第 33 卷　第 11 期　　计　算　机　工　程　　2007 年 6 月
Vol.33　No.11　　Computer Engineering　　June 2007
·网络与通信·　　文章编号：1000—3428(2007)11—0107—03　　文献标识码：A　　中图分类号：TP393.02

基于接入汇聚路由器的IPTV组播的实现

张佳杰，扈红超，刘　强，汪斌强
(国家数字交换系统工程技术研究设计中心，郑州 450002)

摘　要：依据国家“863”计划中 ACR 的规范，提出了基于 FPGA 的 IPTV 组播实现方案，完成了 10Gb 接口板 IPTV 组播的 FPGA 程序设计。测试表明，该方案能够满足 ACR 路由器对 10Gb 接口 IPTV 性能的设计要求。

关键词：接入汇聚路由器；IPTV 组播；10Gb 接口板；FPGA

Implementation of IPTV Multicast Based on Access Convergence Router

ZHANG Jiajie, HU Hongchao, LIU Qiang, WANG Binqiang
(Center of National Digital Switching System Engineering & Technological R&D, Zhengzhou 450002)

【Abstract】This article is based on the requirement of access convergence router of national 863 plan. It puts forward an IPTV multicast implementation scheme based on FPGA, and designs the FPGA program of 10Gb card about IPTV multicast. The test results indicate that the scheme can fulfill the requirement of access convergence router scheme of 10Gb card IPTV performance.

【Key words】Access convergence router(ACR); IPTV multicast; 10Gb card; FPGA

1 概述

在互联网日趋成熟的今天，人们已不再满足于网速的提高，提出了网络电视(IPTV)、VPN、视频会议等多样化的需求。这些业务的共同特点就是交互性和实时性。目前采用的“LAN/DSLAM＋3 层交换机＋BAS”的接入架构，难以从根本上解决交互性、实时性的需求与网络传输延时与抖动之间的矛盾。

大规模接入汇聚路由器是国家数字交换工程技术研究中心根据国家“863”计划中“大规模接入汇聚路由器总体规范”设计的具有大规模用户接入汇聚功能的高性能双栈核心路由器。大规模接入汇聚路由器由 ACR 交换主机(ACR-S)、ACR 设备管理服务器(ACR-DMS)、ACR-Portal-server(ACR-Portal)、ACR DHCP-server(ACR-DHCP)和远端模块(RM)组成，其中远端模块包括扩展用分复用单元(EMD)和远端接口单元(RIU)。ACR 路由器的接入汇聚功能主要是由 RM 完成的，ACR 总体结构如图 1 所示。

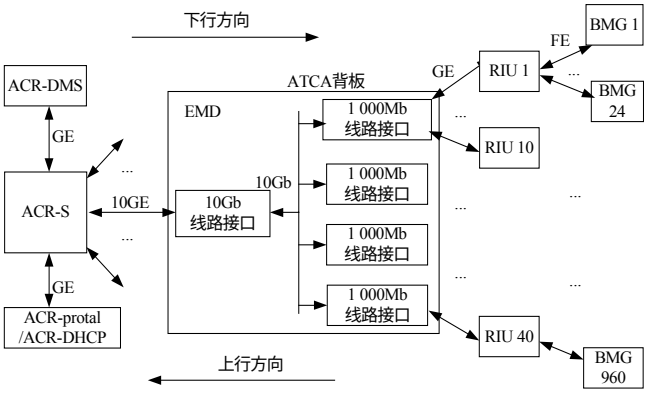

图 1 ACR总体结构

EMD 由一块 10Gb 接口板和 4 块 1 000Mb 接口板组成，每块 1 000Mb 接口下接 10 个 RIU，每个 RIU 下接 24 个 BMG，这样每个 EMD 设备最多允许 960 个用户接入，并在确保用户服务质量的条件下，每个用户可观看 1 路 HDTV 和 1 路 SDTV，还可以拥有 9Mb 的独享上网带宽。EMD 的功能包括：(1)把 ACR-S 送入的下行报文根据查表结果从 EMD 相应的 GE 接口发送到 RIU；(2)将 RIU 送入的上行报文完成相应的认证等处理后送到 ACR-S 或者软件；(3)接入认证；(4)流量统计与计费；(5)组播复制和单、组播查表；(6)IPTV 组播列表管理；(7)IGMP SNOOPING 等。本文主要探讨功能(5)、功能(6)的解决方案和实现。

