摘 要

HiNOC（High performance Network Over Coax）网络是在当前 FTTB 已经存在和普遍应用的前提下，利用有线电视网已有的同轴电缆线路和分配网络，组建而成的最后 100 米范围内的宽带接入网。该技术完全利用现有有线电视网同轴电缆的网络布线，仅增加 HiNOC Bridge (HB)和 HiNOC Modem (HM)等相关设备，实现高速和高质量多业务接入，可提供用户包括 IPTV、SDTV/HDTV、上网等宽带应用。

本文首先对 HiNOC 网络及 HiNOC 系统 MAC 协议进行了分析，在此基础上，重点研究了 MAC 协议中的汇聚子层功能；设计了针对 HiNOC 网络的基于 IGMP Snooping 的二层组播方案，通过实验证明了该方案的可行性；将 VLAN 技术引入 HiNOC 网络，提出了基于 MAC 地址的 VLAN 划分和基于用户标识的 VLAN 划分机制；最后对 HiNOC MAC 协议性能进行了测试，研究了汇聚子层的设计机制对 HiNOC 性能的影响。

关键字：HiNOC　汇聚子层　组播　VLAN

Abstract

HiNOC (High performance Network Over Coax) bridges the last 100 meters from fiber node to home since fiber is widely deployed and used. HiNOC is a home broad-band access network communicating over exiting coaxial cable. HiNOC technology completely use the cable of wired TV, only adding the devices of HM (HiNOC Modem) and HB (HiNOC Bridge) without changing the current wire line. HiNOC provides high-speed and high-quality access for multi-services, such as voice, IPTV, SDTV/HDTV and data.

Firstly, the HiNOC network and the MAC protocol are analyzed. Then the functions of the Convergence Sublayer (CS) in the MAC protocol are studied. And then the link layer multicast protocol based on IGMP Snooping is designed. The protocol is validated by experiment.VLAN technology is put forward into the HiNOC network ,and two mechanisms are proposed to assign VLAN based on MAC address and user identity .At last, a test is given about the performance of HiNOC MAC protocol, and the effect that mechanism of the convergence sublayer to the network performance of HiNOC is studied.

Keywords: HiNOC Convergence Sublayer Multicast VLAN

西安电子科技大学
学位论文独创性（或创新性）声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切的法律责任。

本人签名：赵英华　　日期2009年3月11日

西安电子科技大学
关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。同时本人保证，毕业后结合学位论文研究课题再撰写的文章一律署名单位为西安电子科技大学。

（保密的论文在解密后遵守此规定）

本学位论文属于保密，在___年解密后适用本授权书。

本人签名：赵英华　　日期2009年3月11日

导师签名：　　日期09-3-11

第一章 绪论

1.1 HiNOC 网络的背景及意义

三网融合已成为全球范围内不可阻挡的趋势。我国十一五规划纲要中明确指出了积极推进“三网融合”的发展目标。三网融合当前主要是指高层业务的融合。近来出现的数字电视、IPTV 和移动电视等业务应用，被业界普遍认为是各大网络走向三网融合的切入点。

当前的骨干网带宽可以基本满足网络融合的需要，而接入网速率限制成为影响业务融合发展的瓶颈。以一路数字化标清电视（SDTV）和高清电视（HDTV）节目分别需要 6Mbps 和 25Mbps 左右的带宽计算，若一户家庭同时收看 2 套 SDTV 和 1 套 HDTV 节目，加上音频和部分上网带宽，则每户所需接入速率至少为 40Mbps。现有的接入技术由于带宽、成本、重新布线以及用户接受程度等原因都难以提供可行的解决方案。由于光结点到用户的距离越来越近，在已经实现光纤到楼（FTTB）的情况下，如何提供最后 100 米的宽带接入方案将直接影响着三网融合的推进。

我国现有 1.3 亿有线电视网用户，分布广泛的楼道和户内分配网络具有得天独厚的频带宽、容量大、抗干扰能力强等优点。其中 860MHz 以下为广播电视频道，860MHz 以上频带没有使用，本文称这一频段为带外信道。据初步测算，在带外信道内有超过 2Gbps 的可用物理带宽。在这段有线电视网分配网络内没有放大器等有源设备，无需进行改造就可以进行双向通信。如果能利用带外信道提供宽带接入解决方案，则无需重新布线或改造便可提供多业务宽带接入。

1.2 HiNOC 网络概述

有线电视网在我国是接入覆盖最大、用户数量最多的一张网络，同时也是目前我国最高效廉价的综合网络，它具有频带宽，容量大，多功能、成本低、抗干扰能力强、支持多种业务连接千家万户的优势。目前有线电视网主要用来向用户传送模拟电视节目，宽带双向的点播电视(VOD)及通过有线电视网接入 Internet 开展双向业务是有线电视网的发展方向，最终目的是使有线电视网走向宽带双向的多媒体通信网。

HiNOC（High performance Network Over Coax）网络是在当前 FTTB 已经存在和普遍应用的前提下，利用有线电视网已有的同轴电缆线路和分配网络，组建而成的最后 100 米范围内的宽带接入网。该技术完全利用现有有线电视网同轴电缆

的网络布线，仅增加 HiNOC Bridge (HB)和 HiNOC Modem (HM)等相关设备，实现高速和高质量多业务接入，可提供用户包括 IPTV、SDTV/HDTV、上网等宽带应用。HiNOC 技术主要用于 FTTB＋Cable 的应用环境，根据需求，该技术组网应用的典型情况如图 1 所示。

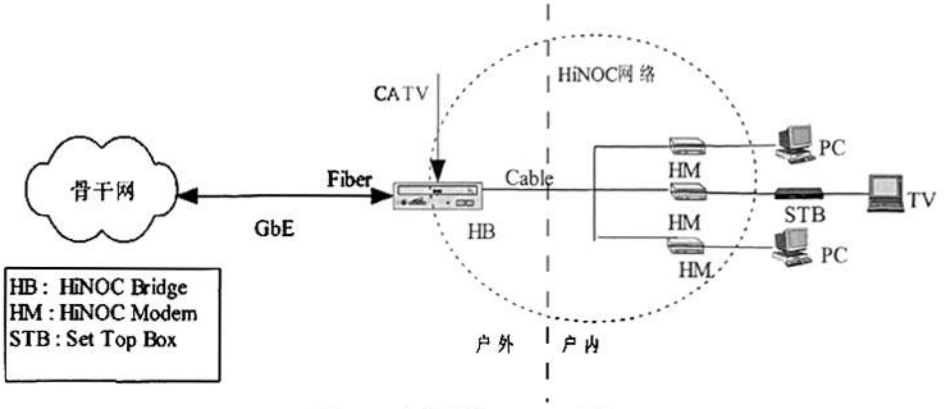


图 1 一个典型的 HiNOC 网络

如图 1 所示，HiNOC 网络内存在两种设备，分别为 HB（HiNOC Bridge）和 HM（HiNOC Modem）。GbE（Giga bit Ethernet）信号通过 FTTB 方式到达用户楼门口，经 HB 调制到同轴电缆带外频带的一个工作信道内，并通过楼内分配网络和电缆终端盒传输到位于用户家内的 HM，经 HM 解调后传送给 PC、机顶盒（STB）等设备。原有的 CATV 信号（860MHz 以下）在 HB 和 GbE 信号混合进入电缆分配网络，因此不影响原有电视的收看。一个 HB 和若干个 HM 以及中间的同轴电缆分配网络构成 HiNOC 网络，工作于同轴电缆的带外信道。

HiNOC 技术允许用户在已有的有线电视电缆上建立起带宽高达 40Mbps MAC 层速率的接入网络。同时，由于高速数据业务工作在 860M 以上的频带，因此可与现有的电视信号共存。用户可以同时享用 HiNOC 带来的高速网络及网上冲浪的感受，又能继续收看现有的模拟/数字电视节目。由于 HiNOC 提供超大带宽的数据接入，可以支持高速上网、基于 IP 的网络电视、互动电视、视频点播、时移电视、视频电话、视频会议和网络电话等多种增值应用。

经过计算，同轴电缆的带外信道所能支持的数据传输速度可达到 2Gbps 以上，能在相当长时间内满足业务发展的带宽要求，FTTB+Cable 的组网方案是在目前最理想的组网方式。一，网络建设的成本比较低，采用现有的同轴电缆进行入户，完全符合现有有线网络分支分配网的网络拓扑，能最大程度上使用已有线路，工程量小，不扰民；二，HiNOC 产品的目标成本为每户小于 600 元，设备成本相当低，和 VDSL 等技术基本相当，而且其一对多的网络结构，具有很大的灵活性；三，HiNOC 能够提供非常高的用户接入带宽，这样的带宽水平远远超过 VDSL[1]、ADSL[2][3]、Cable Modem[4]等，能够满足 IPTV 等高清晰流媒体业务的传输。

HiNOC 技术的主要技术指标是：单信道模拟带宽为 16MHz，单用户的 MAC 层接入速率可达到 40Mbps 以上；MAC 层网络拓扑支持点到多点的组网方式，单信道支持最大结点数为 32 个；网络覆盖范围不小于 100 米，可实现包括 HDTV/SDTV 等流媒体在内的多业务宽带接入。

1.3 本文主要工作和组织结构

从协议分层模型看，HiNOC 技术主要涉及 HiNOC MAC 层和物理层（PHY）的设计。其中 MAC 层从下到上又可分为公共部分子层（CPS Common Part Sublayer)和汇聚子层(CS Convergence Sublayer)两部分。本文工作重点是对 HiNOC 网络 MAC 协议中的汇聚子层技术进行研究。汇聚子层，主要完成上层业务的封装与转发，使得上层业务可以通过 HiNOC 网络透明传输。

本文的主要工作包括：首先对 HiNOC 网络及 HiNOC 系统 MAC 协议进行了简要分析；重点分析了 MAC 协议中的汇聚子层功能，在此基础上设计了针对 HiNOC 网络的基于 IGMP Snooping 的二层组播方案，通过实验证明了该方案的可行性；将 VLAN 技术引入 HiNOC 网络，提出了基于 MAC 地址的 VLAN 划分机制和基于用户的 VLAN 划分机制；最后对 HiNOC MAC 协议的性能进行了测试，研究了汇聚子层对 HiNOC 网络性能的影响，通过测试可以看到汇聚子层的设计对于 HiNOC 网络性能有显著影响。

本文组织结构为：

第二章分析了 HiNOC 网络的 MAC 层协议，主要包括 HiNOC 接入网的拓扑结构、协议栈、信道分配及 MAC 层机制，MAC 层机制又分为：网络搜索、网络接纳和链路维护机制等。

第三章重点研究了 HiNOC 网络的汇聚子层功能和机制，包括通过地址学习构建转发表、数据帧的封装及打包、数据帧的转发以及对业务优先级的划分。根据汇聚子层的特点及组播原理，提出了基于 IGMP Snooping 的组播方案，在 HiNOC 网络中实现了二层组播。将 VLAN 技术引入 HiNOC 网络，提出了基于 MAC 地址和基于用户的 VLAN 划分机制。

第四章主要对 HiNOC 网络的性能进行测试，通过对吞吐量及时延的测试，可以看出汇聚子层对 HiNOC 网络性能的影响，从而说明汇聚机制对 HiNOC 网络的重要性。

第五章总结论文工作，指出不足，提出未来工作思路。

第二章 HiNOC 网络 MAC 层协议概述

HiNOC 网络由一个 HB 和若干个 HM 以及中间的同轴电缆分配网络所构成，工作于同轴电缆的带外信道，利用现有有线电视网同轴电缆的网络布线，仅增加局端的相关设备，实现高速和高质量多业务接入，可提供用户包括 IPTV、SDTV/HDTV 和上网等宽带应用。

2.1 HiNOC 系统及 MAC 层概述

HiNOC 网络逻辑拓扑为星型结构，即点对多点结构，如图 2.1 所示。HB 作为中心结点可以连接若干个子结点 HM，HM 相互之间不能互相通信，只能和 HB 通信。HB 发给 HM 的帧称为下行帧，HM 发给 HB 的帧称为上行帧。

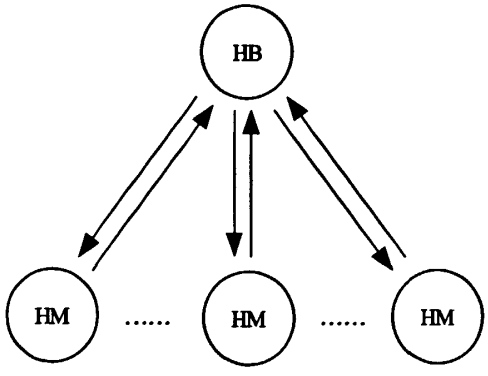


图 2.1 HiNOC 网络星型逻辑结构

在单个 HiNOC 信道内，HB 和多个 HM 设备组成总线型的物理拓扑结构，如图 2.2 所示。

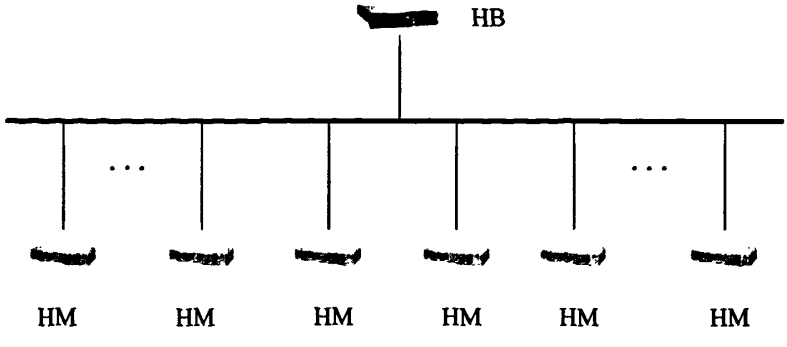


图 2.2 HiNOC 网络总线型物理拓扑结构

HiNOC 网络采用 TDD 方式实现双工通信，下行采用 TDM 方式，可以提高传

输速率，实现简单，并且在主结点的同意调度下可以非常灵活的实现上、下行带宽的分配。上行采用 TDMA 多址方式，可以动态分配信道资源，TDMA 机制需要各结点较为严格的时钟同步，在 HiNOC 网络中由于电缆接入网范围非常小，通过主结点 HB 定期发送系统时间就可以保证各结点的同步。

HiNOC 网络中 HB 和 HM 系统的协议栈如图 2.3 所示，分别为物理层、媒体接入层和高层。其中媒体接入层（MAC）又分为汇聚子层（CS）和公共部分子层（CPS）。

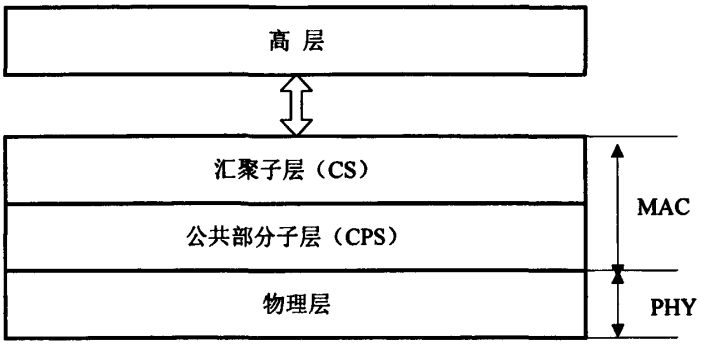


图 2.3 HiNOC 网络系统协议栈

CS 层是 MAC 层的一个功能子层，主要功能是承载上层业务，实现 MAC 层核心功能与高层功能的适配，包括地址学习、数据的封装与解封、帧的转发。其中上层业务主要是指话音、视频和数据业务，而数据主要是高层协议数据单元（PDU），包括 IP 分组、以太网帧等。在本文中主要研究的以太网帧，称为以太网 MAC 帧（EMAC 帧）。以太网 MAC 帧在 HiNOC 网络中要被封装成 HiNOC MAC 帧（简称 HMAC）进行传输。CPS 子层提供 MAC 层的核心功能，包括结点接纳控制、接入控制与信道（带宽）分配、链路维护、QoS 保证等。传输数据时，CPS 子层接收来自 CS 子层的数据封装为 HMAC PDU，然后通过 PHY 层发送。

HMAC 帧由帧首部、帧载荷和尾部（CRC 校验）组成，如图 2.4 所示，MAC 帧首部是由发送时钟、帧类型、源结点 ID、目的结点 ID、帧长度、帧子类型、版

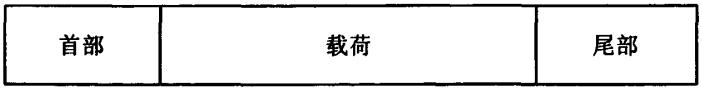


图 2.4 HiNOC MAC 帧结构

本号、子帧个数、首部校验和构成。结构如图 2.5 所示。其中目的结点和源结点 ID 是收发该帧 HiNOC 网络结点地址；子帧个数表示该帧载荷域内包含的 HMAC 业务数据单元的数目，通过这一功能 MAC 层可以支持打包功能，将多个高层 PDU 组合为一个 MAC PDU，这样可以提高传输效率。载荷域存放上层业务信息或控制信息。尾部是帧校验序列，用于实现对载荷域的校验。

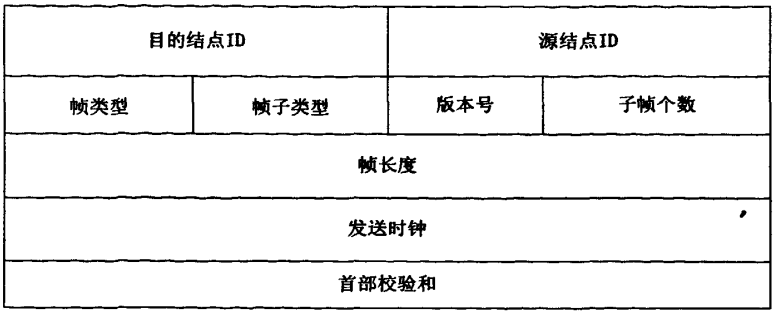


图 2.5 MAC 帧首部结构

从功能上分，MAC 帧可以分为控制帧和数据帧两类。数据帧用于承载上层业务信息（如以太网帧）。控制帧用于实现协议的接纳、调度、预约、维护和信道探测等功能。主要的控制帧见表 2-1。

表 2-1 HMAC 控制帧

帧类型	发送方向，方式	功能
MAP 帧	主站—>子站，广播方式	调度和发布（已接纳结点）各帧的发送时机
预约请求（R）帧	子站—>主站，单播方式	子站向主站发送预约，请求安排发送时机
链路控制帧（多种）	双向，单播或广播	用于传送接纳、维护、信道探测过程中的控制信息
Beacon 帧	主站—>子站，广播方式	用于控制和调度未接纳结点的信息传送

所有的 MAC 帧都是通过物理层（PHY）发送的，但是 MAC 层不产生和发送用于信道训练的探测帧（Probe）和功率控制帧，这两种帧是由物理层产生的，但是这些帧的发送必须由 MAC 层为其调度发送时机。

2.2 信道分配

HiNOC 网络采用全协同的 TDM/TDMA 机制实现对共享信道的访问，信道分配由 HB 决定。HB 通过周期性发送 MAP(Media Allocation Plan)帧向各结点发布当前这一周期内的信道分配方案。一个 MAP 帧所控制的一段时间称为 MAP 周期。每个 MAP 周期时长可变。MAP 周期的安排主要是考虑提高线路利用率，使开销达到最小。为了避免发射机在收发状态之间频繁转换，把 MAP 周期中的上行数据区间和下行数据区间分开。MAP 周期的安排及其必须遵守的时间关系如图 2.6 所示。图中 Tup 为两个上行帧之间的间隔，Tdown 为两个下行帧之间的间隔，TINT 是一个 MAP 周期的结束与下一个 MAP 周期的开始之间的间隔。

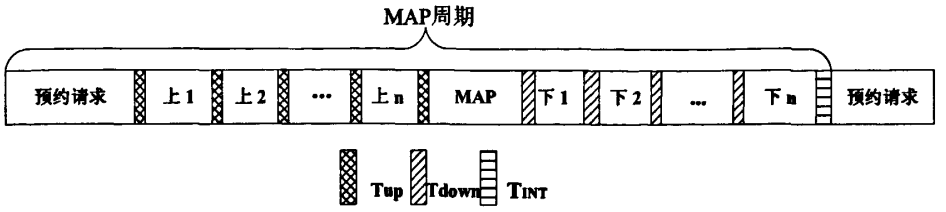


图 2.6 MAP 周期

MAP 周期的使用由 MAP 帧分配。MAP 帧包括以下域：MAP 帧头、零个或多个分配单元（AU）。AU 用来给 MAC 帧（如 Beacon、MAP、Probe、控制、预约请求、数据等）的发送分配传输带宽；MAP 帧的结构如图 2.7 所示。

