

# MIPv6 在分布式体系结构路由器中的设计与实现

郑超, 赵靓, 何红永, 郭云飞

(解放军信息工程大学国家数字交换系统工程技术研究中心, 郑州 450002)

**摘要:** 提出了移动IPv6 (MIPv6) 协议在分布式体系结构路由器上的实现方案。结合MIPv6 协议的需求与分布式体系结构路由器的特点, 分别提出了快速通道MIPv6 实现方案和慢速通道MIPv6 实现方案, 比较了两种方案的优缺点。基于慢速通道MIPv6 实现方案进行了测试, 测试结果表明, 慢速通道MIPv6 方案可以简单可靠地实现移动IPv6 功能。

**关键词:** 分布式体系结构路由器; 移动IPv6; 家乡代理

## Design and Implementation of MIPv6 in Distributed Router

ZHENG Chao, ZHAO Liang, HE Hong-yong, GUO Yun-fei

(National Digital Switching System Engineering and Technology Research Center, PLA Information Engineering University, Zhengzhou 450002)

**【Abstract】** This paper presents implementation of mobile IPv6 in distributed router. Taking both function requirement of MIPv6 protocol and structure of distributed router, the fast channel MIPv6 solution and the slow channel MIPv6 scheme are proposed, both their advantage and disadvantage are compared. The slow channel MIPv6 scheme is tested. Test results show that the slow channel MIPv6 scheme ensures router perform mobile IPv6 function simply and credibly.

**【Key words】** distributed router; mobile IPv6; home agent

### 1 概述

随着互联网业务的进一步发展以及移动电话、PDA 和笔记本等移动数据通信终端数量的增加, 将会有越来越多的互联网用户希望能在任何地点、以任意接入方式(有线或无线)、用固定的 IP 地址和网络配置参数连接到 Internet 或企业网络, 并且在移动过程中不中断正在进行的通信。移动 IP 的设计目标就是移动节点在改变网络接入点时不必改变其 IP 地址, 能够在移动过程中保持通信的连续性, 对上层协议保持透明性。IETF 在 1996 年就制定了基于 IPv4 的移动 IP 协议, 即移动 IPv4, 但是, 移动 IPv6 在继承了移动 IPv4 协议众多特性的同时, 还解决了三角路由问题, 而且随着 IPv4 地址资源日渐紧缺、向 IPv6 网络过渡已成必然趋势, 基于 IPv6 的移动 IP 协议(移动 IPv6), 简称 MIPv6, 受到业内的一致看好, 有望成为推动下一代互联网技术应用的一个亮点。

移动 IPv6 协议<sup>[1]</sup>定义了 3 个实体: 移动节点, 通信对端节点和家乡代理。本文将重点讨论移动 IPv6 协议中家乡代理功能的实现, 家乡代理实际上就是移动节点家乡链路上的一个路由器。通常路由器按照性能可以粗略分成高性能路由器和中低端路由器。为了消除处理能力和存储带宽这两个瓶颈, 高性能路由器采用分布式转发结构, 这种结构可以实现线速转发。分布式体系结构路由器最主要的特点是数据平面与控制平面相分离<sup>[3]</sup>, 即采用分布式硬件转发、集中式控制与处理, 其基本构成为: 线路接口板, 主控单元和交换网络。其中, 主控单元是路由器的控制中心, 负责运行协议、生成硬件转发数据包所需要的转发表, 并下发给各线路接口板。线路接口板收到数据包后, 硬件实现转发表的线速查找, 将其经高速交换网络送往相应的线路接口板输出。运行在主控单元的主控软件由嵌入式 Linux 实时操作系统及运行在操作系

统之上的各种基本协议和路由协议组成, 本文将讨论移动 IPv6 在高性能分布式体系结构路由器中的设计与实现。

### 2 实现 MIPv6 协议的功能需求与关键问题

基于 Linux 开放源码, MIPv6 模块的功能主要是在 Linux 内核中实现的, 作为 MIPv6 协议中的家乡代理节点, 协议处理的主要工作就是对绑定更新消息的处理, 家乡代理如果收到绑定更新消息, 需要对要绑定的家乡地址进行重复地址检测(DAD), 检测成功的地址在本地建立绑定缓存列表。除此之外, 主控还要建立邻居发现(ND)代理功能等, ICMP 协议处理模块也需要做适当的修改, 以便增加对移动 IPv6 的支持。

