

一种基于 VLAN 划分的 MIPv6 仿真实验环境

刘 轶, 张宝强, 肖凯平

(西安交通大学计算机系, 西安 710049)

摘 要:介绍了一种移动 IPv6 实验系统。该系统建立在局域网环境中, 移动节点漫游切换的实现采用了不同于传统手工拔插线或修改 ESSID 的方法, 即通过在交换机中划分 VLAN 来实现。介绍了实验系统的构成、移动节点切换以及系统控制程序的实现。实验结果表明, 与传统切换方式相比, 采用基于 VLAN 的切换方式不仅使 MN 的切换更加方便、快速, 而且使 MN 切换过程中的丢包和时延也都得到了改善。

关键词: MIPv6; VLAN; 网络仿真

MIPv6 Testbed Environment Based on VLAN-partition

LIU Yi, ZHANG Baoqiang, XIAO Kaiping

(Department of Computer, Xi'an Jiaotong University, Xi'an 710049)

【Abstract】This paper presents a mobile IPv6(MIPv6) testbed, which is built in a LAN. Handover of mobile node(MN) is implemented by re-partitioning VLAN in switches, which is different from traditionally using “pull out and plug in” of cable or modification of wireless ESSID. The organization, handover implementation and control program of the testbed are introduced. Emulation results show that compared with traditional handover methods, the VLAN-partition based method is more convenient and faster.

【Key words】MIPv6; VLAN; Network emulation

未来的网络协议必须要支持移动性, 允许主机在不同的 IP 网络间无缝隙地漫游。因此, IETF 的 Mobile IP 工作组在研究 MIPv4(Mobile IPv4)的同时, 投入了大量的精力重点研究 MIPv6(Mobile IPv6)。MIPv6 是在 MIPv4 的基础上发展而来, 较之 MIPv4, MIPv6 具有许多新的特性, 如邻居发现、地址自动配置、优化路由、QoS 改善等。这使得对 MIPv6 的研究已成为当前 Internet 技术研究的热点之一。

目前, 在已有的 MIPv6 实验系统中, 大多使用了 PC 路由器(多网卡)作为家乡代理和接入路由器, 采用无线网络或有线局域网组成家乡链路和外地链路, 而移动节点的切换均使用拔插线(有线情形)或者改变节点的 ESSID(无线情形)两种方式, 可控性不强, 操作很不方便。本文提出了一种基于虚拟局域网(Virtual LAN, VLAN)划分搭建 MIPv6 实验系统的新方法。在此方法中, 家乡代理和接入路由器都只需一块网卡; 采用基于端口的 VLAN 划分方式, 无需考虑物理连线, 即可简洁、灵活地在交换机上划分出家乡链路以及外地链路; 移动节点的切换, 只需通过动态改变节点所属的 VLAN 便可方便快速地完成, 而且切换过程中的丢包和时延也都因此得到了改善。

1 Mobile IPv6 简介

移动 IP 协议设计的主要目的是为了保证无论移动节点是在家乡链路还是在外地链路, 都可以通过家乡地址被访问, 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。但由于移动 IPv4 是在 IPv4 的基础上提出的, IPv4 的固有弱点也必然存在于移动 IPv4 中, 如地址空间严重不足、“三角路由”的路由方案、IPSec 只是可选功能等。为了更好地满足社会发展的需求, IETF 提出了下一代网络协议 IPv6 协议。IPv6 的主要特性对于未来移动无线网络的发展至关重要, 这

些特性包括: 足够多的 IP 地址, 安全数据包头的实现, 目的选项提高了路由效率、地址自动配置等。IETF 于 2004 年 6 月正式推出了 RFC3775 和 RFC3776, 它们定义了移动 IPv6(MIPv6, Mobile IPv6)及其安全性方面的要求, 为移动 IPv6 的实际应用走出了关键性的一步。

