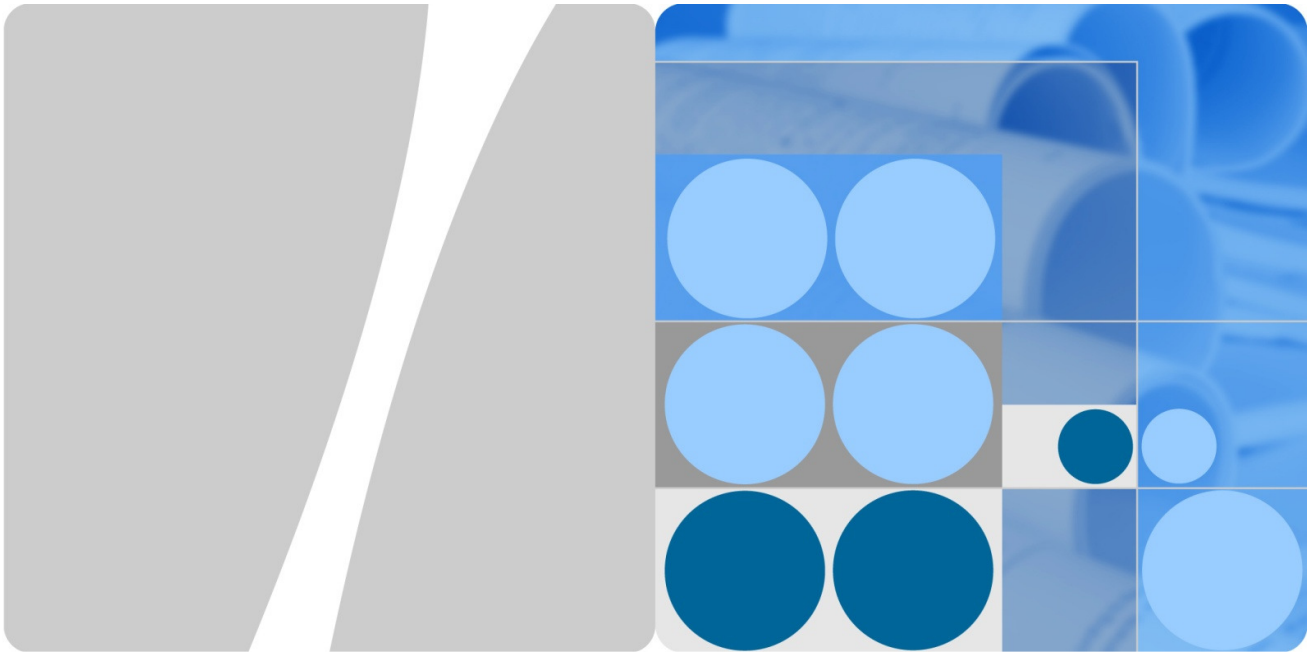


CGN 技术白皮书

文档版本　01

发布日期　2011-09-30

华为技术有限公司

HUAWEI

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

CGN技术白皮书

关键词：NAT, CGN, IPv6, 过渡技术, BNG, 电信级NAT, 溯源, 热备

摘　要：

CGN 即运营商级 NAT，为适应运营商的大规模商业部署，CGN 需要在并发用户数、性能、溯源等方面大幅提升；CGN 可以应用于多个场景，例如 NAT444、DS-Lite 等。本文主要介绍 NAT 原理、NAT 报文处理流程、CGN 的隧道方式以及 CGN 的部署方案。

1 CGN 介绍

电信运营商一方面在大力部署移动互联网、拓展宽带用户并积极推进三网融合，另一方面却不得不面对 IPv4 地址枯竭所带来的现实问题。IANA 的全球 IPv4 地址在 2011 年 2 月 3 日已全部分配出去，全球电信运营商已面临 IP 地址短缺的问题，目前有两种主要的解决思路：

- **引入 IPv6**：可从根本上解决地址耗尽问题，但因目前大部分内容和应用还是基于 IPv4，硬性的全面切换到 IPv6 可能将面临现有业务无法继承的风险；
- **地址转换技术 NAT**：延续使用 IPv4 发展业务，通过规模化部署 IPv4 私有地址，以达到对目前公网 IPv4 地址的统计复用，从而可以在相当长的时间内解决 IPv4 地址问题。

对于运营商来说，同时考虑这两种方向的研究和部署都是必不可少的。

NAT 方案由于无需更换家庭网关设备，大大降低了运营商的投资成本。如果产品设备可以充分“利旧”，将会进一步保护投资。前期采用成熟商用的 NAT444 方案解决地址短缺问题，可以有效缓解运营商眼下的困难。IPv4 向 IPv6 的演进是一个较为长期的过程，通过发掘 IPv4 的地址空间，延续 IPv4 业务的平滑发展，从用户感知度、技术成熟度和部署难易度等方面考虑，NAT444 是目前最佳的选择方案。

运营商级 NAT（Carrier Grade NAT， CGN）又称作大规模部署 NAT（Large Scale NAT ，LSN），与普通 NAT 相比，CGN 主要在支持并发用户数、性能、溯源等方面有很大提升，以适应运营商的大规模商业部署，快速解决 IPv4 地址短缺的急迫问题。CGN 有多个应用场景，例如 NAT444、DS-Lite 等。