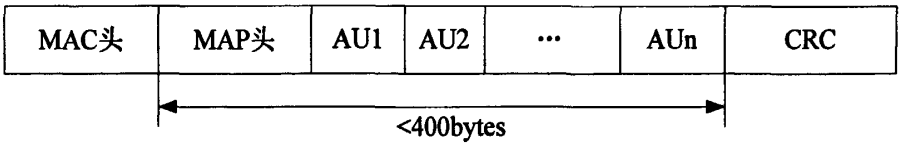


图 2.7 MAP 帧格式

对于 MAP 周期和 MAP 帧，需要注意的是：HB 在产生 MAP 帧时，必须保证 MAP 周期不重叠并且填满整个时间轴。在每个 MAP 帧内，HB 最多只允许调度一次探测码的传输。一个 MAP 周期的持续时间可变，其持续时间由网络中的结点数目和被调度的业务总量决定。每个 MAP 周期最多允许 4 个结点传输预约请求帧。并且，在一个 MAP 周期内 HB 为每个结点调度预约请求帧的次数不允许超过一次。每个周期都要允许 HB 发送下行帧，但是不用给 HB 预约请求机会。对于来自某个 HM 带宽请求，HB 必须在该结点发送下一个预约请求帧之前在 MAP AU 中准许其带宽请求，或将带宽请求丢弃。如果某个结点在发送下一个预约请求帧之前还没有收到来自 HB 的发送时机准许，则该结点就认为前一次发送的预约请求被 HB 丢弃了，此时该结点需要重新发送该预约请求帧。对于相同优先级类型的所有业务，HB 必须保证对来自同一结点的带宽请求的许可顺序与该结点预约请求帧中请求单元的顺序相同。当 HB 准许某结点的带宽请求时，HB 必须保证 MAP AU 中的请求 ID 域的值与该预约请求单元中请求 ID 域的值相同。因此，一个结点需要确保这个请求 ID 值在该结点内是唯一的。如果 HB 给某个 HM 发送预约请求帧的机会，则不论该 HM 是否有待发送的数据，都必须返回一个预约请求帧（如果没有待发送的数据，则预约请求帧可以只包含一个预约请求帧头而不包含预约请求单元）。如果该结点在连续 60 个预请求帧的传输时机中都没有对 HB 进行响应，则 HB 必须将该 HM 从网络中删除。

2.3 MAC 层机制

MAC 层协议按照工作过程分为结点接纳前、接纳及接纳后三个阶段。结点接

纳前需要进行网络搜索；之后通过结点接纳过程控制结点合法接纳到网络，并对信道状况进行探测以便进行有效的信息传输；结点接纳后各结点采用预约/许可机制在 HB 结点的调度下接入信道实现数据通信，并在必要时对链路状况进行维护。

2.3.1 网络搜索

HB 在加电后等待 HM 加入网络的过程，以及 HM 加电以后搜索 HB 的过程称为网络搜索。网络搜索是网络初始化的过程。HB 网络搜索主要是发送 Beacon 信号，使得 HM 可以收到 Beacon 并获得接纳时机从而加入网络；HM 结点网络搜索主要是找到 Beacon，获得发送接纳请求的时机从而进入网络接纳过程。HB 和 HM 应分别在什么频率上发送 Beacon 或去搜寻 Beacon，有两种方法——指定频率网络搜索和变换频率网络搜索。

（1） 指定频率网络搜索，系统配置 HB 和 HM 的工作频率。

- ◇ HB 网络搜索：HB 加电之后指定工作在系统配置的频率上，如果监听之后发现该信道可用，则开始发送 Beacon 帧，以提供给 HM 发送接纳请求的机会，Beacon 帧中包含自己的网络 ID。HB 持续发送 Beacon 帧，如果收到 HM 的接纳请求，则进入网络接纳过程。如果接纳不成功，HB 一直周期性地发送 Beacon 帧，直到建立网络。建立网络之后 HB 也会一直发送 Beacon 帧以提供给其他结点加入网络的机会。
- ◇ HM 网络搜索：HM 加电之后工作在系统配置的频率，一直监听 Beacon 帧，直到收到 Beacon 帧，且 Beacon 帧中的网络 ID 与自己的一致，发送接纳请求，进入网络接纳过程。如果接纳不成功，继续在此频率侦听 Beacon 帧。直到加入网络。

（2） 变换频率网络搜索。这里所用的频率是上次工作频率 LOF（Last Operational Frequency），只要 HB 建立网络或者 HM 成功加入网络后必须记录网络工作频率，即 LOF，以便下次加电后网络搜索时使用。结点必须有非易失性存储器来记录该频率，LOF 必须在结点电源开/关，重启动和重新初始化过程中保持不变。网络搜索过程中频率切换的过程如图 2.8 所示：结点从 LOF 开始，每隔一次切换到非 LOF。结点在网络搜索时，可以选择的最低非 LOF 频率是 F1，其次是 F2，F3 等。由图 2.8 可以看出，结点搜索时尽量在 LOF 上建立或加入网络；在信道搜索时，低频信道更经常的被试探；结点以交替的上升和下降频率搜寻信道。

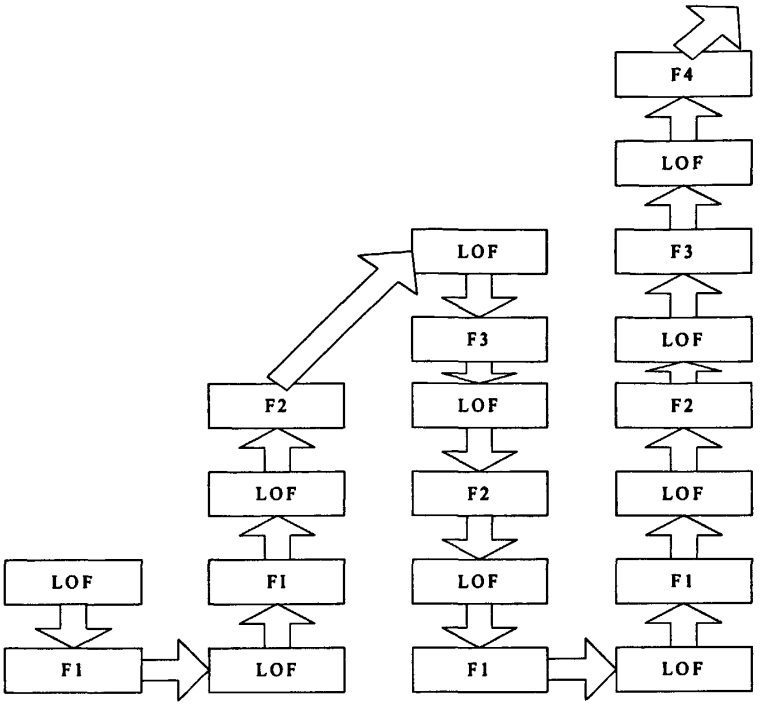


图 2.8 网络搜索过程的频率变换顺序

◇ HB 网络搜索：HB 加电之后工作在 LOF 上。HB 调谐到一个频率之后，首先会设一个定时器监听 Beacon 帧。如果 HB 收到 Beacon 帧则切换频率。定时器超时后没有收到 Beacon 帧，则 HB 会开始在此频率发送 Beacon。发送 50 个 Beacon 帧后，若没有收到接纳请求，则切换频率。重复上述过程。发送 Beacon 帧之后收到接纳请求，进入接纳过程。若没有成功建立网络，则 HB 继续网络搜索过程。HB 建立网络之后仍然会发送 Beacon 帧，以便其他的结点加入。

◇ HM 网络搜索：HM 加电之后工作在 LOF 上。HM 调谐到一个频率之后首先监听 Beacon 帧并设定定时器。如果定时器超时后收不到 Beacon 帧，则 HM 切换频率。如果收到 Beacon 帧，但是网络 ID 不符，HM 切换频率。 如果 HM 收到 Beacon 帧并且 Beacon 帧中网络 ID 与自己相符，如果 Beacon 帧中接纳控制帧的类型域不是 0x00，表明正在有结点接纳，则 HM 继续等。HM 看到 Beacon 帧中接纳控制帧的类型是 0x00 时发送接纳请求。

2.3.2 网络接纳

网络接纳过程是一个新的 HiNOC 设备上电之后，在 HB 的调度下加入到现有 HiNOC 网络的过程。

在没有接入网络前，新结点只能接收 Beacon 帧。主站 HB 周期性发送 Beacon 帧，在 Beacon 帧中为新结点安排了接纳请求帧的发送时机。所有子结点在启动后，必须先发送接纳请求以申请接入加入网络。如有两个（或以上）新结点同时利用 HB 安排的请求时机发送接纳请求，则可能发生碰撞，此时 HB 不能正确接收到请求，因此不能向任何子站回复接纳响应；新结点不能预期接收 HB 的接纳响应，认为发生碰撞，于是采用截断的二进制指数退避算法在之后 Beacon 帧中安排的时机内再次尝试加入网络。一旦收到 HB 发来的接纳响应，则子站接下来利用周期性 Beacon 帧安排的发送时机进行信道探测和功率调整，最终获得双向信道调制参数和功控参数。一旦完成，则新结点接纳过程结束。结点接纳到网络中后，可以接收到 HB 周期性发送的 MAP 帧，之后的通信过程中，新子站和主站之间利用已获得的调制参数和功控参数在 MAP 帧调度下进行通信。以上过程可以简单用图 2.9 描述：

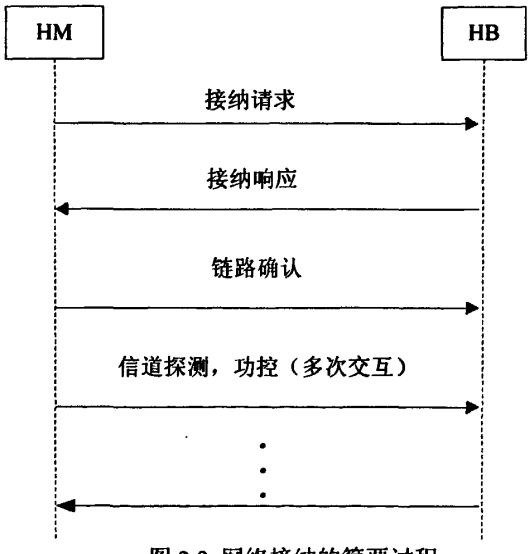


图 2.9 网络接纳的简要过程

2.3.3 数据通信

数据通信也称为正常通信，指的是结点 HM 被接纳进入网络后，在网络中进行信息交互的过程，主要涉及信道的接入控制问题。HM 何时发送信息由 HB 通过 MAP 帧进行调度；因此接入网络后各结点在信道上的信息传送是无冲突的。

MAP 帧机制将整个信道的发送时间划分为 MAP 周期（MAP Cycle）。一个 MAP 周期指的是信道上由一个 MAP 帧安排的一段时间。因此信道上的时间由一个个连续的 MAP 周期组成，每个 MAP 周期内各结点的发送时机由该 MAP 周期对应的 MAP 帧规定。如图 2.10 所示。

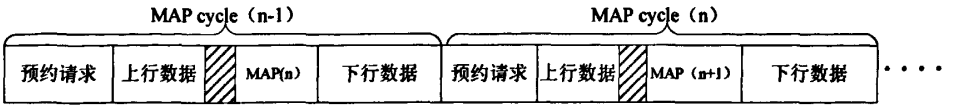


图 2.10 MAP 调度

图中 MAP(n)帧具体规定了 MAP cycle(n) 中各帧的发送时机，MAP(n+1)帧规定了 MAP cycle(n+1) 中各帧的发送时机，依此类推。图中的预约请求时段为被允许的 HM 发送预约请求帧的时段、上行和下行数据分别表示各 HM 和 HB 发送各 MAC 帧的时段。

某 HM 有信息要发送时，先按照 MAP 帧的规定，利用分配给自己的预约请求时机发送预约请求帧，HB 收到该预约请求后，若同意其发送，则在后续的 MAP 帧中为其分配发送时机，HM 收到后，在规定的时机向 HB 发送信息。HB 发送信息的时机也是在 MAP 中定义的，所以对应的接收信息的 HM 通过 MAP 知道在何时接收信息。

MAP 周期的时长目前是固定的，但是标称值可变，建议可为 4ms，5ms 或 6ms，具体取值根据测试情况而定，以方便实现。

另外，Beacon 帧的发送也在 MAP 帧的调度之下，即 Beacon 帧属于上图中的下行数据时段。同样，在 Beacon 帧中也规定了下一个 MAP 帧的发送时机。这样，通过 Beacon 帧可以找到 MAP 帧，反之亦然。

2.3.4 链路维护

为保证 HiNOC 网络在信道条件变化的情况下仍然运行正常，MAC 协议需要进行定期或不定期的信道维护规程。协议可以对信道特性重新进行探测，并且可以灵活设置固定的信道维护周期，也可以由 PHY 层根据信道误码等特性的动态变化触发 MAC 层的信道维护规程。

HiNOC 网络的链路维护分为两种情况：一种是周期性的维护，HB 重新获得与各个 HM 之间的信道参数，同时 HM 也重新获得与 HB 之间的信道参数。一种是不定期的维护，是在当信道条件恶化的情况下，仅维护恶化的信道。

◇ 周期性的链路维护就是周期性进行的链路维护，维护周期为 600s。HB 将设定定时器为 600s，一旦该定时器超时，HB 将选择一个 HM 开始链路维护过程。在与某个 HM 进行链路维护期间，HB 到 HM 的双向信道都将重新探测以获得新的双向信道参数。当 HB 进行完与一个 HM 之间的链路维护过程后，将在稳态停留 1s，然后重新选择另外一个 HM 进行它与该 HM 之间的链路维护。直到 HB 完成与所有 HM 之间的链路维护工作，它将在稳态停留 1s，然后计算新的最大调制物理层简表，并将其广播给网络中所有的 HM。

◇ 不定期的维护是对某条信道恶化情况下的链路维护。当 HM 的 PHY 发现 HB 到 HM 的链路有问题时，应向其 MAC 层发送原语指示。HM 的 MAC 向 HB 的 MAC 发送一个链路维护帧，通知其进行 HB 到 HM 链路的维护，之后 HB 利用 MAP 通知链路进入链路维护状态。当 HB 发现 HB 到 HM 的链路有问题时，应向其 MAC 层发送原语指示。HB 此时将通过 MAP 通知进入链路维护状态。哪条信道有问题，就进行哪条信道的维护，不考虑其他信道。但 HB 要重新计算广播帧的调制模式并向所有结点通告。

2.4 本章小结

本章着重介绍了 HiNOC 网络和 HiNOC 网络的 MAC 层协议，包括 HiNOC 接入网拓扑结构、协议栈、MAC 层机制等。对 MAC 层的介绍是为研究 HiNOC 网络汇聚子层做基础，因为汇聚子层是 MAC 层的一个功能子层。接下来本文将对 HiNOC 网络汇聚子层进行研究。

第三章 HiNOC 网络汇聚子层的研究

汇聚子层在逻辑上位于 MAC 层公共部分子层之上，主要功能是承载上层业务，实现 MAC 层核心功能与高层功能的适配，包括地址学习、数据的封装与解封、帧的转发。其中上层业务主要是指话音、视频和数据业务。

3.1 HiNOC 网络汇聚子层

HiNOC 网络汇聚子层是 MAC 层的一个功能子层，位于 HiNOC 的 MAC 层与高层相接的位置，负责从高层接收高层 PDU,使之成为 HMAC 的业务数据单元（SDU)；按照高层 PDU 的参数，构造或者更新转发表；依照高层 PDU 的参数，查询转发表，根据查询结果和业务优先级将高层 PDU 分类映射并转发到不同的 HiNOC 发送队列；也可以将多个 HMAC SDU 进行打包，并将一个或多个 HMAC SDU 传送到公共部分子层。

汇聚子层的主要功能为负责接收高层的协议数据单元，并将高层的协议数据单元映射到 MAC 的公共部分子层（CPS)，以及进行相反的操作。高层的协议数据单元（PDU）主要是以太网的 MAC 帧，EMAC 帧承载话音，视频和数据应用等业务。

图 3.1 给出了 HiNOC 网络承载以太网业务示意图。图中的虚线内为 HiNOC 网络，从图中可以看出，在进出 HiNOC 网络前后，网络接口都是以太网口，传输的都是以太网 MAC 帧。当以太网 MAC 帧进入 HiNOC 网络发送结点后，需要结点将 EMAC PDU 帧封装为 HMAC PDU 在 HiNOC 网络中传输，通过 Cable 到达 HiNOC 网络中的目的结点。

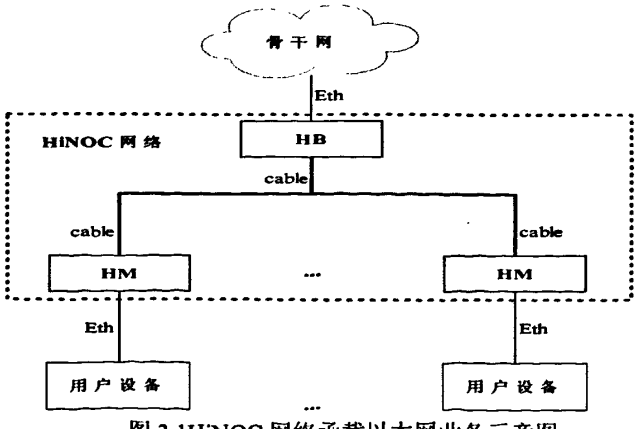


图 3.1HiNOC 网络承载以太网业务示意图

从图 3.1 中可以看到，由 HB 和多个 HM 构成的 HiNOC 网络可以被近似等效

为一个以太网交换机：HB 的以太网端口类似以太网交换机的上行端口，而各 HM 的以太网端口则类似该交换机的普通端口，HiNOC 网络内 HB 和各 HM 之间的通信则实现了以太网交换机的内部交换功能。因此从 HiNOC 网络外部（即上层以太网协议角度）看，HiNOC 网络应当具有以太网交换机支持的所有功能。

在完成上述功能的过程中，汇聚子层需要实现的机制主要包括：通过地址学习构建转发表、地址查找和数据帧（即 EMAC 帧）转发、数据帧打包/拆包及对业务优先级的支持。下面对主要机制的实现原理进行较详细的讨论。

3.1.1 地址学习构建转发表

HiNOC 网络采用点到多点结构，对于 HB 来说，收到 Ethernet MAC 帧后，需要判断该 Ethernet MAC 帧的目的地址位于 HiNOC 网内哪个 HM 结点之下，以实现 HiNOC 网内的正确定向发送。否则，所有 Ethernet MAC 帧只能采用广播方式向所有结点发送，而广播调制方式比单播调制方式效率低，因此这无疑会造成网络性能的下降。同样，HB 接收到来自网内某个 HM 发送的帧时，也要判断将该帧转发到以太网接口还是其他 HM 结点。对于 HM 来说，虽然它只有一入一出两个逻辑接口，但也需要对来自某一侧接口的 EMAC 帧地址进行判断，以便过滤掉那些不需要向另一侧接口转发的 EMAC 帧。这就需要构建转发表，转发表反映了 EMAC 地址与 HiNOC 结点地址的映射关系，是实现 EMAC 帧在 HiNOC 网内正确转发和高效传送的依据。该转发表是通过地址学习方式完成的。

在汇聚子层构建转发表，需要考虑上层以太网业务的具体实现机制。从目前的应用场景看，HiNOC 所承载的以太网除了具有基本的单播和广播传输能力外，通常还具备组播和 VLAN（Virtual LAN，虚拟局域网）支持能力。为此，HiNOC 汇聚子层也需要提供相应的对组播和 VLAN 的支持能力。

（1） 单播转发表的构建

单播转发表为标准的单播 EMAC 帧在 HiNOC 网内的正确转发和高效传送提供依据。转发表中每一项反映了 EMAC 地址和 HiNOC 结点地址的对应关系。构建单播转发表可以采用地址学习机制，即各 HiNOC 结点（HB 或 HM）在每次收到 EMAC 帧（无论来自哪个接口）时都读取 EMAC 源地址，并查找单播转发表，如果表中没有相关表项就生成一个新表项，如果有则对该表项信息（如老化时间）进行更新。通过对 EMAC 帧的源地址进行不断自学习，从而构造出一张反映 EMAC 地址和 HiNOC 结点地址映射关系的单播转发表。

