分运营商倾向于选择双栈方案来部署 IPv6，并采用 NAT 解决地址短缺问题；也有部分运营商有意选择 DS-Lite 作为演进方案，但目前尚无商用部署案例；而极少数采用 6to4、6RD 做早期实验部署的运营商，也逐步减少了对 6RD 的投入。

选择双栈+NAT 的运营商主要的考虑因素是：地址短缺已非常严重急需解决，而选择双栈无需对现网设备大量改造，终端 CPE 也不急于更换，可渐进部署，初始部署成本较低，适合绝大多数对资本支出比较敏感的运营商。

DS-Lite 的推动者大多是已在 DS-Lite 领域研究投入较多，拥有专利和标准影响力的运营商。例如 FT 选择三层到边缘的网络架构，所以希望采用纯 IPv6 以简化管理。而且其现网已经部署的 CPE 具备软件升级支持 DS-Lite 的能力，成本增加相对较小。

部分运营商正计划对双栈+NAT 和 DS-Lite 进行试验，在试商用中评估选择最合适的演进方案。

7 总结

本文讨论了运营商网络从 IPv4 演进到 IPv6 的难点，重点分析了地址转换技术。从 IPv4 到 IPv6 是一个长期渐进的过程，网络架构也是渐进迁移的过程。主流的两种演进路线是双栈和 DS-Lite，都可保证业务平滑过渡，现阶段大部分运营商倾向于选择双栈来部署 IPv6。

作为 IPv6 产业的领导者，华为可提供从终端到网络到业务的端到端全场景的 IPv6 解决方案。基于 NE40E 路由平台的 CGN 可支持 NAT44、DS-Lite、NAT64 等多种 IPv6 演进方案，而创新的 BNG/CGN 融合技术也为运营商提供了一种简明的 NAT +IPv6 部署方案。

附录 A 参考资料

(1) Nordmark, E. and Gilligan, R., "Basic Transition Mechanisms for IPv6 Hosts and Routers",RFC4213, Oct. 2005

(2) Carpenter B, et al. Connection of IPv6 Domains via IPv4 Clouds[S]. RFC 3056.

(3) Aoun, C. and E. Davies, "Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status", RFC 4966, July 2007

(4) Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", RFC 2766, February 2000.

(5) J Wu, et al, "4over6 Transit Solution Using IP Encapsulation and MP-BGP Extensions", RFC5747, Mar 2010.

(6) S. Jiang, D. Guo, and B. Carpenter, "An Incremental Carrier-Grade NAT (CGN) for IPv6 Transition" draft-jiang-v6ops-incremental-cgn, work in progress, May 2009

(7) A. Durand, R. Droms, B. Haberman, and J. Woodyatt, "Dual-stack lite broadband deployments post IPv4 exhaustion", draft-ietf-softwire-dual-stack-lite-00, work in progress, March 2009

(8) X. Li, S. Dawkins, D. Ward, and A. Durand, "Softwire Problem Statement", RFC 4925, July 2007

(9) Bagnulo, M., Matthews, P., and I. Beijnum, "NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", draft-bagnulo-behave-nat64-03 (work in progress), March 2009

(10) Bagnulo, M., Sullivan, A., Matthews, P., and I. Beijnum, "DNS64: DNS extensions for NetworkAddress Translation from IPv6 Clients to IPv4 Servers", draft-ietf-behave-dns64-00 (work in progress), July 2009

(11) Nordmark, E., "Stateless IP/ICMP Translation Algorithm (SIIT)", RFC 2765, February 2000.

(12) Li, X., Bao, C., Chen, M., Zhang, H., and J. Wu, "The CERNET IVI Translation Design and Deployment for the IPv4/IPv6 Coexistence and Transition", draft-xli-behave-ivi, (work in progress),January 2010.

(13) Templin F, Gleeson T, Thaler D. Intra-Site Automatic Tunnel Addressing Protocol (ISATAP). RFC 5214 March 2008

(14) Nordmark E. Stateless IP ICMP Translation Algorithm (SIIT) [S]. RFC2765, February 2000.

