

中华人民共和国通信行业标准

YD/T 3186—2016

运营级网络地址翻译（NAT）技术要求 NAT64

**Carrier grade network address translation technical
requirement for NAT64**

2016-10-22 发布

2017-01-01 实施

中华人民共和国工业和信息化部 发布

目 次

前言.....	II
1 范围.....	1
2 规范性引用文件.....	1
3 术语、定义和缩略语.....	2
3.1 术语和定义.....	2
3.2 缩略语.....	3
4 概述.....	4
4.1 NAT64 部署方式.....	4
4.2 NAT64/DNS64 工作流程.....	6
5 NAT64 处理流程.....	7
5.1 NAT64 处理流程概述.....	7
5.2 BIB.....	7
5.3 会话表.....	8
5.4 输入元组的提取.....	10
5.5 过滤和更新绑定和会话信息.....	11
5.6 输出元组的计算.....	21
5.7 报文的翻译.....	22
5.8 发卡行为的处理.....	30
6 安全性.....	30
6.1 端到端的安全性.....	30
6.2 过滤.....	31

前 言

本标准按照 GB/T1.1-2009 给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由中国通信标准化协会提出并归口。

本标准牵头起草单位：中国联合网络通信集团有限公司、中兴通讯股份有限公司、华为技术有限公司。

本标准参加起草单位：中国信息通信研究院。

本标准主要起草人：马季春、何晓峰、谢梦楠、宋 盈、马高峰、傅 瑜。

运营级网络地址翻译（NAT）技术要求

NAT64

1 范围

本标准确立了 NAT64 的技术内容和使用范围，规定了 NAT64 实现中 NAT64/DNS64 的模型与工作流程，以及 NAT64 的技术要求，用于部署大规模 V6-V4 互通网关解决纯 IPv6 客户端访问 IPv4 业务的场合，主要包括数据的处理流程、相关信息和资源的更新以及一些安全性的考虑，对 IPv6 演进具有指导意义。

本标准适用于有 DNS64 配合的场景。DNS64 是一种将 DNS 查询信息中的 A 记录合成 AAAA 记录的机制。AAAA 记录中的 IPv6 地址是由 IPv4 地址和分配给 NAT64 的 IPv6 前缀生成的。在支持 DNS64 机制的域名系统的配合下，只需要部署支持 NAT64 机制的设备连接纯 IPv6 网络和纯 IPv4 网络，就可以实现纯 IPv6 客户端发起通信，访问 IPv4 服务器的场景。

2 规范性引用文件

下列文件对于本文件的应用时必不可少的，凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

- IETF RFC 793 传输控制协议
- IETF RFC 2460 IPv6
- IETF RFC 4443 ICMPv6
- IETF RFC 4787 单播 UDP 中 NAT 的行为需求
- IETF RFC4884 支持多部分信息的扩展
- IETF RFC 5382 TCP 中 NAT 的行为需求
- IETF RFC 5508 ICMP 中 NAT 的行为需求
- IETF RFC 6052 IPv4/IPv6 翻译器中的 IPv6 寻址
- IETF RFC 6145 IP/ICMP 翻译算法
- IETF RFC 6146 IPv6 客户端到 IPv4 服务器网络地址和协议的翻译（有状态 NAT64）
- IETF RFC 6147 IPv6 客户端到 IPv4 服务器的网络地址转换的 DNS 扩展（DNS64）
- YD/T 2956-2015 DNS64 技术要求

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本文件。

3.1.1

三元组 3-Tuple

这个三元组由（源 IP 地址，目的 IP 地址，ICMP 标识符）组成。一个三元组用来表示一个 ICMP 查询会话。

3.1.2

五元组 5-Tuple

这个五元组由（源 IP 地址，源端口，目的 IP 地址，目的端口，传输层协议）组成。一个五元组用来表示一个 UDP/TCP 会话。

3.1.3

端点无关映射 Endpoint-Independent Mapping

只取决于内部 IP 和内部源端口的 NAT 转换。即不管内部目的地址和内部目的端口是什么，同一个内部源 IP 地址和内部源端口映射到同一个外部源 IP 地址和外部源端口。

3.1.4

端点无关过滤 Endpoint-Independent Filtering

NAT64 只过滤不到内部 IP 地址和端口的报文。换言之，NAT64 可以转发内部到外部任何 IP 的报文，并允许任何回到内部端口的报文。即 NAT64 对于从外部来的报文是否过滤只取决于内部 IP 地址和端口，只过滤不发往内部 IP 地址和端口的报文。

3.1.5

地址相关过滤 Address-Dependent Filtering

NAT64 过滤掉不到内部 IP 地址和端口的报文。此外，如果 X:x 没有向 Y:any 发送报文，则 NAT64 也过滤掉来自 Y:y，目的是 X:x 的报文；换言之，为了接收到来自某外部端口的报文，内部端点必须先向该外部端点发送报文。

3.1.6

发卡行为 Hairpinning

NAT64 之后的两台主机(主机 A 和主机 B)交换数据时，NAT64 为他们分别分配外网 IP 地址和端口，当主机 A 需要和主机 B 联系时，主机 A 的报文会发往主机 B 的外网地址，这个报文经过 NAT64，NAT64 设备把这个数据发给主机 B。

3.1.7

ICMP 查询报文 ICMP Query Packet

ICMP 查询报文是指除了 ICMP 差错报文的 ICMP 报文。对于 ICMPv6 来说, ICMPv6 查询报文是 ICMPv6 信息报文[RFC4443]。对于 ICMPv4, ICMPv4 查询报文是所有除了 ICMPv4 差错报文的 ICMPv4 报文。

3.1.8

会话 Session

两个不同主机之间的数据流, 可以是 TCP, UDP 或 ICMP 查询。

3.1.9

绑定信息库 Binding Information Base

绑定信息库是 NAT64 保存的一张绑定信息表。NAT64 为每个被翻译的协议生成一个绑定信息库。本文档中, 分别为 TCP、UDP 和 ICMP 查询协议生成了绑定信息库。如果需要, 也可以为其他协议添加绑定信息库, 如 SCTP 等。

3.1.10

传输地址 Transport Address

一个 IPv4 或者 IPv6 地址与一个端口的组合。

3.1.11

IPv4 转换的 IPv6 地址 IPv4-Converted IPv6 Address

在 IPv6 网络中表示 IPv4 节点的 IPv6 地址, 由 IPv4 地址和 IPv6 前缀构成。

3.1.12

AAAA 记录 AAAA Record

DNS 协议中的一个资源记录类型用来表示 IPv6 地址。

3.1.13

A 记录 A Record

DNS 协议中的一个资源记录类型用来表示 IPv4 地址。

3.2 缩略语

以下缩略语适用于本文件。

AAA	Authentication, Authorization, Accounting	认证、授权和计费
BIB	Binding Information Base	绑定信息库
DCCP	Datagram Congestion Control Protocol	数据报拥塞控制协议
DF	Don't Fragment	不分片
DNS	Domain Name System	域名系统

DoS	Denial of Service	拒绝服务
ESP	Encapsulating Security Payload	封装安全负载
FQDN	Fully Qualified Domain Name	完全合格域名
ICMP	Internet Control Message Protocol	互联网控制报文协议
ICMP_DEFAULT	ICMP query session timer	ICMP 查询会话计时器
IGMP	Internet Group Management Protocol	互联网组管理协议
IP	Internet Protocol	互联网协议
IPSec	Internet Protocol Security	互联网协议安全
MLD	Multicast Listener Discovery	组播侦听发现协议
MTU	Maximum Transmission Unit	最大传输单元
NAT	Network Address Translation	网络地址转换
STE	Session Table Entry	会话表项
TCP	Transmission Control Protocol	传输控制协议
TCP_EST	TCP Established connection idle timeout	TCP 完成建立超时
TCP_INCOMING_SYN	TCP Incoming Synchronization idle timeout	TCP 入方向同步超时
TCP_TRANS TCP	TCP Transitory connection idle timeout	短暂连接超时
TOS	Type of Service	服务类型
TTL	Time To Live	生存时间
UDP	User Datagram Protocol	用户数据报协议
UDP_DEFAULTUDP	Default value of UDP mapping timer	映射计时器缺省值
UDP_MIN	Minimum value of UDP mapping timer	UDP 映射计时器最小值

4 概述

4.1 NAT64 部署方式

4.1.1 NAT64 部署方式概述

随着 IPv4 公有地址耗尽，运营商可以通过 IPv6[RFC2460]方式提供用户接入服务，但是，目前互联网上绝大多数业务仍是纯 IPv4 业务，为了在 IPv4 向 IPv6 迁移过程中，实现纯 IPv6 主机与纯 IPv4 服务器之间的通信，可以采用 NAT64 方案，其部署方式如图 1 所示。

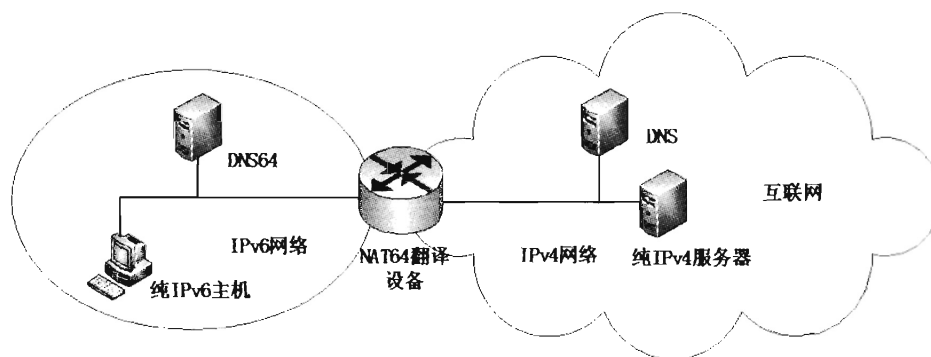


图1 NAT64 部署方式

实现 NAT64 机制的设备至少有两个接口，一个 IPv4 接口连接到 IPv4 网络，一个 IPv6 接口连接到 IPv6 网络。从 IPv6 网络产生的需要到达 IPv4 网络接收端的报文会在 IPv6 网络中被路由到 NAT64 设备，NAT64 翻译并转发这些报文到 IPv4 网络中的 IPv4 接收端。

NAT64 解决方案主要由 NAT64 翻译机制和 DNS64[RFC6147]机制组成。

4.1.2 NAT64 翻译

NAT64 翻译是 NAT64 解决方案的主要组成部分，它由地址翻译机制和协议翻译机制两部分构成。

从 IPv4 报头向 IPv6 报头翻译，或者相反方向的翻译，是通过 IP/ICMP 翻译算法[RFC 6145]实现的。

地址翻译机制主要负责 IPv6 传输地址和 IPv4 传输地址之间的相互映射。为了实现映射功能，NAT64 需要两个地址池：一个 IPv6 地址池（代表 IPv6 网络中的 IPv4 地址）和一个 IPv4 地址池（代表 IPv4 网络中的 IPv6 地址）。

IPv6 地址池是由一个或多个分配给 NAT64 的 IPv6 前缀组成。此后，将 IPv6 地址池称作前缀 $\text{Pref64}::/n$ 。根据 RFC6052 定义，有两种前缀可以用作 $\text{Pref64}::/n$ ， $\text{Pref64}::/n$ 可以是知名前缀 $64::ff9b::/96$ ，知名前缀是一种具有全球意义的前缀； $\text{Pref64}::/n$ 也可以是一个特定网络前缀，特定网络前缀是由组织分配的。配置给 NAT64 的 $\text{Pref64}::/n$ 与 DNS64 相同。NAT64 利用 $\text{Pref64}::/n$ 生成 IPv4 转换的 IPv6 地址。由于 IPv6 巨大的地址空间，可以分配一个或多个 $\text{Pref64}::/n$ ，每个 $\text{Pref64}::/n$ 所能表示的地址范围都可能大于或者等于整个 IPv4 空间。因此，通过简单地连结一个 $\text{Pref64}::/n$ 前缀、一个 IPv4 地址和一个后缀，就可以将一个 IPv4 地址映射成不同的 IPv6 地址。同时 NAT64 向 IPv6 网络发布路由，目的地址包含这些 IPv6 前缀的报文会被路由到 NAT64 进行处理。

