

# 路由设备 IPv6 网络隧道功能驱动研究

蒋建峰<sup>1,2</sup>

(1. 苏州工业园区服务外包职业学院, 江苏 苏州 215123;

2. 南京邮电大学 计算机学院, 江苏 南京 210003)

**摘要:**因特网的发展导致 IP 地址的资源耗尽, IPv6 网络将逐渐取代 IPv4 网络的主体地位。IPv4 网络向 IPv6 网络过渡过程中, 出现了很多 IPv6 网络的孤岛, 隧道技术是实现 IPv6 孤岛之间联通的关键过渡技术。IPv6 隧道技术是将 IPv6 报文作为要发送给隧道另一端的 IPv4 报文的净荷数据, 然后经由 IPv4 骨干网络传输。在网络数据的转发过程中, 基于硬件的数据转发效率要远高于软件驱动的数据转发效率。依据 IPv6 隧道技术设计了一套驱动流程, 指导隧道报文在转发芯片上的转发。介绍了系统结构图以及 NP 芯片的转发流程和驱动指令, 基于三层单播转发将数据包在隧道中封装, 在 OUTINFO 表项内添加隧道标记和信息, 将各种指令信息下发 NP 芯片, 生成 FIB 转发表指导数据转发方案。实验表明该驱动流程可以很好地指导 6to4 隧道报文正常转发, 提高网络的数据转发速率, 降低网络延迟, 特别是当数据流量增大到 300 Mbps 以上时, 网络延迟和抖动性能可以提升 100% ~ 400%, 具有较强的实用性。

**关键词:**隧道; 6to4 中继; NP 芯片; 驱动指令; 转发信息库

**中图分类号:** TP393

**文献标识码:** A

**文章编号:** 1673-629X(2022)08-0129-06

doi: 10.3969/j.issn.1673-629X.2022.08.021

## Research on Function Drive of IPv6 Network Tunnel of Routing Equipment

JIANG Jian-feng<sup>1,2</sup>

(1. Suzhou Industrial Park Institute of Services Outsourcing, Suzhou 215123, China;

2. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** The development of the Internet has led to the exhaustion of IP address resources, and IPv6 networks will gradually replace the dominant position of IPv4 networks. During the transition from the IPv4 network to the IPv6 network, many isolated islands of the IPv6 network have appeared, and the tunnel technology is the key transition technology to realize the connectivity between the IPv6 isolated islands. The IPv6 tunnel technology uses IPv6 packets as the payload data of the IPv4 packets to be sent to the other end of the tunnel, and then transmits them via the IPv4 backbone network. In the process of network data forwarding, hardware-based data forwarding efficiency is much higher than software-driven data forwarding efficiency. Based on the IPv6 tunnel technology, a set of driving procedures is designed to guide the forwarding of tunnel messages on the forwarding chip. In this study, we introduce the system structure diagram and the forwarding process and driving instructions of the NP chip. Based on the three-layer unicast forwarding, the data packet is encapsulated in the tunnel, the tunnel mark and information are added to the OUTINFO entry, and then various instruction information is issued to the NP chip to generate the FIB forwarding table to guide the data forwarding scheme. Experiments show that the driving process can well guide the normal forwarding of 6to4 tunnel messages, increase the data forwarding rate of the network, and reduce network delay. When the data stream increases to more than 300 M, the network delay and jitter performance can be improved by 100% ~ 400%.

**Key words:** tunnel; 6to4 relay; NP chip; drive instruction; forward information database

## 0 引言

随着智能终端、移动设备的不断增加, IPv4 网络的地址空间已经面临耗尽的危险。IPv6 网络已经在运营商核心网、教育网的核心设备上成功部署。但是

在很长一段时间内 IPv4 网络仍然占据着一席之地, 因此产生了大量的 IPv6 孤立网络群, 在下一代通信网络中 IPv6 网络与 IPv4 网络共存是一个常态。

基于业务的融合, IPv4 网络与 IPv6 网络无法相互

收稿日期: 2021-09-13

修回日期: 2022-01-18

基金项目: 江苏省博士后研究基金(2018K009B); 江苏省专业带头人高端研修项目(2020GRFX074); 江苏省青蓝工程项目(202010)