无论 HB 还是 HM 都构建各自的单播转发表。对于 HB，当收到一个来自以太网接口侧的 EMAC 帧时，建立（或更新）EMAC 帧源地址与该以太网接口的映射关系；当收到来自 Cable 接口某个 HM 的 HiNOC 数据帧时，首先解析出该帧承载

的 EMAC 帧，而后建立该 EMAC 帧源地址与 HM 结点地址的映射关系。对于各 HM 结点，当收到以太网接口的以太网帧时，HM 记录 EMAC 源地址与以太网接口的映射关系（这样做可以限制以太网接口下联的各主机间的内部流量被不必要的转发到 HB。当 HM 以太网接口下联集线器时，就会出现这种情况）；而对于来自 Cable 接口侧的以太网帧，HM 记录 EMAC 源地址与 Cable 接口的映射关系。久而久之，HB 或 HM 通过反向 EMAC 源地址学习就会各自建立一个单播转发表。

地址学习过程中，为提高转发表的查找速度，可以采用哈希（Hash）表结构[6][7]来实现转发表。转发表的各表项应支持动态更新，设有老化时间，一旦超时，应将该表项从表中删除。转发表中每个表项的构成，除用于 Hash 表结构的字段外，其核心信息包括：EMAC 地址、HiNOC 结点地址以及更新时间（age）。具体过程如下

◇ 地址表结构。为了实现高效快速插入和查找地址对应项，HiNOC 网络单播地址转发表采用哈希表结构。哈希表又称为散列表，是一种重要的存储方式，也是一种快速的查找方法。哈希算法的核心是在记录的存储位置和它的关键字之间建立一个确定的对应关系 H，使每个关键字与结构中唯一的一个存储位置相对应。因而，在查找时，只要根据这个对应关系 H，就可以找到给定值 key 的映像 H(key)。若结构中存在关键字和 key 相等的记录，则必定在 H(key)的存储位置上。由此，不需要进行比较便可直接取得所查记录。这个对应关系 H 称为哈希函数。按这个思想建立的表称为哈希表或散列表。

哈希函数是从关键字集合到地址集合的压缩映像。不同的关键字可能映像到同一个哈希地址，这种现象称为冲突。生成同一哈希地址的关键字称为同 义词。 常用的冲突处理方法有开放定址法、再哈希法以及链地址法。链地址法就是将同义词映射到一个线性链表，查找时只要顺着链表顺序查找即可。链地址法适用于增加或删除动作较多的动态链表管理中。可以取 48 位 MAC 地址作为关键字，计算 hash 值，即存储位置值。解决冲突采用链地址法，即将相同 hash 值的 MAC 地址对应项链接为一个双向链表。哈希表结构如图 3.2 所示。

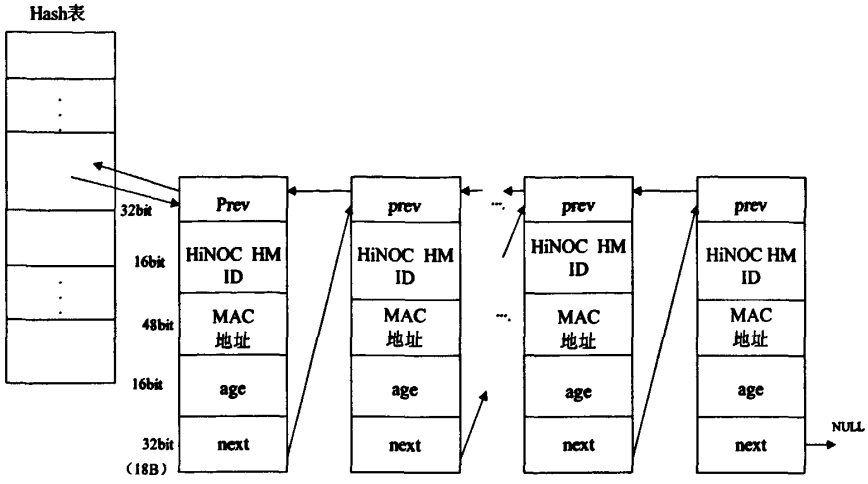


图 3.2 HB 单播地址转发表结构

Hash 表每项包括一个 MAC 地址、一个结点 ID 以及一个 age 域，age 域主要用来更新表项。Hash 值的计算是将 MAC 地址逐字节左移两位，然后与下一字节值求异或，完成之后，再将高 8 位和低 8 位再异或，最后将值与 hash 表大小相与，将 hash 值限定在指定范围之内。哈希表每个表项存储的信息为一个 MAC 地址、对应的 HiNOC ID 以及 age 域。

◇ 地址表查找。Hash 表查找与插入采用同样的算法，首先由 MAC 地址计算 hash 值，根据 hash 值查找 hash 表对应位置，顺着链表查找关键字相同的表项，返回对应的结点 ID。

◇ 地址表更新。哈希表的每一项包含一个 8 比特的 age 域。首次插入该地址项时 age=1，如果插入时发现该表项已存在，则将 age 域加 1，模 256；另外设有定时器，周期性遍历整个 hash 表，将每个 age 域值减 1，当 age 域小于 0 时，即该地址项已经很久没有用，则将该地址项删除。如果插入新的表项时转发表缓存区已满，则遍历整个 hash 表选择 age 域值最小的一项（最早以前用过的）删除，释放的空间以便记录新的表项

◇ 地址表大小。每个 HB 最多连接 32 个 HM，每个 HM 连接 4 个以太网设备，上行方向 HB 可以最大学习到 4×32 = 128 个地址对应项；假设每个以太网设备连接到 8 个外部 MAC 地址，则下行方向 HB 可以学到 4×32×8 = 1024 个地址对应项。总转发表大小为 128+1024 = 1032 个。若每个表项大小为 18B(如图 3.2 所示)，则 HB 的转发表容量估计为：1032×18B = 18,576 字节。对于 HM，只学习上行方向数据包，每个 HM 连接 4 个以太网设备（MAC 地址），因此表项为 4 个，单播地址表容量为 4×18 = 72 字节。

（2） 组播转发表的构建

组播转发表的建立和单播转发表的建立是不同，不能通过源地址反向学习来

实现，因为没有哪个数据包的源地址会是组地址。组播转发表的建立和具体的实现机制有关系，可以采用IGMP Snooping[5]机制来构建组播转发表。IGMP Snooping的任务是根据IGMP（Internet Group Management Protocol，Internet组管理协议）[8][10]报文中的组播组信息和报文来源的HiNOC网络ID，维护组播组和网络ID的对应关系，建立组播转发表，并且根据IGMP协议中主机的加入和离开消息同步更新组播表。详细过程将在3.2节进行介绍。

（3） 对 VLAN 的支持

VLAN（Virtual Local Area Network）[9]即虚拟局域网，是一种通过将局域网内的设备逻辑地划分成一个个网段从而实现虚拟工作组的新兴技术。IEEE 于 1999 年颁布了用以标准化 VLAN 实现方案的 802.1Q 协议标准草案。

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，一个广播域就是一个 VLAN。每一个 VLAN 都包含一组有着相同需求的计算机工作站，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，所以同一个 VLAN 内的各个工作站无须被放置在同一个物理空间里。不同 VLAN 之间的通信需要使用路由功能，主要由路由器提供。

采用 VLAN 技术可以将物理拓扑上属于一个整体的 LAN 划分为与物理位置无关的逻辑上相互隔离的多个 LAN，从而改善因过大的广播域而造成的性能和安全问题。当前的以太网普遍支持 VLAN 功能。因此，要求承载以太网业务的 HiNOC 网络也相应的提供对 VLAN 的支持，以实现对属于不同 VLAN 的以太网业务的隔离和正确转发。

实际应用中，通常利用 VLAN 标记[11]不同的用户或业务。用户业务包括宽带上网、IPTV 和 VoIP 等业务。可根据不同策略进行 VLAN 规划，如按用户划分 VLAN、按业务类型划分 VLAN、同时考虑用户和业务类型划分 VLAN 等。无论采用何种划分方式，都需要对 HiNOC 结点进行相应配置，以便与 VLAN 划分相匹配。配置可以用静态方式完成，在用户安装 HiNOC 设备或申请相关业务后进行。配置过程就是确定 HB 和 HM 的各个接口与 VLAN ID（标识）的映射关系。这样，当 HB 收到来自上联以太网接口（或 Cable 接口）带有 VLAN ID 的 EMAC 帧后，就可以正确转发到属于同一 VLAN 的 HM 结点（或上联以太网接口）。HM 有类似的机制。

HiNOC 网络结点为正确传递 VLAN 帧，需要支持 VLAN Trunk（中继）机制[11]。VLAN Trunk 支持在同一条链路上传递多个 VLAN 的流量。实际应用中，在网络结点与其相连的以太网设备之间的链路上，以及在网络中 HB 和各 HM 结点之间的链路上，都可能需要传递多个 VLAN 的流量，因此在这些链路需要使用 VLAN Trunk 功能。

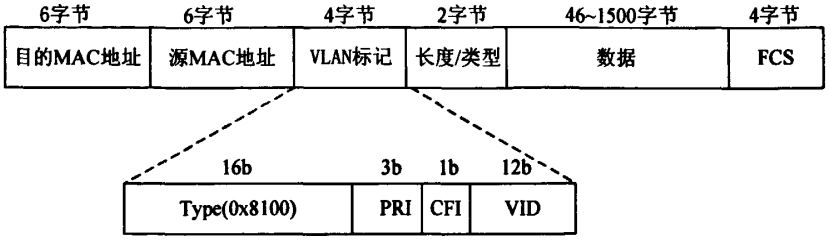


图 3.3 支持 VLAN 的以太网帧格式

VLAN Trunk 功能由 IEEE 802.1Q 协议规定，该协议规定了包含 VLAN 标识的以太网帧格式，如图 3.3 所示。该帧格式在传统以太网帧格式中增加了 4 个字节的 VLAN 标记字段。其中，TYPE 字段占 2 个字节，用来指出该以太网帧的类型，目前统一设为 0x8100。PRI 字段占 3 比特，用来表示该以太网帧的优先级（Priority），以提供一定的服务质量要求，3 比特共可表示 8 种优先级。优先级的相关定义在 IEEE 802.1p 中规定。CFI 字段占 1 比特，表示规范格式指示符。VID 字段即 VLAN ID，占 12 比特，用来表示该以太网帧所属的 VLAN。

结合以上单播、组播和 VLAN 等情况，最终将构建完整的转发表。

3.1.2 数据帧的封装及打包

在 HiNOC 网络承载的是以太网数据，而 HiNOC 网络中传输的是 HMAC 帧，以太网帧要通过 HiNOC 网络到达目的结点，必须封装成 HiNOC 网络的 MAC 帧格式传输，并且在对等层进行解封装。为提高 HiNOC MAC 层传输效率，降低协议开销，协议支持将多个 EMAC PDU（帧）统一打包（packing）并封装为一个 HMAC PDU（帧）。

在基于预约/许可的 TDM/TDMA 机制下，预约、调度等帧要周期性占用信道，基于物理层算法和硬件性能而选择的帧间隔取值也基本固定，它们都会对协议性能带来较大影响。要保证 HMAC 层速率达到 40Mbps，就必须尽可能增大 HMAC 帧中净荷长度。考虑到 EMAC PDU 最大只有 1518 字节（采用 VLAN 后为 1522 字节），因此 HMAC 协议支持多个 EMAC 帧的打包。

打包后 HMAC 帧的长度取值需要在信道传输效率和因打包产生的排队时延之间进行折中。只有发往同一目的结点、同一优先级的多个 EMAC 帧才能打包成一个 HMAC 帧。多个 EMAC 帧打包后再封装为 HMAC PDU，其结构如图 3.4 所示。

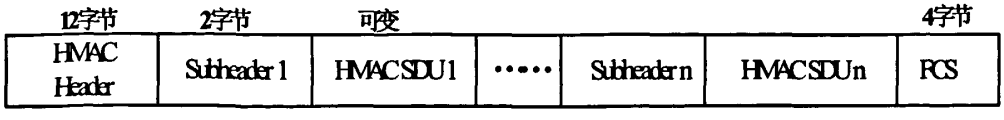


图 3.4 打包后的 HMAC PDU 结构

通过打包，一个 HMAC PDU 可以包括多个 SDU。当承载 EMAC 帧时，一个 EMAC PDU 对应一个 HMAC SDU。每个 SDU 前带有一个 2 字节子帧头，表示其后 SDU 长度与子帧头长度之和（即 SDU 长度+2，字节为单位）。在 HMAC 帧头中有一个子帧个数域，指示该帧中所含 SDU 的个数。协议规定 HMAC PDU 最大长度为 4588 字节，此时一个 HMAC 帧中恰好可以包含 3 个最大长度（1522 字节）的 EMAC 帧。

为进行打包，需要将到达 HiNOC 结点的多个 EMAC 帧（HMAC SDU）进行排队，这会导致 EMAC 帧通过 HiNOC 网络的时延增大。为避免因打包而导致 SDU 在 HiNOC 网内排队造成的过大时延，需要对 HMAC 帧的最大长度进行限制，并根据业务特性对各类业务通过 HiNOC 网络的时延进行规定。因此，在实现打包时，只要符合以下两个条件之一，都应将 SDU 队列中已有的各 SDU 立即打包构成一个 HMAC PDU：一是当多个 SDU 打包构成的 HMAC PDU 超过最大长度限制；二是 SDU 队列内各 SDU 的最大排队时延超过设定门限。各时延门限的取值，与业务的类型有关；实时高优先级业务取值小，非实时低优先级业务取值大。

在接收端，CS 子层在收到 HMAC PDU 时，负责实现拆包，取出各个 SDU 然后交付给上层。

为保证 HiNOC 信令的及时传送，规定对 HiNOC 控制帧不进行打包，每个控制帧单独构成一个 HMAC PDU 传送。

3.1.3 数据帧的转发

数据帧的转发是根据转发表来实现的，CS 根据数据帧的类型来实现在 HiNOC 网络中的单播，组播和广播的转发。转发过程如下：

（1）HB 结点的帧转发策略

1) 当 HB 收到一个来自以太网接口的下行 EMAC 帧：

- ◇ 若是单播帧，则若目的 EMAC 地址对应某 HM 所连接的 EMAC 地址，则应向该 HM 转发；若目的 EMAC 地址对应 HB 设备本身（如远程网管信息），则交由本结点 CPU 处理；若目的 EMAC 地址未在转发表中，则向 HiNOC 网络广播；
- ◇ 若是广播帧，则向 HiNOC 网络广播，同时转交本地 CPU 处理；
- ◇ 若是组播帧，则若为 IGMP 报文，则向 HiNOC 网络广播；若转发表中存在该帧目的组播地址的转发项，则向 HiNOC 网络转发该帧。转发时，若该目的组播地址表项只对应一个 HM，则向该 HM 单播转发；若该目的组播地址表项中对应有多个 HM，则将该帧向 HiNOC 网络广播；

◇ 其余情况，则丢弃该数据帧。

2)当 HB 收到来自 Cable 接口的上行 HMAC 数据帧，解析出 EMAC 帧：

◇ 若是以太网单播帧，则若目的 EMAC 地址对应 HB 本身的 EMAC 地址，则转给本结点 CPU 处理；若目的 EMAC 地址对应上行以太网接口侧的 EMAC 地址，则向上行以太网接口转发；若目的 EMAC 地址对应某个 HM 所连接的 EMAC 地址，则向该 HM 转发；若目的 EMAC 地址未在转发表中，则向上行以太网接口转发的同时，也向 HiNOC 网络广播转发；

◇ 若是以太网广播帧，则向上行以太网接口转发，同时向 HiNOC 网络广播转发，并转交本地 CPU 处理；

◇ 若是组播帧，则向以太网接口转发；

◇ 其余情况，则丢弃该数据帧。

（2）HM 结点的帧转发策略

1)当 HM 收到来自 Cable 接口的下行 HMAC 数据帧，解析出 EMAC 帧：

◇ 若 EMAC 帧源地址对应 HM 以太网接口侧的地址或者 HM 自身设备地址，则丢弃该数据帧；(这说明该帧是之前从本 HM 发送到 HB，又被 HB 用广播帧发送给其他 HM 时，“反弹”到本 HM 的。)

◇ 若是以太网单播帧，则若目的 EMAC 地址为本 HM 自身地址，则交由本结点 CPU 处理；若目的 EMAC 地址对应来自以太网接口的 EMAC 地址，则向以太网接口转发；若目的 EMAC 地址不在转发表中，则向以太网接口转发；

◇ 若是以太网广播帧，则向以太网接口侧转发，同时转交本地 CPU 处理；

◇ 若是以太网组播帧，则若为 IGMP 报文，则向以太网接口转发该帧；若转发表中存在目的组播地址对应的表项，则向以太网接口转发该帧；

◇ 其余情况，则丢弃该数据帧。

2)当 HM 收到来自与主机相连的以太网接口的 EMAC 帧：

◇ 若是单播帧，则若目的 EMAC 地址对应 Cable 接口侧地址，则应向 HB 转发；若目的 EMAC 地址对应本 HM 自身地址，则转交本结点 CPU 处理；若目的 EMAC 地址未在转发表中，则向 Cable 接口转发；

◇ 若是广播帧，则向 Cable 接口转发，同时转交本地 CPU 处理；

◇ 若是组播帧，则向 Cable 接口转发；

◇ 其余情况，则丢弃该数据帧。

(3) 若采用 VLAN 机制，则帧转发过程中，所有 EMAC 帧的转发都被限制只在同一个 VLAN 内进行，HiNOC 结点对属于不同 VLAN 的 EMAC 帧不进行转发。此时，可以针对不同的 VLAN 分别构建转发表，并分别执行上述转发策略。

3.1.4 业务优先级的划分

为了使 HiNOC 网络能提供基于优先级的 QoS 保障机制，汇聚子层需要完成以下的工作：

- ◇ 根据上层 EMAC 帧内指示的优先级、业务类型或其他参数，将上层以太网业务映射为 HiNOC 网络内的相应优先级；
- ◇ 对于发往同一 HiNOC 结点的上层 EMAC 帧，要能根据映射到 HiNOC 网络内的不同优先级分别排队、打包和封装，只有属于同一优先级的 EMAC 帧才被打包封装成一个 HMAC 帧。

经过 CS 子层的优先级映射后，在 CPS 子层各 HM 向 HB 预约信道时，其预约请求帧内携带请求发送 HMAC 帧的优先级标识，HB 根据此标识进行资源调度和信道分配，实现基于优先级的 QoS 保障。

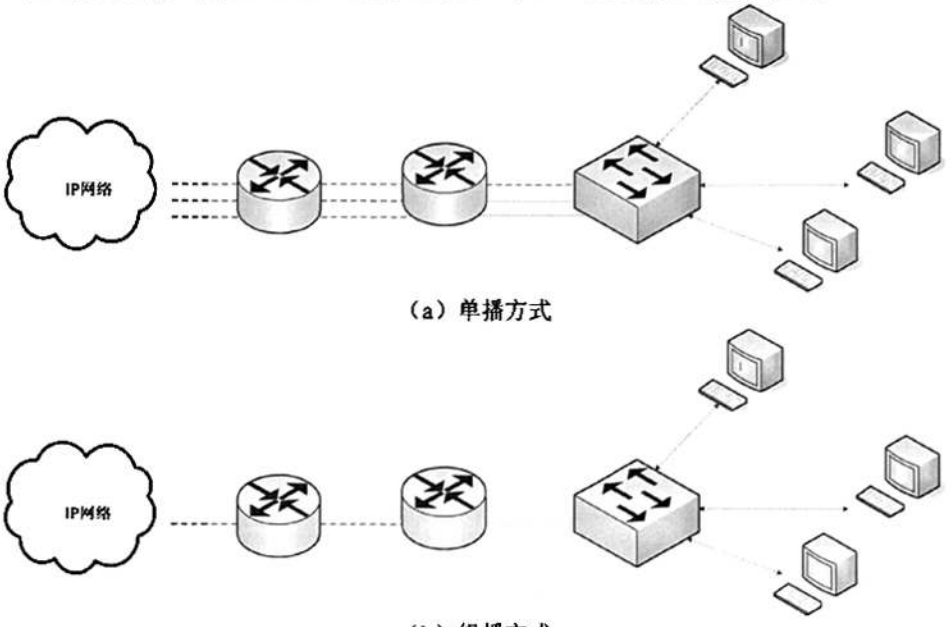
HiNOC 协议对其承载的上层业务规定了 3 种优先级，分别对应以 VoIP 为代表的迟延敏感的实时交互性业务（优先级类型 2）、以 IPTV 为代表的流媒体业务（优先级类型 1）以及 BE 业务（优先级类型 0）。协议规定 HiNOC 控制帧具有最高优先级（优先级类型 3），以保证控制帧的及时传送。