IPv4 地址池是由一系列 IPv4 地址组成。NAT64 用这些 IPv4 地址与 IPv6 地址映射，表示 IPv6 地址在 IPv4 网络中的映射地址。由于 IPv4 地址资源有限，不能与 IPv6 地址形成固定的一对一映射关系，因此除了静态或手动建立的映射，使用 IPv4 地址池的映射都是动态生成和释放的。此外，为了提高 IPv4 地址的利用率，NAT64 通常绑定 IPv4 传输地址和 IPv6 传输地址，而不是直接绑定 IPv4 的 IP 地址和 IPv6 的 IP 地址。这表明 NAT64 实施的是地址和协议的翻译。NAT64 同时向 IPv4 网络发布路由，目的地址为 IPv4 地址池中地址的报文会被路由到 NAT64 进行处理。

协议翻译机制主要负责 IPv4 报头和 IPv6 报头之间的相互转化，如图 2 所示。除了 IP 报头报头的翻译，还需要对 TCP、UDP 和 ICMP 进行翻译。对于 TCP 和 UDP 协议，由于 IP 地址和端口发生

改变，因此 TCP 和 UDP 的报头报头也需要翻译，如果存在校验和，则需要重新计算；对于 ICMP 协议，主要是 ICMPv4 报头和 ICMPv6[RFC4443]报头之间的翻译；对于其他数据部分，则不做改变。

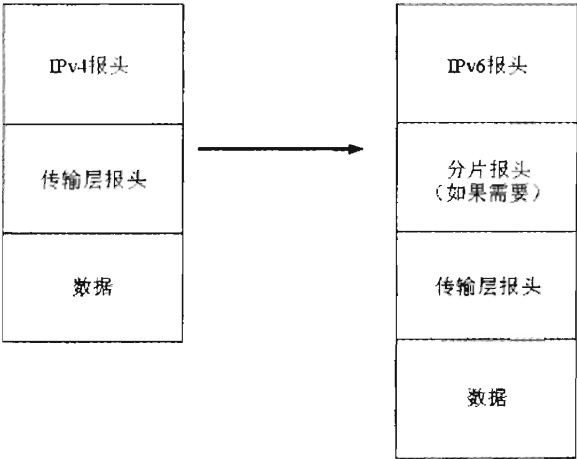


图 2 IPv4 到 IPv6 之间的翻译

4.1.3 DNS64

DNS64 是一种从 A 记录合成 AAAA 记录的方法。DNS64 配合 NAT64 使得纯 IPv6 客户端获得表示纯 IPv4 服务器的 IPv6 地址，从而能够在两者之间建立连接。DNS64 相关规定参考标准 YD/T2956-2015。

4.2 NAT64/DNS64 工作流程

NAT64 解决纯 IPv6 网络中的 IPv6 主机向 IPv4 互联网的 IPv4 服务器发起通信的场合，如图 3 所示：

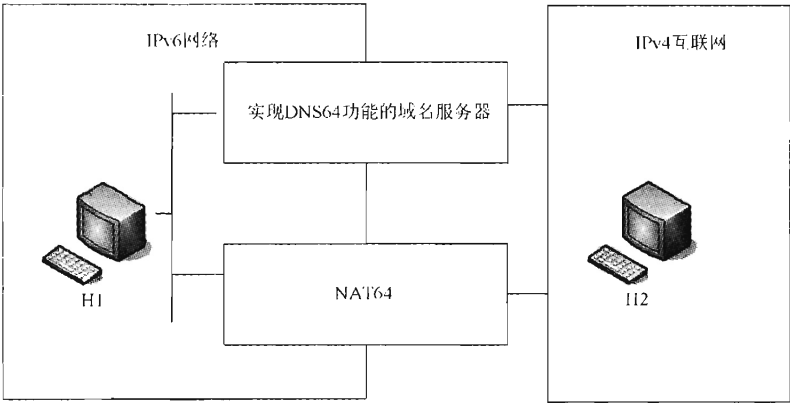


图 3 NAT64/DNS64 工作流程

当主机 H1 创建和服务 H2 的通信时，其步骤如下：

- a) IPv6 主机 H1 向其 IPv6 DNS 服务器发起 AAAA 域名请求；
- b) DNS64 服务器接收到 IPv6 主机 H1 的域名请求后，查询本地域名系统，如发现有 AAAA 记录，则返回该域名对应的地址给 IPv6 主机 H1。当整个系统未发现有 AAAA 记录时，则向 IPv4 网络发起 A 域名请求；

c) 位于 IPv4 域的 DNS 服务器将 A 的查询响应返回给 DNS64;

d) DNS64 接收到 A 域名响应后, 将该 A 域名响应进行 DNS64 的转化, 将 IPv4 服务器 H2 的 IPv4 地址嵌入到配置的 Pref64 前缀地址中, 将 A 记录转换为 AAAA 记录, 并将该 AAAA 记录响应给 IPv6 主机 H1;

e) IPv6 主机 H1 获得 AAAA 响应后, 向 IPv4 服务器 H2 对应的 IPv6 地址发起连接请求, 该报文的源地址为 IPv6 主机地址, 目的地址为 AAAA 响应中的地址。请求报文被路由转发到 NAT64 后, NAT64 为该 IPv6 主机 H1 分配 IPv4 地址, 对其进行 IPv6-IPv4 地址转换和 IPv6 协议-IPv4 协议转换。完成转换后将转换后的 IPv4 报文发送到 IPv4 网络, 最终到达 IPv4 服务器 H2;

f) IPv4 服务器 H2 对其连接建立请求进行响应;

g) IPv4 服务器 H2 的 IPv4 响应报文到达 NAT64 后, NAT64 识别目的地址是 NAT64 地址, 进行 NAT64 映射查找, 并对 IPv4 报文进行 IPv4-IPv6 地址转换和 IPv4 协议-IPv6 协议转换。完成转换后将转换后的 IPv6 报文发送到 IPv6 网络, 最终到达 IPv6 客户端。连接建立完成。

5 NAT64 处理流程

5.1 NAT64 处理流程概述

NAT64 设备作为网关连接 IPv4 与 IPv6 网络, 会通过连接端口接收到报文, 这些报文可以是 IPv6 或者 IPv4 报文, 本标准目前只规定了三种报文处理, 包括 TCP、UDP 和 ICMP 报文。

NAT64 对收到的数据包处理流程如下:

- a) 输入元组的提取;
- b) 过滤和更新绑定和会话信息;
- c) 输出元组的计算;
- d) 翻译数据包;
- e) 发卡行为的处理。

5.2 BIB

NAT64 利用 BIB 记录映射绑定信息。NAT64 维护三种 BIB, 分别为: TCPBIB, UDPBIB 和 ICMP 查询 BIB。

对于 TCPBIB 和 UDPBIB, 每个 BIB 表项指定了在 IPv6 传输地址和 IPv4 传输地址之间的一个映射:

$$(X', x) \longleftrightarrow (T, t) \quad (1)$$

式中:

X'——IPv6 地址;

T——NAT64 分配的 IPv4 地址;

x、t——端口号。

TCP 或 UDP BIB 维护的具体信息见表 1。

表 1 TCP 或 UDP BIB

字段名称	长度（比特）	是否强制
源IPv6地址	128	是
翻译后的源IPv4地址	32	是
源传输端口	16	否
翻译后的源传输端口	16	否

BIB 中的信息还可以用来实现端点无关过滤。

对于 ICMP 查询 BIB，每个 ICMP 查询 BIB 表项指定了在（IPv6 地址，ICMPv6 标识符）和（IPv4 地址，ICMPv4 标识符）之间的一个映射：

$$(X', i1) \longleftrightarrow (T, i2) \quad (2)$$

式中：

X'——IPv6 地址；

T——NAT64 分配的 IPv4 地址；

i1——ICMPv6 标识符；

i2——ICMPv4 标识符。

ICMP 查询 BIB 维护的具体信息见表 2。

表 2 ICMP 查询 BIB

字段名称	长度（比特）	是否强制
源IPv6地址	128	是
ICMPv6标识符	16	是
源IPv4地址	32	是
ICMPv4标识符	16	是

三种 BIB 的表项可以动态的通过数据流来建立，也可以通过管理员手动配置来建立，NAT64 的三种 BIB 的表项都应该支持手动建立。动态建立的 BIB 表项在最后一个会话从会话表中移除后随之删除，手动配置的 BIB 表项则不会随着会话表项的移除而删除，只能由管理员来删除。

5.3 会话表

NAT64 维护三种会话表：TCP 会话表，UDP 会话表和 ICMP 查询会话表，每个 STE 记录了相关会话的状态信息。

在 TCP 或 UDP 会话表中，每个表项指定了一对 IPv6 传输地址和一对 IPv4 传输地址之间的映射关系，其关系如下所示：

$$(X', x), (Y', y) \longleftrightarrow (T, t), (Z, z) \quad (3)$$

式中：

X'、Y'——IPv6 地址；

T、Z——IPv4 地址；

x、y、z、t——端口号；

T——NAT64 分配的 IPv4 地址；

Y'——IPv4 地址 Z 的 IPv6 地址，Y'可以根据 NAT64 使用的算法由 IPv4 地址 Z 产生，其具体过程要求遵循 RFC6052。

每个 UDP/TCP STE 维护的信息见表 3。

表 3 TCP 或 UDP STE

字段名称	长度（比特）	是否强制
生存时间	64	是
源IPv6地址	128	是
翻译后的源IPv4地址	32	是
源传输端口	16	是
翻译后的源传输端口	16	是
目的IPv6地址	128	是
翻译后的目的IPv4地址	32	是
目的传输端口	16	是
翻译后的目的传输端口	16	是

在 ICMP 查询会话表中，每个表项指定了一个三元组（IPv6 源地址，IPv6 目的地址，ICMPv6 标识符）和一个三元组（IPv4 源地址，IPv4 目的地址，ICMPv4 标识符）之间的映射，其关系如下所示：

$$(X', Y', i1) \longleftrightarrow (T, Z, i2) \quad (4)$$

式中：

X'、Y'——IPv6 地址；

T、Z——IPv4 地址；

i1——ICMPv6 标识符；

i2——ICMPv4 标识符；

T——NAT64 分配的 IPv4 地址；

Y'——IPv4 地址 Z 的 IPv6 地址，Y'可以根据 NAT64 使用的算法由 IPv4 地址 Z 产生，其具体过程应遵循 RFC6052。

每个 ICMP 查询 STE 维护的信息见表 4。

表 4 ICMP 查询 STE

字段名称	长度（比特）	是否强制
生存时间	64	是
源IPv6地址	128	是
目的IPv6地址	128	是
ICMPv6标识符	16	是
源IPv4地址	32	是
目的IPv4地址	32	是
ICMPv4标识符	16	是

5.4 输入元组的提取

5.4.1 输入元组的提取概述

输入元组的提取主要是将输入的 IP 报文和一个输入的元组联系起来以便后续处理。对于 TCP、UDP 或者 ICMP 差错报文，这个输入的元组是一个五元组（源 IP 地址，源端口，目的 IP 地址，目的端口，传输协议）；对于 ICMP 查询报文，这个输入的元组是一个三元组（源 IP 地址，目的 IP 地址，ICMP 标识符）。