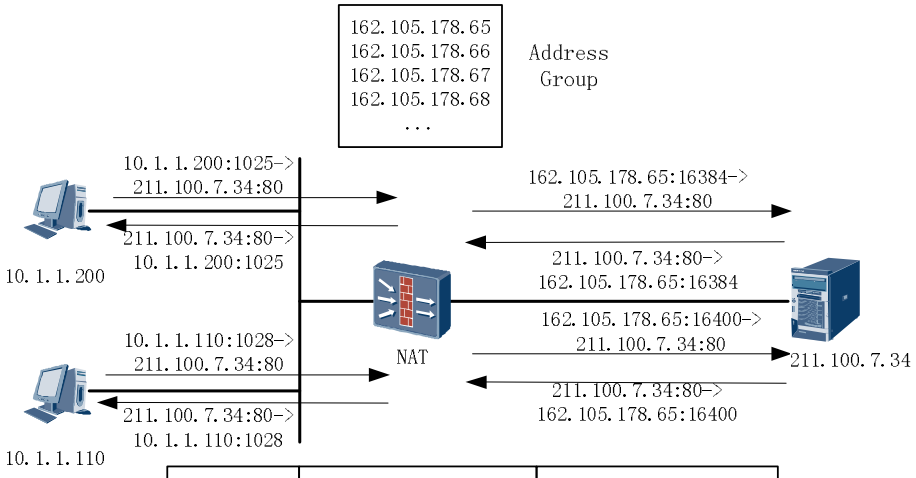
CGN 的部署面临自身的一些问题，如地址转换增加了报文传输时延，用户溯源难，某些应用穿越 NAT 难等。在 CGN 的开发部署中，这些问题被作为重点逐个解决。

2 CGN 方案介绍

2.1 CGN 原理

地址转换有两种方式：NAT 和 NAPT（Network Address and Port Translation）。NAT 方式只转换 IP 地址，对端口号不处理。在 NAPT 方式时，NAT 设备进行地址转换的时候，不仅要对 IP 报文中的地址进行转换，还要对报文中的端口号进行转换。应用 NAT 时，每个私网 IP 都需要一个公网地址，对地址同样比较浪费，因此，实际应用中多采用 NAPT 方式。

图 1 所示为 NAPT 方式的基本原理。其对报文的处理过程如下：



NAT Table	before NAT	after NAT
outbound	10.1.1.200:1025	162.105.178.65:16384
inbound	162.105.178.65:16384	10.1.1.200:1025
outbound	10.1.1.110:1028	162.105.178.65:16400
inbound	162.105.178.65:16400	10.1.1.110:1028

图1 NAPT 方式基本原理

- NAT 设备接收到私网用户发出的访问公网报文。
- 如果是私网用户对外发起一个新连接，NAT 设备从地址池中选择一个空闲的公网地址和端口号、建立 NAT 转换表项。
- 根据私网源 IP 地址、目标 IP 地址、源端口号和目标端口号查找 NAPT 表项，根据查表结果转换报文，向公网侧发送。

- NAT 设备接收到公网侧的回应报文，根据目的 IP 地址和端口号查找反向 NAPT 表项，根据查表结果转换报文，向私网侧发送。

NAPT 方式同时转换 IP 地址和端口号，可以更加充分地利用 IP 地址资源，实现更多的内部网主机对 Internet 的同时访问。

2.2 CGN 报文处理流程

从私有网络到公有网络的报文处理，接口卡针对所有的流量都开启 ACL。现在的 ACL 一般都是使用 TCAM 查找，不影响转发性能，可以线速转发，CGN 报文处理流程如图 2 所示：

- 步骤 1：在入接口板，查找 ACL/UCL 表，发现是需要 NAT 转换的流量，将流量发送到 NAT 板；
- 步骤 2：在 NAT 板上根据三元组（源 IP+源端口+协议号）或五元组（源/目的 IP+源/目的端口+协议号）查 IPv4 正向会话表，NAT 板对报文进行 NAT 转换（转换报文的 IP 地址和端口号等）；查 FIB 表转发到接口板下行继续转发到网络侧出口。五元组匹配条件较三元组严格，在考虑安全等因素时使用五元组；
- 步骤 3：出接口板，对已经完成 NAT 的流量转发流量。