可以依据不同的参数实现优先级映射。这些参数可以是 EMAC 帧内包含的 IEEE 802.1p 优先级标识、IEEE 802.1Q 规定的 VLAN ID 等。如果依据 802.1p 进行映射，CS 子层可以将 802.1p 中的优先级 7、6 和 5 映射为 HiNOC 中的优先级类型 2，802.1p 中的优先级 4 和 3 映射为 HiNOC 中的优先级类型 1，其他优先级编码映射为 HiNOC 的优先级类型 0。当 VLAN 被用来区分不同的业务类型时，可以依据 VLAN ID 进行优先级映射，将属于不同 VLAN 的各类型业务映射为不同的 HiNOC 优先级。具体采用哪种优先级映射方法需要根据具体应用场景决定，通过对映射策略进行配置而实现。

3.2 HiNOC 网络的组播方案

在因特网上，多媒体业务（如流媒体、视频会议和视频点播等）正在成为信息传送的重要组成部分。这类业务的传输特性是：单点发送、多点接收。点对点传输的单播方式不能适应这类业务的需要，因为服务器必须为每一个接收者提供一个相同内容的 IP 拷贝，同时网络上也要重复地传输相同内容的报文，占用了大量资源。虽然 IP 广播允许一个主机把一个 IP 报文发送给同一个网络的所有主机，

但是并不是所有主机都需要这些报文，因而浪费了网络资源，而且广播是无法通过路由器转发到其它的网段的。这种情况下组播应运而生，它的出现解决了一个主机向特定的多个接收者发送消息的问题。图 3.5 是组播和单播的区别：



（a）单播方式

（b）组播方式

图 3.5 组播方式和单播方式的比较

3.2.1 组播技术简介

IP 组播是在网络层为实现一对多通信设计的技术，可以有效地完成将相同信息发送到多个特定主机的功能，降低了带宽消耗，提高了网络效率和性能。IP 组播协议框架分为路由器－路由器之间的协议和主机－路由器之间的协议。路由器－路由器之间协议包括各种组播路由协议，主要用在骨干网。主机－路由器协议即组播成员管理协议，包括基于 IPv4 的 IGMP（Internet Group Management Protocol，Internet 组管理协议）和基于 IPv6 的 MLD（Multicast Listener Discovery，组播侦听发现）协议，在接入网中使用。下面详细介绍一下基于 IPv4 的 IGMP，由于 IPv6 是 IPv4 的未来替代协议，并且是 IPv4 发展的必然结果，因此也会详细介绍基于 IPv6 的 MLD。

（1）IGMP 因特网组管理协议

IGMP[12]是为实现对组播组成员的管理而设计的 Internet 协议， 是实现 IP 组播的基础，运行于主机和它所在的子网组播路由器之间，用以支持主机和路由器进行组播。到目前为止，IGMP 有三个版本。IGMPv1 中定义了基本的组成员查询和报告过程；IGMPv2 在 IGMPv1 的基础上增加了组成员快速离开机制，即主机主动发送离开报文给组播路由器，而不是被动的等待组播路由器查询，从而减少系

统处理停止组播的延时。IGMPv3 中增加了“源过滤”功能，成员之间可以进行指定接收或指定不接收某些组播源的报文，从而可以实现数据包在特定的组播成员关系之间进行组播。目前使用最多最广泛的是 IGMPv2，因此会在下面详细介绍一下 IGMPv2。

IGMP 消息在 IP 数据报内传送，用 IP 协议号为 2 标识。IGMPv2 的消息格式如下：

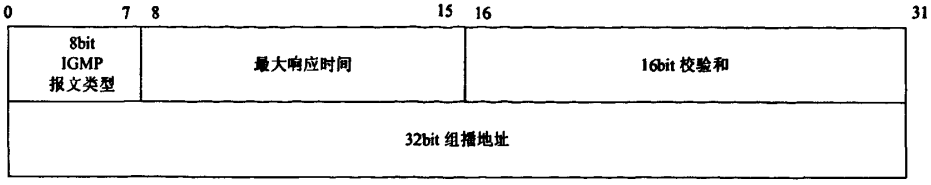


图 3.6 IGMP 消息格式

1） 类型字段：四种消息类型

◇ 成员关系查询（类型代码=0x11）。成员关系查询信息有两种子类型：

普通查询：用于确定在和 IGMPv1 相同的类型中哪些组播组是有效的，普通查询由全零的组地址字段表示。

特定组查询：用于决定特定组组播组是否有活动的接收者，特定组查询的组地址字段为正在查询的组地址。

◇ 版本 1 的成员关系报告（类型代码=0x12），此信息类型用于和 IGMPv1 兼容。

◇ 版本 2 的成员关系报告（类型代码=0x16）

◇ 离开组（类型代码=0x17）

2） 最大响应时间字段：在 IGMPv1 信息中本字段未使用。最大响应时间字段是允许查询路由器为它的查询报文指定准确的查询间隔响应时间。主机在随机选择它们的响应时间值时以此作为上限。这样在查询响应间隔时有助于控制突发响应。

3） 校验和字段：共 16 位，表示的是 IGMP 信息补码之和的补码。该校验和字段在进行校验计算时等于 0。

4） 组地址字段：在发送普通查询时，该字段被设置为零以区别特定组查询，后者包含将要查询的组播组。当成员关系报告或者离开组信息发送时，本字段设置为目标组播组地址。

工作过程：组播路由器周期性地发送通用组查询消息进行成员关系查询；主机发送报告消息来响应查询。主机发送报告消息的时间有随机性，当检测到同一网内有其它成员发送同样的消息时，则抑制自己的成员响应报文。如果有新的主机要加入组播组，不必等待组播路由器的查询消息，可以主动发送报告消息。主

机离开组播组时，主机发送离开组消息；收到离开组消息后，组播路由器发送特定组查询消息来确定是否所有组成员都已经离开。如果所有的组成员都已经离开及所收到的离开组消息是组最后一个成员发送的，则组播路由器就会在组播表中删除改组，停止向该子网转发此组播。

（2）MLD 组播侦听发现协议

MLD(Multicast Listener Discovery)[13]是组播侦听发现协议的简称，是在IGMPv2 基础上改进以支持 IPv6 功能的组播组管理协议。组播组管理协议主要负责对组播组的成员进行管理。一个组播路由器建立路由并传送其组播成员关系信息之前，它必须确定在本地网络上是否有接收主机加入了某个组播组。为此组播路由器和接收主机必须进行组成员关系的交互，这个功能在 IPv6 网络上是通过 MLD 来实现的，在 IPv4 中使用 IGMP 协议来实现。目前 MLD 协议有两个版本，即 v1 和 v2。

MLDv1 协议只能用于 IPv6 环境，与运行在 IPv4 环境下的 IGMP 协议相比，其最大变化在于使用 ICMP 报文承载数据，而不像 IGMP 直接使用 IP 报文承载数据。MLDv1 协议的工作方式和 IGMPv2 协议类似。路由器定期发送查询报文，主要如果想加入组播组，则通过报告报文对路由器进行响应。主机退出组播组时，发送退出报文给路由器。主机之间采用响应抑制来避免发送重复确认。

MLDv2[14]的提出，主要是为了配合源特定组播的实现，非常类似与 IGMPv3 对 IGMPv2 的改进。源特定组播（SSM：Source Specific Multicast)是一种区别于传统组播的新的业务模型，它使用组播组地址和组播源地址同时标识一个组播会话，而不是像传统组播服务那样只使用组播组地址来标识一个组播会话，SSM 保留了传统 PIM-SM 模式中的主机显式加入组播组的高效性，但是跳过了 PIM-SM 模式中共享树和 RP 规程。SSM 特别合适于点到多点的组播服务，例如视频点播、在线网络教学等业务。为了在 IPv6 网络中实施 SSM 组播业务，除了要求网络端到端地支持网络组播和 SSM 模式外，同时还要求网络和应用支持 MLDv2 协议栈。

MLDv2 是一个不对称的协议，它包括分离的两部分：针对组播地址监听者部分（监听组播数据包的主机或者路由器）和组播路由器部分。MLD 协议的目的是使每一个组播路由器知道在本地链路上监听者对哪些组播地址和源地址感兴趣。被 MLD 收集到的信息提供给在路由器上运行的任何组播路由协议，来保证组播数据包被传递到组播组成员。组播路由器只需要知道在一个链路上，至少有一个结点在监听从特定的源发给一个特定组播地址的数据抱，它不需要知道有多少个这样的结点。

MLD 协议使用的是 ICMP 的报文格式而不是 IGMP 报文格式，ICMPv6 协议是为 IPv6 网络定义的一套控制信息协议。在一个 IPv6 数据报文中，通过将 Next Header 字段的值置 58 来表示该报文的数据部分是一个 MLD 协议报文；同时将 Hop

Limit 字段置 1。

MLD 消息的封装格式如下图所示

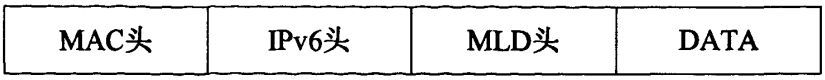


图 3.7 MLD 消息封装格式

MLD 消息格式如下图所示：

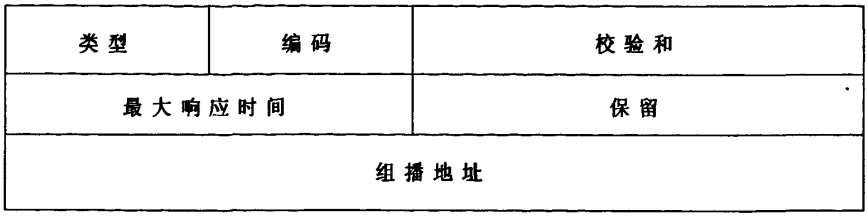


图 3.8 MLD 消息格式

1） MLD 消息类型分为三种：

◇ 组播监听查询（MLD 消息类型值为 130）：包括一般查询和特定组播地址查询；

◇ 组播监听报告（MLD 消息类型值为 131）；

◇ 组播监听完成（MLD 消息类型值为 132）。

2） 编码：发送方将其初始化为 0；在接收端将忽略该值。

3） 校验和

校验和是整个 ICMPv6 报文的一个 16 位字的补数和，是标准 ICMP 校验和，校验和的计算起始于 ICMPv6 的类型字段，再加上一个 IPv6 的为伪首部，在伪首部中下一个首部字段伪 58.为了计算校验和，校验和字段被设置为 0。

4） 最大响应时延

该字段与 IGMPv2 的相应字段类似。只在查询消息中有效，字段值规定了发送响应报告的最大允许时间间隔，单位伪毫秒。

5）组播地址

在查询消息中，若为普通查询该字段置 0；若为特定组查询，该字段为某个 IPv6 组播组地址。在成员报告和成员离开消息中，该字段伪特定的组播组地址，分别为希望加入的组和准备离开的组地址。

对于运行 MLD 协议的路由器，其接口要监听由 IPv6 组播地址产生的所有链路组播地址。路由器为它所在的每一个链路维护一个列表。表项有此链路中存在的组成员的组播地址，以及该地址相应的定时器。路由器周期性发送通用请求，以查询该链路上是否存在某组播地址的组成员。结点收到路由器发送的常规请求后，经过随机时延后发出组播监听报告。这样是为了防止所有的结点都在同一时

间发出报告分组，从而造成网络的突发性阻塞。当路由器收到链路上的报告分组时，如果报告地址不在路由器的列表上，则加入该项，否则计时器重新置位。如果某个地址的计时器过期，则从列表中删除。

当结点要加入一个组播组时，主动发送组播监听报告，向路由器报告组成员的存在。结点退出组播组时，发送完成分组，删除有关路径。当请求状态的路由器从链路上接收到一个完成消息，如果消息中的组播地址在路由器的列表上，路由器发送一个特定组播地址查询。如果一段时延没有报告分组，则认为该组播地址在此链路上没有组成员了。

（3）组播地址

IP 单播地址用于唯一标识单个 IP 主机，而 IP 组播地址则用于标识一组任意的 IP 主机，该组中的 IP 主机已经加入该组播组并且希望接收发往该组播组的信息。组播 IP 地址代表接收者组，而不是某个单独的接收者，因此组播地址只能用作目的地址，而不能出现在数据报的源地址字段。

IP 组播地址为 D 类地址空间，该空间的地址的第一字节的前 4bit 用 1110 表示，因此 IP 组播地址的范围从 224.0.0.0 到 239.255.255.255 。为了实现组播数据在底层的传输，必须将 IP 组播地址映射为 MAC 组播地址。在以太网中，IP 组播帧使用的是以 0x0100.5Exx.xxxx 的 24 位前缀开始的 MAC 层地址。图 3.9 显示了以太网单播和组播地址的区别以及 IP 组播地址到以太网组播地址的映射关系：如果以太网地址的高字节的最低位是 1，则它是一个组播地址；否则就是一个单播地址。单播以太网地址由接口制造商分配，组播地址由网络协议动态分配。

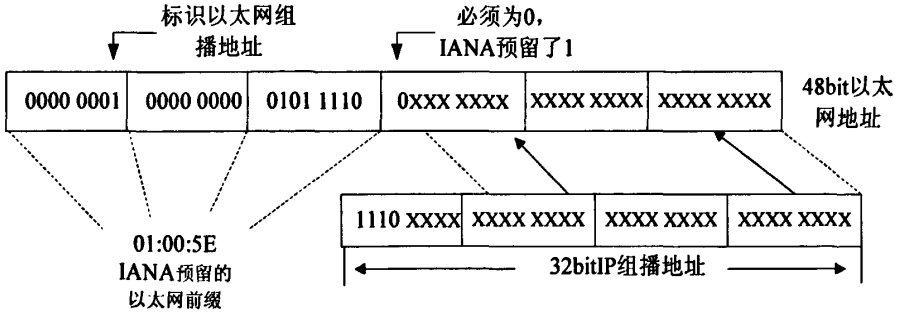


图 3.9 IP 组播地址到以太网 MAC 地址的映射

如图 3.9 所示，从 IP 组播帧到以太网组播帧的映射是一个多到一的映射。IP 组播地址有 28bit 的地址空间，只有 23bit 被映射到以太网 MAC 地址，共有 5bit 重叠，2^5=32，这样将有 32 个 IP 组播地址映射成一个 MAC 组播地址，需要接收者在高层进行过滤。

3.2.2 基于 IGMP Snooping 的组播方案

IGMP 组播成员管理机制是针对第三层设计的，在第三层，路由器可以对组播报文的转发进行控制，只要进行适当的端口配置和 TIL 值的检测就可以了，但是在 HiNOC 网络中，HB 和 HM 是二层设备，类似于以太网中的交换机，无法区分组播和广播报文，组播报文会被当成广播报文转发给网络中的所有设备，而有些设备并不需要这些报文，这样就会浪费大量的资源，影响网络中的正常业务，因此需要在 HiNOC 网络中实现对组播的支持。IGMP Snooping（组播侦听）可以解决这个组播的问题。

（1） 二层组播技术

IGMP Snooping 协议属于二层组播协议，常见的二层组播协议有：

◇ GMRP (GARP Multicast Registration protocol)GARP 组播注册协议

◇ CGMP (Cisco Group Management Protocol)Cisco 组管理协议

◇ IGMP Snooping 组播侦听

GMRP[16]是通用属性注册协议（GARP）的一种应用，主要提供一种类似于 IGMP 探查技术的受限组播扩散功能。是 IEEE802.1D 定义的工业标准协议。允许网桥和终端站向连接到相同局域网段的网桥动态注册组成员信息，并且这些信息可以被传播到支持扩展过滤服务的桥接局域网中的所有网桥系统。GMRP 的操作基于 GARP 所提供的服务，利用 GARP 提供的信息声明和信息传播服务向交换机传播 GMRP 的组成员信息和组服务请求，交换机用这些信息动态更新端口和 MAC 地址的对应关系。优点是不依赖路由器，可扩展性好，转发速度快，支持的组数量多。缺点是在使用 GMRP 时，必须要配置主机和二层的交换机，即主机的网络接口卡和网络的交换机必须提供对 GMRP 的支持，而现有的设备并不是都支持 GMRP 的，并且 GMRP 并不依赖 IGMP。如果要实现二层的组播，还需要实现 IGMP 到 GMRP 的迁移，因为现在的组播软件都是基于 IGMP 的基础上编制的。

CGMP[17]是 Cisco 公司在 1996 年提出的，它是在路由器和交换机之间使用的一种通信协议，必须同时配置在组播路由器和二层交换机上，主要的工作方式是通过 CGMP 信息，交换机可以得到路由器的 IGMP 信息，并根据这些信息进行二层组播的转发。优点是实现简单，交换机不参与组管理。缺点是组播路由器和交换机都需要配置 CGMP 协议，并且要两者相互配合在一起使用。如果采用 CGMP 协议，路由器和交换机可能需要重新配置新的协议，这样不利于对网络拓扑进行多播的扩展。

IGMP Snooping 组播侦听[18][19]，是局域网交换机通过侦听主机和路由器之间不同类型的IGMP报文来动态的维护二层组播组。交换机侦听到特定组播组的IGMP报文时，就在交换机的组播表里记录该主机的端口号，当交换机侦听到主机

的IGMP离开组消息时，就从组播表中消除该主机的端口。IGMP Snooping的实现[20][21]和标准的IGMP协议的实现有相似之处，但IGMP Snooping其实并没有统一的国际标准，所以设计实现起来可以更加灵活高效。

采用IGMP Snooping的好处在于主机和路由器的软硬件设备不需要进行任何修改，只需要在两者之间加一个支持IGMP Snooping的交换机即可提供三层组播在二层的实现，将多播的范围扩展到边缘子网中。正是这个优点使得HiNOC网络可以利用IGMP Snooping来实现二层的组播，只要在HiNOC网络的结点设备上增加一个支持IGMP Snooping的模块即可。

（2） IGMP Snooping的功能

IGMP Snooping的任务是根据IGMP报文中的组播组信息和报文来源的端口，维护组播组和端口的对应关系，并且根据IGMP协议中主机的加入和离开消息同步更新组播表，这样交换机就能根据组播表进行数据转发。

IGMP Snooping的功能主要有：从端口侦听IGMP报文；根据IGMP报文信息创建和管理组播转发表；根据协议状态变化转发IGMP报文，以保持主机IGMP协议实体和路由器IGMP协议实体的状态一致性；解决在交换式结构中所引发的“应答泛洪”问题。

（3） IGMP Snooping协议的运行过程

IGMP Snooping在HiNOC网络中是主要是通过HB和HM来侦听主机和路由器之间的数据包，从而实现组播功能的，根据IGMP的消息类型来进行具体的处理。

1） HB对IGMP消息的处理

◇ HB从cable口收到一个IGMP加入组播组的消息。查找组播表，如果没有相应的组播信息，说明该主机是第一个申请加入的成员，把该组播组插入组播表，并向以太网口转发该信息。如果已经存在该组播组，就把发送该消息的HM的ID号插入组播表，但并不转发该消息。

◇ HB从以太网口收到IGMP成员查询消息。与HB相连的组播路由器会周期地发送IGMP成员查询消息，HB在收到此消息后会将其转发到HiNOC网络中的所有端口。这个组播组的所有主机都会产生成员报告消息，而HB只会向以太网口转发收到的第一个成员报告消息，因为路由器只关心本网段上有没有该组播组的成员而不关心这些成员是谁，这样可以减少带宽的浪费。

◇ HB从cable口收到IGMP离开组消息。HB收到此消息后会查询组播表，如果发现发送该消息的主机是该组播组的最后一个成员，则从组播表中删除该组播组，并向以太网口转发该消息。如果不是该组播组组播组的最后一个成员，则HB只是将该成员的ID号从该组播组成员表中删除，并不将该离开消息转发到以太网口。

2) HM对IGMP消息的处理

◇ HM从以太网口收到主机发送的IGMP加入组消息。HM收到该消息后，把该组播组加入组播转发表，并转发消息给HB.