(15) J. De Clercq, et al, Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE), RFC4798, Feb 2007

(16) J. De Clercq, IPv6 Address Specific BGP Extended Community Attribute, RFC4659,Sep 2006

(17) D. Cheng, NAT44 with Pre-allocated Ports, draft-cheng-behave-nat44-pre-allocated-ports, Mar 2011

(18) P. Srisuresh et al, Traditional IP Network Address Translator (Traditional NAT), RFC3022,Jan 2001

(19) S. Asadullah et alISP IPv6 Deployment Scenarios in Broadband Access Networks, RFC 4779, Jan 2007

(20) D. Cheng, RADIUS Extensions for NAT Forwarding Port, draft-cheng-behave-nat-fwd-port-radius-ext, Feb 2011

(21) M. Boucadair, et al, Port Control Protocol (PCP) NAT-PMP Interworking Function, draft-bpw-pcp-nat-pmp-interworking, Mar 2011

(22) P. Srisuresh, etc., IP Network Address Translator (NAT) Terminology and Considerations, RFC 2663, Aug 1999

(23) T. Hain, Architectural Implications of NAT, RFC 2993, Nov 2000

(24) IPv6 平滑演进策略及方案白皮书, 华为公司文档

(25) IPv6 过渡 CGN 解决方案白皮书, 华为公司文档

附录 B 缩略语

Abbreviations 英文缩写	Full spelling 英文全称	Chinese explanation 中文全称
6PE	IPv6 Provider Edge Routers	IPv6 业务提供商边界路由器
6RD	IPv6 Rapid Deployment	IPv6 快速部署
6vPE	BGP-MPLS IP Virtual Private Network Extension for IPv6 VPN	
AAA	Authentication, Authorization and Accounting	认证授权计费
AFTR	DS-Lite Address Family Transition Router element	地址族转换路由器
ALG	Application Layer Gateway	应用层网关
B4	Base Bridging BroadBand element	基本桥接宽带单元
BGP	Border Gateway Protocol	边界网关协议
BNG	Broadband Network Gateway	宽带网络网关
BRAS	Broadband Remote Access Server	宽带远程接入服务器
CGN	Carrier-Grade NAT	运营商级 NAT
CPE	Customer Premise Equipment	用户驻地设备
DHCP	Dynamic Host Configuration Protocol	动态主机设置协议
DHCPv6	Dynamic Host Configuration Protocol for IPv6	动态主机设置协议版本 6
DNS	Domain Name System	域名系统
DNSsec	Domain Name System Security Extensions	DNS 安全扩展
DPI	Deep Packet Inspection	深度报文检测
DS	Dual Stack	双栈
DSLAM	Digital Subscriber Line Access Multiplexer	数字用户线路接入复用
DS-Lite	Dual Stack Lite	轻量级双栈
HGW	Home Gateway	家庭网关
ICE	Interactive Connectivity Establishment	交互式连通建立
ICMP	Internet Control Message Protocol	互联网控制报文协议
ICMPv6	Internet Control Message Protocol version 6	互联网控制消息协议版本 6

Abbreviations 英文缩写	Full spelling 英文全称	Chinese explanation 中文全称
IGMP	Internet Group Management Protocol	互联网组管理协议
IGP	interior gateway protocol	内部网关协议
IGW	Internet Gateway	因特网网关
IP	Internet Protocol	因特网协议
IPoE	IP over Ethernet	以太网上的 IP 协议
IPSec	Internet Protocol Security	互联网协议安全
IPv6	Internet Protocol Version 6	IP 版本 6
LSN	Large Scale NAT	大容量 NAT
MIB	Management information base	管理信息库
MLD	Multicast Listener Discovery	多播侦听器发现
MP-BGP	Multiprotocol BGP	多协议 BGP
MPLS	Multi-Protocol Label Switching	多协议标签交换
MxU	Multiplexer Unit	复用器单元
NAT	network address translation	网络地址转换
ND	Neighbor Discovery	邻居发现
OLT	Optical Line Terminal	光缆终端设备
ONU	Optical Network Unit	光网络单元
OSS	Operations support system	运营支撑系统
P2P	Peer to Peer	个人到个人
PPPoE	Point-to-Point Protocol over Ethernet	以太网上的点对点协议
QoS	Quality of Service	服务质量
RADIUS	Remote Authentication Dial In User Service	远程用户拨号认证服务
SLA	Service Level Agreement	服务等级协议
SPOP	Service Point of Presence	业务访问点
STUN	Session Traversal Utilities for NAT	会话穿越 NAT 的应用
UPnP	Universal Plug and Play	通用即插即用
VAS	Value Added Service	增值服务