对于回程流量，其报文目的 IP 地址是 NAT 上公网 IPv4 地址，报文被路由到 NAT 单元并做反向 NAT 处理。

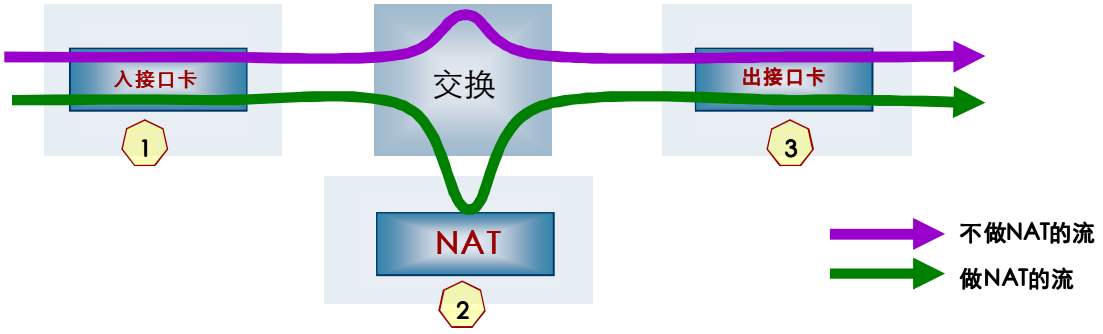


图2 NAT 报文处理流程

2.3 CGN 技术需求

- **性能容量：**运营商环境下，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 kb/s 左右，NAT 设备需要 100G 级别的转发能力。实验表明，WebB2.0 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 个端口配额一般可够用，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 应用环境下需要对用户进行端口配额管理，避免少数用户滥用导致其他共享 IP 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唯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，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 MB/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，从而降低 NAT 日志规模至千分之一或更小；
- **NAT 穿越：**TCP/UDP 的有些应用（例如多媒体会话、文件共享、游戏等“端到端”的应用），报文的负荷部分有 IP 地址或者端口的信息。NAT 做地址转换时，如果不处理负荷中的 IP 地址或者端口，应用程序会失败。3.5 节会介绍如何解决这个问题。

2.4 CGN 应用场景

CGN 与其他技术结合有多种应用场景，如 NAT444、DS-Lite、NAT64 等。

- **NAT444** 包含两次 NAT，分别在 CPE 和运用商网络中的 CGN 网关上做地址转换；
- **DS-Lite** 场景引入两个新概念 AFTR（Address Family Transition Router）和 B4（Basic Bridging BroadBand element），AFTR 是 DS-Lite 网关，负责隧道的封装与解封装以及地址转换；B4 相当于 CPE 的位置，承担隧道的封装与解封装。CPE 和 AFTR 之间单栈 IPv6 网络，B4 向用户分配私有 IPv4 地址支撑 IPv4 业务，IPv4 报文上行时由 B4 将 IPv4 报文封装在 IPv6 隧道中。到达 DS-Lite 网关时，解除外层 IPv6 封装并完成私有 IPv4 地址和公用 IPv4 地址的转换。来自公网的 IPv4 报文的处理过程相反，先将公用 IPv4 地址转成私有 IPv4 地址，并由 IPv6 隧道送达 B4 设备，B4 解封装并转发给目的主机。
- **NAT64** 同时实现 IPv6 与 IPv4 之间的网络地址与协议转换技术。NAT64 一般只支持 IPv6 网络侧用户发起连接访问 IPv4 侧网络资源；如通过手工配置静态映射关系，NAT64 也支持 IPv4 网络主动发起连接访问 IPv6 网络。NAT64 可实现 TCP、UDP、ICMP 协议下的 IPv6 与 IPv4 网络地址和协议转换。DNS64 要配合 NAT64 工作，将 DNS 查询信息中的 A 记录（IPv4 地址）合成到 AAAA 记录（IPv6 地址）中，并将返回合成的 AAAA 记录用户给 IPv6 侧用户。

2.5 CGN 对网络的影响

- **对网络设备的影响：**部署 NAT 需要在网络中增加地址转换功能，对现网调整不大，部分应用系统需要针对私网地址作修改。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；
- **对网络维护的影响：**用户溯源需要配置 NAT 日志服务器以记录用户网络访问情况，可能导致每 NAT 数十 MB/s 的日志流量，需要高性能大存储的日志服务器。增加的 NAT 设备后，故障定位难度增加、用户申告以及溯源的遵从导致网络维护难度和工作量增加；

- **对业务和应用的影响：**NAT 引入的处理时延和某些应用 NAT 穿越困难等都会降低用户业务体验，甚至会影响运营商业务开展，例如如果 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 也会对这类应用产生影响，因为现在地址为多用户共享；对 IP 语音通讯类的应用也往往需要增加应用层网关。

尽管有这些不利影响，NAT 仍然被广泛商用部署，主要原因是地址短缺已经大范围存在，而 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，NAT 相关标准在不断发展完善，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来规避网络层的限制。

3 CGN 部署关键技术

3.1 CGN 形态分类

按 CGN 设备的存在形式可以分为独立式和集成式：

- 独立式 CGN 是一个独立的设备专门承担 CGN 功能；
- 集成式 CGN 也称插卡式，即承担 CGN 功能的一个单板插在其他功能的设备上，如插在 BNG、SR、CR 上。

以 BNG 旁挂独立 CGN 设备和 BNG 集成式 CGN 作对比，独立式 CGN 专门完成地址翻译工作，对 BNG 业务无影响，但是需占用 BNG 转发端口（引流至 CGN）；集成式 CGN 与 BNG 的用户管理相结合，可以提供更好的用户 CGN 管理能力。参见表 1：

表1 独立式 CGN 与集成式 CGN 的对比

	独立式 CGN	集成式 CGN
特点	• 性能强，多核 CPU 处理，整机全部用于 CGN 处理。 • 容量和扩展性较高，可通过级联或多框进行扩展。 • 设备安放较自由。 • 需占用 BNG 的转发端口向独立式 CGN 引流。	• 无需单独设备，空间占地小，功耗散热要求低，投资小。 • 网络层与应用层配合能力强：支持各种路由协议，融合接入和 CGN 管理信息。 • 扩展性好，可在线部署 CGN，无扩展限制。 • 需要占用 BNG 的业务槽位。

按 CGN 设备的部署位置可以分为集中式和分布式：

- 集中式 CGN 的部署位置通常是在网络中比较高的 CR 位置，如国内的城域网 P 路由器、海外骨干网的 P 路由器。可以是独立式 CGN 在 CR 旁，也可以是 CGN 卡集成到 CR 上；
- 分布式 CGN 相对于集中式的网络位置高而言，分布式 CGN 的网络位置较低，通常部署在 BNG 或 SR 所在的位置，可以是独立式 CGN 或集成式 CGN。