在 MIPv6 协议中, 移动节点与通信对端的通信方式有两种: 非优化路由方式和优化路由方式。当通信对端的绑定缓存中不存在与移动节点的绑定条目时, 即没有进行通信注册时, 移动节点与通信对端的通信必须通过家乡代理中转, 家乡代理在移动节点的家乡链路上截取发给该移动节点的分组, 对它们进行 IPv6 封装, 然后通过隧道转发到移动节点的转交地址, 这就是非优化路由方式。如果是移动节点向通信对端进行过通信注册, 双方就可以直接通信, 不需家乡代理中转, 这种优化路由方式不属于本文的讨论范围, 本文只讨论家乡注册的非优化路由方式, 即家乡代理需要完成隧道功能。

但是, 作为家乡代理的路由器如果是高性能分布式体系结构路由器, 那么数据流的走向问题就成为实现 MIPv6 的关

**基金项目:** 国家“863”计划基金资助项目(2005AA121210)

**作者简介:** 郑超(1981-), 男, 硕士研究生, 主研方向: 网络协议分析; 赵靓, 硕士、助教; 何红永, 硕士、高级工程师; 郭云飞, 教授、博士生导师

**收稿日期:** 2006-11-19 **E-mail:** zhengchao@mail.ndsc.com.cn

键问题。高性能路由器采用分布式硬件转发、集中式控制处理的分布式结构,在这种结构下,线路接口板上收到的协议报文要被送到主控单元进行处理,而收到的数据报文只需转发模块进行输出,即协议报文通过慢速通道<sup>[3]</sup>进行处理,而数据报文通过快速通道<sup>[3]</sup>处理,这样一来,MIPv6的数据流走向问题将成为关键问题。

### 3 MIPv6 的实现方案

#### 3.1 快速通道 MIPv6 实现方案

快速通道是指分布式体系结构路由器中数据报文所通过的转发-交换-调度-输出的处理路径。快速通道 MIPv6 实现方案就是指 MIPv6 数据报文仍然通过快速通道进行处理,在线路接口板上进行必要的封装解封装,这种方案保留了分布式体系结构路由器的数据流走向,即保证了数据流的处理速度,该实现方案如图 1 所示。

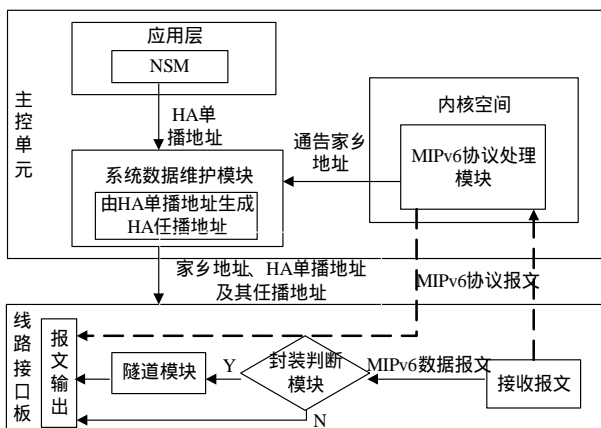


图 1 在线路接口板上实现隧道功能的方案

线路接口板收到 MIPv6 协议报文中交主控单元处理,主控单元的内核空间运行着 MIPv6 协议报文的处理模块,并将处理后的 MIPv6 协议报文下发给线路接口板。在线路接口板上,需要封装解封装的 MIPv6 协议报文和 MIPv6 数据报文要被封装判断模块分拣出来,即需要封装判断模块来判断哪些报文是需要封装解封装的,需要封装解封装的报文要被送往隧道模块处理,处理之后的报文再输出,而不需要隧道封装解封装的报文可直接输出。其中,封装判断模块区分需要封装的 MIPv6 报文,可以依靠如下方法:主控单元的内核空间将成功绑定后的移动节点家乡地址实时通告给系统数据维护模块,再由系统数据维护下发到线路接口板上,根据移动节点的家乡地址,封装判断模块就可以将 MIPv6 报文中需要封装解封装的报文分拣出来。

在线路接口板上实现隧道功能必须用硬件来实现,这种方法虽然保证了数据流的处理速度,对于已经成形却没有实现隧道功能的线路接口板来说,重新设计一块板子无疑是代价过大的,那么如何解决这个问题呢?

