

非对称 IPv6 网络地址空间的半状态翻译方法

闫焘, 马严, 赵钦

(北京邮电大学 网络技术研究院, 北京 100876)

摘要: 提出了一种非对称的 IPv6 地址空间的半状态翻译方法, 可以支持短前缀到长前缀的地址转换, 同时可以尽可能少地记录地址翻译过程中产生的状态信息, 以降低设备存储负载。实验表明, 该方法能够实现任意长度前缀的 IPv6 地址的半状态翻译, 同时可以在传输层保持校验和的中立性, 避免端口转换, 实现对传输层协议和上层应用的透明。本方法是对 IPv6 前缀无状态翻译的补充, 使得 IPv6 地址翻译方法更为普适。

关键词: IPv6 网络地址翻译; IPv6 网络前缀翻译; 地址映射

中图分类号: TP393.2

文献标识码: A

文章编号: 1000-436X(2013)Z2-0117-04

Partial-state translation method between asymmetric IPv6 address spaces

YAN Shen, MA Yan, ZHAO Qin

(Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: A Partial-state IPv6 NAT method was proposed, which could be used to support the address transition from shorter prefix address to longer one. Meanwhile, this method is able to reduce the recording of the translation information as much as possible in order to decrease the load of translation equipment. The experiments show that the proposed method is able to realize the partial-state translation for the IPv6 addresses with arbitrary length. Moreover, it is able to keep the checksum-neutral at the transport layer and avoid port translation, with which it is able to provide transparency for transport layer and application layer protocols. This method, which is more universal, is complementary solution for IPv6 prefix stateless translation.

Key words: IPv6 network address translation; IPv6 network prefix translation; address mapping

1 引言

虽然 IPv6 巨大的地址空间能够从根本上解决 IP 地址短缺的问题, 但是 NAT 技术仍然拥有广泛的应用, 如多宿主站点^[1,2]、拓扑隐藏、网络重编号和过渡技术^[3]等。一种非常典型的场景是: 某网络需要充足的地址空间来满足用户数与业务量的增长, 然而 ISP 无法提供足够大的公有地址块。此网络可以使用比公有地址块网络前缀更短的私有地址空间, 并运用 NAT 技术实现其转换, 这不仅保证了该网络未来发展和局部地址的需求, 更为选择合适的 ISP 提供了极大的灵活性。现行的 IPv6 前缀转换^[4] (NPT66)

技术却不能支持此类应用场景。

RFC6296 中提出的 IPv6 无状态双向前缀映射算法严格要求私有地址空间和公有地址空间的网络前缀长度完全一致, 在不一致的情况下需要强制扩充为相同长度, 因而不能实现不同大小地址块的无状态映射。

基于传统 NAT44 算法的 IPv6 NAT 运用端口复用技术使所有网络设备共享唯一的公有 IPv6 地址, 这种基于 16 bit 端口的复用方法最多只能使内部网络前缀比公有网络前缀的长度多 16 bit。

目前亟需一种能够实现较大私有 IPv6 地址块到较小公有 IPv6 地址块的转换算法。本文提出的算

收稿日期: 2013-09-04

基金项目: 国家科技重大专项基金资助项目(2012ZX03002016-002)

Foundation Item: The National Science and Technology Major Project (2012ZX03002016-002)

法在无需改变传输层校验和的前提下,尽可能少地在路由器侧记录地址转换信息,以节省存储空间,提高设备的可扩展性。该方法涉及前缀的翻译和接口 ID 的计算操作,为区别于 RFC6296 的名称,以下统一使用地址翻译进行陈述。

2 翻译算法

有状态的地址翻译需要在地址翻译设备对地址翻译信息做出完整的记录,如 NAT44 中私网 IP 地址与对应端口的记录条目。此类方法灵活度高,但相对执行效率较低,且由于使用了端口号,所以无法保证对上层协议的透明。无状态的地址翻译通过特定算法对地址进行翻译操作。该方法的执行效率较高,但对地址空间有相对严格的要求,即一般来讲只能做出 1:1 或 1:N 的地址翻译,而无法简单地完成 N:1 的翻译。