集中式 CGN 和分布式 CGN 的对比见表 2：

表2 集中式 CGN 与分布式 CGN 的对比

	集中式 CGN	分布式 CGN
特点	• 适用于整个城域网 CGN 用户比较分散的场景。 • 易于部署，增加节点少。 • 对设备性能要求高;设备故障影响范围大。 • 城域网内，SR/BNG 以上需要引入私网路由。 • 可持续演进性差；随着 CGN 用户量加大，将来还要演进到分布部署。 • 集中式可以减少地址池碎片,提高地址的利用率。	• 特别适合于 CGN 用户集中在某个区域。 • CGN 用户分散情况，部署点较多。 • 设备故障影响范围小。 • 同 SR/BNG 融合后，SR/BNG 同时具有用户信息和 CGN 信息，可以基于用户来管理 CGN，方便解决溯源等由 CGN 引入的问题。

3.2 CGN 接入场景分类

方式一：三层隧道接入

如图 3 所示，终端用户通过 IPv6 隧道接入，CPE 负责将用户的 IPv4 报文封装在 IPv6 隧道中发送到隧道端点 CGN 网关。CGN 拆 IPv6 隧道封装，并做 IPv4 地址转换处理。CPE 作为 IPv4 私网地址 DHCP Server，给家庭网络的用户终端分配 IPv4 私有地址，CPE 本身不做地址转换，只做 IPv4 地址转发。典型应用场景为 DS-Lite。

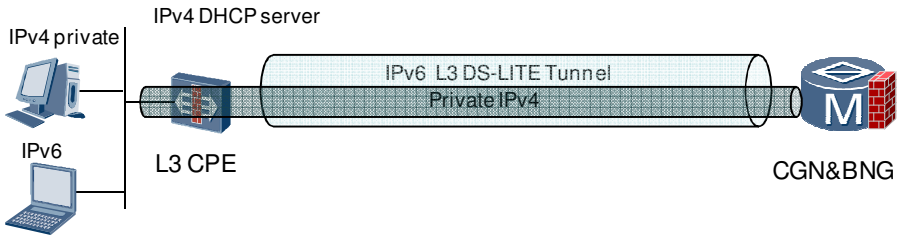


图3 CGN 三层隧道接入

方式二：二层隧道接入，NAT444

如图 4 所示，CPE 作为 IPv4 私网地址 DHCP Server，给终端网络用户分配 IPv4 私有地址，CPE 本身做一级 NAT，CGN 网关做第二级 NAT 处理。CPE 和 CGN 之间通过二层隧道（VLAN/PPPoE）连接。二层通道内仅有一个 IP 会话，二级 NAT 中的私网地址可以采用 Shared Address。

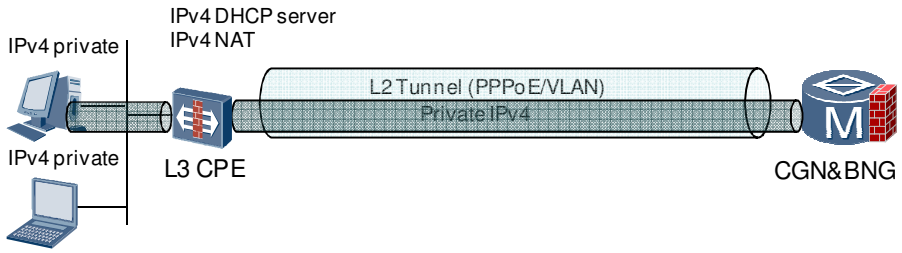


图4 CGN 二层隧道接入，NAT444

方式三：二层隧道接入，一级 NAT，任意地址接入

如图 5 所示，终端用户通过二层隧道（VLAN/PPPoE）接入，用户 IPv4 报文通过二层通道方式传递到 CGN 网关做 NAT 处理，CPE 作为 IPv4 私网地址 DHCP Server 给用户终端分配 IPv4 私有地址，CPE 本身不做 NAT，所有的终端网络的私网地址报文直接通过二层方式接入，二层通道内具有多个 IP 会话，终端网络用户通过二层通道信息来区分标识，如 PPP Session ID/VLAN 等。

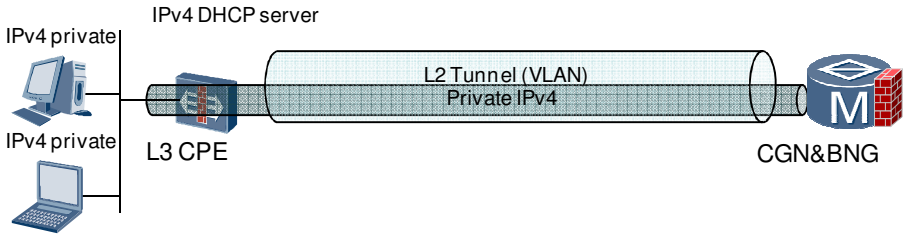


图5 二层隧道接入，一级 NAT，任意地址接入

方式四：二层隧道接入，一级 NAT，地址管理分配

如图 6 所示，终端用户通过二层隧道（VLAN/PPPoE）接入，用户的 IPv4 报文通过二层通道方式传递到 CGN 网关做 NAT 处理,CPE 不具备地址分配能力,BNG 作为 IPv4 私网地址 DHCP Server 给用户终端分配 IPv4 私有地址。CPE 不做 NAT，所有的终端网络的私网地址报文直接通过二层方式接入。二层通道内具有多个 IP 会话，终端网络用户通过二层通道信息来区分标识，如 PPP Session ID/VLAN 等。