#### 3.2 慢速通道 MIPv6 实现方案

慢速通道<sup>[3]</sup>是指分布式体系结构路由器中协议报文通过从线路接口板上交主控单元的处理路径。为了避免改变线路接口板硬件结构的代价,可以考虑让 MIPv6 的数据报文走慢速通道,送往主控单元进行处理,即在主控单元中实现隧道功能。但是,在主控单元上实现隧道功能存在哪些呢?由于线路接口板收到数据报文以后,直接交给转发模块进行输出,这样需要封装解封装的 MIPv6 数据报文无法到达主控单元的隧道模块。现在要解决的问题就是如何把 MIPv6 数据

报文送到主控单元中去。

要想让线路接口板把收到 MIPv6 数据报文交付给主控单元,可以把移动节点的家乡地址配置到线路接口板上,即主控单元根据 MIPv6 协议生成移动节点家乡地址和转交地址绑定缓存,并将成功绑定后的移动节点家乡地址向系统数据维护实时通告,再由系统数据维护下发线路接口板,通过这样的配置,线路接口板可以将目的地址是指定地址的 IPv6 数据报文中交给主控单元。对于需要上交的 MIPv6 协议报文,可以在支持家乡代理功能的接口板上配置家乡代理地址以及该地址对应的任播地址,其中,家乡代理地址是在初始化阶段由主控应用层 NSM 模块下发给系统数据维护,再由系统数据维护生成任播地址(生成规则参见文献[2]),然后将家乡代理地址及其任播地址一并下发给接口板,这样就可使接口板将所有需要家乡代理处理的 MIPv6 协议报文转交主控处理。该方案如图 2 所示。

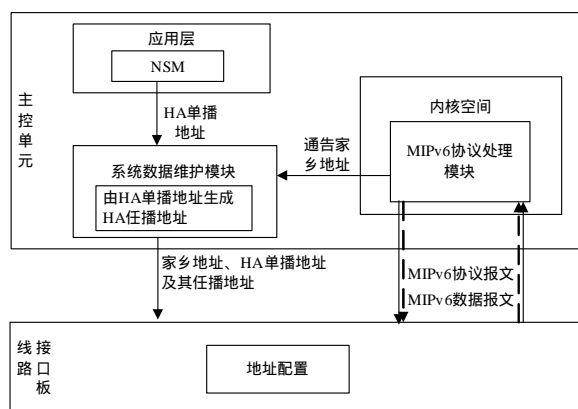


图 2 在主控板上实现隧道功能的方案

该方案可描述如下:主控内核空间运行 MIPv6 报文的处理、移动节点家乡地址和转交地址绑定缓存的维护等功能,可利用 Linux 开放源代码来实现这些功能;此外,内核还要将成功绑定后的移动节点家乡地址实时通告给系统数据维护模块,再由系统数据维护下发到线路接口板,使得接口板能够将 MIPv6 数据报文中交主控进行处理。同时,接口板上配置家乡代理地址及该地址对应的任播地址,这样,需要家乡代理处理的 MIPv6 协议报文将被线路接口板接收并转交主控处理。线路接口板还要将主控生成的 MIPv6 协议报文、隧道封装/解封装后的 MIPv6 数据报文转发给移动节点或通信对端。

该方案建立的模型可以描述为:主控内核对系统数据维护的输出——移动节点的家乡地址;系统数据维护对接口板的输出——移动节点家乡地址及其对应的任播地址;接口板与主控内核的交互——MIPv6 协议报文及数据报文。