◇ HM从cable口收到IGMP成员查询消息。HB会周期的转发IGMP成员查询消息。HM收到该消息后会从以太网口转发给主机，并等待主机的成员报告消息，如果主机想加入改组，就发送成员报告消息。HM收到该主机的成员报告消息就转发给HB。

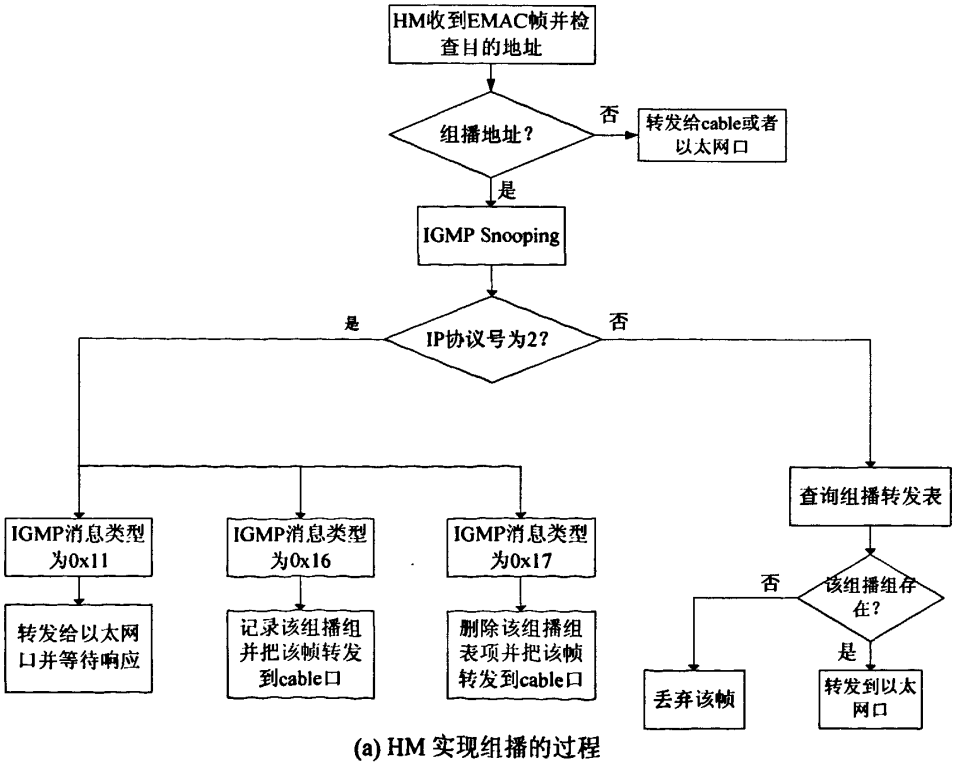
◇ HM从cable口收到IGMP离开组消息。HM收到主机发送的IGMP离开组消息就从组播表中删除该组，并把该消息转发到HB。

需要注意的是，在IGMP Snooping方式下，HB和HM必须检测每个组播数据包以防止漏掉IGMP控制信息，因为IGMP控制信息作为组播报传输，仅依赖二层的头信息无法区分和其他组播包。

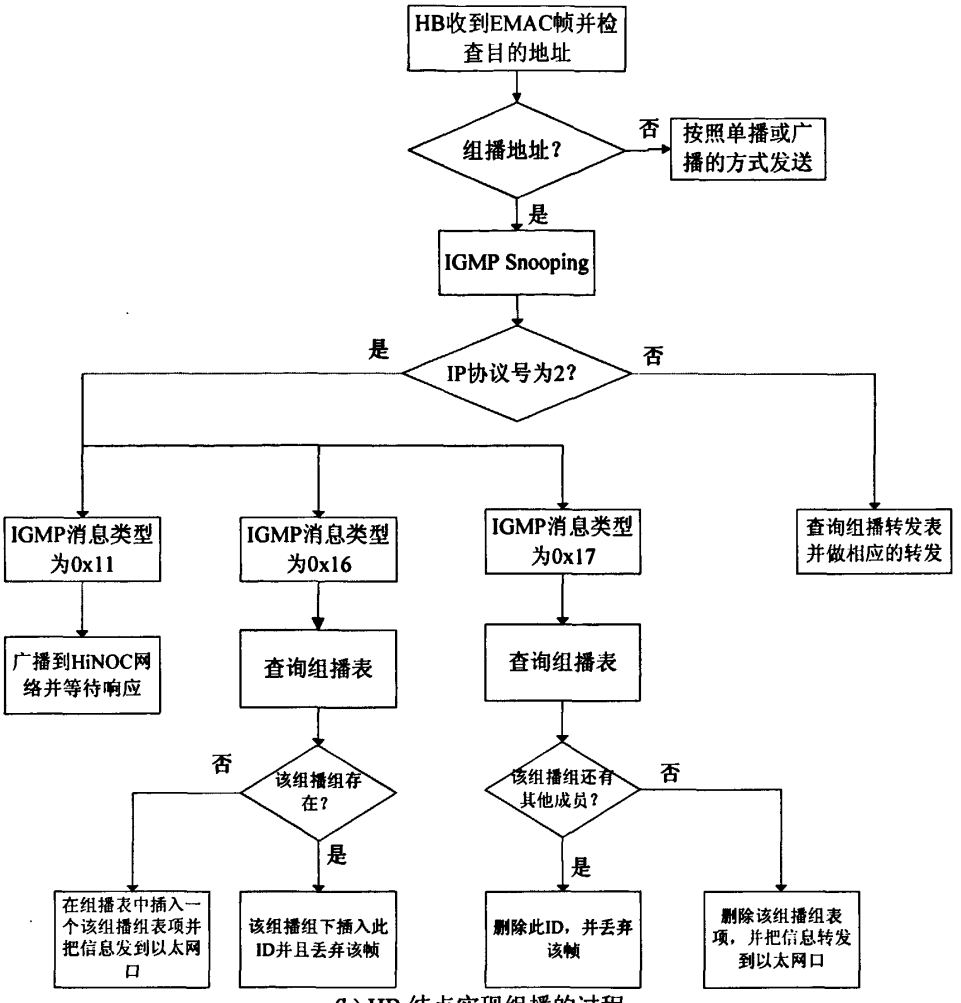
3.2.3 组播的实现

（1）组播实现描述

组播在 HiNOC 网络中的实现主要是 HiNOC 结点对组播的支持。由于 HB 和 HM 结点在网络中的功能不同，对组播的支持也是不同的。图 3.10 就是网络各结点实现组播的过程。



(a) HM 实现组播的过程



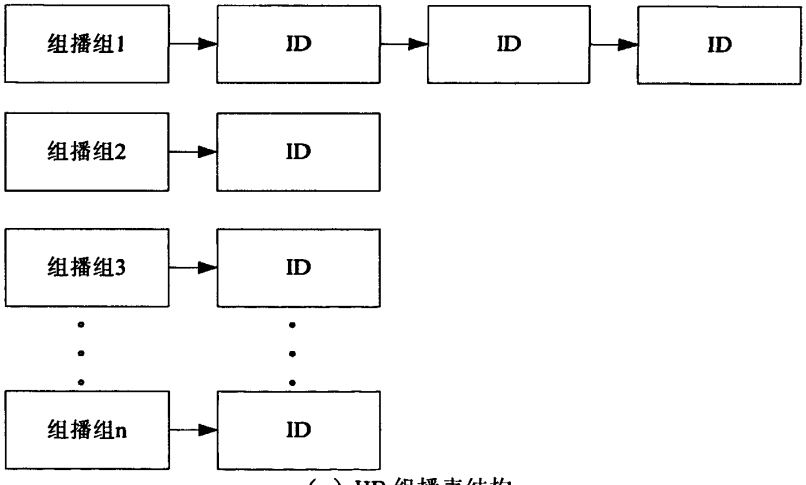
(b) HB 结点实现组播的过程

图 3.10 HiNOC 网络中结点对组播的支持的流程图

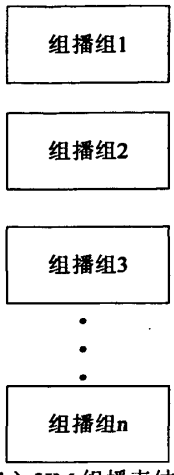
如图 3.10(a)和(b)所示，无论是结点 HB 还是 HM 收到 EMAC 帧以后都要检查 EMAC 帧的目的地址。如果目的地址是组播地址，需要将需要交给 IGMP Snooping 模块去处理。IGMP Snooping 会探测每个 EMAC 帧中的 IP 报文，检测 EMAC 帧头找出使用 IP 协议的 EAMC 帧，检查 IP 包头中的协议标示符，过滤出 IP 协议号为 2 的 IP 包即 IGMP 报文，然后根据 IGMP 报文中的消息类型进行相应的处理。不同的结点对同一个消息类型会进行不同的处理，如图 3.10(a)所示，当 HM 收到消息类型为 0x11 的 IGMP 报文即成员关系查询消息时，会把报文转发到以太网口即传送到与其相连的主机。而 HB 会把该帧封装以后广播到 HiNOC 网络中如图 3.10(b)所示。当收到消息类型为 0x16 的成员关系响应 IGMP 报文时，HM 只是记录下该组播组并把消息转发到另一侧的 cable 口，HB 则需要查找组播表，然后决定是插入该组播组表现还是只是插入成员的 ID 号。如果收到的报文是消息类型为

0x17 的离开组 IGMP 报文，HM 只需要删除该组播组并转发给消息到 cable 口即可，但是 HB 却需要查找组播表来决定是删除该组播组表项还是删除发送该消息的 ID 号。但是对非 IGMP 报文的组播数据帧，HB 的处理就会简单一些，只是查找组播表，并做相应的转发，而 HM 不但要查找组播表，还要根据组播表决定是转发还是丢弃该帧。HM 这个功能可以为主机正确接收组播信息提供一个重要的保证，因为主结点 HB 在转发表中一个组下有多个 ID 号的时候，组播消息就会被当成广播消息发送到 HiNOC 网络中的各个 HM 上，HM 则可以根据自己的组播表来确定是丢弃还是转发这些消息，这样加入组的主机就可以收到自己希望收到的组信息，而未加入组的主机就不会收到该组的信息，这样既减少了网络上的负担，也减少了主机处理无用消息的负担。

从以上的叙述可以看出 HM 中对组播的处理相对 HB 是比较简单的，这是因为 HiNOC 网络是星型结构，在 HiNOC 网络中 HM 只和 HB 直接通信，这样在建立组播表的时只考虑组播组。而 HB 不仅需要和网络中各个 HM 结点通信，还需要和外网进行通信，所以在实现组播上比较复杂。HB 在构建组播表的时候不仅要考虑组播组，还要考虑组播组下的成员。组播表的结构如下图所示：



（a）HB 组播表结构



（b）HM 组播表结构

图 3.11 HiNOC 网络中的组播表结构

链表的每一个结点表示一个组播组的信息，包括的信息有：组播组地址，每一个多播组都有各自的组播组地址。ID 号用来标识在这个组播组里面有哪些 ID 号属于组成员。因此可以定义组播表的结构体为：

```
typedef  struct  MultiAddrTab
{
    unsigned  long  MultiAddr; //组播地址
    unsigned  long  HiID; //HiNOC 网络 ID 号
    MultiAddrTab  *next;
} MultiAddrTab_S;
```

（2）IGMP Snooping 模块使用的主要函数

1） IGMP 报文结构：

由图 3.6 可知 IGMP 的报文格式 ，可定义 IGMP 的结构如下：

```
typedef struct
{
unsigned char Type; //IGMP 报文的类型
unsigned char MaxResponseTime; //最大响应时间值
unsigned short Checksum; //校验和
unsigned long GroupAddress; //组地址
} IGMPMessage;
```

说明：此结构体定义了 IGMP 报文的结构。IGMP 报文由 IGMP 报文的类型、最大响应时间值、校验和、组地址四部分组成。类型是用于区分 IGMP 报文的类别；最大响应时间值只有在 IGMP 报文类型是 IGMP 成员关系查询时才有效，它表示组播路由器从发出查询到收到响应的时间间隔；校验

和的计算方法和 IP 报文的校验和的计算方法相同；组播地址，在通用查询时此字段为零，其余均为有效的 IP 组播地址。

2) 插入组播地址表的函数：

int InsertMultiAddrTab(unsigned long MultiAddr,
unsigned short ID)

功能：用于创建或者插入组播地址表项，也可以对组播组下的成员进行修改。

参数：

unsigned long MultiAddr：EMAC 帧中的 MAC 组播地址。

unsigned short ID ：发送 EMAC 帧的 HM 的 ID 号。

返回值：返回 1 表示插入成功，返回 0 表示不成功。

调用者：被 IGMP Snooping 模块调用。

3) 查询组播地址表的函数：

unsigned short SearchMultiAddrTab（unsigned char*multi_addr_）

功能：该函数用来完成在组播地址表中查找组播地址与 ID 的对应关系，根据组播地址进行查找后返回 ID 号。

参数：

unsigned char*multi_addr_ ：指向组播地址的指针。

返回值：

unsigned short ：取值为 16bit 的 HiNOC 结点 ID

调用者：由 IGMP Snooping 模块调用

4) 删除组播地址表的函数：

int DelMultiAddrTab（unsigned long MultiAddr，unsigned short HiID）

功能：该函数用于在组播地址表中删除相应得组播地址表项，同时也可以将一个组播组中的成员进行删除，即删除该组播组下的成员的 ID 号。

参数：

unsigned long MultiAddr：EMAC 帧中的 MAC 组播地址。

unsigned short ID ：发送 EMAC 帧的 HM 的 ID 号。

返回值：

int：删除成功返回 1，不成功返回 0

调用者：被 IGMP Snooping 模块调用

5) 两个组播地址比较的函数：

int multiaddr_compare (unsigned char maddr_a[],
unsigned char maddr_b[])

功能：该函数完成两个组播地址值的比较

参数： unsigned char maddr_a[],unsigned char maddr_b[] 是两个需要比较的组播地址

返回值：

int：两个组播地址相等时返回 1，不相等时返回 0

调用者：

由 unsigned short SearchMultiAddrTab（unsigned char*multi_addr_）函数、unsigned long InsertMultiAddrTab(unsigned long MultiAddr,unsigned short ID)函数及 int DelMultiAddrTab（unsigned long MultiAddr，unsigned short HiID）函数调用，即在组播地址表中插入组播地址表项、查找组播地址及删除组播地址表时调用。插入时如果发现地址表中已经有相等的组播组地址，则只是该组播组下插入成员的 ID 号，否则插入新的组播组地址到组播地址表中；查找时如果找到相等的组播组地址，则返回该组播组下所有 HiNOC ID；删除时如果找到相等的组播组地址，则查看组播组下是否还是其他成员，若有其他成员则删除该离开组成员的 ID,若没有其他成员即该成员是此组播组的最后一个成员，则删除该组播组。

6) 接收报文的函数：

Void RecvMessage（unsigned short HiID，IP_S IpPacket）

功能：该函数用于接收各种 IGMP 报文并作相应的处理

参数：

unsigned short HiID ：16bit 的 HiNOC 网络 ID 号

IP_S IpPacket ：IP 报文

返回值：无

调用者：被 IGMP Snooping 模块调用

7) 查询子模块接收输入报文的函数：

Void ProcessQuery (IGMPMessage IGMPPacket)

功能：该函数用于接收 IGMP 查询报文并作相应的处理（处理过程如图 3.10 所示）

参数：

IGMPMessage IGMPPacket：IGMP 报文

返回值：无

调用者：被函数 Void RecvMessage（unsigned short HiID，IP_S IpPacket）调用

8) 报告子模块接收输入报文的函数：

Void ProcessReport(unsigned short HiID,IGMPMessage IGMPPacket)

功能：该函数用来接收 IGMP 成员响应报告的报文并做响应的处理（处理过程如图 3.10 所示）

参数：

unsigned short HiID：16 bit HiNOC 网络 ID

IGMPMessage IGMPPacket：IGMP 报文

返回值：无

调用者：被函数 Void RecvMessage（unsigned short HiID，IP_S IpPacket）调用

9) 离开组子模块接收报文的函数：

Void ProcessLeave(unsigned short HiID, IGMPMessage IGMPPacket)

功能：该函数用来接收 IGMP 离开组报文并作相应的处理（处理过程见图 3. 10）

参数：

unsigned short HiID：16 bit HiNOC 网络 ID

IGMPMessage IGMPPacket：IGMP 报文

返回值：无

调用者：被函数 Void RecvMessage（unsigned short HiID，IP_S IpPacket）调用

（3）IGMP Snooping 的测试方案

根据 HiNOC 网络结构组建实验网：4 台主机和 4 个网络结点。其中主机 3 作为组播源发送组播信息。HiNOC 网络中的设备结点可采用 IXP425 开发板[22][23]，简称 MAC 板。IXP425 开发板是根据 HiNOC 网络而开发的，可以运行 HiNOC 系统的 MAC 协议。这 4 个 MAC 板中，一个作为主结点 HB，其他 3 个作为结点 HM。并且设置结点 HB 的 ID 号为 0 ，HM 的 ID 号分别为 1，2，4 。组播在结点 HB 和 HM 实现是相似的，且主结点 HB 中会复杂一些，所以我们将以 HB 中实现为例进行说明。

1) 在结点中加入 IGMP Snooping 模块，加入一个组

主机 1 希望加入组播组 224.1.3.2，因此组播一个主动提供的 IGMP 成员关系报告给该组，该报告中的 MAC 目的地址为 0x0100.5e01.0302 。最初组播表上没有这个组播 MAC 地址的项（见图 3.12）。当 IGMP Snooping 模块收到主机 1 组播的 IGMP 报告时，会根据报告中的信息建立一个组播表项，如表 3-1 所示，该项包括主机 1 所连接的 HM 的 ID 号，组播组地址。

表 3-1 主机 1 加入后的转发表

目的地址	ID
01-00-5e-01-03-02	1

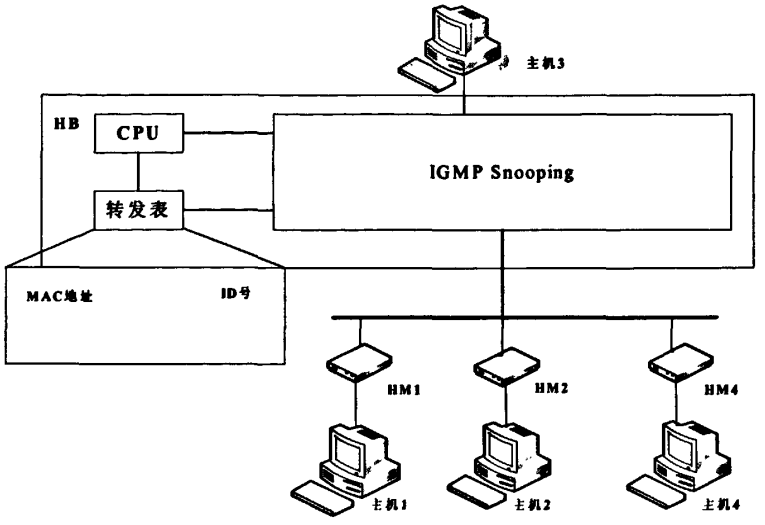


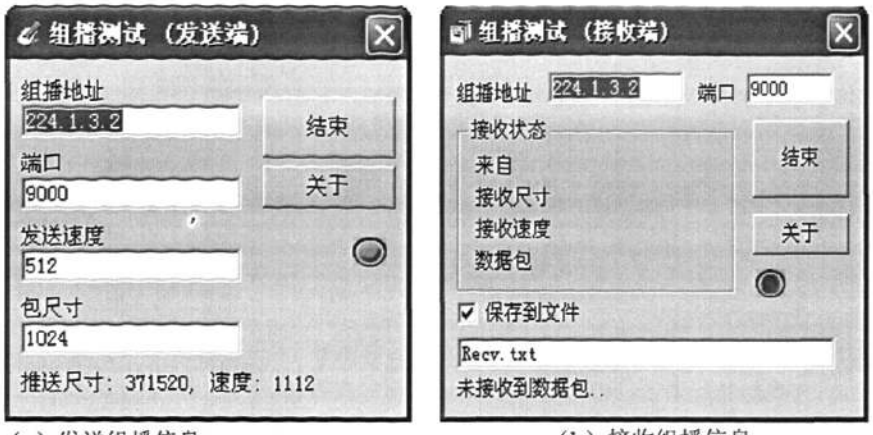
图 3.12 加入 IGMP Snooping 模块的 HiNOC 网络

如果主机 4 也想加入该组，并主动地发一个 IGMP 报告给该组，表 3-2 显示了主机 4 加入该组后的转发表的情况，报告经过 HM4 转发到 HB 上，交给 IGMP Snooping 模块进行处理，由于转发表中已经有该组播组表项，因此只在该组播组项后面加入与主机 4 相连的 HM 的 ID 号 4，由于路由器已经知道该网段上有该组播组成员了，因此 IGMP Snooping 模块不需要把该报告转发到路由器了。