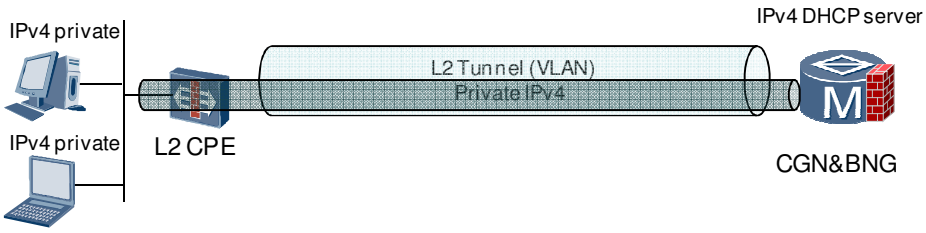


图6 二层隧道接入，一级 NAT，地址管理分配

3.3 BNG+CGN 融合的 NAT 用户策略控制

电信级 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。CGN 的管理体系架构一般以基于源地址网段/CGN 实例等做到针对用户组的策略控制，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，将两者融合实现 BNG+CGN 的控制体系架构一致，可在目前的 BNG 接入的网络体系架构下，以最好的兼容性实现用户的 CGN 控制策略引入。

如图 7 所示，CGN 控制策略包括 NAT 端口分配策略/端口段范围、NAT 会话数、合法监听策略、ALG 能力、QoS/SLA 服务质量、ACL 访问列表等。CGN 控制策略管理通过 Profile 控制策略实例来管理，控制策略 Profile 内容可以在设备本地进行配置，也可以通过策略服务器或者认证系统 RADIUS 下发。

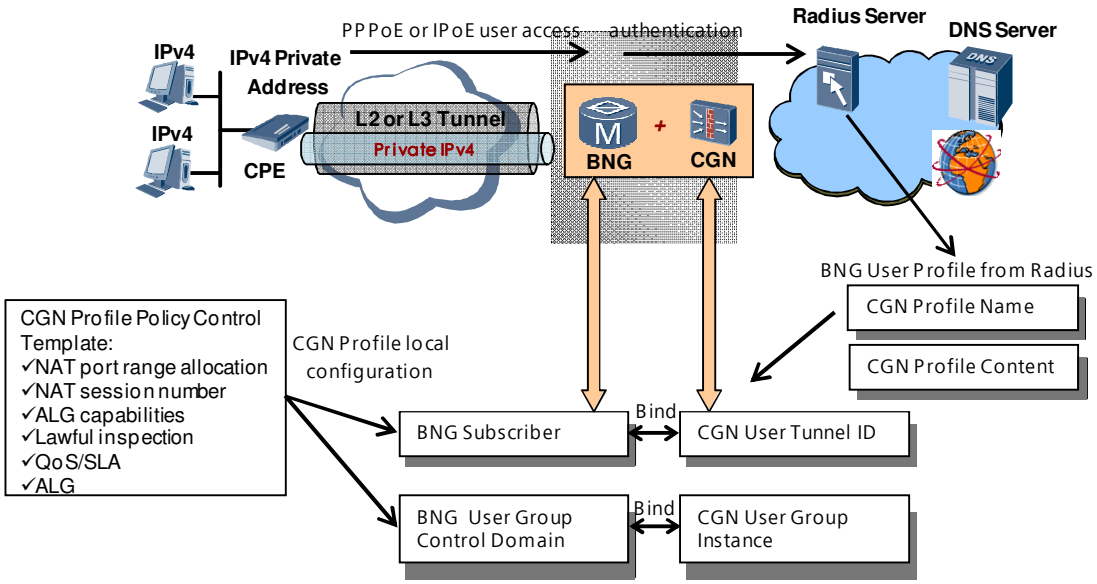


图 1 BNG 与 CGN 融合用户策略控制

- **针对单一用户**，CGN 可以通过唯一性的用户标识来区分用户，并且在用户通过 BNG 接入的认证鉴权的时候，和用户标识绑定，CGN 控制策略可以通过用户帐户策略作用于接入用户，每个用户维护用户管理控制表项，实现用户接入标识和控制策略的绑定关系。BNG/CGN 用户接入时，通过认证流程进行权限认证，获得用户的 CGN 策略，可以通过 RADIUS 或者策略服务器下发给 BNG/CGN 设备，控制策略下发可以将策略的内容直接下发给设备，也可以下

发控制策略 Profile Name，然后在设备本地配置的 Profile 实例中获得策略内容。

- **针对用户组**，CGN 通过将配置的 CGN 实例和 BNG 的控制域进行关联， BNG 特定用户属性的用户归属于统一的控制域，控制域绑定相应的 CGN 实例，实现相应的用户组和控制策略的绑定关系。BNG/CGN 用户接入时，通过认证流程进行权限认证，通过控制域的归属关系找到绑定的 CGN 实例，获得相关联的控制策略。