半状态的地址翻译方法^[5]只记录部分与地址翻译相关的必要信息,减轻地址翻译设备的存储压力,并且由于需要存储和处理的信息相对较少,所以在执行效率上较之于有状态的地址翻译较高。

2.1 设计原则

为了兼顾执行效率和低存储压力,半状态的 IPv6 地址翻译方法有以下设计原则:

- 1) 只存储必要信息,而非全部 IPv6 地址;
- 2) 尽量不改变传输层校验和;
- 3) 对于任意前缀长度的子网均应适用;
- 4) 与现有 RFC6296 设计兼容。

根据 RFC4291^[3]中所述,主机的 IPv6 地址可大致被分为 3 个区域,即路由前缀、子网号和接口 ID。如前文所述,当某网络内部所使用的前缀长度较短,而对外的前缀长度较长时,RFC6296 无法给出较好的解决方案。

如图 1 所示,内网地址 IN 的路由前缀长度为

x bit,子网号长度为 y bit,接口 ID 长度为 z bit,则 $x+y+z=128$ 。外部地址 OU 的路由前缀长度为 a bit,子网号长度为 b bit,接口 ID 长度为 c bit,则 $a+b+c=128$ 。

若任意一内网地址 IN_r 均能够在外网地址中找到唯一的映射 OU_R ,则需要有不少于 $y+z$ 长度的空间来存储相对应的信息。外网地址 OU 能够提供的空间为 $b+c$,当 $(y+z)-(b+c)>0$ 时,外网地址空间无法满足一一映射的条件,需要额外的空间记录该信息。令: $I=(y+z)-(b+c)$ 来表示因地址空间不足而所需额外空间的长度。

2.2 由内向外

当数据分组从内网向外网发送时,源 IPv6 地址被边缘地址翻译设备翻译为外网前缀的 IPv6 地址,操作如下:

- 1) 将内网和外网前缀分别用最少的 0 填充至 16 的整数倍长度(如 31 bit 前缀用 1 bit 0 在末尾填充为 32 bit);
- 2) 用 RFC1071^[6]中所使用的 1 的补码运算计算两前缀的校验和,并计算外网前缀校验和对内网前缀校验和的差值 $Diff$;
- 3) 将内网 IPv6 地址的子网号和接口 ID 部分由低位至高位,以每 16 bit 为单位逐次向外网 IPv6 地址除前缀以外的空间由低位至高位进行填充,直至外网 IPv6 地址可填充空间不足 16 bit 为止;
- 4) 将外网 IPv6 地址不足 16bit 的空间均填充为 0,并在翻译表项中的 B 列记录尚未填充的值 Rem ,在其最高位前端进行 0 填充至 16 的整数位,并计算其校验和 $RemCS$;
- 5) 从第 4)步填充完毕后的外网 IPv6 地址最后一组 16 bit 位置加上 $Diff-RemCS$;
- 6) 在 Rem 所对应的条目中的 A 列记录经由 3)、