对于报文是否存在分片，NAT64 提取相应元组时的处理方式不同。

5.4.2 TCP 输入元组的提取

TCP 存在非分片和分片报文。对于非分片报文，处理如下：

如果收到的报文是一个完整的 TCP 报文，那么可以通过提取报文中的字段计算出输入的五元组（源 IP 地址，源端口，目的 IP 地址，目的端口，TCP）。

对于分片报文，处理如下：

a) NAT64 在处理 TCP 分片时，特别是没有按顺序到达的分片时，NAT64 应限制用来储存分片报文的资源，防止 Dos 攻击；只要 NAT64 有可用的资源，那么 NAT64 应允许分片在一段时间间隔内到达，该时间间隔可以配置并且默认值应至少设置为 FRAGMENT_MIN（2 秒）；NAT64 可以要求 TCP 的报头完整的包含在第一个分片中。

b) 对于包含 TCP 分片且其校验和不为零的报文，NAT64 可以选择在分片到达的时候对它们进行排序，然后同时翻译所有分片，这种情况下，可以通过非分片报文输入元组的提取方式提取分片报文的输入五元组。此外，NAT64 也可以通过存储用来计算分片的五元组的信息，在分片到达时就进行翻译。

5.4.3 UDP 输入元组的提取

UDP 存在非分片和分片报文。对于非分片报文，处理如下：

如果收到的报文是一个完整的 UDP 报文，那么可以通过提取报文中的字段计算出输入的五元组（源 IP 地址，源端口，目的 IP 地址，目的端口，UDP）。

对于分片报文，处理如下：

a) NAT64 在处理 UDP 分片时，特别是没有按顺序到达的分片时，NAT64 应限制用来储存分片报文的资源，防止 Dos 攻击；只要 NAT64 有可用的资源，那么 NAT64 应允许分片在一段时间间隔内到达，该时间间隔可以配置并且默认值应至少设置为 FRAGMENT_MIN（2 秒）；NAT64 可以要求 UDP 的报头完整的包含在第一个分片中。

b) 对于包含 UDP 分片且其校验和不为零的报文，NAT64 可以选择在分片到达的时候对它们进行排序，然后同时翻译所有分片，这种情况下，可以通过非分片报文输入元组的提取方式提取分片报文的输入五元组。此外，NAT64 也可以通过存储用来计算分片的五元组的信息，在分片到达时就进行翻译。

c) 对于包含 UDP 分片且其校验和为零的报文，如果 NAT64 有足够的资源，那么 NAT64 应重新组装报文并且计算校验和；如果此时没有足够的资源，那么 NAT64 应默默丢弃这些报文。

5.4.4 ICMP 输入元组的提取

ICMP 存在非分片和分片报文。对于非分片报文，处理如下：

- a) 如果收到的报文是一个完整的 ICMP 查询报文，那么可以提取出输入的三元组（源 IP 地址，目的 IP 地址，ICMP 标识符）。
- b) 如果收到的 IP 报文是一个包含 TCP 或 UDP 报文的完整的 ICMP 差错消息，那么可以通过提取 ICMP 差错消息中嵌入的 IP 报文的字段计算出输入的五元组（源 IP 地址，源端口，目的 IP 地址，目的端口，传输协议）。如果嵌入的源地址变成输入五元组中的目的地址，嵌入的源端口变为输入五元组中的目的地址，嵌入的目的地址变成输入五元组中的源地址，嵌入的目的端口变为输入五元组中的源端口，那么，交换提取得到的源和目的信息在五元组中的位置。如果无法提取输入的五元组（可能由于 ICMP 消息中嵌入的报文信息不足），那么应默默丢弃收到的 IP 报文。
- c) 如果收到的 IP 报文是一个包含 ICMP 差错信息的完整的 ICMP 差错消息，那么应默默丢弃收到的 IP 报文。
- d) 如果收到的报文是一个包含 ICMP 查询信息的完整的 ICMP 差错消息，那么可以通过提取 ICMP 差错消息中嵌入的 IP 报文的字段计算出输入的三元组（源 IP 地址，目的 IP 地址，ICMP 标识符）。如果嵌入的源地址变成三元组中的目的地址，嵌入的目的地址变成三元组中的源地址，嵌入的 ICMP 标识符用作三元组中的 ICMP 标识符，那么，交换提取得到的源和目的信息在三元组中的位置。如果无法提取输入的三元组（可能由于 ICMP 消息中嵌入的报文信息不足），那么应默默丢弃收到的 IP 报文。

对于分片报文，处理如下：

NAT64 在处理 ICMP 分片时，特别是没有按顺序到达的分片时，NAT64 应限制用来储存分片报文的资源，防止 Dos 攻击；只要 NAT64 有可用的资源，那么 NAT64 应允许分片在一段时间间隔内到达，该时间间隔可以配置并且默认值应至少设置为 FRAGMENT_MIN（2 秒）；NAT64 可以要求 ICMP 的报头完整的包含在第一个分片中。

5.4.5 其他协议的处理

如果 NAT64 收到一个 IPv4 报文，但该报文所包含的协议不是 TCP、UDP 或 ICMPv4，那么 NAT64 应该丢弃这个报文；如果安全策略允许，NAT64 应向报文的源地址发送一个代码值为 2 的 ICMPv4 目的不可达的差错消息，表示协议不可达。

如果 NAT64 收到一个 IPv6 报文，但该报文所包含的协议不是 TCP、UDP 或 ICMPv6，那么 NAT64 应该丢弃这个报文；如果安全策略允许，NAT64 应向报文的源地址发送一个代码值为 4 的 ICMPv6 目的不可达的差错消息，表示端口不可达。

5.5 过滤和更新绑定和会话信息

5.5.1 过滤和更新绑定和会话信息的概述

获得报文的输入元组后，NAT64 会对绑定和会话信息进行更新。如果需要，NAT64 会对接收的报文进行过滤。

NAT64 对绑定和会话信息进行更新和过滤时，其具体细节依赖于承载的协议，如 TCP、UDP 或 ICMP。但无论哪种传输协议，NAT64 应默默丢弃源地址包含前缀 Pref64::/n 的 IPv6 报文。这样做是为

了防止发卡行为的环路。此外，对于收到的 IPv6 报文，NAT64 只处理目的地址包含前缀 Pref64::/n 的报文。同理，对于 IPv4 报文，NAT64 只处理目的地址属于分配给 NAT64 的 IPv4 地址池的报文。

NAT64 支持的过滤方式有以下两种：

a) 外部端点无关过滤：NAT64 只过滤不到内部 IP 地址和端口的报文。换言之，NAT64 可以转发内部到外部任何 IP 的报文，并允许任何回到内部端口的报文。即 NAT64 对于从外部来的报文是否过滤只取决于内部 IP 地址和端口，只过滤不发往内部 IP 地址和端口的报文。

b) 地址相关过滤：NAT64 过滤掉不到内部 IP 地址和端口的报文。此外，如果 X:x 没有向 Y:any 发送报文，则 NAT64 也过滤掉来自 Y:y，目的是 X:x 的报文；换言之，为了接收到来自某外部端口的报文，内部端点必须先向该外部端点发送报文。

NAT64 推荐支持外部端点无关过滤，因为存在一些游戏和 P2P 的应用需要 NAT64 支持外部端点无关过滤来穿越 NAT，这样可以使 NAT64 最小化对一些应用程序的连接破坏，但同时外部端点无关过滤也增加了 NAT64 被攻击的危险。

5.5.2 UDP 会话的处理

5.5.2.1 UDP 会话的处理过程

当接收到 IPv6 报文时，可以得到相应的五元组(X',x,Y',y,UDP)，其处理过程如下：

a) NAT64 检索 UDP BIB 表项，该表项应包含匹配 IPv6 源传输地址(X',x)的 BIBIPv6 传输地址。如果该 BIB 表项不存在，那么 NAT64 会生成一个新的 BIB 表项（如果资源和策略允许）。IPv6 报文的源传输地址(X',x)被用作 BIB 的 IPv6 传输地址，NAT64 分配的 IPv4 传输地址 (T,t) 被用作 BIB 的 IPv4 传输地址，这样一个 BIB 表项就成功建立了，如式 (1) 中所示：

b) NAT64 根据得到的五元组(X',x,Y',y,UDP)检索相应的 STE，如果该 STE 没有找到，那么 NAT64 会生成一个新的 STE（如果资源和策略允许）。会话表项中包含如下信息：

1) STE 的源 IPv6 传输地址（即源 IPv6 地址和源传输端口）被设置为(X',x)，也就是接收到的报文的源 IPv6 传输地址；

2) STE 的目的 IPv6 传输地址（即目的 IPv6 地址和目的传输端口）被设置为(Y',y)，也就是接收到的报文的目的 IPv6 传输地址；

3) 从相关 UDPBIB 表项中提取的 STE 的源 IPv4 传输地址（即翻译后的源 IPv4 地址和翻译后的源传输端口）被设置为(T,t)；

4) STE 的目的 IPv4 传输地址（即翻译后的目的 IPv4 地址和翻译后的目的传输端口）被设置为(Z,z)，其中，端口号 z 的值与目的 IPv6 传输地址的端口号 y 一致，地址 Z 则通过反向算法从目的 IPv6 地址 Y' 获得，也可以表示为 Z(Y')，具体算法过程见 5.5.5 节。

这样，一个 STE 就产生了：

$$(X',x),(Y',y) \longleftrightarrow (T,t),(Z(Y'),y) \quad (5)$$

NAT64 设置（或重置）这个 STE 计时器的值为最大会话生存时间。这个最大会话生存时间是可以配置的，默认值至少为 UDP_DEFAULT（5 分钟）[RFC4787]，但不能小于 UDP_MIN（2 分钟）[RFC4787]。

当接收到 IPv4 报文时，其源 IPv4 传输地址为(W,w)，目的 IPv4 传输地址为(T,t)，可以得到相应的五元组(W,w,T,t,UDP)，其处理过程如下：

a) NAT64 检索 UDP BIB 表项, 该表项应包含匹配(T,t)的 BIB IPv4 传输地址, 也就是 IPv4 报文中的 IPv4 目的传输地址。如果该 BIB 表项不存在, 那么 NAT64 应该丢弃该报文, 并且可能向报文的源地址发送一个类型值为 3 (目的不可达) 的 ICMP 差错信息。

b) 如果 NAT64 在 IPv4 接口上应用地址相关过滤策略, 那么 NAT64 会根据地址相关过滤策略检查是否允许接收报文。为了实现该行为, NAT64 会检索 STE, 要求该 STE 的源 IPv4 传输地址为(T,t), 也就是报文的目的 IPv4 传输地址, 并且要求该表项的目的 IPv4 地址为 W, 也就是报文的源 IPv4 地址。如果存在这样的 STE, 则进行后续处理。如果不存在这样的 STE, 则丢弃该报文, 并且可能向报文的源地址发送一个类型值为 3 (目的不可达)、代码值为 13 (管理上禁止通信) 的 ICMP 差错信息。

c) 如果报文在前面的过程中没有被丢弃 (因为 NAT64 没有过滤或者因为报文符合地址相关过滤策略规则), 那么 NAT64 会检索 STE, 要求该 STE 的源 IPv4 传输地址为(T,t), 并且要求该 STE 的目的 IPv4 传输地址为(W,w)。如果不存在这样的 STE, 那么 NAT64 会生成一个新的 STE (如果资源和策略允许)。一旦新的 STE 产生, 应包含如下信息:

1) STE 的源 IPv6 传输地址 (即源 IPv6 地址和源传输端口) 可以从相关的 UDPBIB 表项中得到;

2) STE 的目的 IPv6 传输地址 (即目的 IPv6 地址和目的传输端口) 被设置为(Y'(W),y), 其中端口号 y 的值与收到 IPv4 报文的源端口地址 w 相同, Y' 从收到 IPv4 报文的源地址 W 中通过相应的算法产生, 可表示为 Y'(W), 具体算法过程见 5.5.5 节;