该方案具体实现时的数据流如图 3 所示。最初(移动节点尚未在家乡代理中注册过),接口板只截获目的地址是家乡代理地址及其任播地址的分组并送交主控。分组到达 IP 层时,要进行 ND 代理检测。文献[1]中指出,在支持移动 IP 功能的路由器中,必须支持代理 ND 功能。由于代理 ND 功能主要是在 IP 层数据报文的接收过程中完成的,因此 IP 层收到 IPv6 报文后,检测 IPv6 报文是否代理节点地址的 ND 请求消息(即 ICMP 报文,其 code 域值为 135),如果是就直接交本地协议栈处理;如果不是则查找路由,根据 IPv6 报文目的地址判断:如果是需要本地处理的数据报文就直接发送本地协议栈处理,如果目的地址不是本地地址则进行路由查表并转发输出。

交由本地处理的 IPv6 报文,进一步分流,在扩展头分析处理后,交由不同的协议模块处理。

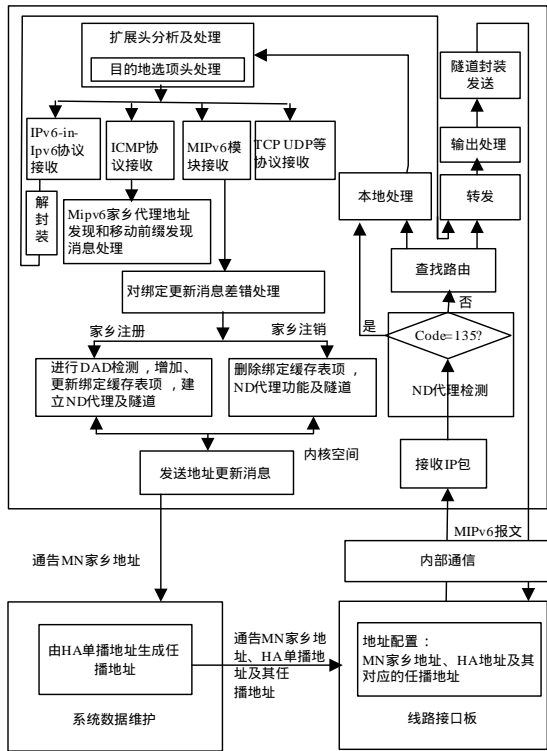


图3 MIPv6数据流

在判断收到的 IPv6 报文是 MIPv6 协议报文时,交由 MIPv6 协议模块进行处理。作为家乡代理节点,只可能处理绑定更新消息以及与“动态家乡代理地址发现机制”相关的协议报文。在移动节点知道家乡代理的 IP 地址(单播地址)的情况下发送绑定更新消息,家乡代理收到绑定更新消息后,对要绑定的移动节点家乡地址进行 DAD 检测,检测成功的地址在本地要建立移动节点家乡地址和转交地址的绑定缓存项,并将成功绑定后的移动节点家乡地址通过 Linux 内核提供的 Netlink 机制中的 Rtnetlink 消息向系统数据维护实时通告。此时,配置了移动节点家乡地址、家乡代理地址及其任播地址的接口板就可以接收所有与移动 IPv6 相关的协议报文和数据报文并上交主控进行处理。

在 ICMP 协议处理过程中,需要调用 MIPv6 模块的 ICMP 处理子模块,主要是针对 MIPv6 家乡代理地址发现和移动前缀发现消息处理。而 IPv6-in-IPv6 隧道协议也要配合 MIPv6 模块将移动节点通过家乡代理发往通信对端的数据报文和部分协议报文(如 HOTI、COTI 等)进行解封装处理,再交转发输出模块输出。

由上述分析可见,慢速通道 MIPv6 实现方案虽然增加了主控数据流的注入,但是,对于已经成形却没有实现隧道功能的线路接口板来说,避免了重新设计线路接口板的代价,仅仅对相关部件增加适当的软件功能和进行正确配置即可,其中包括在系统数据维护的内核中增加移动 IP 协议处理功能、增加与系统数据维护模块的交互以及对线路接口板进行地址实时配置等。

#### 4 功能测试

现对慢速通道 MIPv6 方案进行测试,为了完成该方案的功能测试,先构建如图 4 所示的简单测试环境。MN 是移动节点, CN 是通信对端, R<sub>1</sub>、R<sub>2</sub>、R<sub>3</sub> 是路由器,其中, R<sub>1</sub> 是具有