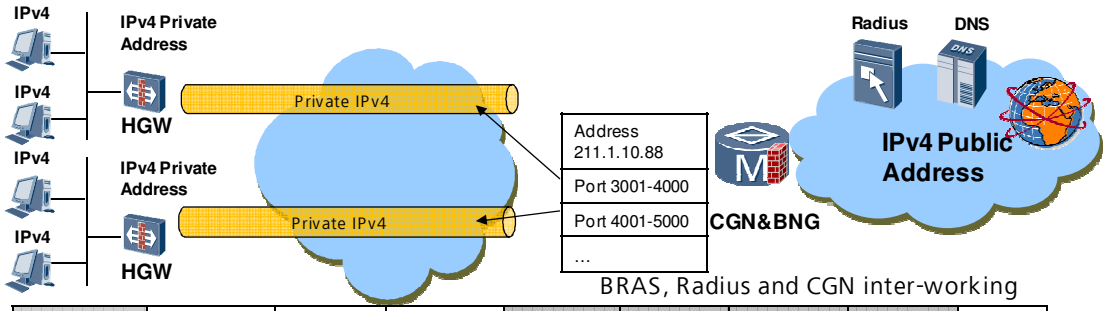
CGN 策略架构融入 BNG 用户策略架构，能提供可运营的 NAT 管理策略控制能力，实现用户的地址和端口资源的有序化管理和分配，最大化的实现资源的有效利用，并且用户差异化的 NAT 业务运营和 NAT 策略统一控制集中管理，可有效的降低运维成本，实现电信级的 NAT 运营能力。

3.4 有序化端口预分配管理

通过多个用户共享 IP 地址来实现 IP 地址的统计复用，最大化的利用有限的公网地址资源。传统的 NAT 机制一般为按需分配，先来先得，这样的分配机制造成多个管理上的问题，如用户的端口资源占用缺乏合理控制，少部分用户可能会占用大量公网端口；端口分配耗尽后需改变 IP 地址，造成业务访问问题；用户的地址端口分配不连续，会造成用户朔源困难，需要大量的日志来记录用户的端口分配情况，造成系统负荷增加。

通过有序的端口预分配机制，可以有效地提高地址端口的使用效率，提高管理运营效率。用户在接入网络时，根据用户绑定的控制策略分配私网地址，同时指定分配用户的公网地址和端口范围，所有信息记录在用户管理控制表中。

如图 8 所示，两个的 HGW 分别分配私有 IPv4 地址，192.168.1.1 和 192.168.1.2。两个 HGW 共享同一个公有 IPv4 地址 211.1.10.88，通过不同的 TCP/UDP 端口段区分是哪个 HGW 的报文，公有 IP 地址 211.1.10.88 加端口 3001-4000 和 6001-7000 的报文对应私有地址 192.168.1.1，公有 IP 地址 211.1.10.88 加端口 4001-5000 和 7001-8000 的报文对应私有地址 192.168.1.2。



User Identify	Private IPv4	Profile	Public IPv4	Start port 1	End port 1	Start port 2	End port 2	...
Tunnel ID 1	192.168.1.1	Profile 1	211.1.10.88	3001	4000	6001	7000	...
Tunnel ID 2	192.168.1.2	Profile 2	211.1.10.88	4001	5000	7001	8000	...

图 2 有序化 CGN 端口预分配管理

用户通过私网地址访问业务时，NAT 根据用户的已分配公网地址和端口实现地址的翻译转换。用户的识别通过用户接入管道来唯一标识，属于同一用户的所有终端需共享该用户分配的地址和端口资源，受限于该用户的 CGN 控制策略。

给用户分配公网 IPv4 地址的外部端口，按端口范围规则有序的分配给用户，也可分配给用户组，如果端口范围内端口使用完，可根据实际需要追加分配端口给终端用户。

端口范围分配机制实现了将 IP 地址和端口资源的有序化分配管理，限制少量用户过渡消耗地址和端口资源，易于对用户实现溯源追踪管理，并且大幅减少 CGN 海量会话的日志存储量，降低了系统负荷压力。

3.5 NAT 穿越

NAT 有效的解决了 IP 地址资源的问题，但同样也带来了地址转换问题。NAT 可以实现 IP 层地址及 UDP 或 TCP 头端口的转换，但某些应用在 TCP/UDP 负载中也带地址信息，通常的标准 NAT 设备并不修改 TCP/UDP 负载中的内容。这样经过 NAT 后，报文的 IP 层地址与 UDP/TCP 负载中地址信息不一致，应用程序就会出错，FTP、RTSP、SIP、DNS 等协议都存在这个问题。

两种解决办法，一种是 ALG（Application Layer Gateways）。ALG 是 NAT 设备具备特定应用协议的解析能力。应用程序在负载中填写的其自身地址，可以被 NAT 的

ALG 模块修改为 NAT 外部地址，这样就保证了报文的 IP 报头与应用负载中地址的一致性，应用程序就不会出错。另一种方法的思路是，应用程序先通过某种机制预先得到其地址对应在 NAT 上的外部地址，这样在负载中就直接填写 NAT 上的对外地址，这样负载中的内容在经过 NAT 时就无需被修改，如 STUN 协议。