3) STE 的源 IPv4 传输地址 (即翻译后的源 IPv4 地址和翻译后的源传输端口) 被设置为(T,t), 也就是收到的 IPv4 报文的目的 IPv4 传输地址;

4) STE 的目的 IPv4 传输地址 (即翻译后的目的 IPv4 地址和翻译后的目的传输端口) 被设置为(W,w), 也就是收到的 IPv4 报文的源 IPv4 传输地址。

d) NAT64 会设置 (或重置) 这个 STE 计时器的值为最大会话生存时间。这个最大会话生存时间是可以配置的, 默认值至少为 UDP_DEFAULT (5 分钟), 但不能小于 UDP_MIN (2 分钟)。

5.5.2.2 UDP 连接中 IPv4 传输地址分配规则

在一个 UDP 连接中, 当需要为 IPv6 源传输地址(S',s)建立一个 UDP BIB 表项时, NAT64 分配 IPv4 传输地址时需要注意以下几点:

a) NAT64 必须支持端点无关映射方式, 即 NAT64 转换只取决于内部源 IP 地址和内部源端口, 即不管内部目的 IP 地址和内部目的端口是什么, 同一个内部源 IP 地址和内部源端口映射到同一个外部源 IP 地址和外部源端口;

b) 如果已经存在其他的 BIB 表项, 其源 IPv6 地址为 S', 并且映射为 IPv4 地址 T, 那么 NAT64 使用这个地址 T 作为映射地址。如果不存在这样的 BIB 表项, 那么 NAT64 从 IPv4 地址池中取出任意一个 IPv4 地址用作翻译后的 IPv4 地址;

c) 如果源端口 s 是一个位于 0~1023 范围内的知名端口, 那么 NAT64 会为其分配一个相同范围的映射端口 t, 如果 NAT64 在相同范围内无法提供可用的端口, 那么 NAT64 可以在其他范围分配一个可用端口;

d) 如果源端口 s 是一个位于 1024~65535 范围内的端口, 那么 NAT64 会为其分配一个相同范围的映射端口 t , 如果 NAT64 在相同范围内无法提供可用的端口, 那么 NAT64 可以在其他范围分配一个可用端口;

e) 在分配端口时, 奇数端口分配奇数外部端口, 偶数的端口分配偶数外部端口;

f) 任何情况下, 已经分配的 IPv4 传输地址 (T, t) 不能再分配给同类型 BIB 的其他表项, 但可以分配给其他类型的 BIB。

如果 NAT64 无法分配 IPv4 传输地址或创建 BIB 记录, 那么丢弃收到的报文, 然后 NAT64 发送一个代码值为 3 (地址不可达) 的 ICMPv6 目的不可达差错消息。

5.5.3 TCP 会话的更新与处理

5.5.3.1 状态定义

TCP 会话处于空闲期的生存时间至少为 2 小时 4 分钟, 因此, 为了保证资源的利用, 每一个存在的 TCP 会话都需要对应一个 TCP STE。为了实现这个目的, 在每个 TCP 会话建立后, 需要通过下面的状态机来跟踪 TCP 连接状态:

关闭状态: 此状态与 RFC793 的关闭状态相似, 表示一个虚拟的状态, 它表示没有某个五元组的状态, 所以没有连接。

IPv4 初始化状态: NAT64 收到一个 IPv4TCP 同步报文, 意味着从 IPv4 这边发起一个 TCP 连接。NAT64 正等待来自另一个方向的相匹配的 IPv6 TCP 同步报文。

IPv6 初始化状态: NAT64 收到、翻译并转发一个 IPv6TCP 同步报文, 意味从 IPv6 这边发起一个 TCP 连接。NAT64 正等待来自另一个方向的相匹配的 IPv4 TCP 同步报文。

完成建立状态: 代表一个打开的连接, 这个状态下, 数据可以在这个连接的双方向上传输。

接收到 IPv4 结束标志状态: NAT64 接收到一个 IPv4TCP 结束报文, 这个状态下, 数据仍然能够在此连接上传输, NAT64 正在等待来自另一个方向的相匹配的 IPv6 TCP 结束报文。

接收到 IPv6 结束标志状态: NAT64 接收到一个 IPv6TCP 结束报文, 这个状态下, 数据仍然能够在此连接上传输, NAT64 正在等待来自另一个方向的相匹配的 IPv4 TCP 结束报文。

接收到 IPv4 和 IPv6 结束标志状态: NAT64 收到同一个连接的 IPv4TCP 结束报文和 IPv6TCP 结束报文。此时 NAT64 还会保持该连接一小段时间, 使剩下的报文 (如, 确认报文) 能够完成双向传输。

过渡状态: 由于 NAT64 收到连接的 TCP 重置报文, 或者连接的生存时间减少至只剩下 TCP_TRANS, 使连接状态的生存时间被设置为 TCP_TRANS, NAT64 将保持该状态 TCP_TRANS 时间。如果在这段时间内 NAT64 没有接收到此连接的其他报文, 那么这个连接将会中断。

5.5.3.2 NAT64 中 TCP 状态机的处理流程

NAT64 使用状态机来描述 TCP 会话的处理。每个通过 NAT64 建立的 TCP 连接都对应着一个状态机。NAT64 引导启动后, 所有的 TCP 会话处于关闭状态。当状态机收到报文时, 可以得到输入报文的五元组与其中一个状态机相匹配。

NAT64 在 IPv6 接口收到一个包含同步标志的 TCP 分段，可以称为 IPv6 同步标志，同样的，还有以下一些标志：IPv4 同步标志，IPv4 结束标志，IPv6 结束标志，IPv4 和 IPv6 结束标志，IPv6 重置标志，IPv4 重置标志。

通过 TCP 状态机的状态变化，可以得到相应的 BIB 表项和 STE 的生成和更新处理，具体描述如图 4 所示。

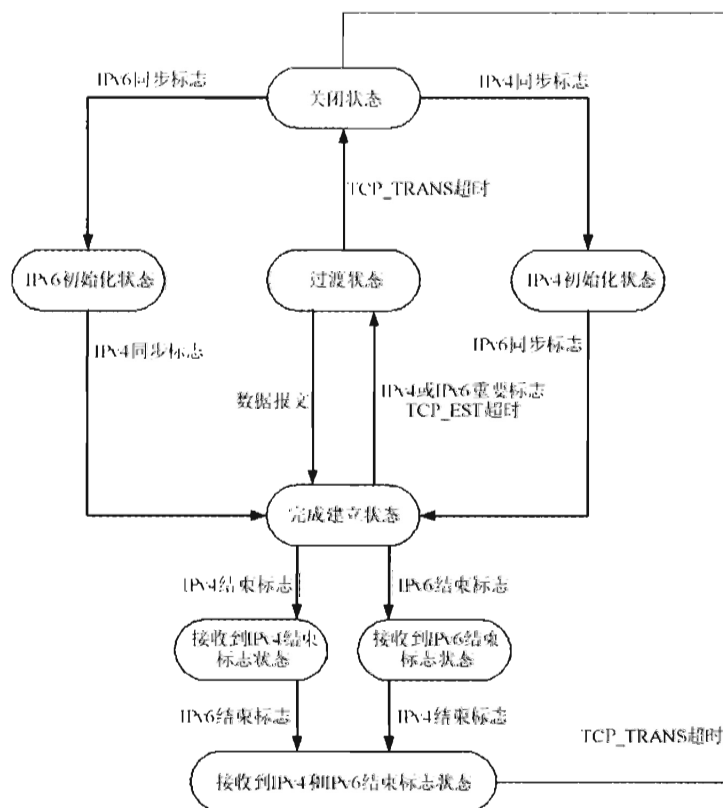


图 4 TCP 连接中 NAT64 状态机

a) 关闭状态：

当 TCP 连接从 IPv6 端发起时，如果 NAT64 收到一个 IPv6 同步报文，得到相应的输入五元组信息为 (X', x, Y', y, TCP) ，其处理过程如下：

1) NAT64 会查找匹配 IPv6 源传输地址 (X', x) 的 BIB 表项，如果相匹配的 BIB 表项不存在，那么 NAT64 会生成一个新的 BIB 表项（如果资源和策略允许），BIB 的 IPv6 传输地址设置为 (X', x) ，NAT64 分配的地址 T 与端口 t 设置为 BIB 的 IPv4 传输地址 (T, t) 。

2) 然后 NAT64 会建立一个新的 TCP STE 表项：

STE 的源 IPv6 传输地址（即源 IPv6 地址和源传输端口）设置为 (X', x) ，这两个信息可以从接收到的 IPv6 同步报文中得到。

STE 的目的 IPv6 传输地址（即目的 IPv6 地址和目的传输端口）设置为 (Y', y) ，这两个信息也可以从接收到的 IPv6 同步报文中得到。

STE 的源 IPv4 传输地址（即翻译后的源 IPv4 地址和翻译后的源传输端口）设置为(T,t)，这两个信息来自对应的 TCP BIB 表项。

STE 的目的 IPv4 传输地址（即翻译后的目的 IPv4 地址和翻译后的目的传输端口）设置为(Z,z)，其中端口号 z 的值与目的传输端口的值相同，地址 Z 则通过相应的算法从目的 IPv6 地址 y 得到，可以表示为 Z(Y')，具体算法过程见 5.5.5 小节。

STE 的生存时间设置为 TCP_TRANS。

3) TCP 会话状态迁移到 IPv6 初始化状态。

4) NAT64 会翻译并转发这个报文。

当 TCP 连接从 IPv4 端发起时，如果 NAT64 收到一个 IPv4 同步报文，得到相应的输入五元组信息为(Y,y,X,x,TCP)。如果安全策略不允许外部发起的连接，那么这个默默丢弃这个报文，否则，处理这个报文，其处理过程如下：

1) 如果输入的 IPv4 同步报文中的目的传输地址(X,x)在 TCP BIB 表项中没有存在相关匹配。那么 NAT64 会在 TCP 会话表中创建一个新的 STE：

STE 的源 IPv4 传输地址（即翻译后的源 IPv4 地址和翻译后的源传输端口）设置为(X,x)；

STE 的目的 IPv4 传输地址（即翻译后的目的 IPv4 地址和翻译后的目的传输端口）设置为(Y,y)；

STE 的源 IPv6 传输地址（即源 IPv6 地址和源传输端口）在本文档中不做说明；

STE 的目的 IPv6 传输地址（即目的 IPv6 地址和目的传输端口）设置为(Y'(Y),y)，其中端口号 y 的值与收到 IPv4 报文的源端口地址 y 相同，Y'从收到 IPv4 报文的源地址 Y 中通过相应的算法产生，可表示为 Y'(Y)，具体算法过程见 5.5.5 小节；

TCP 会话状态迁移到 IPv4 初始化状态；

STE 的生存时间设置为 TCP_INCOMING_SYN（6 秒）[RFC5382]，并且保存这个报文，这样 NAT64 不会丢弃这个报文同时也不会生成新的 BIB 记录，主要是为了支持两端同时打开 TCP 连接这种情形。

2) 如果输入的 IPv4 同步报文中的目的地址信息(X,x)在 TCP BIB 表项中存在相关匹配，那么 NAT64 会在 TCP 会话中创建一个新的 STE：

STE 的源 IPv4 传输地址（即翻译后的源 IPv4 地址和翻译后的源传输端口）设置为(X,x)。

STE 的目的 IPv4 传输地址（即翻译后的目的 IPv4 地址和翻译后的目的传输端口）设置为(Y,y)。

STE 的源 IPv6 传输地址（即源 IPv6 地址和源传输端口）设置为相关的 TCP BIB 表项的源 IPv6 传输地址。

STE 的目的 IPv6 传输地址（即目的 IPv6 地址和目的传输端口）设置为(Y'(Y),y)，其中端口号 y 的值与收到 IPv4 报文的源端口地址 y 相同，Y'从收到 IPv4 报文的源地址 Y 中通过相应的算法产生，可表示为 Y'(Y)，具体算法过程见 5.5.5 小节。