表 3-2 主机 4 加入后的转发表

目的地址	ID
01-00-5e-01-03-02	1,4

此时我们可以让组播源主机 3 向目标组 224.1.3.2 发送信息，会发现在主机 2 上接收不到该组播信息如图 3.13 中(b)所示。由此可以得出结论，在 HiNOC 网络的结点中加入 IGMP Snooping 模块后，可以实现二层组播，组播不再以广播的形式来处理。



（a）发送组播信息 （b）接收组播信息

图 3.13 组播测试

2) 在 IGMP Snooping 方式下离开一个组

主机 1 接收数据完毕要离开组播组 224.1.3.2，通过向 224.0.0.2（MAC 地址 0x0100.5e00.0002）组播一个离开组消息声明它正在离开组。该消息经过 HM 时被提交到 IGMP Snooping 模块，经过相应的处理后被转发到 HB,HB 收到该消息后也把该消息提交到 HB 中的 IGMP Snooping 模块中，发现是离开组消息，就查询转发表，若该组播组下还有其他成员就删除与主机 1 相连的 HM1 的 ID 号，然后丢弃该消息。如表 3-3 所示，表中没有了 HM1 的 ID 号 1。这里和普通的交换机处理离开组消息不同的是 HB 不需要向接收到该消息的端口发送一个 IGMP 通用查询来作为对离开组消息的响应，这是因为每个 HM ID 下只有一个用户设备。

表 3-3 主机 1 离开后的转发表

目的地址	ID
01-00-5e-01-03-02	4

从表 3-3 中可以看出，组播组 224.1.3.2 中主机 1 所连接的 HM 的 ID 号 1 已经被删除。此时在主机 1 上测试组播，发现主机 1 上已无法收到组 224.1.3.2 的数据包了。

总之，在 HiNOC 网络的各结点中加入 IGMP Snooping 模块后，可以实现对二层组播的支持，组播数据包不再以广播的形式传送，这对 HiNOC 网络宝贵的网络带宽和主机资源来讲，有着重要的意义。二层组播在 HiNOC 网络中的实现，为多媒体信息清晰，流畅的在 HiNOC 网络传输奠定了坚实的基础。

3.3 VLAN 在 HiNOC 网络中的实现方案

3.3.1 VLAN 概述

（1）VLAN 概述

VLAN（Virtual Local Area Network）即虚拟局域网，是一种通过将局域网内的设备逻辑地划分成一个个网段从而实现虚拟工作组的新兴技术。它建立在局域网交换机的基础之上，通过 VLAN 用户能方便地在网络中移动和快捷地组建宽带网络，而无需改变任何硬件和通信线路。这样，网络管理员就能从逻辑上对用户和网络资源进行分配，而无需考虑物理连接方式。 VLAN 充分体现了现代网络技术的重要特征：高速、灵活、管理简便和扩展容易。网络的虚拟化是未来网络发展的潮流。

传统的抑制广播风暴的基本方法是隔离广播域。显而易见的解决方法是限制以太网上的结点，这就需要对网络进行物理分段。将网络进行物理分段的传统方法是使用路由器[24][25]，如图 3.14 所示。路由器的基本作用是只发送和接收来往于不同物理网段的信息。

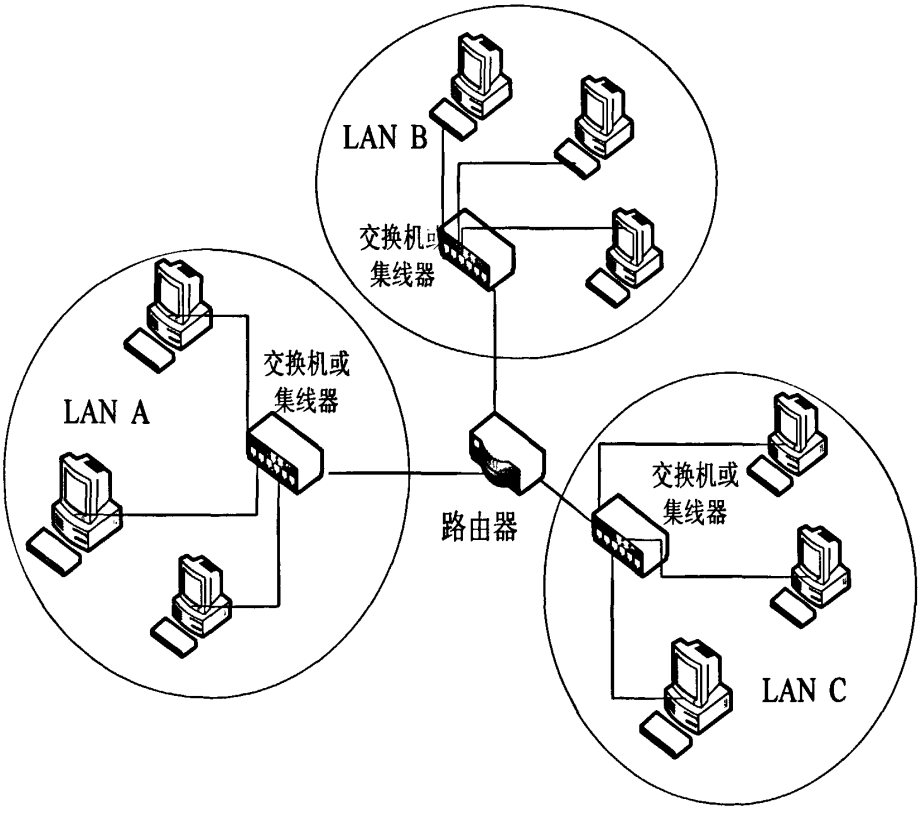


图 3.14 传统的 LAN

VLAN[26]是一种不用路由器解决隔离广播域的网络技术。VLAN 概念的引入，使交换机代替路由器承担了网络的分段工作，如图 3.15 所示。VLAN 打破了传统网络的许多固有观念，使网络结构变得灵活、方便、随心所欲。VLAN 不必考虑用户的物理位置，根据功能、应用等因素，将用户从逻辑上划分为一个个功能相对独立的工作组，每一个 VLAN 都可以对应于一个逻辑单位，如部门、项目组等。

同一个 VLAN 中的成员都共享广播，而不同 VLAN 之间广播信息是相互隔离的。这样，将整个网络分割成多个不同的广播域。例如，网络管理员可以把相关的客户和服务器分别构成不同的 VLAN，同一 VLAN 内客户和服务器可以方便地频繁通信，在同一个 VLAN 中的用户相互存取网络资源就如同在使用传统的局域网一样。

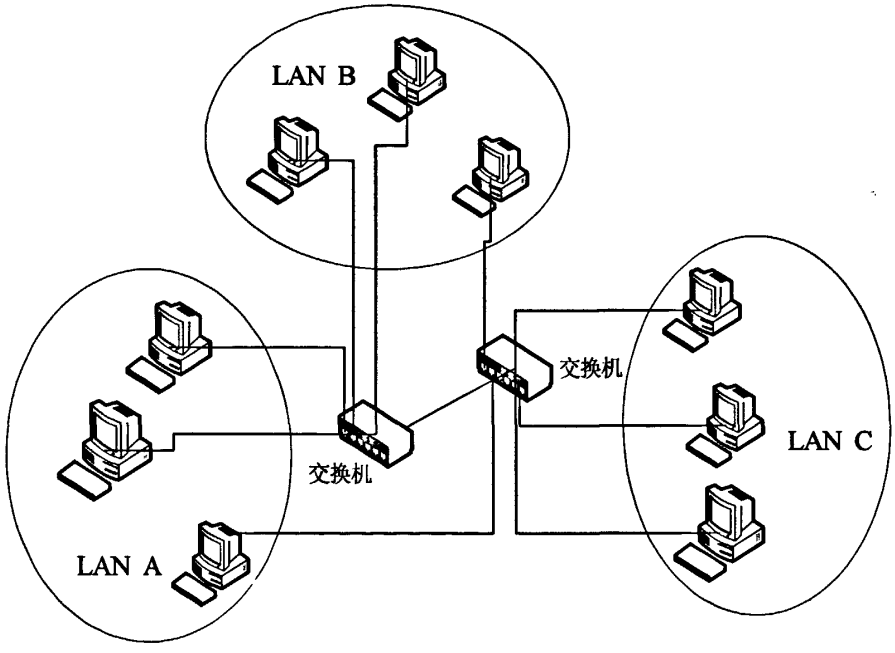


图 3.15 VLAN

（2）VLAN的优点

1) 控制广播风暴。在一个拥有上千台计算机的LAN中，大量的广播信息极有可能造成网络的阻塞,而由于不同的VLAN都处于不同的广播域，而广播只能在本地VLAN内进行，从而大大减少了广播对网络带宽的占用，提高网络传输效率，并可以有效的避免广播风暴的产生。
2) 方便管理。由于VLAN中的服务器和工作站可以不受地理位置的限制，对网络的管理和维护带来了方便，如果用户改变位置,只需更改用户所属的网络，不必更换用户所在的端口和联线，只须在定义VLAN 的交换机上重新定义VLAN即可。对于网络管理而言,可以轻松的完成变更。

3) 增强网络的安全性。由于交换机只能在同一VLAN内的端口之间交换数据，即便在同一交换机上，处于不同VLAN 的端口也不能互相通信，并且如果没有路由器，一个VLAN的数据包不会发到另一个VLAN，确保了该VLAN信息不被其它VLAN 窃听。因此，通过划分VLAN可以提高网络的安全性。

4) 增强网络连接的灵活性。借助VLAN技术，将不同地点、不同网络、不同用户组合在一起形成一个虚拟的局域网环境，在不改动物理连接的前提下可以将工作站在任意子网间移动，就像使用本地 LAN 一样方便、有效。

（3）VALN 的划分

交换机的端口，可以分为接入链接(Access Link)和汇聚链接(Trunk Link)。接入链接，指的是只属于一个 VLAN，且仅向该 VLAN 转发数据帧的端口。在大多数情况下，接入链路所连的都是客户机。如何设定接入链接，是 VLAN 应用的关键问题。接入链接的设定可是事先固定的，也可以是根据所连的计算机而动态改变设定。前者被成为静态 VLAN，后者为动态 VLAN。

静态 VLAN，又被称为基于端口的 VLAN，根据以太网交换机的端口来划分，这些属于同一 VLAN 的端口可以不连续，如何配置，由管理员决定， 优点是定义 VLAN 成员时非常简单，只要将所有的端口都指定义一下就可以了。缺点是如果 VLAN 的用户离开了原来的端口，到了一个新的交换机的某个端口，那么就必须重新定义。

动态 VLAN 则是根据接入交换机的计算机来决定这个端口是工作在哪个 VLAN 中的可根据 OSI 参考模型的层次分为 3 类。

◇ 基于 MAC 地址的 VLAN（MAC Based VLAN）

◇ 基于网络层的 VLAN（Subnet Based VLAN）

◇ 基于用户标识的 VLAN（User Based VLAN）

基于 MAC 地址的 VLAN[27] ,根据每个主机的 MAC 地址来划分，即对每个 MAC 地址的主机都配置它属于哪个组。这种划分 VLAN 的方法的最大优点就是当用户物理位置移动时，即从一个交换机换到其他的交换机时，VLAN 不用重新配置，所以可以认为这种根据 MAC 地址的划分方法是基于用户的 VLAN，这种方法的缺点是初始化时，所有的用户都必须进行配置，如果有几百个甚至上千个用户的话，配置是非常麻烦的。而且这种划分的方法也导致了交换机执行效率的降低，因为在每一个交换机的端口都可能存在很多个 VLAN 组的成员，这样就无法限制广播包了。另外，对于使用笔记本电脑的用户来说，他们的网卡可能经常更换，这样，VLAN 就必须不停的配置。

基于网络层的 VLAN[28]，根据每个主机的网络层地址或协议类型划分的，虽然这种划分方法是根据网络地址，比如 IP 地址，但它不是路由，与网络层的路由

毫无关系。优点是用户的物理位置改变了，不需要重新配置所属的 VLAN，而且可以根据协议类型来划分 VLAN，这对网络管理者来说很重要，还有，这种方法不需要附加的帧标签来识别 VLAN，这样可以减少网络的通信量。缺点是效率低，因为检查每一个数据包的网络层地址是需要消耗处理时间的，一般的交换机芯片都可以自动检查网络上数据包的以太网帧头，但要让芯片能检查 IP 帧头，需要更高的技术，同时也更费时。

基于用户标识的 VLAN[29]，根据交换机各端口所连接的计算机上当前登陆的用户来决定该端口属于哪个 VLAN。这些用户识别信息，一般是计算机操作系统登录的用户，也可以是用户事先申请的帐户。这些用户名信息，属于 OSI 第四层以上的信息。

而汇聚链接指的是能够转发多个不同 VLAN 的通信的端口。汇聚链路流通的是数据帧，都被附加了用于识别属于哪个 VLAN 的特殊信息，汇聚链接所连的都是支持 VLAN 的设备，一般用于交换机之间的互连。

3.3.2 基于 MAC 地址的 VLAN 划分机制

在上一节所述的动态 VLAN 的划分中，最流行的就是基于 MAC 地址定义的，它是在 OSI 的第二层设定访问连接的，根据用户端设备的 MAC 地址来定义 VLAN 成员资格的，即当设备连入一个交换机端口时，该交换机必须查询它的一个数据库以建立 VLAN 的成员资格。这种方式的优势很大，但是创建 MAC 数据库是一项比较艰苦和繁琐的工作。

（1） 工作原理

在形成 VLAN 的过程中，最重要的就是端口和 VLAN 之间的映射，在基于 MAC 地址的 VLAN 中，这个映射取决于管理员创建的数据库。当一个 PC 连接到交换机上时，交换机就缓存初始帧的源 MAC 地址，即 PC 的 MAC 地址。然后交换机便向 VMPS（Vlan Membership Policy Server ，VLAN 成员策略服务器）服务器发出请求，VMPS 服务器中包含一个文本文件，文件中存有 VLAN 和 MAC 地址的对应关系表。根据对应关系表，决定 PC 所连接的端口的归属和状态。

（2） 工作过程

在 HiNOC 网络中，HB 和 HM 都可以看成是二层的交换机，也可以从网外看，把两者当成一个交换机。由于 HiNOC 网络在物理上是总线型结构，在具体的实现中还是和普通的交换以太网还是有区别的。图 3.16 是基于 MAC 地址的 VLAN 的网络组件及工作过程：

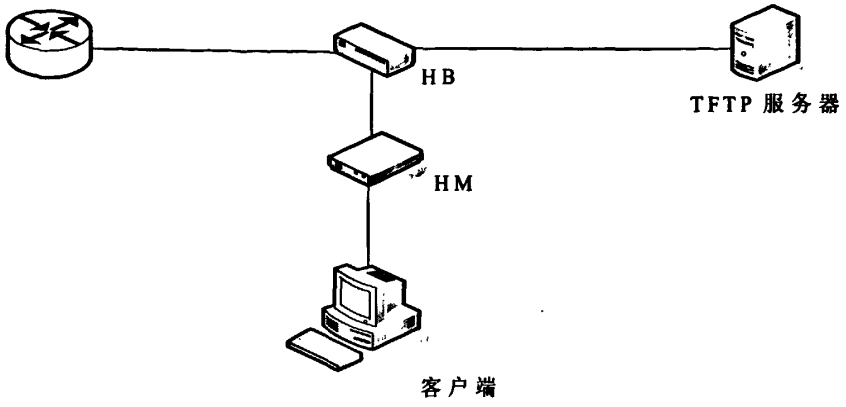


图 3.16 基于 MAC 地址的 VLAN 需要的网络组件

过程描述：

1) PC 机连接到终端结点设备 HM 的端口上，HM 通过 PC 机所发的帧记下该 PC 的 MAC 地址。
2) HM 向 HB 发送信息，HB 便向 VMPS 服务器请求下载 VLAN 和 MAC 地址对应的关系表。
3) HB 对 PC 的 MAC 地址进行查询比较，根据该 MAC 地址对应的 VLAN，把与该 MAC 地址相连的网络结点 HM 分配到该 MAC 地址对应的 VLAN 中，并形成 VLAN 和 HiNOC 网络 ID 之间的对应关系；如果没有该 MAC 地址的对应 VLAN，就不把与该 MAC 地址相连的 HM 划分到任何的 VLAN 中。

（3） 实现方案

1) 创建 VMPS 数据库

实现基于 MAC 地址的 VLAN 的第一步就是创建 VMPS 数据库，即建立 VLAN 和网内 MAC 地址的对应关系表，并把其存放在 TFTP 服务器或者装有 VMPS 软件的服务器的一个文本文件中，配置为 VMPS 服务器的设备或者计算机能够访问到该文件。数据库文件包括全局设置及 MAC 地址和 VLAN 的对应关系表。

2) 在 HB 中配置 VMPS 服务器

在 HB 中加入模块，使其可以配置成 VMPS 服务器，当启动该模块时，它可以将从 TFTP 服务器上下载 MAC 地址和 VLAN 的对应关系文件到 HB 上。在 HB 上本身要构建一张转发表，是 VLAN 和网络节点 HM 的 ID 号的对应关系。这张表是为了在 HiNOC 网络中实现正确的数据转发而建立的。当HB收到带有VIAN标识的EMAC帧时，就给查询该表，找到相应的VLAN号，就把该帧以广播的形式在 HiNOC 网络中发送。否则丢弃该帧。

3) 在 HM 中建立转发表

由于 HiNOC 网络是在物理上是总线型结构，而 HB 将带有 VLAN 标识的帧以广播的形式发送，这就意味着所有的 HM 都可以收到该帧。为了实现帧的正确转发，需要在 HM 中建立另一个转发表。这个转发表可以过滤一些不需要发给主机的信息。当 HB 根据主机的 MAC 地址和服务器中的 VLAN 和 MAC 地址的对应关系表，将与主机相连的 HM 划分到某个 VLAN 中后就会发送一个响应信息给 HM,该 HM 记录该 VLAN 号。如果 HM 收到下行方向带 VLAN 标识的 EAMC 帧，HM 就会查询自身的 VLAN 表，若有该 VLAN 号，就把该帧转发给主机；否则丢弃该帧。这样既可以减少主机的负担，又可以增加网络的安全性，保证同一个 VLAN 中的信息不会被转发出去。

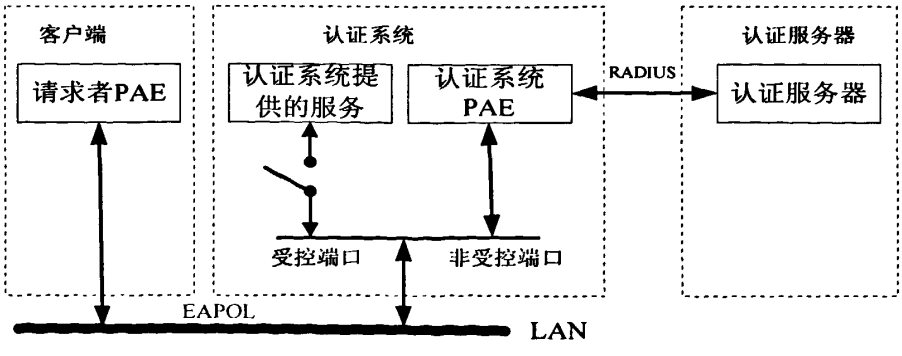
由于在基于 MAC 地址的 VLAN 中任何使用该主机的用户都可以获得该主机所属 VLAN 的服务，会造成信息的不安全性。因此提出了基于用户标识的 VLAN 的划分机制。