在 MIPv6 协议中定义了 3 类重要的功能实体: 移动节点(Mobile Node, MN), 家乡代理(Home Agent, HA), 通信对端节点(Correspondent Node, CN)。这种分类只是概念上的, 并非对 IPv6 节点的分类定义。比如, 移动节点同时可能是其它节点的通信对端节点, 家乡代理也可能是移动节点的通信对端节点等。

当移动节点在家乡链路时, 其工作方式如位置固定的主机, 移动 IPv6 不需要进行任何特别的操作。当漫游到外地链路后, 利用 IPv6 的“邻居发现机制”, 移动节点可从接入路由器(Access Router, AR)的路由通告中, 得到“在线子网前缀”以及“路由器地址”, 进而完成地址的无状态自动配置, 获得一个或多个转交地址。当然, 也可以通过有状态的地址自动配置获得转交地址。

移动节点在外地链路获得转交地址后, 首先向家乡代理发送“绑定更新”申请注册, 家乡代理则发送“绑定确认”对移动节点的请求进行应答。这样, 在家乡代理上, 建立了移动节点的家乡地址和转交地址的“绑定”。随后, 家乡代理使用“代理邻居发现”机制, 在家乡链路上截取目的地址

基金项目:国家自然科学基金资助项目(60673180); 华为科技基金资助项目(YJCB2006032WL)

作者简介:刘 轶(1968 -), 男, 博士、副教授, 主研方向: 计算机网络, 嵌入式系统; 张宝强、肖凯平, 硕士生

收稿日期: 2006-02-16 **E-mail:** liuyi97@263.net

是移动节点家乡地址的所有 IPv6 分组,通过隧道将它们转发到移动节点的转交地址。

在向家乡代理注册的同时,移动节点也向通信对端发送“绑定更新”,告知通信对端当前的转交地址,使通信对端能够修改自己的“绑定缓存”,进而可以使用“第 2 类路由头”,直接发送分组到移动节点,完成“优化路由”,减轻网络负担。

2 仿真实验环境的构成

2.1 实验系统的拓扑结构及配置

实验系统的逻辑拓扑结构如图 1(b)所示,而实际的物理拓扑如图 1(a)所示。为了简化实验系统,家乡代理 HA 同时也是通信对端 CN。在实验系统中,使用具有可划分 VLAN 功能的二层交换机,采用基于端口的 VLAN 划分方式。“控制台”运行我们设计的系统控制程序,其主要功能是自动配置、管理该交换机,使 VLAN 的划分和 MN 的切换更加容易和快速。

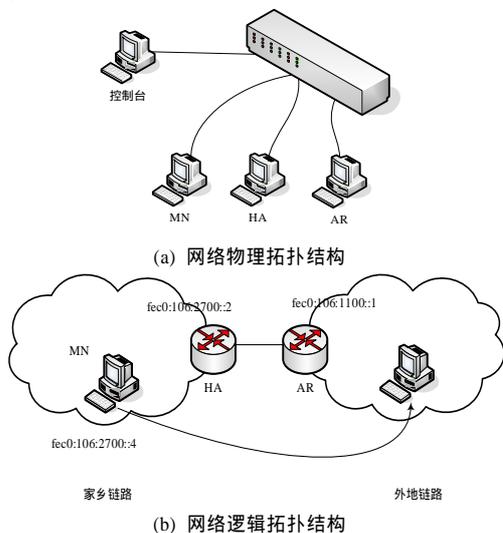


图 1 MIPv6 仿真实验环境

VLAN 技术允许网络管理者将一个物理的 LAN 逻辑地划分成不同的广播域(或称虚拟 LAN, VLAN),VLAN 与物理上形成的 LAN 有着相同的属性。但由于它是逻辑地而不是物理地划分,因此同一个 VLAN 内的各个工作站不需要被放置在同一个物理空间内,即这些工作站不一定属于同一个物理 LAN 网段。一个 VLAN 内部的广播和单播流量都不会转发到其他 VLAN 中。利用 VLAN 的这一特性,将 HA 置于 VLAN A 中,AR 置于 VLAN B 中。当 MN 位于 A 时,移动节点处于家乡链路;MN 位于 B 时,移动节点处于外地链路。HA、AR 都安装了“路由公告守护进程(Router ADvertisement Daemon, RADVD)”模块,它根据 IPv6 的邻居发现协议,周期性地组播路由公告。通过获取该路由公告,MN 能够自动检测并配置其 IPv6 地址以及默认路由。