然后 TCP 会话状态迁移到 IPv4 初始化状态。

如果此时 NAT64 支持的是地址相关过滤策略，那么 STE 的生存时间设置为 TCP_INCOMING_SYN，此时主要是为了支持两端同时打开 TCP 连接时的情形；如果 NAT64 支持的是如端点无关过滤策略，那么生存设置为 TCP_TRANS。

对于其他任何属于此连接的报文，如果在 TCP BIB 中不存在相关的表项，那么默默丢弃这个报文；如果存在，并且在安全策略允许的情况下，那么翻译并转发这个报文。

b) IPv4 初始化状态：

在此状态时，如果 NAT64 收到一个 IPv6 同步报文，假设此数据包相关五元组为(X',x,Y',y,TCP)，那么设置相关 TCP STE 的生存时间为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST (2 小时) [RFC5382]，然后 NAT64 翻译、转发报文，TCP 状态迁移到完成建立状态。

如果生存时间超时，那么 NAT64 会发送一个包含之前保存的 IPv4 同步报文的类型值为 3，代码值为 3 的 ICMP 端口不可达差错报文到发送这个 IPv4 同步报文的源，并且 STE 会被删除，TCP 状态迁移到关闭状态。

对于其他报文，NAT64 翻译并转发报文，但不会改变 TCP 的状态。

c) IPv6 初始化状态：

在此状态时，如果 NAT64 收到一个 IPv4 同步报文，假设此数据包相关五元组为(Y,y,X,x,TCP)，那么设置相关 TCP STE 的生存时间为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST。然后 NAT64 翻译、转发报文，并迁移到完成建立状态。

如果生存时间超时，这个 STE 会被删除，TCP 状态迁移到关闭状态。

如果此时收到一个 IPv6 同步报文，那么这个报文会被翻译并转发，TCP STE 的生存时间被设置为 TCP_TRANS，TCP 状态不改变。

对于其他报文，如果安全策略允许，报文会被翻译并转发，但不会改变 TCP 状态。

d) 完成建立状态：

在此状态时，如果 NAT64 收到一个 IPv4 结束报文，这个报文会被翻译并转发，TCP 进入接收到 IPv4 结束标志状态。

如果收到一个 IPv6 结束报文，这个报文会被翻译并转发，TCP 进入接收到 IPv6 结束标志状态。

如果收到一个 IPv4 重置或者 IPv6 重置报文，这个报文会被翻译并转发，相关 STE 的生存时间被设置为 TCP_TRANS，TCP 迁移到过渡状态。

如果收到其他报文，报文会被翻译并转发，并且相关的 TCPSTE 的生存时间为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST，默认值为 TCP_EST。TCP 仍然保持完成建立状态。

如果生存时间超时，那么 NAT64 应该发送一个探测报文到这个 TCP 连接中的至少一个端点。这个探测报文是与这个连接相关的 TCP 分段，但是不包含任何数据，其序列号和确认序号都为 0，所有的标志字段除了确认字段都为 0，然后 TCP 迁移到过渡状态。

e) 接收到 IPv4 结束标志状态：

处于此状态时，如果 NAT64 收到一个 IPv6 结束报文，这个报文会被翻译并且转发，相关 STE 的生存时间被设置为 TCP_TRANS。TCP 状态进入接收到 IPv4 和 IPv6 结束标志的状态。

如果收到非 IPv6 结束报文时，报文会被翻译并转发，并且相关的 STE 的生存时间为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST，默认值为 TCP_EST。TCP 仍然保持接收到 IPv4 结束标志的状态。

如果生存时间超时，那么这个 STE 会被删除，TCP 状态转入关闭状态。

f) 接收到 IPv6 结束标志状态：

处于此状态时，如果 NAT64 收到一个 IPv4 结束报文，这个报文会被翻译并且转发，相关 STE 的生存时间被设置为 TCP_TRANS。TCP 状态进入接收到 IPv4 和 IPv6 结束标志的状态。

如果收到非 IPv4 结束报文时，报文会被翻译并转发，并且相关的 STE 的生存时间为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST，默认值为 TCP_EST。TCP 仍然保持接收到 IPv6 结束标志的状态。

如果生存时间超时，那么这个 STE 会被删除，TCP 状态迁移到关闭状态。

g) 接收到 IPv4 和 IPv6 结束标志状态：

所有的报文会被翻译并转发。

如果生存时间超时，相关 STE 会被删除，TCP 状态迁移到关闭状态。

h) 过渡状态：

如果收到一个非重置报文，那么这个 TCP 连接相关的 STE 的为最大会话生存时间，这个值是可以配置的，但是不能小于 TCP_EST，默认值为 TCP_EST。TCP 状态迁移到完成建立状态。

如果生存时间超时，相关 STE 会被删除，TCP 状态迁移到关闭状态。

5.5.3.1 TCP 连接中 IPv4 传输地址的分配规则

在一个 TCP 连接中，当需要为 IPv6 源传输地址(S',s)建立一个 TCPBIB 记录时，NAT64 分配 IPv4 传输地址时需要注意以下几点：

a) NAT64 必须支持端点无关映射方式，即 NAT64 转换只取决于内部源 IP 地址和内部源端口，即不管内部目的 IP 地址和内部目的端口是什么，同一个内部源 IP 地址和内部源端口映射到同一个外部源 IP 地址和外部源端口；

b) 如果已经存在其他的 BIB 表项，其 IPv6 源地址为 S'，并且映射为 IPv4 地址 T，那么 NAT64 使用这个地址 T 作为映射地址。如果不存在这样的 BIB 记录，那么 NAT64 从 IPv4 地址池中取出一个 IPv4 地址用作翻译后的 IPv4 地址；

c) 如果源端口 s 是一个位于 0~1023 范围内的知名端口，那么 NAT64 会为其分配一个相同范围的映射端口 t，如果 NAT64 在相同范围内无法提供可用的端口，那么 NAT64 可以在其他范围分配一个可用端口；

d) 如果源端口 s 是一个位于 1024~65535 范围内的端口，那么 NAT64 会为其分配一个相同范围的映射端口 t，如果 NAT64 在相同范围内无法提供可用的端口，那么 NAT64 可以在其他范围分配一个可用端口；

g) 在分配端口时，奇数端口分配奇数外部端口，偶数的端口分配偶数外部端口；

h) 任何情况下，已经分配的 IPv4 传输地址 (T,t) 不能再分配给同类型 BIB 的其他表项，但可以分配给其他类型的 BIB。

如果 NAT64 无法分配 IPv4 传输地址或创建 BIB 记录，那么丢弃收到的报文，然后 NAT64 发送一个代码值为 3（地址不可达）的 ICMPv6 目的不可达差错消息。

5.5.6 ICMP 查询会话的更新与处理

当接收到 ICMPv6 信息报文时，其源 IPv6 地址和目的 IPv6 地址分别是 X' 和 Y' ，ICMPv6 标识符为 $i1$ ，可以得到相应的三元组 $(X', Y', i1)$ ，其处理过程如下：

a) 如果本地安全策略认为应该过滤这个 ICMPv6 信息报文，那么 NAT64 默默丢弃这个报文。否则，NAT64 检索与 $(X', i1)$ 相匹配的 ICMP 查询 BIB 表项。如果这个 BIB 表项不存在，那么 NAT64 会生成一个新的 BIB 表项（如果资源和策略允许），该表项包含如下信息：

BIB 的源 IPv6 地址被设置为 X' ，也就是 IPv6 报文的源 IPv6 地址。

BIB 的 ICMPv6 标识符被设置为 $i1$ ，也就是 ICMPv6 报文的标识符。

如果已经存在其他的 BIB 表项，包含相同的源 IPv6 地址为 X' ，并且映射为 IPv4 地址 T ，那么 NAT64 使用这个地址 T 作为新 BIB 表项的 IPv4 地址。否则，从分配给 IPv4 接口的地址中选择一个分配给新 BIB 表项。

BIB 的 ICMPv4 标识符可以是任意值，只要不与现有的 ICMPv4 标识符冲突。

b) 然后 NAT 会根据得到的三元组 $(X', Y', i1)$ 检索匹配的 ICMP 查询 STE。如果没有找到匹配的记录，那么 NAT64 会生成新的 STE（如果资源和策略允许）：

STE 的源 IPv6 地址被设置为 X' ，也就是 IPv6 报文的源 IPv6 地址。

STE 的目的 IPv6 地址被设置为 Y' ，也就是 IPv6 报文的目的 IPv6 地址。

STE 的 ICMPv6 标识符被设置为 $i1$ ，也就是 IPv6 报文的标识符。

STE 的源 IPv4 地址被设置为相应的 BIB 表项的源 IPv4 地址。

STE 的 ICMPv4 标识符被设置为相应的 BIB 表项的 ICMPv4 标识符。

STE 的目的 IPv4 地址通过相应的算法从 IPv6 报文的目的 IPv6 地址 Y' 得到，可以表示为 $Z(Y')$ ，具体算法过程见 5.5.5 小节。

c) NAT64 设置（或重置）STE 的计时器的值为最大会话生存时间。这个计时器的值是可以配置的，默认情况下，最大会话生存时间为 ICMP_DEFAULT[RFC5508]。然后 NAT64 翻译并转发报文。

当接收到 ICMPv4 查询报文时，其源 IPv4 地址和目的 IPv4 地址分别是 T 和 Z ，ICMPv4 标识符为 $i2$ ，可以得到相应的三元组 $(T, Z, i2)$ ，其处理过程如下：

a) NAT64 检索 ICMP 查询 BIB 表项，该表项应包含 IPv4 地址 T 和标识符 $i2$ 的 BIB 记录，如果不存在该表项，丢弃这个 ICMPv4 查询报文。同时，NAT64 会向发送这个报文的源发送一个类型值 3，代码值为 1（主机不可达）的 ICMP 差错报文。如果 NAT64 在 IPv4 接口上进行过滤，那么 NAT64 会根据地址相关过滤策略检查是否允许接收报文。因此，NAT64 检索相关的 STE，该 STE 应包含源 IPv4 地址 T ，目的 IPv4 地址 Z ，ICMPv4 标识符 $i2$ 。如果这样的 STE，则进行后续处理。如果不存在匹配的 STE，那么丢弃这个报文，并且向发送这个报文的源发送一个类型值为 3（目的不可达），代码值为 13（管理上禁止通信）的 ICMP 差错报文。

b) 如果 ICMP 报文没有被丢弃（因为 NAT64 没有过滤或者因为报文符合地址相关过滤策略规则），那么 NAT64 会检索 STE，该 STE 应包含源 IPv4 地址 T ，目的 IPv4 地址 Z ，和 ICMPv4 标识符 $i2$ 。如果没找到相匹配的表项，那么 NAT64 会生成一个新表项（如果资源和策略允许），该表项包含如下信息：

STE 的源 IPv4 地址被设置为 T ，也就是 IPv4 报文的源 IPv4 地址。

STE 的 ICMPv4 标识符被设置为 i2，也就是 IPv4 报文的标识符。

STE 的目的 IPv4 地址被设置为 Z，也就是 IPv4 报文的目的 IPv4 地址。

STE 的源 IPv6 地址被设置为相应的 BIB 表项的源 IPv6 地址。

STE 的 ICMPv6 标识符被设置为相应的 BIB 表项的 ICMPv6 标识符。

STE 的目的 IPv6 地址通过相应的算法从 IPv4 报文的目的 IPv4 地址 Z 得到，可以表示为 Y'(Z)，具体算法过程见 5.5.5 小节。

d) NAT64 设置（或重置）STE 的计时器的值为最大会话生存时间。这个计时器的值是可以配置的，默认情况下，最大会话生存时间为 ICMP_DEFAULT。然后 NAT64 翻译并转发报文。