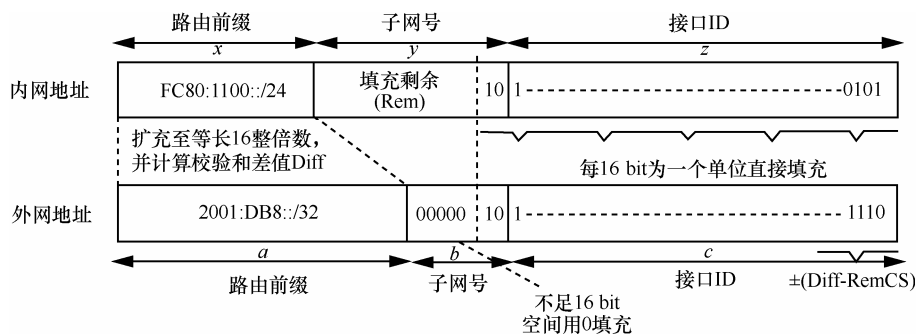


图 1 半状态 IPv6 地址翻译方法

4)、5)操作之后被修改的填充值。

2.3 由外向内

当数据分组从外网向内网发送时, 目的 IPv6 地址被边缘地址翻译设备翻译为内网前缀的 IPv6 地址, 操作如下。

1) 将内网和外网前缀分别用最少的 0 填充至 16 的整数倍长度 (如 31 bit 前缀用 1 bit 0 在末尾填充为 32 位)。

2) 用 RFC1071 中所使用的 1 的补码运算计算两前缀的校验和, 并计算外网前缀校验和对内网前缀校验和的差值 $Diff$ 。

3) 查找表项中 A 列与目的地址末端相同的条目, 获得对应的 Rem 值。在其最高位前端进行 0 填充至 16 的整数位, 并计算其校验和 $RemCS$ 。

4) 将 B、A 两列值依顺序填入内网 IPv6 前缀的后半部分。

5) 从第 4)步已经填充好的内网 IPv6 地址最后 16 bit 位置减去 $Diff-RemCS$ 。

2.4 细节讨论

上述地址翻译方法从理论角度上讲存在 2 个内网地址被翻译为同一个外网地址的可能性。设内网地址 IN_1 的最后 16 bit 值为 a_1 , 填充剩余位的校验和为 $RemCS_1$; 内网地址 IN_2 的最后 16 bit 值为 a_2 , 填充剩余位的校验和为 $RemCS_2$ 。 IN_1 与 IN_2 的其他位完全一致。当且仅当 $a_1 - RemCS_1 = a_2 - RemCS_2$ 时, 翻译结果 OU_1 与 OU_2 完全相等。由于接口 ID 部分的现行标准大多采用 MAC 地址填充和随机生成 2 种, 重复的可能性几乎不存在, 所以无需对此种情况设计更为复杂的算法和机制保障。如在特定场景需要, 可以对该处做出独立于以上算法的设计。

2.2 节与 2.3 节讨论的内容基于 2.1 节中 $I > 0$ 的情况, 即内网前缀长度小于外网前缀长度。当 $I \leq 0$ 时, 则可以直接使用 RFC6296 中的算法, 无需因空间不足而额外记录地址信息, 也就是 RFC6296 所述的内容作为本文方案的一个特例出现, 保持了与 RFC6296 算法的一致性。

本文所述半状态翻译方法不要求内网前缀和外网前缀之间的长短关系, 特别针对内网前缀长于外网前缀的应用场景考虑。基于算法设计, 要求外网前缀最长为 112 bit, 即至少有 16 bit 用来修正前缀校验和差值; 内网前缀则无过多要求。

另外, 外网前缀的长度应足够短, 以保证有足够的冗余空间消除地址翻译结构相同的可能。一般来讲, 实际场景中的前缀长度足以满足避免重复的需要。

设翻译设备计算 16 bit 数据所需时间为 t , 复制及比对 16 bit 数据所需时间为 s , 则将一个 2.1 节所述内网前缀长为 x 的地址翻译为一个外网前缀长为 a 的地址所需时间为

$$T_{Diff} = \left\lfloor \frac{x}{16} \right\rfloor t + \left\lfloor \frac{a}{16} \right\rfloor t + t$$

$$T_{填充} = \left\lfloor \frac{(128-a)}{16} \right\rfloor s$$

$$T_{RemCS} = \left\lfloor \left(128 - x - \left\lfloor \frac{(128-x)}{16} \right\rfloor \right) / 16 \right\rfloor t + t$$

$$= \left\lfloor 9 - \frac{x}{16} - \left\lfloor \frac{(128-x)}{16} \right\rfloor \right\rfloor t$$