移动 IPv6 功能的分布式体系结构路由器。路由器 R<sub>1</sub> 的 Eth0 网卡所连接的链路作为家乡链路,家乡链路的网络前缀是 3ffe:3240:8007:1004::/64,路由器 R<sub>3</sub> 的 Eth3 网卡所连接的链路作为外地链路,外地链路的网络前缀是 3ffe:3240:8007:1002::/64,移动节点在这两个链路之间切换,实现移动 IPv6 功能。

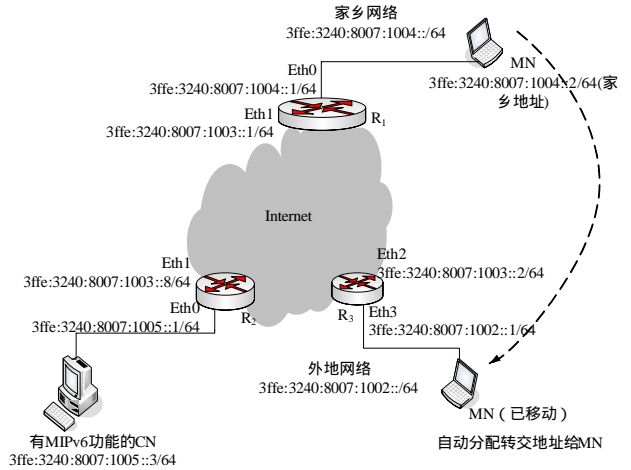


图4 方案测试环境

在测试中使用的协议分析软件是 Ethereal。下面对测试作一个简单的描述。首先,查看移动节点有没有收到路由通告,如果没有收到路由通告,移动 IP 将不能正常工作。当移动节点在外地链路时,通过收到的路由通告确定自己发生了移动,然后自动生成一个转交地址,并通过邻居发现消息确定这个地址的唯一性,接着就向家乡代理发送绑定更新消息,家乡代理收到绑定更新消息后会发送绑定应答消息,移动节点收到绑定应答之后,绑定已建立,此时从通信节点 Ping 移动节点就可以正常收发数据包。

```

Internet Protocol version 6
Version: 6
Traffic class: 0x00
Flowlabel: 0x000000
Payload length: 40
Next header: IPv6 routing (0x2b)
Hop limit: 254
Source address: 3ffe:3240:8007:1004::1
Destination address: 3ffe:3240:8007:1002:20c:ceff:fe16:6104
Routing Header, Type 2
Next header: Mobile IPv6 (0x87)
Length: 2 (24 bytes)
Type: 2
Segments left: 1
Home Address: 3ffe:3240:8007:1004::2 (3ffe:3240:8007:1004::2)
Mobile IPv6
Payload protocol: IPv6 no next header (0x3b)
Header length: 1 (16 bytes)
Mobility Header Type: Binding Acknowledgement (6)
Reserved: 0x00
Checksum: 0xb7d4
Binding Acknowledgement
Status: Binding Update accepted (0)
0... .. = Key Management Compatibility (K) flag: No Key Management Mobility
Sequence number: 0
Lifetime: 250 (1000 seconds)
Mobility Options
PadN: 4 bytes

0000 00 0c ce 16 61 04 00 0c 85 9c 3d 8e 86 dd 60 00 .....a.....
0010 00 00 00 28 2b fe 3f fe 32 40 80 07 10 04 00 00 ...(+?.28.....
0020 00 00 00 00 01 3f fe 32 40 80 07 10 02 02 0c .....?.28.....
0030 ce ff fe 16 61 04 87 02 02 01 00 00 00 00 3f fe .....a.....?..
0040 32 40 80 07 10 04 00 00 00 00 00 00 02 3b 01 28.....;..
0050 06 00 b7 d4 00 00 00 00 fa 01 02 00 00 .....

```

图5 软件抓屏和具体数据包

根据移动节点与家乡代理之间的所有测试报文可以看出:在慢速通道 MIPv6 方案简单可靠地实现了高性能分布式 (下转第 130 页)