5.5.7 地址翻译算法

NAT64 支持多种算法实现 IPv4 与 IPv6 地址之间的相互翻译，这些算法需要满足以下条件：

a) 算法应该是可逆的，即可以通过 IPv4 地址的 IPv6 表示获得原始的 IPv4 地址。

b) 算法的输入应限制为 IPv4 地址，用于 IPv6 表示的 IPv6 前缀（Pref64::/n），这个前缀与 DNS64 中所用前缀相同，以及一些在 NAT64 中配置的固定参数（如，用作后缀的固定字符串）。如果用 n 来表示 IPv6 前缀的长度，n 必须小于或等于 96。如果 NAT64 配置了 IPv6 前缀，缺省算法必须使用这个前缀。如果没有可用的前缀，算法应使用知名前缀(64:ff9b::/96)。

NAT64 支持多种地址翻译算法，并且应默认支持以下地址翻译算法：

1) IPv4 地址到 IPv6 地址的翻译：

将 IPv6 前缀、32 位的 IPv4 地址与后缀（如果需要）连接起来，可以构成 128 位 IPv6 地址，如图 5 所示。

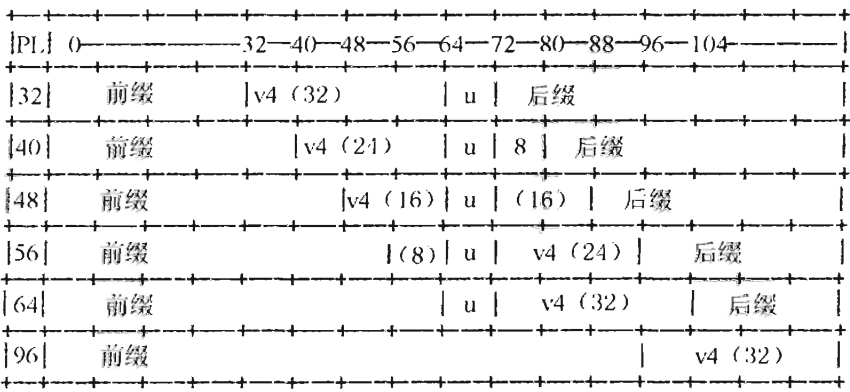


图 5 嵌入 IPv4 地址的 IPv6 地址

其中 PL 表示前缀长度。前缀为知名前缀或特定网络前缀，其中知名网络前缀的值固定为 64:ff9b::/96，特定网络前缀的长度为以下一种：32、40、48、56、64 或 96；8 位 u 是保留位，置为 0；后缀也是保留位，用于未来的扩展，置为 0。构成 IPv6 地址时应按照如下规则：

如果前缀是 32 位，那么将 IPv4 地址嵌入至第 32 到第 63 位。

如果前缀是 40 位，那么将 IPv4 地址前 24 位嵌入至第 40 到第 63 位中，后 8 位嵌入至第 72 到第 79 位。

如果前缀是 48 位，那么将 IPv4 地址前 16 位嵌入至第 48 到第 63 位中，后 16 位嵌入至第 72 到第 87 位。

如果前缀是 56 位，那么将 IPv4 地址前 8 位嵌入至第 56 到第 63 为中，后 24 位嵌入至第 72 到第 95 位。

如果前缀是 64 位，那么将 IPv4 地址嵌入至第 72 到第 103 位。

如果前缀是 96 位，那么将 IPv4 地址嵌入至第 96 到第 127 位，这种情况下形成的 IPv6 地址不包含保留位 u 以及后缀。

2) IPv6 地址到 IPv4 地址的翻译：

如果前缀长度为 96 位，那么提取 IPv6 地址的后 32 位可以得到 IPv4 地址。

如果前缀长度为其他长度，那么首先移除保留位 u ，得到一个 120 位序列，然后从前缀后面提取出 32 位 IPv4 地址。

5.6 输出元组的计算

5.6.1 计算输出元组的概述

通过翻译相应输入元组的地址、端口号或 ICMP 标识符可以计算得到输出元组。输出元组分为输出五元组和输出三元组，表示了 NAT64 与目的地之间的一个会话连接。

5.6.2 输出五元组的计算

输出五元组的计算主要针对 TCP、UDP，和包含 TCP 或 UDP 的 ICMP 差错报文。

a) 当翻译方向为 IPv6 到 IPv4 时：

1) 假设输入五元组的源 IPv6 传输地址为 (S',s) ，输入五元组的目的 IPv6 传输地址为 (D',d) ，那么输出五元组可以通过如下方式计算得到：如果 BIB 存在表项 $(S',s) \leftrightarrow (T,t)$ ，那么输出五元组的源 IPv4 传输地址为 (T,t) ；输出五元组的目的 IPv4 地址可以根据 5.5.5 节描述的算法，从输入五元组的的目的 IPv6 地址 D' 得到。

2) 输出五元组中的传输协议字段信息与输入五元组的传输协议字段信息相同。

b) 当翻译方向为 IPv4 到 IPv6 时：

1) 假设输入五元组的源 IPv4 传输地址为 (S,s) ，输入五元组的目的 IPv4 传输地址为 (D,d) ，那么输出五元组可以通过如下方式计算得到：如果 BIB 存在表项 $(X',x) \leftrightarrow (D,d)$ ，那么输出五元组的目的 IPv6 传输地址为 (X',x) ；输出五元组的源 IPv6 地址可以根据 5.5.5 节描述的算法，从输入五元组的源 IPv4 地址 S 得到。

2) 输出五元组中的传输协议字段信息与输入五元组的传输协议字段信息相同。

5.6.3 输出三元组的计算

输出三元组的计算主要针对 ICMP 查询报文和包含 ICMP 查询信息的 ICMP 差错报文。

a) 当翻译方向为 IPv6 到 IPv4 时：

假设输入三元组的信息为(S',D',i1),那么输出三元组可以通过如下方式计算得到:如果存在 BIB 表项(S',i1)←→(T, i2),那么输出三元组的源 IPv4 地址为 T, ICMPv4 标识符为 i2;输出三元组的目的 IPv4 地址可以根据 5.5.5 节描述的算法,从输入三元组的目的地址 D'得到。

b) 当翻译方向为 IPv4 到 IPv6 时:

假设输入三元组的信息为(S,D,i2),那么输出三元组可以通过如下方式计算得到:如果存在相应的 BIB 表项(X',i1)←→(D,i2),输出三元组的目的 IPv6 地址为 X',输出三元组的 ICMPv6 标识符为 i1;输出三元组的源 IPv6 地址可以根据 5.5.5 小节描述的算法,从输入三元组的源 IPv4 地址 S 得到。

5.7 报文的翻译

5.7.1 IPv4 到 IPv6 的翻译

5.7.1.1 IPv4 到 IPv6 的翻译概述

当 NAT64 收到发往 IPv6 网络的 IPv4 报文时,会将报文的 IPv4 报头翻译到 IPv6 报头,如图 2 所示。如果 NAT64 支持传输层校验和的更新,则进行更新。原先 IPv4 报文的数据部分则不进行改变,然后 NAT64 会根据翻译得到的 IPv6 报文的目的地址进行转发。

在 IPv6 中路径 MTU 发现是强制的,而在 IPv4 中是可选的。IPv6 路由器不能对报文进行分片,只有发送者可以对报文进行分片。

当一个 IPv4 节点进行路径 MTU 发现时(通过在报头中设置 DF 位),可以越过 NAT64 进行端到端的路径 MTU 发现。这时,IPv6 路由器或者 IPv4 路由器(包括 NAT64)都可能返回 ICMP 报文过大消息给发送者。当 IPv6 路由器发送这个 ICMPv6 差错报文至 IPv4 发送者时,报文在经过 NAT64 时需要翻译为相应的 ICMPv4 差错报文。因此,只有当 IPv4 报文已经分片的情况下,才会产生 IPv6 分片扩展头。

然而,当 IPv4 发送者没有设置 DF 位时,NAT64 应保证报文不超过 IPv6 网络的路径 MTU,这主要通过通过对 IPv4 报文进行分片来实现。但是 IPv6 分片扩展头在实际操作中受限于防火墙是否支持分片功能,存在着操作困难。因此,NAT64 可以提供给网络管理员一个配置功能,能够调节最小 IPv6 MTU 的阈值,该阈值可以反映网络中最小 IPv6 MTU 的真实值。这样可以减少在翻译后的报文中包含分片扩展头的可能性。

当 IPv4 发送者没有设置 DF 位时,NAT64 应包含一个分片报头表明发送者允许分片。NAT64 也可以提供一个配置功能对于非分片的 IPv6 报文不包含分片扩展头。

当 IPv4 路由器或者发送者对报文进行分片时,16 位分片标志被端到端承载,保证报文能够正确的重新组装。

NAT64 应确保属于同一个流的报文以到达 NAT64 的相同顺序离开 NAT64。

5.7.1.2 IPv4 报头到 IPv6 报头的翻译

如果 IPv4 报文没有设置 DF 位,这个报文将导致 IPv6 报文大于 IPv6 的 MTU (1280 字节),那么这个报文需要被分片,从而使翻译后的 IPv6 报文(每个分片都包含分片扩展头)小于或等于 1280 字节。例如,如果报文在到达 NAT64 之前被分片,那么分片后的 IPv4 报文除了 IPv4 报头,最大长度为 1232 字节。如果管理员已经知道网络中最小 IPv6 MTU 的真实值,NAT64 可以提供给网络管理员一个配置

功能，用来调节最小 IPv6 MTU 的值大于 1280 字节。然后通过使用如下方法，可以实现各个分片被独立翻译：

如果 DF 位被设置，并且 NAT64 下一跳接口的 MTU 小于 IPv4 报文长度加上 20，那么 NAT64 应该发送一个“需要分片”的 ICMPv4 差错消息到 IPv4 源地址。

如果 DF 位被设置，并且报文不是一个分段报文，那么 NAT64 不应该在翻译后的报文中添加分片扩展头。

IPv6 报头字段应按照如下描述进行设置：

版本：6。

流量类型：默认情况下，与 IPv4 报文中的服务类型一致。在某些情况下，NAT64 应支持忽略 IPv4 的服务类型，而将 IPv6 的流量类型设置为 0。

流标签：0。

净荷长度：该值为 IPv4 报头中的总长度值减去 IPv4 报头和 IPv4 选项的长度。

下一报头：如果 IPv4 报文中包含 ICMPv4 (1)，那么下一报头为 ICMPv6 (58)，其他情况下，协议字段应该从 IPv4 报头中复制得到。

跳数限制：该值来自于 IPv4 报头中生存时间的值。由于 NAT64 也是一个路由器，在转发报文时需要减少 IPv4 的生存时间或者 IPv6 的跳数限制。当 NAT64 减少生存时间或跳数限制时，需要检查值是否为 0，如果为 0，则发送 ICMPv4 的“生存时间超时”差错报文或者 ICMPv6 “跳数限制超限”差错报文。

源地址：从输出元组的源 IPv6 地址获得。

目的地址：从输出元组的目的 IPv6 地址获得。

对于任何出现在 IPv4 报文中的 IPv4 选项，都应该被忽略。但是如果一个未到期的源路由选项出现，那么这个报文应该被丢弃，并且发送一个“目的不可达，源路由失败”（类型值为 3，代码值为 5）的 ICMPv4 差错报文给发送者。