2 表项设计

IPTV 的实现主要包括 2 个部分：IPTV 权限管理和 IPTV 组播复制。为了满足 IPTV 对交互性和实时性的要求，这 2 部分都被设计放在 EMD 及下游设备中实现。EMD 10Gb 接口在接入汇聚设备 RM 中的顶层地位，IPTV 权限管理放在 EMD 10Gb 中便于管理和维护。对于 IPTV 组播复制，极限情况下 1：960 的复制在现有硬件条件下不可能一次完成，本文引入了 EMD 10Gb、EMD 1 000Mb、RIU 3 级复制的概念。EMD 10Gb 负责组播查表及组播的第 1 次复制，并将其后两级复制的出接口列表封入组播包包头，供后两级组播复制、格式封装时使用。

本文设计了如图 2 所示的 IPTV 组播表和 IPTV 权限表两个表，组播表是 IPTV 组播组地址到 960 个用户出接口的映射，权限表是 960 个用户到 IPTV 组播组的映射。EMD 10Gb 接口在下行方向上查组播表，将查表结果封装在组播数据包

基金项目：国家“863”计划基金资助项目(2004AA103130)
作者简介：张佳杰(1982－)，男，硕士，主研方向：高性能路由器；扈红超、刘　强，硕士；汪斌强，教授、博导
收稿日期：2007-01-17　**E-mail：**zjj@mail.ndsc.com.cn

中，然后进行第一次组播复制；在上行方向上，10Gb 接口响应用户换台请求，进行用户的权限审核，如审核通过，修改 IPTV 组播表。根据 ACR 设计目标，最多支持 100 个 IPTV 组播台，顺序编号为 0~99；最多支持 960 个用户，顺序号为 0~959。

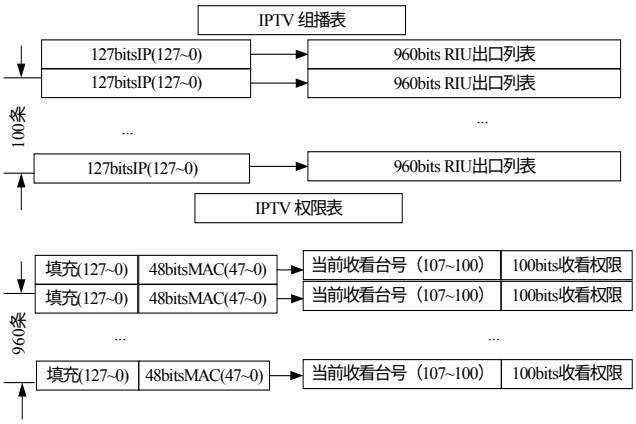

图 2 IPTV 组播表和 IPTV 权限表

3 硬件实现

要实现 IPTV 组播，首先要解决如下 2 个问题：

(1)查表效率问题。组播表表项过宽，达到了 960 位，使用传统的表存取方式效率低，周期浪费严重。

(2)组播表访问冲突的问题。下行处理流程需要查组播表获取组播复制标志；上行处理流程需要查组播表分类上行数据报文，获取 IPTV 台号，根据用户换台请求更改组播复制标志。如果上下行都访问组播表，在数据流中就存在冲突。

为了解决查表效率低下问题，本文设计了 FPGA->TCAM->FPGA->SRAM->FPGA 的查表流水线。传统的 FPGA->TCAM-> SRAM->FPGA 查表流水线，表项宽度较窄的时候效率较高，但是当表项宽度较宽时，效率明显较低。原因是 TCAM 送出的控制 SRAM 连续读写的 burst 信号仅能维持 5 个周期，而对于数据位宽为 32 位的 SRAM，需要最少 30 个周期才能完成 IPTV 组播表的读写操作，也就是最少进行 6 次 TCAM 查表；TCAM 的查表操作从数据送入到查询结果送出需要约 10 个时钟周期，去掉读 SRAM 的 5 个周期，多一次查表就浪费 5 个时钟周期，也就说一次组播表项读写总共需要 60 个周期，其中有 25 个周期被浪费。而如果用 FPGA->CAM->FPGA->SRAM->FPGA 这种查表方式实现该表项，虽然流水线长度较长，但是一次表项的读写仅需要 10 个 TCAM 周期、30 个 SRAM 周期，40 个周期就完成了表操作。目前的 TCAM，SRAM 的时钟频率是 200MHz，也就说一个周期是 5ns。传统流水线表项条目的一次读写需要 10*5*6=300ns，浪费 125ns。而本文设计的查表流水线表项一次读写仅需要 10*5+30*5=200ns，比传统的方式效率提高 50%。