由于二层交换机没有路由模块,为了使分别处于 VLAN A、VLAN B 内的终端能相互通信,做如下设置:

- (1)HA、AR 同时属于 VLAN C;
- (2)HA、AR 上的网卡设置为支持转发;
- (3)修改 HA、AR 的路由表,使得 VLAN A 内所有发往 VLAN B 的数据包由 HA 转发,VLAN B 内所有发往 VLAN A 的数据包由 AR 转发。

2.2 移动节点切换的实现

当 MN 在家乡链路和外地链路之间切换时,MIPv6 才能发生作用。当前的切换方式多是采用拔插线(有线情形)或者

改变节点的 ESSID(无线情形)两种方式。在我们搭建的实验系统中,采用了一种基于 VLAN 的切换方式。初始时,移动节点 MN 位于家乡链路,属于 VLAN A,如果要仿真其漫游至外地链路时的情形,可通过重新划分 VLAN,将 MN 从 HA 所在的 VLAN 切换到 AR 所在 VLAN。如图 2 所示,先将 MN 从 VLAN A 中删除,然后再将其加入 VLAN B,这样便可快速完成 MN 从家乡链路到外地链路的切换。实际操作时,只需在控制台处输入 MN 将要到达的目的 VLAN 号,系统控制程序通过 SNMP 协议自动完成以上过程,即“一键切换”。

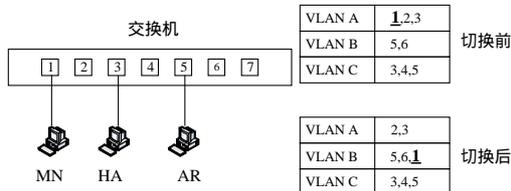


图 2 MN 的切换过程(1, 2, ...为交换机端口编号)

2.3 系统控制程序的设计与实现

系统控制程序主要有 2 个功能:(1)管理和维护资源情况,主要是交换机的使用信息;(2)通过 SNMP 协议完成交换机上 VLAN 的划分。该程序的结构如图 3 所示。

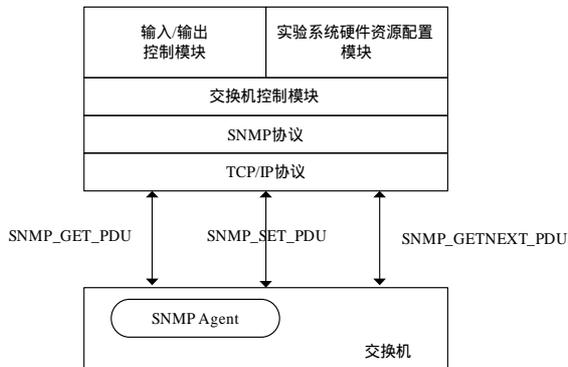


图 3 系统控制程序结构

(1)输入/输出控制模块是提供给用户的接口。输入控制包括基本的输入错误检测,如数据类型不正确、数据溢出、不存在的 VLAN 号等不合法输入;输出控制主要功能是控制输出格式;

(2)实验系统硬件资源记录模块,记录实验系统的硬件资源(包括主机、交换机)的详细信息。记录主机的 CPU、内存、IP 地址、网卡类型以及该主机与交换机相连的端口号;记录交换机的端口、VLAN 的使用情况;

(3)交换机控制模块是读取、修改交换机当前配置。该模块通过调用 SNMP 协议的接口构造 SNMP 的协议数据单元,与交换机的 SNMP Agent 通信,要求该 Agent 处理 SNMP 请求消息。

该程序是在 Win2000 平台上,采用 VC++6.0 编写,通过调用 mgmtapi.lib 和 snmpapi.lib 两个库文件实现 SNMP 协议访问接口。