如果翻译时需要添加分片扩展头，除了以下几个特例，IPv6 报头字段的设置和前面描述的设置一致：

IPv6 字段：

净荷长度：该值为 IPv4 报头中的总长度值加上 8 字节的分片扩展头长度，再减去 IPv4 报头和 IPv4 选项的长度。

下一报头：分片扩展头 (44)。

分片扩展头字段：

下一报头：如果 IPv4 报文中包含 ICMPv4 (1)，那么下一报头为 ICMPv6 (58)，其他情况下，协议字段应该从 IPv4 报头中复制得到。

分片偏移：从 IPv4 报头的分片偏移中复制得到。

更多分片标志位：从 IPv4 报头的更多分片位中复制得到。

标识：该标识的低 16 位从 IPv4 报头的标识字段中复制得到，高 16 位设置为 0。

5.7.1.3 ICMPv4 报头到 ICMPv6 报头的翻译

在进行 ICMPv4 报文到 ICMPv6 报文的翻译时，由于 ICMPv6 包含一个伪报头的校验和，因此 ICMPv6 的校验和字段也需要进行翻译。

此外，应该翻译所有 ICMPv4 报文的类型字段，而且对于包含在 ICMPv4 差错报文中的 IP 报头也应该进行翻译。

以下描述了如何将各种 ICMPv4 消息进行翻译：

a) ICMPv4 查询报文：

回显和回显应答（类型值分别为 8 和 0）：调整回显和回显应答的类型值分别为 128 和 129，考虑到报文类型值的变化和 ICMPv6 的伪报头，应同时重新计算 ICMPv6 的校验和。

信息请求/应答（类型值分别为 15 和 16）：在 ICMPv6 中作废不用，默默丢弃该报文。

时间戳请求和时间戳应答（类型值分别为 13 和 14）：在 ICMPv6 中作废不用，默默丢弃该报文。

地址掩码请求/应答（类型值分别为 17 和 18）：在 ICMPv6 中作废不用，默默丢弃该报文。

ICMP 路由器通告（类型值为 9）：单跳信息，默默丢弃该报文。

ICMP 路由器请求（类型值为 10）：单跳信息，默默丢弃该报文。

未知的 ICMPv4 类型：默默丢弃该报文。

IGMP 信息：由于 IPv6 协议中的 MLD 信息是 IPv4 IGMP 信息的对等信息，因此，所有的 IGMP 信息是单跳信息，并且应该被 NAT64 默默丢弃。

标识符：如果存在标识符字段，则这个字段的值从输出三元组的标识符中复制得到。

b) ICMPv4 差错报文：

目的不可达（类型值为 3）：应按照如下描述翻译代码值，设置类型值为 1，并且考虑到报文类型值、代码值的变化和 ICMPv6 的伪报头，应重新计算 ICMPv6 的校验和。代码值的翻译如下：

代码值 0、1（网络不可达，主机不可达）：设置代码值为 0（没有到目的的路由）。

代码值 2（协议不可达）：翻译为 ICMPv6 参数问题（类型值为 4，代码值为 1），同时将指针指向 IPv6 的下一报头字段。

代码值 3（端口不可达）：设置代码值为 4（端口不可达）。

代码值 4（需要分片，但设置不分片比特）：翻译为 ICMPv6 报文过大信息（类型值为 2，代码值为 0）。由于 IPv4 报头和 IPv6 报头长度存在差异，MTU 字段也需要进行调整。

代码值 5（源路由失败）：设置代码值为 0（没有到目的的路由）。

代码值 6、7、8：设置代码值为 0（没有到目的的路由）。

代码值 9、10（管理上禁止与目标主机的通信）：设置代码值为 1（管理上禁止与目标通信）。

代码值 11、12：设置代码值为 0（没有到目的的路由）。

代码值 13（管理上禁止通信）：设置代码值为 1（管理上禁止与目标通信）。

代码值 14（主机越权）：默默丢弃该报文。

代码值 15（优先中止生效）：设置代码值为 1（管理上禁止与目标通信）。

其他代码值：默默丢弃该报文。

重定向（类型值为 5）：单跳信息，默默丢弃该报文。

交替的主机地址（类型值为 6）：默默丢弃该报文。

源端被关闭（类型值为 4）：在 ICMPv6 中作废不用，默默丢弃该报文。

超时（类型值为 11）：设置类型值为 3，代码值不变，考虑到报文类型值的变化和 ICMPv6 的伪报头，应重新计算 ICMPv6 的校验和。

参数问题（类型值为 12）：应按照如下描述翻译代码值，设置类型值为 4，并且考虑到报文类型值、代码值的变化和 ICMPv6 的伪报头，应重新计算 ICMPv6 的校验和。代码值的翻译如下：

代码值 0（指针指示错误）：设置代码值为 0（错误的报头字段），并且根据表 5 更新指针（如果没有列出原始 IPv4 指针的值，或者翻译后 IPv6 指针的值不存在，则默默丢弃该报文）。

表 5 IPv4 到 IPv6 的指针值翻译

原始IPv4指针的值		翻译后IPv6指针的值	
0	版本/互联网报头长度	0	版本/流量类型
1	服务类型	1	流量类型/流标签
2, 3	总长度	4	净荷长度
4, 5	标识	不存在	
6	标识/分片偏移	不存在	
7	分片偏移	不存在	
8	生存时间	7	跳数限制
9	协议	6	下一报头
10, 11	头部校验	不存在	
12~15	源地址	8	源地址
16~19	目的地址	24	目的地址

代码值 1（丢失一个要求的选项）：默默丢弃该报文。

代码值 2（错误长度）：设置代码值为 0（错误的报头字段），并且根据表 5 更所示新指针（如果没有列出原始 IPv4 指针的值，或者翻译后 IPv6 指针的值不存在，默默丢弃该报文）。

其他代码值：默默丢弃该报文。

未知的 ICMPv4 类型：默默丢弃该报文。

ICMP 错误净荷：如果收到的 ICMPv4 报文包含 ICMPv4 扩展[RFC4884]，那么翻译这个 ICMPv4 报文将导致翻译后的 ICMPv6 报文的长度改变。当这种情况发生时，ICMPv6 扩展的长度属性应该进行相应的调整。如果 ICMPv4 扩展超过了出接口上 ICMPv6 消息的最大长度，ICMPv4 扩展应该被简单截断。如果某些 ICMPv4 的扩展在 RFC4884 中没有定义，NAT64 将扩展视为不透明的字符串不进行翻译，导致包含 IPv4 地址的数据不会翻译成包含 IPv6 地址的数据，这会在处理 ICMP 扩展时引发问题。

5.7.1.4 ICMPv4 差错报文到 ICMPv6 差错报文的翻译

ICMPv4 差错报文和 ICMPv6 差错报文的格式存在着一些差异。如图 6 所示，包含在 ICMP 差错报文中的错误报文应该像普通 IP 报文一样进行翻译。如果翻译错误报文时改变了数据报的长度，那么应该更新外层 IPv6 报头中的总长度字段。

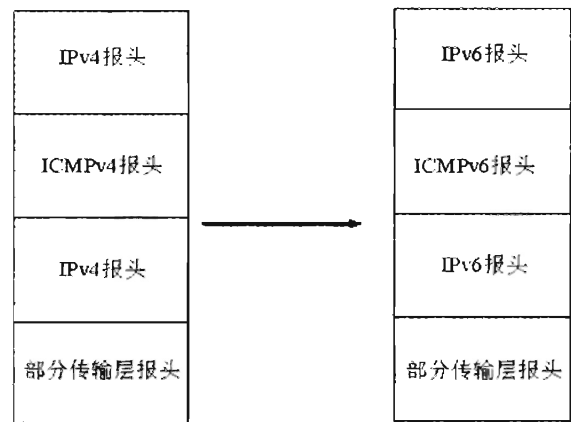


图 6 ICMPv4 差错报文到 ICMPv6 差错报文的翻译

内层 IP 报头的翻译过程与外层 IP 报头的翻译过程相同，该过程应在处理完第一个内层 IP 报头就停止。如果报文内嵌多个内层 IP 报头，则会丢弃这个报文。

5.7.1.5 ICMPv4 差错报文的产生

如果 IPv4 报文被丢弃，那么 NAT64 应该向报文的发送者返回一个 ICMPv4 差错报文，除非这个丢弃的报文自身是个 ICMPv4 报文。除非本文档和 RFC6146 中明确说明，这个返回的 ICMPv4 差错报文的类型值应该为 3（目的不可达），代码值为 13（管理上禁止通信）。NAT64 应允许管理员配置是否发送 ICMPv4 差错报文，是否对 ICMPv4 差错报文进行限速。

5.7.1.1 传输层报头的翻译

NAT64 进行地址翻译时，传输层报头也需要进行更新。

- a) 对于 TCP、ICMP 和包含校验和的 UDP 报文，源传输端口更新为输出五元组中的源传输端口，目的传输端口更新为输出五元组中的目的传输端口，校验和也需要进行相应的更新。其他字段和数据则保持不变。
- b) 对于不包含校验和的 UDP 报文，源传输端口更新为输出五元组中的源传输端口，目的传输端口更新为输出五元组中的目的传输端口。通过输出五元组得到的地址和端口，计算校验和并转发这个报文。
- c) 选择性支持其他协议（如，DCCP）。

5.7.2 IPv6 到 IPv4 的翻译

5.7.2.1 IPv6 到 IPv4 的翻译概述

当 NAT64 收到发往 IPv4 网络的 IPv6 报文时，会将报文的 IPv6 报头翻译到 IPv4 报头，如图 7 所示。移除报文原始的 IPv6 报头，并通过 IPv4 报头替代。由于 ICMPv6、TCP、UDP 和 DCCP 的报头包含覆盖 IP 报头的校验和，在翻译前需要对校验和重新计算，并且 ICMP 和传输层的报头也需要更新。由于 IPv6 与 IPv4 存在着差异，比如分段、MTU 等，这些都会影响翻译。

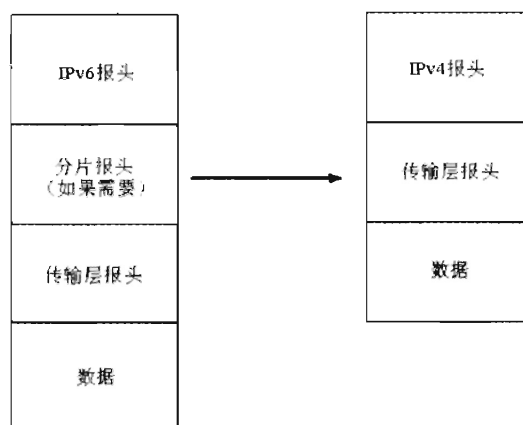


图7 IPv6 到 IPv4 的翻译

5.7.2.2 IPv6 报头到 IPv4 报头的翻译

如果 IPv6 报文没有分片扩展头，那么 IPv4 报头应按照如下描述进行设置：

版本：4。

互联网报头长度：5（不包含 IPv4 选项）。

服务类型：默认情况下，应该从 IPv6 报头中的流量类型字段复制得到。在某些情况下，NAT64 应支持忽略 IPv6 的流量类型，而将 IPv4 的服务类型设置为特殊值。

总长度：该值为 IPv6 报头中净荷长度的值加上 IPv4 报头的长度。

标识：设置为全 0。

标志：更多分片标志设置为 0，不分片标志设置为 1。

分片偏移：设置为全 0。

生存时间：生存时间的值从 IPv6 报头中的跳数限制的值复制得到。由于 NAT64 也是一个路由器，在转发报文时需要减少 IPv6 的跳数限制或者 IPv4 的生存时间。当 NAT64 减少跳数限制或生存时间时，需要检查值是否为 0，如果为 0，则发送 ICMPv4 的“生存时间超时”差错报文或者 ICMPv6 “跳数限制超限”差错报文。

协议：IPv6 分片扩展头（44）按照本节中的后续描述进行处理。ICMPv6（58）翻译为 ICMPv4（1），净荷按照 5.7.2.3 的描述进行翻译。在翻译过程中忽略 IPv6 逐跳选项（0），IPv6 路由扩展头（43）和 IPv6 目的选项（60）等 IPv6 扩展头，由于这些值在 IPv4 中没有意义。除了以上情况，IPv4 报头中协议字段的值从 IPv6 的下一报头中复制得到。