作者简介: 蒋建峰(1983-), 男, 硕士, 副教授, CCF 会员(87087M), 研究方向为网络技术、虚拟化、云计算技术。

分离,当前双栈,隧道技术能够很好地在 IPv4 网络上对 IPv6 业务的承载,保证业务的共存和过渡,已定义的隧道技术种类很多,主要包括手工配置隧道、兼容地址自动配置隧道、6over4、6to4 隧道等<sup>[1-3]</sup>。

国内外学者对于 IPv6 网络过渡技术的研究很多,主要有三种主流的过渡技术:双栈、网络地址转换和隧道技术。而这三种过渡技术都是基于软件实现数据包的封装,这就要求网络设备具有良好的 CPU 计算性能,而在数据包的封装过程中会影响设备的数据转发效率。由于硬件条件的限制,通过硬件驱动的方法来提升隧道数据包的转发效率的研究还很少。

当前利用隧道技术进行数据转发的行为,多数网

络设备厂家只是注重软件的驱动,虽然使用软件的驱动能够解决数据的封装,但是对于网络的延迟和通信速率没有很好的性能保障。该文通过驱动指令指导数据发送,在性能方面有明显的优势。

## 1 隧道技术

### 1.1 手工隧道

手工隧道<sup>[4-5]</sup>是实现 IPv6 孤岛连接的一个基本方法,通过手工指定隧道的源和目的,在 IPv4 主干网络中建立一条永久虚链路。手工隧道主要有 GRE 隧道和 IPv6 手动隧道。手工隧道转发的原理如图 1 所示。

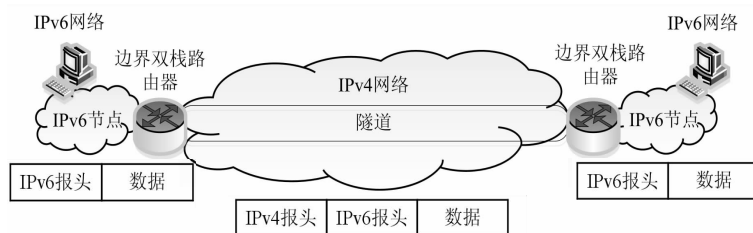


图 1 手工隧道原理

### 1.2 自动隧道

自动隧道是目前实现 IPv6 孤岛连接的主要隧道技术,包括 IPv4 兼容 IPv6 隧道、ISATAP 隧道<sup>[6-8]</sup>、6to4 隧道、6 to 4relay 等。

### 1.3 6to4 隧道

6to4 隧道是一种自动连接隧道<sup>[9-10]</sup>,通过在 IPv6 地址中嵌入 IPv4 地址信息,自动获取隧道的目的方,6to4 地址格式如表 1 所示。

表 1 6to4 地址格式

3	13	32	16	64 bits
001	0x0002	V4DDR	SLAID	Interface

6to4 隧道的地址是一种特殊的 IPv6 地址,6to4 地址以 2002 为固定的前 16 位前缀,紧接着的 32 位是由 IPv4 地址转换而来的地址,其格式为:2002:abcd:efgh::子网号::接口 ID/64。6to4 隧道可以实现利用 IPv4 网络完成 IPv6 网络的互连,克服了 IPv4 兼容 IPv6 自动隧道使用的局限性。

6to4 隧道的地址只能是以 2002 前缀的网络,所以如果要和其他的 IPv6 网络通信,必须有一台 6to4 路由器作为网关转发到 IPv6 网络的报文,这台路由器就叫做 6to4 中继(6to4 relay)路由器。6to4 中继路由器负责转发去往其他 IPv6 网络的数据,所以,当边缘设备需要和外界 IPv6 网络通信时,必须要配置到底中继的静态路由。

因此,可以看出 6to4 隧道不仅可以克服手动隧道配置的点对点单一性,而且其扩展隧道 6to4relay 能够与 IPv6 网络进行互联,所以该文主要对 6to4 隧道的原理以及驱动实现进行研究。

## 2 系统结构图及芯片转发流程

驱动是平台操作系统和硬件通信的程序,可以说相当于硬件的接口,操作系统只能通过这个接口,才能控制硬件设备的工作,基于硬件的数据转发比软件操作的速度要快很多。隧道驱动模块系统结构如图 2 所示。