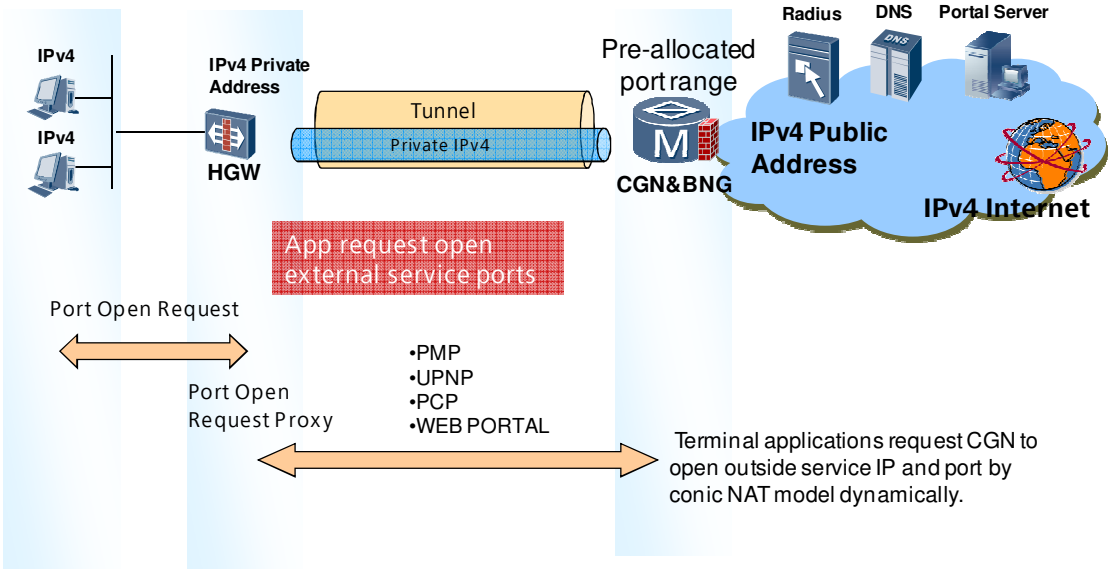


图 3 规模化部署 CGN 业务支持 NAT 穿越

具备 STUN 类型穿越方式的应用，通常需要 NAT 设备具备非对称 NAT 能力，并且对于连接的 TCP 方式也不合适，因此目前大多数 P2P 流媒体应用需要 NAT 设备支持外部地址端口开放能力，即 UPNP/PMP 方式。UPNP 方式可以通过发送控制信息在 NAT 网关上添加端口映射，来实现可以从外部访问 NAT 内部的能力，从而完成 NAT 的穿越。另外对于两个都处于NAT之后的节点需要通信时，互相无法知道对端的外部服务端口，通常某些应用软件通过端口试探或猜测的方式来完成 NAT 穿越，但是这样的方式效率低且成功率不能保证，所以通常需要中继服务器帮助建立连接，实现通信双方在中继服务器的信息中转。但是在 NAT 被规模部署后，大多数通信双方都处于 NAT 后的情况会比较普遍，这时容易造成中继服务器的性能压力瓶颈，通过中继服务器的中转方式将不可行，这时即需要通过外部端口开放协议进行动态的外部端口开放，又需要具备非对称 NAT 的 NAT Server 服务，以增加安全性，这样通信双方可以直接建立连接，不再需要通过中继服务的中转服务。

电信级 CGN 需要支持大多数应用使用的动态端口开放能力，即 PMP、UPNP、PCP 等动态端口映射开放方式，结合 STUN 方式，实现规模化部署 NAT 的业务 NAT 穿越。这样的组合方式可实现 P2P 单、多通道业务的 NAT 支持，不用受限于特殊应用的 ALG 能力限制。类似的 PCP 和 Web Portal 方式也可以协商和管理外部端口，在某些应用下可满足特殊的要求。

3.6 CGN 业务级不中断可靠性

CGN 在部署 DS-Lite 场景时通常在网络核心位置，对于可靠性和扩展性要求较高，为了达到电信级业务不中断，需要进行高可靠的冗余设计。单台 CGN 设备可通过多块 CGN 业务板卡实现板卡间的负荷分担和冗余备份，提高整机的设备处理性能和扩展性。对于多台设备的冗余和负荷分担，需要设备间的冗余备份机制来保证，CGN 设备间的备份机制如图 10 所示：