报头校验和：当得到翻译后的 IPv4 报头时，计算校验。

源地址：从输出五元组的源 IPv4 地址得到。

目的地址：从输出五元组的目的 IPv4 地址得到。

如果收到的 IPv6 报文包含一个分片扩展头，除了以下几个特例，IPv4 报头字段的设置和前面描述的设置一致：

总长度：该值为 IPv6 报头中的净荷长度的值减去 8 字节分片扩展头的长度再加上 IPv4 报头的长度；

标识：从分片扩展头中的标识字段的低 16 位复制得到。

标志：IPv4 的更多分片标志位从 IPv6 的分片扩展头中更多分片标志位复制得到；IPv4 的不分片标志设置为 0。

分片偏移：从 IPv6 的分片扩展头中的分片偏移字段中复制得到。

协议：对于下一报头为 ICMPv6（58），设置为 ICMPv4（1）；其他情况下，忽略扩展头，下一报头字段从最后一个 IPv6 报头的下一报头字段中复制得到。

如果翻译后的报文 DF 位为 1，其长度将超过下一跳接口的 MTU，那么 NAT64 会丢弃这个报文并且发送一个 ICMPv6 报文过大的差错报文（类型值为 2，代码值为 0）给 IPv6 发送者，并且调整了 ICMPv6 差错信息的 MTU。

5.7.2.3 ICMPv6 报头到 ICMPv4 报头的翻译

如果使用的是一个非校验和中性的翻译地址，那么在翻译 ICMPv6 到 ICMPv4 时，应该更新其校验和字段，因为 ICMPv6 包含着像 TCP 和 UDP 那样的伪报头。

此外，应该翻译所有 ICMPv6 报文的类型字段，而且对于包含在 ICMPv6 差错报文中的 IP 报头也应该进行翻译。

以下描述了如何将各种 ICMPv6 消息进行翻译：

a) ICMPv6 信息报文：

回显和回显应答（类型值分别为 128 和 129）：调整回显和回显应答的类型值分别为 8 和 0，考虑到报文类型值的变化和不包含 ICMPv6 伪报头，应同时调整 ICMPv4 的校验和。

MLD 组播侦听查询/报告/完成（类型值分别为 130、131 和 132）：单跳信息，默默丢弃该报文。

标识符：如果存在标识符字段，那么这个字段的值从输出三元组的 ICMP 标识符中复制得到。

邻居发现信息（类型值从 133 到 137）：单跳信息，默默丢弃该报文。

未知的消息类型：默默丢弃该报文。

b) ICMPv6 差错报文：

目的不可达（类型值为 1）：应按照如下描述翻译代码值，设置类型值为 3，考虑到报文类型值/代码值的变化以及不包含 ICMPv6 伪报头，应同时调整 ICMPv4 的校验和。代码值的翻译如下：

代码值 0（没有到目的的路由）：设置为 1（主机不可达）。

代码值 1（管理上禁止与目标通信）：设置为 10（管理上禁止与目标主机通信）。

代码值 2（超出源地址范围）：设置为 1（主机不可达）。

代码值 3（地址不可达）：设置为 1（主机不可达）。

代码值 4（端口不可达）：设置为 3（端口不可达）。

其他代码值：默默丢弃该报文。

报文过大（类型值为 2）：翻译为 ICMPv4 目的不可达差错报文（设置类型值为 3，代码值为 4），考虑到报文类型值的变化和不包含 ICMPv6 伪报头，应同时调整 ICMPv4 的校验和。由于 IPv4 报头和 IPv6 报头长度存在差异，MTU 字段也需要进行调整。

超时（类型值为 3）：设置类型值为 11，考虑到报文类型值的变化和不包含 ICMPv6 伪报头，应重新计算 ICMPv4 的校验和，代码值不变。

参数错误（类型值为 4）：按照如下描述翻译类型值和代码值，考虑到报文类型值和代码值的变化以及不包含 ICMPv6 伪报头，应重新计算 ICMPv4 的校验和。类型值和代码值的翻译如下：

代码值 0（错误的报头字段）：设置类型值为 12，代码值为 0，并且按照表 6 对指针进行更新（如果没有列出原始 IPv6 指针的值，或者翻译后 IPv4 指针的值不存在，则默默丢弃该报文）。

表 6 IPv6 到 IPv4 的指针值翻译

原始IPv6指针的值		翻译后IPv4指针的值	
0	版本/流量类型	0	版本/互联网报头长度，服务类型
1	流量类型/流标签	1	服务类型
2, 3	流标签	不存在	
4, 5	净荷长度	2	总长度
6	下一报头	9	协议
7	跳数限制	8	生存时间
8~23	源地址	12	源地址
24~39	目的地址	16	目的地址

代码值 1（下一报头类型无法识别）：翻译为 ICMPv4 协议不可达（设置类型值为 3，代码值为 2）。

代码值 2（IPv6 选项无法识别）：默默丢弃该报文。

未知的差错信息：默默丢弃该报文。

ICMP 错误净荷：如果收到的 ICMPv6 报文包含 ICMPv6 扩展[RFC4884]，那么翻译这个 ICMPv6 报文将导致翻译后的 ICMPv4 报文的长度改变。当这种情况发生时，ICMPv6 扩展的长度属性应该进行相应的调整。如果某些 ICMPv6 的扩展在 RFC4884 中没有定义，NAT64 将扩展视为不透明的字符串不进行翻译，导致包含 IPv6 地址的数据不会翻译成包含 IPv4 地址的数据，这会在处理 ICMP 扩展时引发问题。

5.7.2.4 ICMPv6 差错报文到 ICMPv4 差错报文的翻译

ICMPv4 差错报文和 ICMPv6 差错报文的格式存在着一些差异。如图 8 所示，包含在 ICMP 差错报文中的错误报文应该像普通 IP 报文一样进行翻译。翻译 ICMP 差错报文中的错误报文时很可能改变报文的长度，因此应该更新外层 IPv4 报头中的总长度字段。

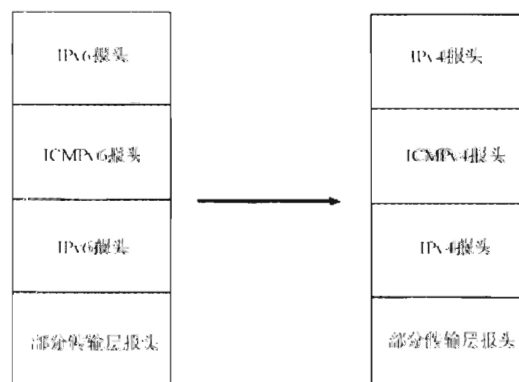


图 8 ICMPv6 差错报文到 ICMPv4 差错报文的翻译

5.7.2.5 ICMPv6 差错报文的产生

如果 IPv6 报文被丢弃，那么 NAT64 应该向报文的发送者返回一个 ICMPv6 差错报文，除非这个丢弃的报文自身是个 ICMPv6 报文。除非本文档和 RFC6146 中明确说明，这个返回的 ICMPv6 差错报文的类型值应该为 1（目的不可达），代码值为 1（管理上禁止与目标通信）。NAT64 应允许管理员配置是否发送 ICMPv6 差错报文，是否对 ICMPv6 差错报文进行限速。

5.7.2.6 传输层报头的翻译

NAT64 进行地址翻译时，传输层报头也需要进行更新。

a) 对于 TCP、ICMP 和包含校验和的 UDP 报文，源传输端口更新为输出五元组中的源传输端口，目的传输端口更新为输出五元组中的目的传输端口，校验和也需要进行相应的更新。其他字段和数据则保持不变。

b) 对于不包含校验和的 UDP 报文，源传输端口更新为输出五元组中的源传输端口，目的传输端口更新为输出五元组中的目的传输端口。通过输出五元组得到的地址和端口，计算校验和并转发这个报文。

c) 选择性支持其他协议（如 DCCP）。

5.8 发卡行为的处理

如果翻译后的报文的地址是一个分配给 NAT64 自身的 IPv4 地址，那么这个报文是一个发卡报文，这种发卡报文的处理过程如下：

- a) 输出五元组变为输入五元组；
- b) 该报文被认为是从 NAT64 的输出接口上收到的；
- c) 过滤和更新相关绑定和会话信息；
- d) 翻译报文并发送到目的地。

6 安全性

6.1 端到端的安全性

任何保护 IP 报头的协议与 NAT64 都不相容，这意味着当使用认证报头或者 ESP 时，端到端的 IPSec 检验可能失败。这是网络层翻译机制存在的缺陷，当然，端到端的 IPSec 保护也可以通过使用 UDP 封装来重建。

在使用 NAT64 进行通信时，NAT64 对地址、端口号及协议进行转换，NAT64 破坏了端到端的通信模型，互联网上的主机直接与 NAT64 设备通信，而并非与专用网络中的实际的主机通信，这样，在公网 IPv4 地址被复用的情况下，为用户溯源带来困难。在部署 NAT64 技术的网络中，为了在公网 IPv4 地址被复用的情况下实现用户溯源，增强网络中端到端的透明性，减少网络通信的安全隐患，NAT64 可以与 AAA 及网管平台相配合实现在 NAT 环境下的用户溯源，其中 NAT64 根据映射表项的状态生成相应的 NAT 日志信息；AAA 及网关平台作用为存储用户的验证、授权和记账（AAA）记录信息，并

且接收 NAT 日志信息。NAT64 将 NAT 日志信息上传到 AAA 及网管平台，AAA 及网管平台将此日志信息与存储的用户 AAA 记录信息结合生成用户溯源数据库；这样，在需要进行用户溯源时，可以通过查找用户溯源数据库实现用于用户溯源，从而解决部署在 NAT64 技术的网络中由于复用 IPv4 地址带来的用户溯源问题。

6.2 过滤

NAT64 使用从 IPv6 端到 IPv4 端的报文流来创建绑定状态。在创建绑定状态时，NAT64 必须支持“端点无关映射”，也就是说：对于任何一个源为($S',s1$)，目的为($Pref64::D1,d1$)的 IPv6 报文，NAT64 会产生一个到($S1,s1v4$),($D1,d1$)的外部映射，对于后面来自源($S'1,s1$)，目的为($Pref64::D2,d2$)的 IPv6 报文会产生一个到($S2,s2v4$),($D2,d2$)的外部映射，如果 NAT64 支持“端点无关映射”，那么对于任何($D2,d2$)，其外部映射中($S1,s1v4$)=($S2,s2v4$)。

NAT64 的安全性受到过滤配置和过滤行为的影响，而不是地址映射行为。按照当前说明，NAT64 只要求基于五元组进行过滤。在某些情况下，如静态配置的映射，使攻击者非常容易猜到。攻击者不再需要猜测其他字段，就可以得到通过 NAT64 的报文。一旦这样的流量被目的地丢弃，NAT64 就不能减少针对内部网络的带宽或 CPU 的攻击。为了避免这种情况，NAT64 可以跟踪 TCP 报文的序列号来核实分段的正确顺序，特别是 TCP 协议的同步标志和结束标志。

中华人民共和国通信行业标准
运营级网络地址翻译（NAT）技术要求
NAT64

YD/T 3186—2016

*

人民邮电出版社出版发行
北京市丰台区成寿寺路 11 号邮电出版大厦

邮政编码：100064

北京康利胶印厂印刷

版权所有 不得翻印

*

开本：880×1230 1/16

2017 年 6 月第 1 版

印张：2.25

2017 年 6 月北京第 1 次印刷

字数：57 千字

15115 • 1216

定价：25 元

本书如有印装质量问题，请与本社联系 电话：(010)81055492