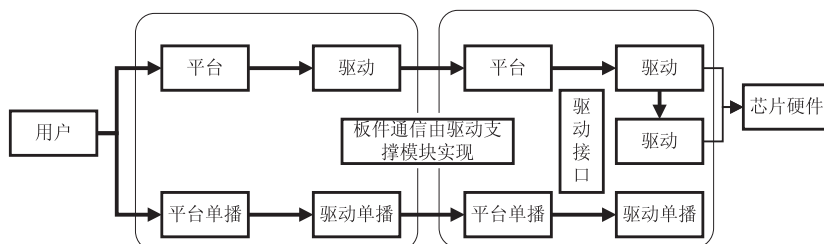


图 2 系统结构

平台模块负责与用户交互,当用户进行配置时,创

建隧道命令时,平台模块将这个指令传递给驱动相关

模块,主控板和接口板的平台是一套代码,维护相同的资源。

驱动的实现设计主控板和接口板,需要两套代码对应不同的处理方式,维护的资源也不相同,主控板主要负责数据层面的处理,接口板负责转发层面的处理。平台下发指令给驱动时先与主控板的驱动模块进行交互,主控板驱动进行相应处理后同步下发指令给各个接口板进行处理。最终驱动模块将相关信息下发芯片后,隧道报文从接口板上的端口进入后芯片就会正确指导报文的转发。

NP 芯片 (network processor) 编程模式简单,可以很方便地通过微码编程进行实现,是该研究所使用的芯片,NP 芯片三层单播转发流程如图3所示。

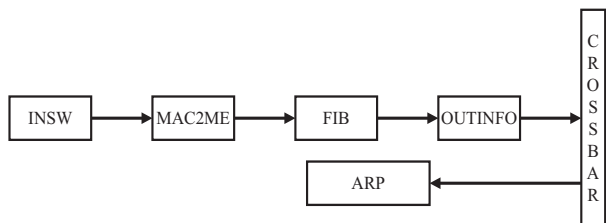


图3 NP 芯片单播三层转发流程

(1) 报文从入接口进入芯片,首先查询 INSW (in logical switch) 表,对于 IPv6 报文,INSW 中 L3 和 IPv6 相关标记置 1,允许报文进行三层转发。

(2) 报文在芯片中继续查询 MAC2ME 表项,如果报文中以太头的 mac 地址与接口的 mac 相同,说明报文是要本机处理走三层转发。

(3) 使用报文 IP 头中的 DIP 作为 KEY 查询 FIB (forward information base) 表项,进行最长匹配查找,得到的 KEY 为 OUTINFO 的 ID,如果是等价路由则会得到一个 BASEID 和 ECMP 标记,进行等价负载分担。

(4) 通过 OUTINFO ID 查询 OUTINFO (out interface info) 表项,这个表项中存储着报文的出端口信息以及 ARP INDEX。

(5) 报文出 crossbar 后,使用 ARP INDEX 查询 ARP 表项,得到目的 MAC 地址,封装 MAC 头后从接口转发出去。

### 3 系统关键技术

#### 3.1 芯片隧道转发流程

隧道技术的关键是封装与解封装过程<sup>[11-12]</sup>,在三层单播转发基础上实现隧道功能即在入隧道时可以对报文进行封装,由于 OUTINFO 表项内存大,易于扩展,所以在 OUTINFO 表中可以添加隧道标记以及隧道信息,如果为隧道报文就会以封装的新的 IP 头的目的 IP 重新查找 IPv4 FIB 表进行正常的转发,具体如图4所示。

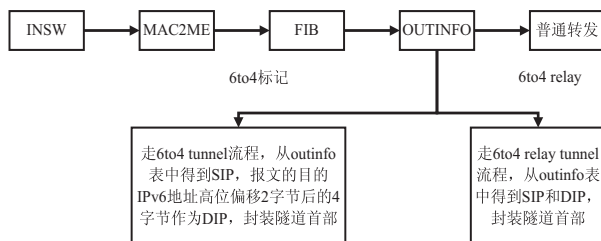


图4 入隧道报文转发流程

出隧道流程是对报文的解封装过程,在 MAC2ME 后增加对报文类型判断,如果为隧道报文,就会以 sip+dip+protocol+payload 为 key 查询 TNLEND 表项,如果是在本机终结则执行解封装过程,随后进行正常的 IPv6 报文转发,如图5所示。