3 实验及结果分析

3.1 实验系统的具体配置

在所搭建的 MIPv6 仿真实验环境中,使用了一台锐捷 Star-S1926F+ 二层交换机^[3],该交换机配置 24 个 10/100 Mbps 网络端口,同时支持 Port VLAN 和 Tag VLAN 两种工作模式。在实验中采用基于 Port VLAN 的工作模式。HA、MN、AR 采用 3 台配置相同的 PC:P4 2.4GHz,256MB DDR 内存,10/100M

D-Link自适应网卡, 操作系统为Linux Red Hat 9.0(内核版本Linux-2.4.26), MIPL(Mobile IPv6 for Linux) 版本为MIPv6-1.1-v2.4.26, RADVD版本为radvd-0.8。

3.2 实验系统的连通性测试

使用 ping6 程序进行测试。首先在 HA 上启动 RADVD 和 MIPL, 在 AR 上启动 RADVD。随后在 MN 上启动 MIPL, 让 MN 持续的 ping6 家乡代理 HA: ping6 fec0:106:2700::2。在 HA 上用 tcpdump 监听网络上的数据分组, 此时会看到如图 4 中所示的 icmp6 分组信息, 表明 MN 在家乡 VLAN 工作正常。

在控制台, 将 MN 从家乡 VLAN 一键切换到外地 VLAN。观察 ping6 结果, 会发现 ping6 略有中断, 这是由于切换造成的。在 HA 上, 用 tcpdump 会很清楚地看到切换这一过程。如图 4 所示: 是 MN 在家乡 VLAN 的最后一次 ping6 信息; 表明, MN 被切换到外地 VLAN, 获得转交地址 fec0:106:1100:0:20d:88ff:fe48:1803, 并且在 HA 处成功完成注册; 表明, ping6 恢复, MN 在外地 VLAN 工作正常。

```

16:34:36.445444 fec0:106:2700::4 > fec0:106:2700::2: icmp6: echo request (len 64, hlim 63)
16:34:36.445484 fec0:106:2700::2 > fec0:106:2700::4: icmp6: echo reply (len 64, hlim 64)
16:34:36.725513 fec0:106:2700::4 > fec0:106:2700::2: DSTOPT (padn)(homeaddr:
    fec0:106:1100:0:20d:88ff:fe48:1803)ip-proto-60 40 (len 40, hlim 254)
16:34:36.725562 fec0:106:2700::2 > fec0:106:1100:0:20d:88ff:fe48:1803: srct (len=2, type=2,
    seqleft=1)[srct]ip- proto-43 40 (len 40, hlim 255)
16:34:37.445528 fec0:106:1100:0:20d:88ff:fe48:1803 > fec0:106:2700::2: icmp6: echo request (len 64, hlim 63)
16:34:37.445568 fec0:106:2700::2 > fec0:106:1100:0:20d:88ff:fe48:1803: icmp6: echo reply (len 64, hlim 64)
    
```

图 4 切换前后, HA 上 tcpdump 输出结果

综上所述, 当 MN 在家乡 VLAN 与外地 VLAN 之间切换时, 在切换造成的短暂中断后, MN 与 HA 之间的通信恢复正常, MN 在家乡代理成功地完成注册, ping6 继续工作, 说明在我们搭建的 MIPv6 仿真实验环境中, MIPv6 协议能正常运行。

3.3 切换过程中的丢包和时延测试

使用 ping6 程序, 对移动切换所造成的丢包和时延进行测试。方法仍如上, 让 MN 持续 ping6 HA, 控制台控制切换。在 MN 上观察 ping6 结果, 计算切换丢包数。同时, 在 HA 上用 tcpdump 监听网络上的数据分组到达时间, 计算出切换时延。

(上接第115页)