3.3.3 基于用户标识的 VLAN 划分机制

基于用户标识的 VLAN 划分机制[30]是一种以用户鉴别和鉴权为基础的 VLAN 访问方式，工作于 IEEE802.1x 的基础之上，针对不同的用户为交换机端口划分不同的 VLAN。克服了静态 VLAN 和基于 MAC 地址的动态 VLAN 所存在的安全问题。

IEEE802.1x 协议[31][32]是 IEEE 为了解决基于端口的接入控制（Port-Based Access Control）而定义的一个标准。起源于 802.11 协议，802.11 协议是标准的无线局域网协议，802.1x 协议的主要目的是为了解决无线局域网用户的接入认证问题。现在已经开始被应用于一般的有线 LAN 的接入。

(1) 实现机理



PAE：认证机制中负责处理算法和协议的实体

EAP：Extensible Authentication Protocol

图 3.17 基于用户标识的 VLAN 机制的认证体系

1) 客户端

一般是位于局域网链路终端的用户设备，是需要接入 LAN，及享受 switch 提供服务的设备（如 PC 机），客户端需要支持 EAPOL（Extensible Authentication Protocol over LAN）协议，客户端必须运行 802.1X 客户端软件，如：802.1X-complain，Microsoft Windows XP。用户通过启动客户端软件发起 IEEE802.1x 认证。

2) 认证系统

认证系统通常为支持 802.1 x 协议的网络设备(比如交换机)，它为请求者提供服务端口，该端口可以是物理端口也可以是逻辑端口，一般在用户接入设备(如 LAN Switch)上实现 802.1x 认证。它能有效地控制终端的访问权限并根据服务器的信息为相应的端口划分 VLAN。

3) 认证服务器

通常为 RADIUS[33] (Remote Authentication Dial In User Service)服务器，是为认证系统提供认证服务和 VLAN 划分的实体，该服务器存储有关用户的信息，比如用户的身份标识和密码及用户所属 VLAN 等等。当用户进行认证时，认证系统需要通过认证服务器来验证用户是否合法。如果用户合法，认证服务器通知认证系统，用户已经通过认证，把可控端口打开。此后，用户就可以正常地通过端口访问认证系统所提供的服务。认证系统和 RADIUS 服务器之间通过承载于 RADIUS 协议之上的 EAP 协议进行通信。

(2) 认证过程

IEEE802.1x 采用 EAP 点对点协议认证，EAP 消息被封装在 IEEE802.1x 消息中，称做 EAPOL。

EAP 消息有 EAP 请求，EAP 响应，EAP 成功通知和 EAP 失败通知四种类型。其中，只有 EAP 请求消息是可以在认证系统和接入客户端系统之间直接发送，其它的消息都是从认证服务器发给客户端系统，或者从客户端系统发给鉴权服务器的，认证系统只是完成中转和协议转换(认证系统与客户端系统之间用 EAPOL 承载，认证系统与认证服务器之间用其它高层协议，如 RADIUS)。

为了在 HiNOC 网络中实现基于用户标识的 VLAN，HB 和 HM 必须支持 802.1x 协议，从而完成对用户的认证。图 3.18 是基于用户标识的 VLAN 所需的网络组件，其中 HB 和 HM 是作为一个认证系统使用的。认证过程如图 3.19 所示。

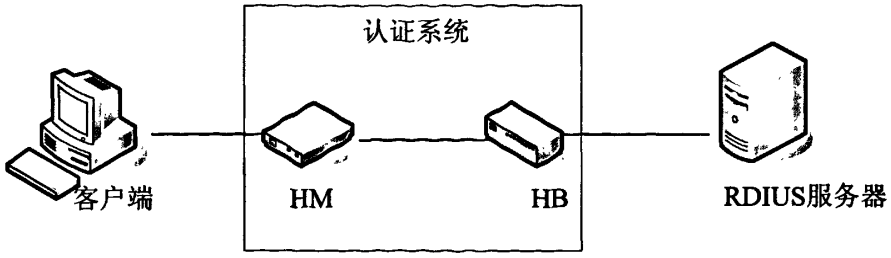


图 3.18 基于用户标识的 VLAN 的网络组件

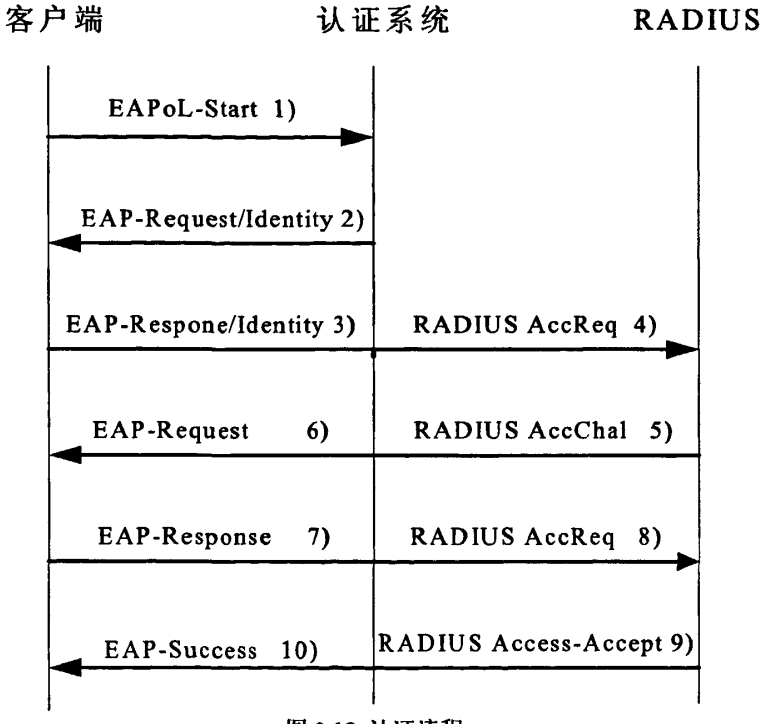


图 3.19 认证流程

流程说明：

1) 客户端向认证系统中 HM 发送一个 EAPoL-Start 报文，开始认证接入；
2) 认证系统向客户端发送 EAP-Request/Identity 报文，要求客户端将用户名送上来；
3) 客户端回应一个 EAP-Response/Identity 给认证系统中的 HM，其中包括用户名；
4) 认证系统中的 HB 将 EAP-Response/Identity 报文封装到 RADIUS Access-Request 报文中，发送给 RADIUS 服务器；
5) RADIUS 服务器产生一个 Challenge ，通过认证系统将 RADIUS Access-Challenge 报文发送给客户端，其中包含 EAP-Request/ Challenge；
6) 认证系统通过 EAP-Request 发送给客户端，要求客户端进行认证；

7) 客户端收到 EAP-Requeste 报文后，将密码和 Challenge 做认证算法后的 Challenge-Password，在 EAP-Response 中回应给认证系统；

8) 认证系统将 Challenge、Challenge-Password 和用户名一起送到 RADIUS 服务器，由 RADIUS 服务器进行认证；

9) RADIUS 服务器根据用户信息做认证算法，判断用户是否合法，然后回应认证成功/失败报文到认证系统。如果成功，携带协商参数、用户所属 VLAN 以及用户相关的业务属性给用户授权。如果认证失败，发送 RADIUS Access-Rejected 报文；

10) 如果认证通过，认证系统打开受控端口，HB 将给 HM 划分到 RADIUS 服务器所指定的 VLAN 中去，发送认证成功报文给客户端。如果认证失败，则发送 EAP-Failure 报文到客户端；

只有认证通过，客户才可访问网络资源，否则，用户需要重新进行认证。这个过程中需要在 RADIUS 服务器上建立用户认证信息时加入用户 VLAN 的信息，当认证通过时即可将用户的 VLAN 信息返回给认证系统，让用户所连接端口的 VLAN 属性更改为用户所属 VLAN。

（3）工作过程

认证完成后，将是正常的工作过程。认证通过后，客户就可以享受自己所属 VLAN 提供的服务。为了保证信息的安全性，在 HiNOC 网络中必须实现数据的正确转发。基于用户标识的 VLAN 在 HiNOC 网络内的数据的转发和基于 MAC 地址的 VLAN 的转发是一样的，不同的是基于用户标识的 VLAN 中的用户，在离开时，HM 会删除自己的 VLAN 表并向 HB 报告一个离开信息，HB 则根据信息把该 HM 的 ID 号从 VLAN 表中删除，如果该 VLAN 中只有一个 ID 号的时候将以单播的方式发送，这样可以节省很多的网络资源。

基于用户标识的 VLAN 综合了 IEEE802.1X 和 VLAN 的优点，优势更为突出：

1) 安全性

基于用户标识的 VLAN 中，终端系统能否获得 VLAN 访问权以及通过 VLAN 所能访问的内容决定于其身份标识。非授权的网络用户即使得到 HM 或主机的使用权，也会因为无法通过身份验证而被拒绝访问网络；而不同的网络授权用户即使使用了相同的 HM 或主机设备，其通过 VLAN 所得到的网络服务也因为这些用户定制的 VLAN 不同而不同。因此可以保证用户访问的合法性和安全性。

2) 可管理性

Radius 服务器在认证成功以后，可发送计费信息包给网络计费服务器实施对该用户的计费，并在用户请求退出网络或某种原因下线时，发送计费结束包给网络计费服务器以停止对该用户的计费。

3.4 本章小结

本章主要对 HiNOC 网络的汇聚子层进行了研究，介绍了汇聚子层的功能：通过地址学习构建转发表、地址查找和数据帧（即 EMAC 帧）转发、数据帧打包/拆包及对业务优先级的支持。根据 HiNOC 网络汇聚子层的特点，提出了基于 IGMP Snooping 的组播支持方案。在 HiNOC 网络中实现了二层组播，组播不再以广播的形式传送，为网络节省了宝贵的带宽资源。将 VLAN 技术引入 HiNOC 网络中，提出了针对 HiNOC 网络的基于 MAC 地址的 VLAN 和基于用户标识的 VLAN 划分机制。VLAN 技术的引入，使得在 HiNOC 网络中可以控制广播风暴，方便管理，并增强了网络的安全性。

通过本章的研究，我们可以看出汇聚子层对于 HiNOC 网络及 MAC 协议的重要性。下一章我们将对 HiNOC 网络性能进行测试，来进一步说明汇聚子层对 HiNOC 网络的重要性。

第四章 汇聚子层对 HiNOC 网络性能的影响

衡量网络性能的主要指标是网络的吞吐量和传输时延，HiNOC 网络也不例外。HiNOC 网络的性能主要取决于 HiNOC MAC 协议的性能，因此衡量 HiNOC MAC 协议的指标为 MAC 层的吞吐量和传输时延。MAC 层吞吐量定义为在 MAC 层测得的上层业务传输速率。传输时延定义为从上层业务帧到达 HiNOC 发送结点的 CS 子层开始，到 HiNOC 接收结点 CS 子层将该帧交付给上层经历的时间间隔。

本章主要对不同的参数取值下 HiNOC 网络的性能进行测试，并对测试结果进行分析，通过分析结果来说明汇聚子层对 HiNOC 网络性能的影响。

4.1 测试环境

HiNOC 网络是星型的网络拓扑结构，使用总线型的共享信道，由于以太网也是总线型拓扑结构，而且具有简便、通用、可扩展的特点，因而可使用以太网模拟 HiNOC 系统物理层的共享信道，搭建测试 HiNOC 网络性能的环境。

HiNOC 网络中的设备结点可采用 IXP425 开发板，简称 MAC 板。IXP425 开发板是根据 HiNOC 网络而开发的，可以运行 HiNOC 系统的 MAC 协议，其主要的硬件配置如下：

- ◇ 处理器：采用 Intel IXP425BD，BGA，533MHz XScale 内核，内部集成了 3 个高性能网络处理器引擎（NPE）。
- ◇ SDRAM：采用 64M×32bit 的 HY561620 SDRAM 内存。
- ◇ NOR FLASH：采用 4M×16bit 的 INTEL NOR FLASH 芯片。
- ◇ NAND FLASH：采用 32M×8bit 的 SAMSUNG NAND FLASH 芯片。
- ◇ LAN：采用 8 个二层线速全交换 10/100M 以太网，与 IXP425 内核的 NPEA 的 MII 接口连接（共 8 个交换网口）。
- ◇ cosolek UART：用做控制端口。
- ◇ RS232：采用 MAX3232，与 IXP425 自带的 2 个 RS232 接口相连。可以采用 com0 或者 com1 作为控制台。

IXP425 开发板的主要软件配置如下：

- ◇ BootLoader 软件：采用 RedBoot-1.94 和 2.02 版本，包含 Big-Endian 和 Little-Endian 两种字节顺序，在 HiNOC 系统 MAC 协议的实现中使用的是 Big-Endian 版本。支持 NPE 网口，支持通过 TFTP 将文件下载到开发板，支持 Linux 操作系统的装载和引导；强大的 Flash 读写操作和内存 Dump 功能；为向开发板上配置、加载、启动操作系统提供了方

便的手段。

- ◇ Linux 开发套件：采用 Linux-2.4.27 和 Linux-2.6.x 两个内核开发版本。全套完全移植好的源码树压缩包。在 HiNOC 系统 MAC 协议的实现中使用 Linux-2.4.27 版本。
- ◇ 开发工具：采用 gcc-4.0.1+gdb-6.4，glibc-2.3/uclibc0.9.29+binutils-2.6.1。
- ◇ 根文件系统：基于 uclibc 和基于 glibc 的根文件系统全套源码及预编译压缩好的映像文件。

图 4.1 为利用以太网作为测试 HiNOC 系统共享信道的组网方式。其中一块 MAC 板作为中心结点 HB，其他的 MAC 板作为结点 HM 来使用。由于 HiNOC MAC 协议规定 HB 下所连接的 HM 最多为 32 个，因此在测试中所用的 MAC 板不能超过 32 个。MAC 板的两侧是以太网口，一个端口连接用户设备另一个端口连接交换机。用户设备是为了提供上层业务来测试网络的性能。用户设备与 MAC 板之间形成了一个以太网，称为 Ethernet1，它所承载的是各用户的数据业务，而各 MAC 板之间通过交换机又形成了一个以太网，称为 Ethernet2。当客户机的数据业务到达 MAC 板后，经过 HiNOC 系统 MAC 协议的处理，形成 HiNOC 网络中的数据帧，在 MAC 板的出口处再次封装为以太网数据帧，该数据帧通过 Ethernet2 传输到对端的 MAC 板，经过相反的解析过程到达对端的客户机，完成一次数据帧的传送。在实际的测试中，将采用 4 个 MAC 板作为网络中的结点，一个为结点 HB，三个为结点 HM。

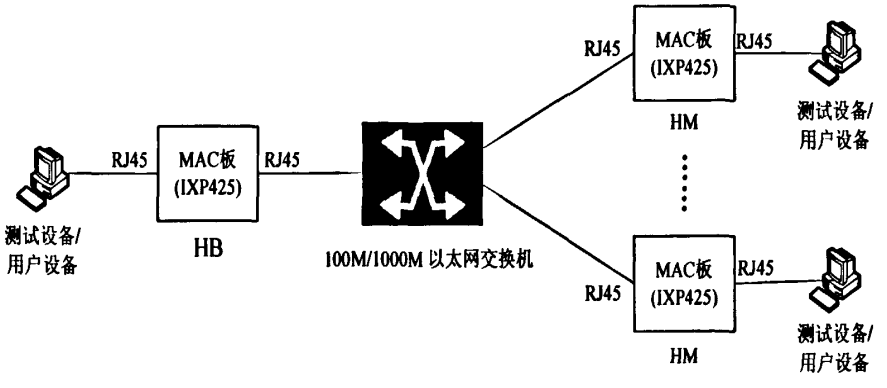


图 4.1 HiNOC 网络测试环境

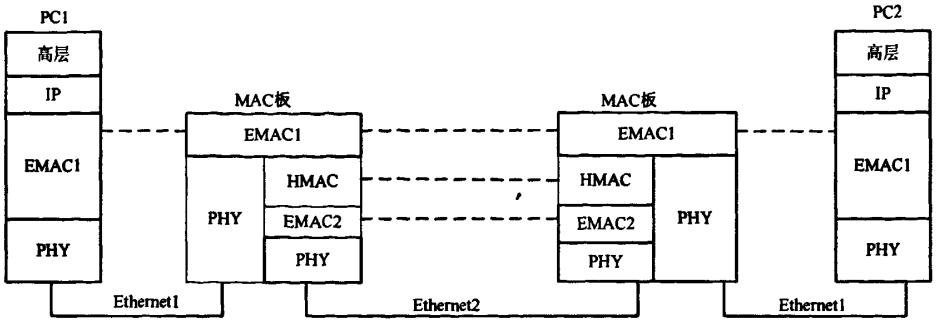


图 4.2 HiNOC 网络测试环境对应的协议栈

如图 4.2 所示，HiNOC 网络测试环境对应的协议栈。在 Ethernet2 共享信道上，协议栈的关系为 EMAC1 over HMAC over EMAC2，其中 EMAC1 表示客户机以太网端口的 MAC 协议，EMAC2 为 MAC 板上以太网端口的 MAC 协议，而 HMAC 为被考察的 HiNOC 系统的 MAC 协议。从 HMAC 角度看，EMAC1 为上层业务，EMAC2 为物理层。EMAC 以及对应的 100M/1000M 以太网的功能是仅仅为 HMAC 协议研究提供了一个模拟的共享物理信道。

4.2 测试的方法

在本测试中采用 SmartBits 捕获数据帧来分析 HiNOC 网络测试环境中各个结点接收和发送 MAC 帧的情况。SmartBits 是一种网络性能分析设备，它能够对流过路由器、中继器、网桥和网卡等网络设备的数据包进行捕获、测试和分析，统计数据包的吞吐量、丢失率、延迟和延迟分布等性能参数。使用 SmartBits 捕获流过交换机的数据帧，并保存为以 pcap 为扩展名的文件，用于后续的分析。在测试中 SmartBits 使用方式为环路模式见图 4.3。

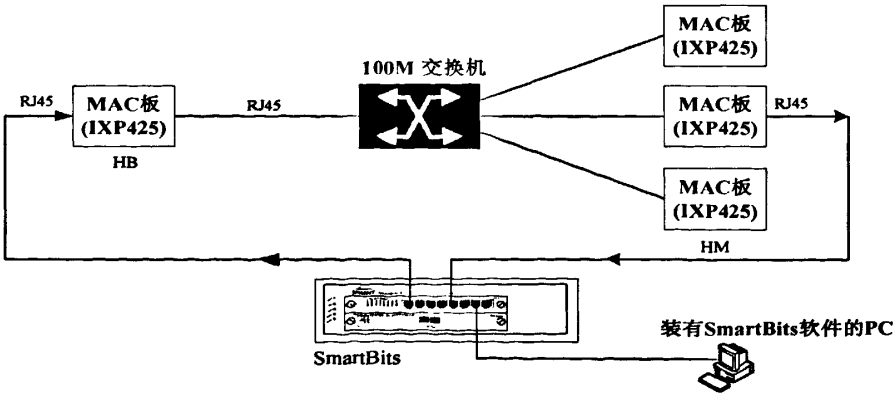


图 4.3 HiNOC 网络性能测试方法

图 4.3 为 HiNOC 网络性能测试方法，图中 SmartBits 的一个端口连接装有

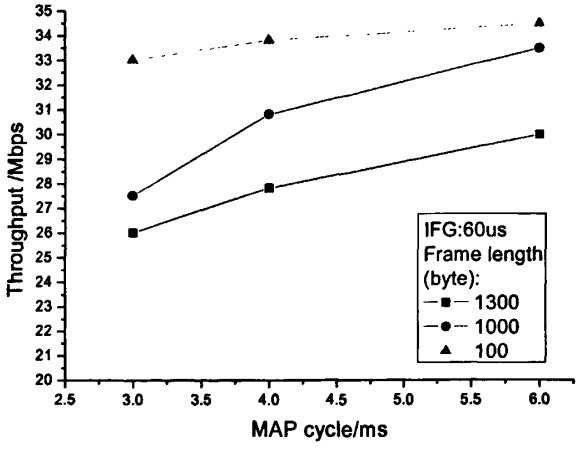
SmartBits 软件的计算机，另外两个端口分别连接两个 MAC 板的以太网端口。SmartBits 产生上层数据业务，通过其中一块 MAC 板发送出去，数据进入调测试验网络，SmartBits 通过另外一块 MAC 板接收相应的数据帧。因为 SmartBits 产生的数据帧包含了特定的标志信息，在发送数据时记录该信息，在接收到数据帧时验证该标志信息，这样就可以统计这两块 MAC 板之间的相关参数。由于 SmartBits 具有多个类似的端口，因此可以同时测试和分析多对 MAC 板间的性能参数，从而得到 HiNOC 网络的性能指标。