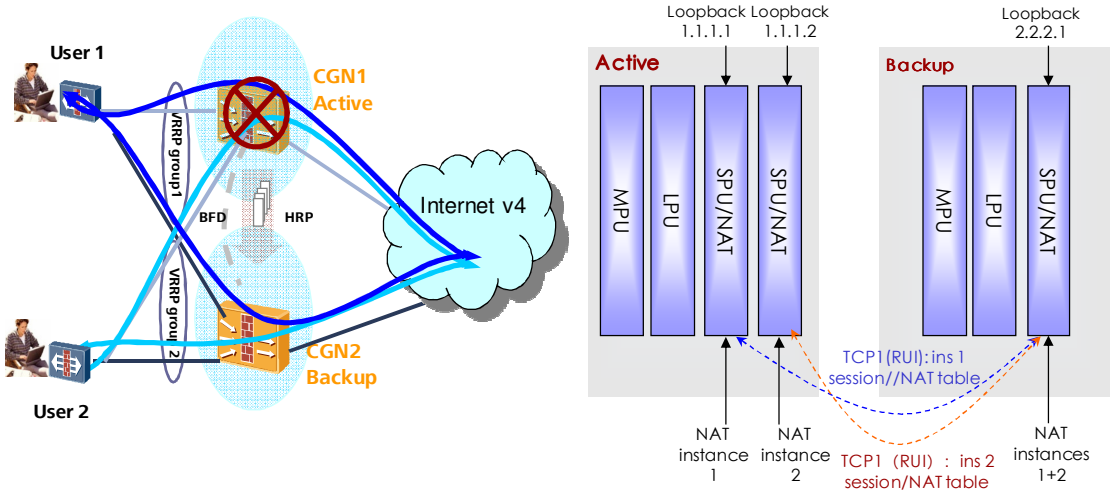


图 4 CGN 备份机制

CGN 设备的可靠性备份机制基于扩展的 VRRP 功能，可以利用 VRRP 进行主备冗余备份的状态管理。VRRP 协议的扩展，用于实现设备间或板卡间多接入接口环境下的接口备份（华为公司的专利），对 VRRP 扩展主要在于传统的 VRRP 实现的是设备级的备份，而扩展后 VRRP 实现的是接口级的备份，可以把一台设备的多个接口配置到同一备份组中，因此扩展后 VRRP 的备份粒度比传统 VRRP 更细，备份方式更加灵活；另外传统的 VRRP 是一个三层协议，而扩展后 VRRP 可以作为二层协议，用在二层用户接入

侧场景下，因此无需配置接入接口的虚拟 IP 地址。CGN 利用扩展的 VRRP 协议，实现设备间或板卡间备份。如果主接口、主接口所连接链路或主接口所在单板发生故障，将会把用户业务切换到备份接口。

主备设备间或板卡间的 CGN 用户信息需要实时同步，对于在主设备或板卡创建的用户管理控制表项信息需要同步到备用设备或板卡。一旦 VRRP 管理的接口或者接口链路发生故障，导致 VRRP 的主备状态发生切换，绑定在此 VRPP 上的用户立刻切换到备用设备上，并保证用户业务不中断、不丢包。设备间的信息备份机制采用 HRP 协议（华为公司的专利），实时同步状态信息。

对于管理 VRRP 组的主备状态检测，CGN 采用 VRRP 关联 BFD 技术，实现电信级的设备或链路故障，主备 VRRP 实体间 BFD 检测可以保证故障检测时间小于 50 毫秒，这样主、备之间的平滑切换，用户感知不到，业务保持不中断。

系统内 VRRP 组数目有限，为了同时达到最大的管理灵活性，需要尽可能的细分管理的颗粒度，这样可以最大化的增加运营管理的灵活性。VRRP 的管理组的备份关系可以关联相关的用户组，即 CGN 的管理实体 CGN 实例或与 CGN 实例关联的 BNG 的控制域，将相应的 CGN 用户组绑定到 VRRP 管理组，一旦 VRRP 的备份组发生主备倒换，相关联的 CGN 用户组 CGN 实例或控制域用户切换到备用设备或板卡上，由备用设备处理后续的业务。

CGN 用户 session 实时在设备间热备同步，故障实时检测，备份关系支持 1:1 或 N:1 模式，以达到电信级 CGN 业务可靠性和扩展性要求。

4 总结

从 IPv4 向 IPv6 演进是一个长期的过程,CGN 是运营商快速解决 IPv4 地址短缺的最成熟最经济的方案。

华为开发的 CGN 设备可提供高并发用户数、高性能和良好用户溯源机制，可适合运营商的大规模商业部署；部署灵活，可独立部署，也可集成在 BRAS 或 CR 设备上部署，可集中式部署，也可分布式部署；CGN 的板卡间热备和设备间热备为高可靠性不间断业务提供保证;而 CGN 与 BNG 的融合更提供了更加灵活的用户控制策略。华为 CGN 解决方案可为运营商网络向 IPv6 的平稳过渡、IPv4 业务的连续性提供了可靠保障。

附录 A 参考资料

(1) RFC2663: IP Network Address Translator (NAT) Terminology and Considerations

(2) RFC2709: Security Model with Tunnel-mode IPsec for NAT Domains

(3) RFC2993: Architectural Implications of NAT

(4) RFC3022: Traditional IP Network Address Translator (Traditional NAT)

(5) RFC3235: Network Address Translator (NAT)-Friendly Application Design Guidelines

(6) RFC3519: Mobile IP Traversal of Network Address Translation (NAT) Devices.

(7) RFC4008: Definitions of Managed Objects for Network Address Translators (NAT)

(8) RFC4787: Network Address Translation (NAT) Behavioral Requirements for Unicast UDP

(9) RFC5135: IP Multicast Requirements for a Network Address Translator (NAT) and a Network Address Port Translator (NAPT)

(10) RFC5382: NAT Behavioral Requirements for TCP

(11) RFC5508: NAT Behavioral Requirements for ICMP

(12) RFC5597: Network Address Translation (NAT) Behavioral Requirements for the Datagram Congestion Control Protocol

(13) RFC6264: An Incremental Carrier-Grade NAT (CGN) for IPv6 Transition

附录 B 缩略语

Abbreviations 英文缩写	Full spelling 英文全称	Chinese explanation 中文全称
CGN	Carrier Grade NAT	运营商级 NAT
CPE	Customer Premises Equipment	用户驻地设备
DHCP	Dynamic Host Configuration Protocol	动态主机设置协议
DS-Lite	Dual Stack Lite	轻型双栈
NAT	Network Address Translation	网络地址翻译
Native IPv6	Native IPv6	本真/原生 IPv6
Radius	Remote Authentication Dial In User Service	远程用户拨号认证系统
VRRP	Virtual Router Redundancy Protocol	虚拟路由器冗余协议