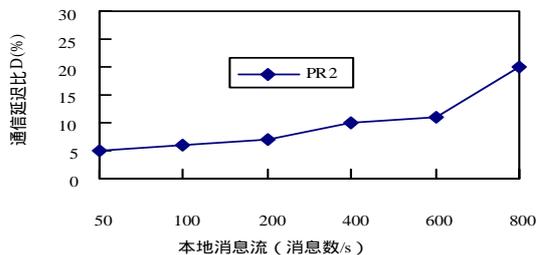


图 4 多对多实验

5 结束语

本文针对网络环境下非专用机群进行分布式并行计算的特点, 分析和评估了网络流量和延迟的变化对并行任务间通信的影响, 实际上, 这类问题根据网络环境的不同, 情况十分复杂, 往往很难做出准确的量化评估。

参考文献

1 Mark B, Rajkumar B. Cluster Computing at a Glance High Performance Cluster Computing: The Architecture and System[M].

作为对比, 对手工插拔线切换方式所造成的丢包和时延也进行了一组测试, 结果如表 1 所示。从数据可以看出, 基于 VLAN 的切换方式与手工插拔线切换方式相比, 切换过程中的丢包和时延都有所改善, 但是基于 VLAN 的切换方式, 可控性更强, 切换更加方便、快速。

表 1 对比测试结果

	平均切换丢包(个)	平均切换时延(ms)
基于 VLAN 的切换	7.33	445.03
手工插拔线切换	7.60	519.99

4 结束语

针对目前大多数 MIPv6 实验系统搭建过程复杂、移动切换可控性较差、操作不方便等问题, 本文提出了一种基于 VLAN 划分搭建 MIPv6 实验系统的新方法。与以往方法相比, 本方法具有如下优点: 组网方便, 只需一台可划分 VLAN 的二层交换机, 采用基于端口的 VLAN 划分方式, 无需考虑物理连线, 即可简洁、灵活地在交换机上划分出家乡链路以及外地链路; 移动切换控制灵活、简便, 只需通过动态改变 MN 所属的 VLAN 便可方便快捷地完成。测试结过表明, 在使用本方法搭建的 MIPv6 实验系统中, MIPv6 协议能够很好地工作, 而且 MN 切换过程中的丢包和时延也都得到了改善。

MIPv6 应用是未来移动领域一个很重要的应用, 因此, 本文所提出的基于 VLAN 划分搭建 MIPv6 实验系统的新方法, 将对 MIPv6 的研究有着积极的意义。

参考文献

1 Johnson D, Perkins C, Arkko J. Mobility Support in IPv6[S]. RFC3775, 2004.

2 Lars Strand[Z]. 2004. <http://www.gnist.org/~lars/work/mip6/Mobile-IPv6-HOWTO.html>.

3 S1926F+产品说明书-V12[Z]. 2003. <http://www.ruijie.com.cn/products/channel>.

4 孙利民, 阚志刚, 郑健平, 等. 移动 IP 技术[M]. 北京: 电子工业出版社, 2003.

5 张云勇, 刘韵洁, 张智江. 基于 IPv6 的下一代互联网[M]. 北京: 电子工业出版社, 2004.

Prentice Hall, 1998.

2 Andrzej G. Distributed Operating System the Logical Design[M]. Addison-Wesley Publisher Ltd., 1991.

3 Atallah M J, Lock C, Marinescu D C, et al. Co-scheduling Compute-intensive Tasks on a Network of Workstation: Models and Algorithms[C]//Proceedings of the 11th International Conference on Distributed Computing System, 1991.

4 Du Xing, Zhang Xiaodong. The Impact of Memory Hierarchies on Cluster Computing[R]. Department of Computer Science, College of William and Mary, Technical Report: TR-99-2, 1999.

5 Shyamal C. The Greedy Load Sharing Algorithm[J]. Journal of Parallel and Distributed Computing, 1990, 9(1): 93-99.

6 Al G, Adam B, Jack D, et al. PVM3 User's Guide and Reference Manual[Z]. Oak Ridge National Laboratory, 1994.

7 Adam. TPVM: Distributed Concurrent Computing with Lightweight Processes[C]//Proceedings of the 4th IEEE International Symposium on High Performance Distributed Computing, 1995.