KEY: SIP+DIP+protocol+Payload type



图5 出隧道报文转发流程

#### 3.2 驱动隧道模块实施方案

##### 3.2.1 初始化阶段

主控板主要负责隧道资源池的初始化,隧道计数的维护,隧道引用者的记录,资源池的初始化状态以及分配回收过程如图6所示。

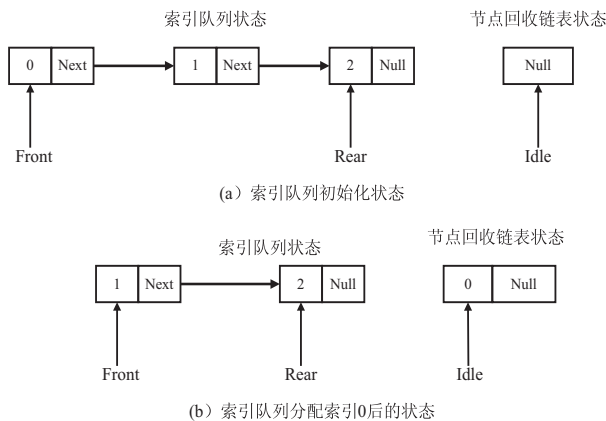


图6 隧道资源索引分配

接口板主要记录隧道属性相关信息,并对这些信息进行维护、更新、删除,下发芯片硬件指导芯片对报文进行转发。

接口板维护的信息包括:

(1) 隧道基本信息。

```
{
    隧道类型;
    隧道目的 IP;
    隧道源 IP;
    OUTINFOID;
    TNLEND 标记;
}
```

隧道驱动信息是以设备支持的最大隧道数目为

组下标,在初始化时申请相应的内存,以后有隧道创建删除更新时直接对驱动隧道信息进行操作。

### (2) TNLEND 信息。

```
{
    隧道目的 IP;
    隧道源 IP;
    隧道协议类型;
    负载类型;
    引用计数;
}
```

TNLEND 信息以链表形式存储,因为多个隧道可能会共用一个 TNLEND,所以初始化时只申请首节点,后续再进行添加删除更新等操作。

### (3) 6to4 relay 信息。

```
{
    6to4 ID;
    6to4Relay ID;
    DIP;
    OUTINFOID;
}
```

由于 6to4 隧道可以扩展出多个 6to4 relay 隧道,6to4 relay 与 6to4 隧道不同的信息在于 DIP 以及 OUTINFO 信息,所以需要建立两者的映射关系,在初始化时,以设备最大支持隧道数目为值,为每个隧道都创建一个链表,但只申请首节点内存,后续有 6to4 relay 隧道创建时再加入链表。

### 3.2.2 隧道创建事件响应

(1) 用户创建 6to4 隧道时,配置隧道的 SIP、DIP、隧道类型、隧道的 IP 地址,主控板平台模块会将这些信息保存下来,然后主控板的平台隧道模块会通知主控板的驱动隧道模块隧道的创建,驱动会从主控板的隧道资源池分配一个索引并且回填到平台隧道信息中,隧道使用计数加一,并且将这个隧道资源的引用者设置为隧道模块。

(2) 主控板的平台隧道模块将隧道信息同步给接口板,接口板的平台隧道模块再通知驱动隧道的建立,驱动隧道模块将平台带来的信息经过处理储存在驱动信息中,并且申请 OUTINFO 资源,将相关信息如 6to4 隧道标示以及 SIP 下发到芯片的 OUTINFO 信息中并且 ID 保存在驱动隧道信息中。

TNLEND 的下发条件比较严格,TNLEND 下发需要隧道为 UP 状态,隧道 UP 的条件如下:

(1) 自动隧道会用 SIP 检查配置该 IP 的接口是否为 UP 状态,如果为 UP,则隧道 UP,手动隧道还需要检查 DIP 是否可达。

(2) 当两条隧道配置相同的 SIP 时,有一个先达到 UP 条件,另一条隧道就不能 UP。

隧道 UP 后,会将 SIP、DIP、隧道协议类型、负载类型作为 KEY 下发到 TNLEND 表中,Result 为隧道终结解封装隧道 IP 头;由于 TNLEND 是以 HashTree 表形式存储在芯片内存上,所以具有相同 KEY 的报文会查询到同一张 TNLEND 表,对于驱动也可以将具有相同属性的隧道 TNLEND 信息保存在驱动 TNLEND 链表中的同一个节点中,只是将引用计数加一。

表项下发好后,还不能实现 6to4 隧道功能,需要在设备上配置一条静态路由,目的 ip 为 2002::掩码为 16 位,下一条为 6to4 tunnel,这样驱动单播模块就会下发一条 fib 表项到芯片,且 fib 表项的 Result 为 6to4 隧道申请的 OUTINFOID,这样就将路由与隧道关联起来,当有目的 IP 地址为 6to4 地址时的报文查询 fib 时,会进入隧道转发流程。

### 3.2.3 Relay 隧道创建事件响应

6to4 relay 隧道比较特殊,由于平台不区分 6to4 隧道与 6to4 中继隧道,所以并不能创建隧道类型为 6to4 relay 的隧道,所以 6to4 relay 隧道创建的流程如下:

(1) 6to4 relay 隧道的创建需要 2001::/64 位网段的用户进行通信,配置一条静态路由,目的 IP 为 2001::/16,下一跳为在连接该网段的 6to4 relay 中继路由器的 6to4 地址下一跳,由于下一跳地址为非直连网段,所以会进行路由迭代使用下一跳的 IP 作为目的 IP 查询路由表,得到下一跳为 6to4 隧道。

(2) 主控板平台模块会通知驱动单播模块下发一条带有隧道标记的路由,这时需要驱动自己判断是否为中继路由,判断条件为路由下一跳地址是 6to4 地址,路由目的地址不是 6to4,路由带有隧道标记,路由下一跳出接口指向 6to4 隧道 ID。当满足以上条件时,驱动单播模块会通知驱动隧道模块进入 6to4 relay 分支。

(3) 主控板的驱动隧道模块响应单播模块的隧道创建事件,分配一个隧道索引给单播模块,并且将这个索引的引用者设置为单播模块,之所以要为每个隧道资源设置占用者是因为有可能会下面这种时序问题情况。由于单播模块和隧道模块处于不同的任务队列,在板件通信的时候可能会出现由于任务异步通信原因发生资源抢占,最终在删除 6to4 relay 隧道时会将 IP 隧道保存好的驱动软件表项以及下发的硬件表项全部删除,导致平台收到隧道创建成功的应答,并且在用户侧成功配置隧道,但是隧道功能却是没有实现,如图 7 所示。

所以为每个隧道资源标记使用者,如果有在申请隧道资源时,在主控板分配资源后,会检查该资源的引用者是否为本模块,如果不是则返回平台信息,要求重新申请隧道资源,这就是重刷机制。

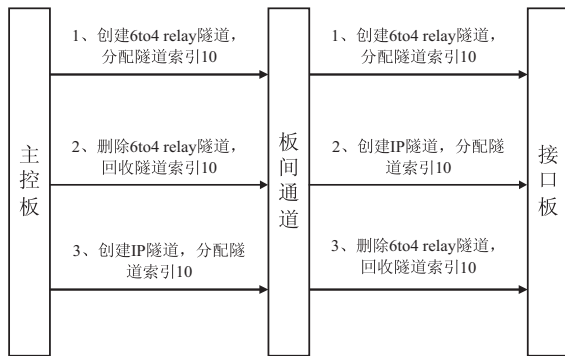


图7 任务异步通信机制导致资源抢占示意图

主控板的驱动单播模块将路由信息以及隧道索引同步给接口板的驱动单播模块,然后接口板的驱动单播模块通知驱动隧道模块 6to4 relay 隧道的建立,并将新申请的隧道索引以及相关关联的 6to4 隧道索引带给驱动隧道模块。

(4) 驱动隧道模块会为 6to4 relay 隧道申请 OUTINFO 资源,然后以相关联的 6to4 隧道信息中的 SIP 和路由下一跳的 6to4 地址高位偏移两字节后的 4