4.3 吞吐量的测试

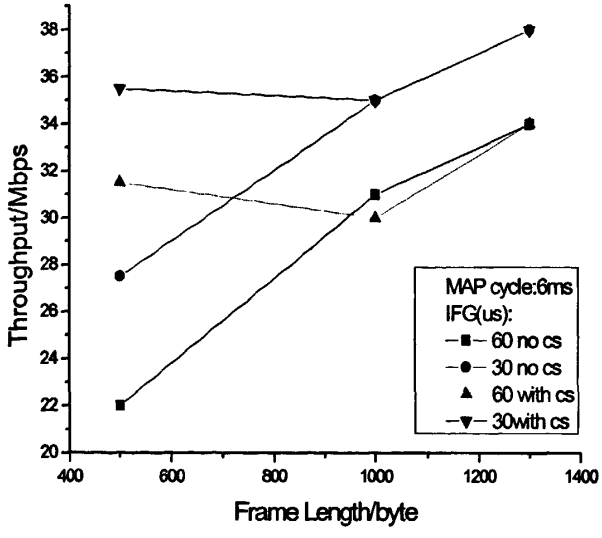
HiNOC MAC 协议中帧长、帧间隔和 MAP 周期长度等参数取不同的值时将会对 HiNOC 网络的吞吐量产生影响，是否采用汇聚机制也会对 HiNOC 网络的吞吐量产生影响。这里所以说的汇聚机制指的是汇聚子层的打包功能。

在以太网协议中规定以太网 MAC 帧的最大长度为 1518 字节，使用 VLAN 机制后的最大长度是 1522 字节（VLAN 标识 4 个字节）。因此我们在测试中，在 Ethernet2 共享信道上传输的帧不能超过这个最大长度限制。这个规定打包后的帧长不能超过 1400 字节，因为还要考虑到 HiNOC MAC 帧成帧时会加上 HiNOC MAC 帧头开销及一些调试信息。

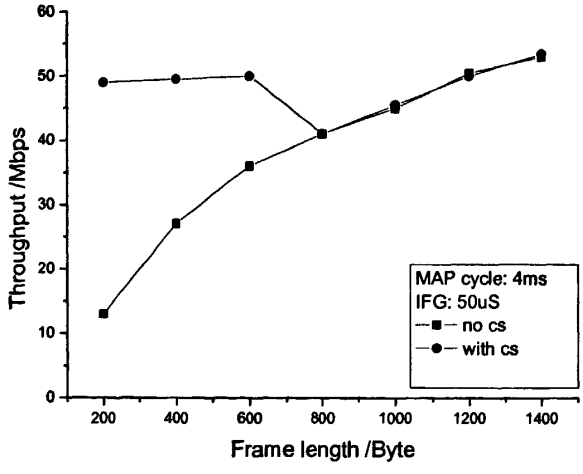
图 4.4 是根据测试结果而得出的，在测试中采用固定和不固定相结合的方法，例如：有固定的帧间隔、不固定的 MAP 周期、固定的帧长在有无汇聚机制下的的吞吐量；固定的 MAP 周期、不固定的帧间隔、不固定的帧长在有无汇聚机制的吞吐量；固定的 MAP 周、固定的帧间隔、不固定的帧长在有无汇聚机制下的吞吐量。



(a)



(b)



(c)

图 4.4 有无汇聚下网络吞吐量的测试结果

图 4.4 中的(a)就是在有固定的帧间隔、不固定的 MAP 周期、固定的帧长有无汇聚机制下测得的网络吞吐量。从图中我们看出帧长为 100 字节的帧在网络中传输时网络的吞吐量明显的很高，这是因为 100 字节的 EMAC 帧经过汇聚，最大可以汇聚成 1400 字节的 HiNOC MAC 帧，因此吞吐量会大于 1000 字节 EMAC 帧的吞吐量。也可以从图中看书在 MAP 周期小于 4ms 时，吞吐量变大比较大，过了 4ms 以后变化就比较小了。图 4.4 中(b)是在固定的 MAP 周期 6ms，不同的帧间隔和不同的帧长有无汇聚机制下测试所得的网络吞吐量。在图中通过对比可知，在有汇聚的情况下帧长越小，网络的吞吐量越大，这是因为帧长越小，打包后越接

近传输帧的极限。而在无汇聚情况下的情景刚好相反，这是因为无法充分利用带宽资源。但是总的来说在帧长小于 1000 字节之前，有汇聚时的吞吐量明显的大于无汇聚时的吞吐量。在帧长大于 1000 字节以后，两者并没有大的区别，这是因为在 1000 字节以后，受底层传输的限制，已经无法打包了，所以两者的吞吐量也就基本一样了。图 4.4(c)是在固定的 MAP 周期 4ms。固定的帧间隔 50us 和不同的帧长有无汇聚的情况下的吞吐量。这个取值是结合图 4.4(a)和(b)，为了取得更好的吞吐量而选择的。从图中可以看出在参数的取值下，测得的结果明显的好于图 4.4(a)和(b)中的结果，并且在汇聚情况的吞吐量更是有了明显的提高。不过在帧长大于 800 字节以后，两者又基本一样。

从测试的结果及分析可以得知汇聚子层对于 HiNOC 网络的重要性，对 HiNOC 网络的吞吐量起着主要的作用。

4.4 传输时延的测试

HiNOC 网络的传输时延包括排队时延和帧发送迟延两部分；其中排队时延又可以分解为打包时延、预约时隙等待时延、预约时延等几个部分。

打包时延指的是为将多个上层数据帧组合成一个 HMAC 帧而产生的时延。它主要由业务到达速率、最大打包帧长以及最大打包等待时间决定。当业务到达速率为 25Mbps（HDTV），打包后最大帧长为 4588Byte 时，打包时延为 1.47ms。为避免因业务帧到达速率较慢而产生过大的打包时延，协议还应规定最大打包等待时间。当开始对一个新 HMAC 帧进行打包时，最大打包等待定时器开始计时；若定时器超时而 HMAC 帧尚未达到最大帧长时，立即将已有的上层数据帧进行封装并发送。最大打包等待时间与上层业务的类型有关，对 VoIP 等实时业务该时间取值较小；具体取值大小需要根据业务类型和实际组网情况进行进一步研究。

预约时隙等待时延是各结点为获得中心结点分配的预约请求时隙而等待的时间。当子结点数目为 32（最大）时，每个 MAP 周期内最多允许 6 个结点发送预约请求，因此某个子结点等待预约时隙的时延最大为 6 个 MAP 周期，即 24ms。

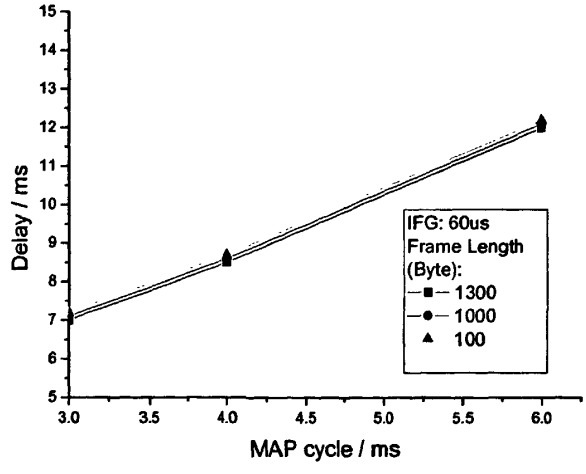
预约时延是子结点从获得预约请求发送时隙，到获得中心结点许可在规定时隙开始发送 HMAC 帧的时延。这一时延与网络中当时到达的总业务量有关。按照协议，若当前 MAP 周期内的一个预约请求被立即许可，则可以在相邻的下一个 MAP 周期内的规定时隙开始发送，此时的预约时延不超过 2 个 MAP 周期即 8ms。当各结点申请发送的业务量较大时，如果某结点的请求在下一个 MAP 周期内未能被许可，则可以在其后的 MAP 周期内被许可（但一定不能晚于同一子结点再次获得预约请求时隙的 MAP 周期），此时迟延会相应增大。

帧发送时延是一个 HMAC 帧发送到信道上所需的时间，是该帧对应的物理帧

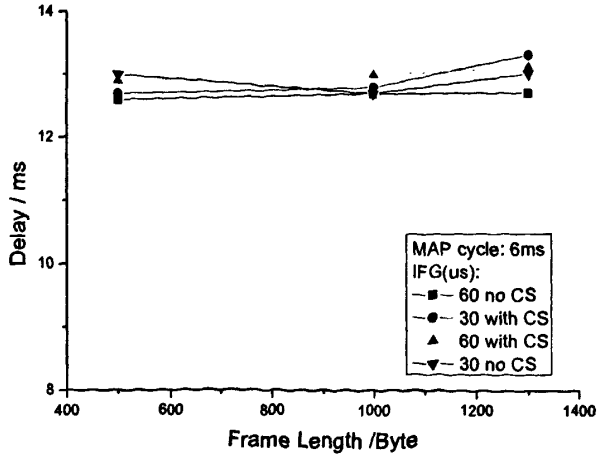
长与信道物理层速率的比值。当 HMAC 帧长为 4588 字节且 OFDM 各子载波采用 256QAM 调制时，帧发送时延小于 1ms。

通过上述分析我们看到，影响网络传输时延的主要因素是预约时隙等待时延和预约时延。打包时延对网络传输时延的影响可以忽略不计，即汇聚子层对网络传输时延不大。也可以从以下的测试结果得到证实。

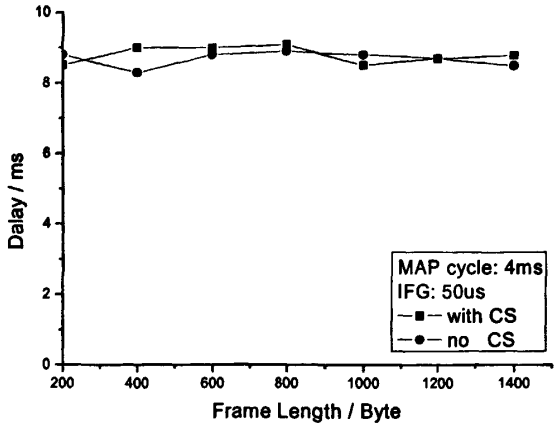
图 4.5 是网络在协议参数的不同取值下测得的传输时延结果，测试方法与吞吐量的测试基本相同，并且采用与测量吞吐量时相同的参数。由图 4.5 可知，在有汇聚机制下网络时延要比没有汇聚机制下网络时延大一些，但是并不明显，基本上可以忽略不计，测试的结果和上述分析基本上是一致的。但是 HiNOC 系统的 MAC 协议采用汇聚机制是必须的，因为和汇聚机制带来的巨大的吞吐量相比，牺牲一些时延也是值得的。



(a)



(b)



(c)

图 4.5 有无汇聚情况网络传送时延的测试结果

4.5 本章小结

网络性能是衡量一个网络好坏的标准，也是衡量网络协议好坏的标准。网络性能的指标数主要有网络吞吐量和传输时延。本章对 HiNOC 网络的吞吐量和时延进行了测试，本次测试主要是在有无汇聚情况下对网络的吞吐量和时延测试的，这里说的是汇聚主要是指汇聚机制中的打包功能。通过对测试结果的分析，可以看出汇聚子层对 HiNOC 网络性能有很大的影响，尤其是对网络的吞吐量，在有汇聚情况下的吞吐量比没有汇聚情况下的吞吐量有很明显的提高，从而说明汇聚子层的打包功能是必须存在的。

第五章 结束语

HiNOC 网络作为宽带家庭接入网络，解决从 FTTH 或 FTTB 的光纤结点到用户的最后 100 米接入，主要作用是承载高层业务，所以须提供对上层业务的支持。从协议分层模型看，HiNOC 系统由 HiNOC MAC 层和物理层（PHY）构成。其中 MAC 层从下到上又可分为公共部分子层（CPS）和汇聚子层（CS）两部分。汇聚子层是 HiNOC MAC 层的一个功能子层，主要处理数据业务，完成数据打包、封装和转发，并区分业务优先级。在 HiNOC 系统中，HB 和 HM 本质上可以看作是以太网交换机，属于二层设备，对组播包和广播包不加区别，HB 在收到组播包后将其广播出去，这样不但浪费了大量的网络带宽，也影响了正常业务，因此需要将二层组播技术引入，在 HiNOC 系统中实现对组播数据的支持。为了进一步优化 HiNOC 网络，将 VLAN 技术引入到 HiNOC 网络中，VLAN 是一种不用路由器解决隔离广播域的网络技术。 VLAN 不必考虑用户的物理位置，根据功能、应用等因素，将用户从逻辑上划分为一个个功能相对独立的工作组，每一个 VLAN 都可以对应于一个逻辑单位，如部门、项目组等。

本文首先对 HiNOC 网络及 HiNOC 系统 MAC 协议进行了分析，在此基础上，重点研究了 MAC 协议中的汇聚子层功能；设计了针对 HiNOC 网络的基于 IGMP Snooping 的二层组播方案，通过实验证明了该方案的可行性；将 VLAN 技术引入 HiNOC 网络，提出了基于 MAC 地址的 VLAN 划分和基于用户标识的 VLAN 划分机制；最后对 HiNOC MAC 协议性能进行了测试，研究了汇聚子层的设计机制对 HiNOC 性能的影响。

本文在设计 HiNOC 网络的组播和 VLAN 方案时，只是考虑如何针对 HiNOC 网络的特点进行支持，而没有考虑一些后续的问题：

1. 组播技术是一种高效利用宽带的网络技术，但是 IP 组播广泛采用数据包协议进行信息传送，加之 IP 网络是一种“尽力而为”的网络技术，并且传统的 IP 组播模型是开放式的，任意主机都可以创建组播组，接收和发送组播数据，这就使得 IP 组播网络在应用中会遇到诸多问题：用户管理、业务管理、访问控制等。这些问题严重制约了组播业务的进展，如何更好的对组播业务进行管理，提高组播业务的实施性能力，加强组播业务的安全性将是目前以及下一步研究的重要课题。
2. VLAN 是一种新兴的网络技术。它为解决网络站点的灵活配置和网络安全性等问题提供了良好的手段。本文只是提出了一些可行方案，并未进行实现，如何在 HiNOC 网络中进行实现，将在以后的工作中进行完成。

致 谢

首先，要特别感谢我的导师张冰老师。无论在科研工作中，还是在论文的选题、资料查询、研究和撰写的每一个环节，都得到了张老师的悉心指导和帮助。两年多来，张老师严谨认真的治学态度，孜孜不倦的工作作风，丰富渊博的专业知识，幽默风趣的谈吐举止都深深影响着我，使我受益匪浅。在此对张老师表示衷心的感谢。

感谢实验室的张奭老师，在科研中和论文进展，都给予及时的建议和指导，使得科研能够顺利进行。

感谢李严梅、崔金、金鑫、冉冬卉等同学和朋友们，在过去的岁月里，他们在学习和生活中给了我很多帮助和鼓励，陪伴我度过了一段快乐的人生路途。还要感谢实验室的其他师兄师弟们。

感谢同一寝室的姐妹们，感谢王淑君、许唐雯、汤少钰和杨红梅给予的鼓励和帮助，和你们我度过了两年多年的美好时光，快乐的完成研究生的工作。

最后，要特别感谢我的家人。血浓于水，在我成长的日子里，他们给予我无限的关心和爱，始终是我最坚强的后盾，为了我的成长默默无闻地奉献着一切，我将继续努力进取，不会让你们失望的！

感谢百忙之中参与评审的各位专家、教授。

参考文献

[1] John A. C. Bingham. ADSL, VDSL, and multicarier modulation . New York : John Wiley & Sons, c2000.

[2] 郭士秋编著. ADSL 宽带网技术. 北京 ：清华大学出版社，2001.

[3] 冯建和, 王岚编著. ADSL 宽带接入技术及应用. 北京：人民邮电出版社，2002.

[4] David Fellows, Doug Jones. DOCSIS Cable Modem Technology. IEEE Communications Magazine, March 2001. 202～209.

[5] 何英, 陆正福, 方刚. 二层交换机中IGMP/MLD探听机制的原理分析. 云南民族大学学报. 2005，1(14). 78-82.

[6] 陈文革，程向前. 高速地址 Cache－－散列表的应用. 计算机应用研究. 2002. 82－84.

[7] 严蔚敏，吴伟民著. 数据结构（C 语言版）. 北京：清华大学出版社，2002.

[8] 岩延，郭江涛等编著. 组播路由协议设计及应用. 北京：人民邮电出版社，2002.

[9] Marina Smith 著, 黄锡伟，王涛译. virtual lans 虚拟局域网. 北京：清华大学出版社, 2003. 29-47.

[10]Fenner, W. Internet Group Management Protocol, Version 2. RFC2236, Nov 1997.

[11]IEEE Std 802.1Q-2005.Local and Metropolitan area networks－Virtual Bridge Local Area Networks, 2005.

[12]Cain B. Internet Group Management Protocol, Version 3. RFC 3376, Oct 2002.

[13]林予松, 贺磊, 汪斌强. IPv6组管理协议MLD研究. 计算机工程与应用. 2004. 36 131-133.

[14]Vida R. Multicast Listener Discovery Version2 (MLDv2) for IPv6. RFC 3810, Jun 2004.

[15]Beau Williamson. IP组播网络设计开发(第一卷). 北京：电子工业出版社 2000, 6. 417-453.

[16]IEEE 802.1D. Media Access Contro(MAC)Bridge. Aug 2001.

[17]Cisco. Multicast in a Campus Network CGMP and IGMP Snooping.

[18]杜旭，张连靖，余江. IGMP Snooping协议实现方案. 计算机应

用. 2004, 6(24). 14-15.

[19]程传庆. IP 组播组管理协议及其在二层的实现. 信息技术. 2003, 7(27). 50-52.

[20]张美枝，贺思德，尹赛娥等. 基于IGMP Snooping组播在校园网多媒体系统中的研究与设计. 西北大学学报. 2004, 10(34). 221-224.

[21]尹海春. 二层以太网交换机上组播协议的实现. 江南大学学报. 2003年 10(2). 361-364.

[22]李福林，李立新，徐开勇等. INTEL 网络处理器 IXP425 应用分析. 电子产品世界，2004，4.

[23]徐军宁，都斯丹，高敦堂. 网络处理器 IXP425 及其应用. 微计算机应用. 第 25 卷第 1 期. 28～31.

[24]谢希仁编著. 计算机网络（第四版）. 北京：电子工业出版社. 2003.

[25]杨义先. 计算机网络. 北京：北京邮电出版社，2000.

[26]孙炜, 张朝辉, 周坤. VLAN技术研究. 通信与计算技术. 2005，第4期 . 26-31.

[27]陈伟川. 基于MAC地址的动态VLAN原理及组网应用. 计算机与现代化. 2007，第4期. 107-109.

[28]夏龄. 利用三层交换机建设校园网VLAN. 计算机应用，2002，22卷第6期. 234-236.

[29]施晓秋, 陈文铺, 张新楚. 基于用户标识VLAN在宽带小区接入中的应用与实施. 计算机工程. 2004 , 8, 30卷. 152-154.

[30]王华, 陈沛, 高传薯. 基于用户鉴别的虚拟网络. 计算机工程与应用. 1999, 8. 84-86.

[31]田辉，张彤. 基于 802.1x 认证技术的应用分析 http://technology.ctl.com.cn/jshlts how.php?id=8788 2005.3。

[32]IEEE std 802.1X-2001,IEEE Standards for Local and Metropolitan Area Networks:Port-Based Network Access Control[S].

[33]IETFRFC2865-2000.Remote Authentication DialIn User Service (RADIUS).

硕士在读期间的研究成果

一、参加科研情况

“利用有线电视带外信道的多业务宽带接入网（HiNOC）技术研究”，国家“863”项目。

二、发表论文情况

杨红梅,胡予濮,赵英华,高玮.《标准模型下的基于口令的群密钥协商协议》，论文已被 《计算机工程》录用，拟定在2009年20期上刊出。