对于每一个参与翻译的地址, 所需翻译设备提供的总存储空间为 $y+z$, 较之于有状态的映射节省至少 50% 的存储开销。

3 实验与测试用例

以下使用具体地址来简单演示整个翻译算法的过程。

3.1 由内向外

设主机内网地址 IN_r 为

FC80:1123:1234:2345:3456:4567:5678:6789

该地址的路由前缀为 FC80:1100::/24, IN_r 的校验和为 0x8523。

依照 RFC3849^[7]对文档中前缀的规定, 设主机所在网络的外网前缀长度为 32 bit, 前缀为

2001:DB8::/32

内网地址的前缀校验和为

$$(0xFC80 + 0x1100)' = 0xF27E$$

外网地址的前缀校验和为

$$(0x2001 + 0xDB8)' = 0xD246$$

所以:

$$Diff = 0xD246 - 0xF27E = 0xDFC7$$

由低位至高位向外网地址 OU_r 填充, 直至所剩空间不足 16 bit, 结果为

2001:DB8:1234:2345:3456:4567:5678:6789

由于没有剩余空间,所以无需进行 0 填充。此处:

$$Rem = 0x0023$$

$$RemCS = 0xFFDC$$

在填充结果中,加上 $(Diff - RemCS)$, 所得到的外网 IPv6 地址 OU_R 为

2001:DB8:1234:2345:3456:4567:5678:4774

校验和为 0x8523。此次翻译过程所产生的翻译表项如表 1 所示。

表 1 地址翻译条目

A 列	B 列
1234:2345:3456:4567:5678:4774	23

此例中, $I=(y+z)-(b+c)=8$, 即 B 列需 8 bit 空间记录信息, B 列中 2、3 各占 4 bit 空间。

3.2 由外向内

地址翻译设备收到的目的地址为 2001:DB8:1234:2345:3456:4567:5678:4774 的数据分组。查找表项 A 列, 找到与该 IPv6 地址后半部分匹配的条目。将 B、A 两列值依顺序填入内网 IPv6 前缀后半部分, 得到

FC80:1123:1234:2345:3456:4567:5678:4774

由内网和外网前缀可以计算出

$$Diff = 0xD246 - 0xF27E = 0xDFC7$$

由表项 B 列可计算出

$$RemCS = 0xFFDC$$

在填充结果中, 减去 $(Diff - RemCS)$, 所得到的外网 IPv6 地址 IN_r 为

FC80:1123:1234:2345:3456:4567:5678:6789

4 结束语

本文所述半状态 IPv6 地址翻译方法可以在尽量少的信息存储的条件下对任意长度前缀的 IPv6 私网地址与公网地址进行翻译, 为 IPv6 地址翻译提供了更为普适的方法。该方法既没有改变传输层校验和, 也没有使用端口复用, 为上层协议和应用提

供了灵活性。该方法在过渡技术、多宿主、拓扑隐藏等场景中将会有重要的应用。

参考文献:

- [1] TROAN E O, MILES D, MATSUSHIMA S, et al. IPv6 Multihoming without Network Address Translation[S]. IETF Draft, 2013.
- [2] HADDAD D S W, HALPERN J. Multihoming in Homenet[S]. IETF Draft, 2013.
- [3] DESPRES R I R. IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)[S]. IETF RFC 5569, 2010.
- [4] WASSERMAN P S M, BAKER F. IPv6-to-IPv6 Network Prefix Translation[S]. IETF RFC 6296, 2011.
- [5] LI C B X. Stateless/Partial-State 1:N Network Address and Protocol Translation between IPv4 and IPv6 Nodes[S]. IETF Draft, 2010.
- [6] BRADEN I R, BORMAN D, PARTRIDGE C. Computing the Internet Checksum[S]. IETF RFC 1071, 1988.
- [7] HUSTON T G, LORD A, SMITH P. IPv6 Address Prefix Reserved for Documentation[S]. IETF RFC 3849, 2004.

作者简介:



闫岫 (1987-), 男, 北京人, 北京邮电大学博士生, 主要研究方向为分布式移动性管理。



马严 (1955-), 男, 北京人, 北京邮电大学博士生导师, 主要研究方向为下一代互联网关键技术。



赵钦 (1982-), 男, 河北张家口人, 北京邮电大学工程师, 主要研究方向为 IPv6 过渡技术、移动互联网。