字节作为 DIP 以及 6to4 relay 标记下发 OUTINFO,将 6to4 relay 信息加入到驱动 6to4 relay 链表中,便于维护,TNLEND 表项下发过程与 6to4 隧道 TNLEND 类似。

(5) 当驱动隧道模块将 6to4 relay 创建好后,会通知单播模块,单播模块这时会将路由表项下发芯片,并且 FIB 表项中的 key 为 6to4 relay 申请的 OUTINFO ID。这时 6to4 relay 隧道功能完成,当与 2001::/64 网段用户通信时,查询 fib 表会进入 6to4 relay 隧道转发流程<sup>[4]</sup>。

## 4 实验结果

RouterA 为 6to4 路由器,网络地址使用 IPv6 6to4 地址。RouterB 是 6to4 中继路由器,连接到网络 2001::/16。需要在 RouterA 和 RouterB 之间配置 6to4 隧道,使 Hosts 可以与网络中的 IPv6 Hosts 通信,网络拓扑如图 8 所示。

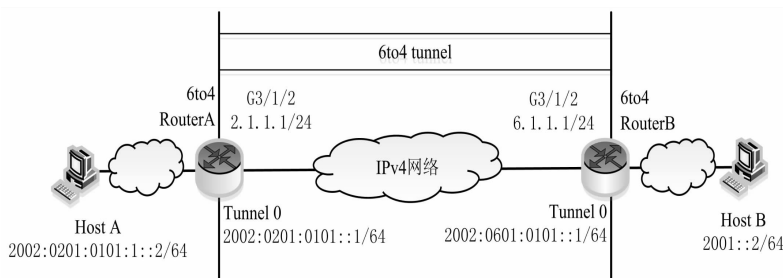


图8 6to4 relay 组网图

设备主要配置部分命令如下:

```
<RouterA>System-view
[RouterA] ipv6
[RouterA - GigabitEthernet3/1/2] ip address 2.1.1.1 255.255.0
[RouterA - GigabitEthernet3/1/2] ipv6 address 2002:0201:0101:1::1/64
[RouterA] interface tunnel 0
[RouterA - Tunnel0] ipv6 address 2002:0201:0101::1/64
[RouterA - Tunnel0] source GigabitEthernet3/1/2
[RouterA - Tunnel0] tunnel-protocol ipv6-ipv4 6to4
[RouterA] ipv6 route-static 2002:0601:0101::64 tunnel 0
[RouterA] ipv6 route-static ::0 2002:0601::1
```

完成配置后,验证配置结果如下所示,Host A 可以 Ping 通 Host B。

```
D: \>ping6 -s 2002:201:101::2 2001::2
Pinging 2001::2
From 2002:201:101:1::2 with 32 bytes of data:
Reply from 2001::2 bytes=32 time=13ms
```

通过仿真软件<sup>[13-15]</sup>连续抓取通信报文统计,当数据发送速率比较低的时候,网络数据包的延迟,语音的

抖动影响很小;当数据包发送速率较大时,通过软件算法实现隧道技术的网络数据延迟较大,而且语音的抖动比较严重,通过硬件驱动算法之后能改善数据包的延迟,并且保证了语音和视频的延迟抖动,充分保证了网络的利用率。

如图 9 和图 10 显示,当发送的数据包数量明显增加时,由硬件驱动的数据转发延迟明显较小,而且随着数量的增加,软件驱动的数据转发延迟增幅较大。实验结果表明,说明该驱动方案能够很好地提升网络性能。

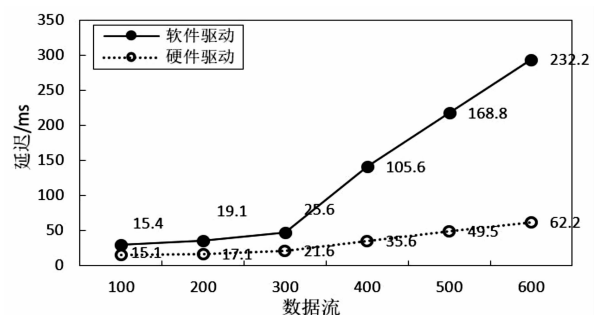


图9 网络延迟性能比较

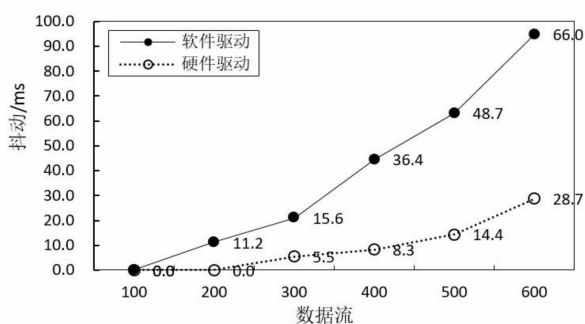


图 10 网络抖动性能比较

## 5 结束语

随着运营商和企业 IPv6 网络的部署,IPv4 和 IPv6 网络将在很长一段时间内共存。IPv6 隧道实现了 IPv4 海洋中 IPv6 孤岛的连接。当 IPv6 网络进行大规模部署时,隧道技术是实现通信网络的关键过渡技术,它将 IPv6 报文封装在 IPv4 报文中实现骨干网络的连接,并且 6to4 隧道能够实现隧道目的 IPv4 地址的自动匹配。依据 IPv6 隧道技术设计了一套驱动流程,基于 NP 芯片实现了驱动的设计指导隧道报文转发。实验结果显示,驱动能够很好的指导报文的转发,并且在网络延迟和抖动方面有着显著的性能优势。

### 参考文献:

- [1] 黄 腾,郑 凯,姜孝伟. 基于网络设备驱动的 IP 隧道应用软件[J]. 电子技术与软件工程,2019(3):21.
- [2] 贾 茹,胡曦明,李 鹏,等. 面向“新工科”的 IPv6 网络 ISATAP 隧道安全实验教学研究[J]. 中国信息技术教育,2020(22):92-98.
- [3] KUMAR C R,GOYAL H. IPv4 to IPv6 migration and performance analysis using GNS3 and Wireshark[C]//2019 international conference on vision towards emerging trends in communication and networking (ViTECoN). India: IEEE, 2019.
- [4] 周 兰. IPv6 无线传感网网络层协议研究[D]. 南京:南京邮电大学,2013.
- [5] 刘清华. 企业 IPv4/IPv6 网络 ISATAP 隧道技术的仿真实现及性能分析[J]. 成都大学学报:自然科学版,2018,37(3):291-295.
- [6] 纪德伟. IPv6 隧道技术实现要点分析[J]. 信息系统工程,2018(7):122.
- [7] HAMARSHEHA. Deploying IPv4-only connectivity across local IPv6-only access networks[J]. IETE Technical Review,2019,36(4):398-411.
- [8] ABDULLA S A. Survey of security issues in IPv4 to IPv6 tunnel transition mechanisms[J]. International Journal of Security and Networks,2017,12(2):83-102.
- [9] SHAH J L,PARVEZ J. Impact of IPSec on real time applications in IPv6 and 6to4 tunneled migration network[C]//International conference on innovations in information, embedded and communication systems. Coimbatore:IEEE,2015:1-6.
- [10] 吴念达. CMNET 城域网 IPv6 过渡技术及部署方案研究[D]. 长春:吉林大学,2019.
- [11] 叶洪涛. 基于 Teredo 隧道的 IPv4-IPv6 过渡机制研究分析[J]. 计算机产品与流通,2019(11):287-288.
- [12] XU Jianqiang,LIU Yaohua. Research on IPv6 network construction and application in higher vocational colleges[C]//2020 IEEE 4th information technology, networking, electronic and automation control conference (ITNEC). Chongqing: IEEE,2020:598-601.
- [13] 曾丽君,闵 芳. 基于隧道技术的 IPv6 企业网的仿真设计[J]. 数字技术与应用,2015(5):38-39.
- [14] JIANG J,LIU Z,GU R. 6to4 tunnel drive research and realization on high-end routers[J]. Advanced Materials Research,2013,838-841:3268-3272.
- [15] 李伟波,杨国良,刘 承. 通过睿哲 6aaS 实现数据中心应用服务 IPv6 快速升级[J]. 科研信息化技术与应用,2018,9(1):92-96.