为了解决表访问冲突的问题，在上行方向设计了一个 IPTV 组播表的副表，此表只存储 IPTV 组播组地址，随 IPTV 组播表同步更新；该表仅供上行处理使用，从而减少了冲突的发生。上下行共用一个 IPTV 组播表的设计方式，虽然可以用，但是由于上下行访问频繁，因此效率很低。在上行方向另外建立一个 IPTV 组播表的副表，只存储 IPTV 组播组地址的表项，称之为组播地址表。该表与 IPTV 组播表都接收单板处理机板发送的维护指令，同步更新。上行处理流程中需要查组播表时就查该副表，需要更新组播表也先查组播地址表，获取更新位置、产生 SRAM 更新指令，不需要查 IPTV 组播表索引项就能直接更新组播表的内容。可以看出，建立组播地址表，可以分离上下行查表操作，减少冲突的发生。

基于以上的讨论，采用 FPGA+ TCAM+SRAM 的方式实现该接口。相对于 EMD 来说，下行方向相当于接口的输入，上行方向相当于接口的输出，称下行查表为输入查表模块，上行查表为输出查表模块。IPTV 组播表放在输入查表模块中。IPTV 权限表，组播地指表放在输出查表模块中。图 3 为实现方案示意图。其中，I 表输入，O 表输出，D 表数据流，M 表管理流。

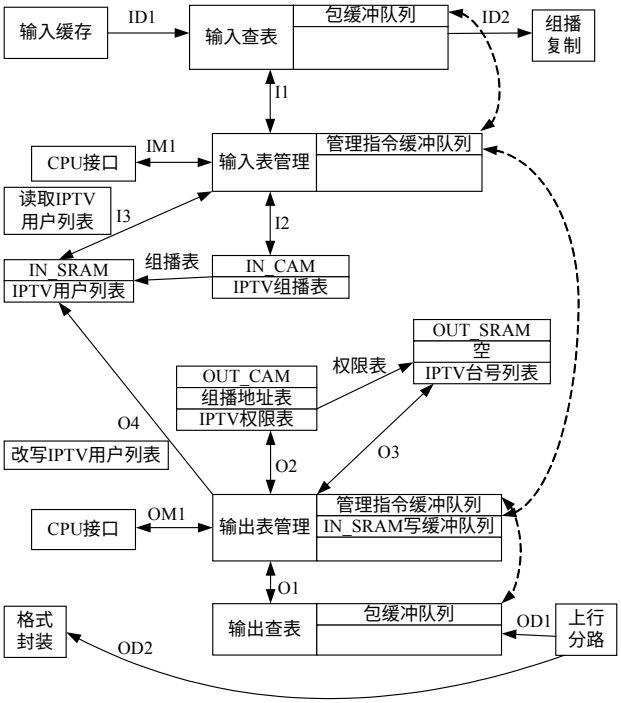

图 3 实现方案

EMD 10Gb 收到用户上行数据包后，首先根据组播地址表分路(OD1)，查表命中的为组播换台请求(O1,O2)，返回欲换台台号，用返回结果查 IPTV 权限表(O2)，将其视为“命中”，修改用户当前收看台号(O3)，将用户原来的收看台号、欲收看台号、用户编号送到输入查表模块管理缓冲队列，输入查表模块根据缓冲队列维护更新表项(O4)，当 IPTV 组播数据包到来时(ID1)，查 IPTV 组播表(I1,I2)，返回更新后的复制标志(I3)，进行第 1 次组播复制(ID2)，则将其送出。

4 FPGA 实现及仿真

在 FPGA 实现及仿真中，本文选用的是 Xilinx 公司的 Virtex-4 系列 XC4VLX160，作为 Xilinx 公司的最新产品的高端型号，该芯片的性能完全满足需求。FPGA 实现如图 4 所示，FPGA 实现难点是：输出表管理模块发送的 IPTV 表项更新指令与输入表管理模块的正常 IPTV 查表操作的冲突。传统的表项实现一般仅有单板处理机进行维护操作，当有维护操作时相当于中断发生，数据方向在当前查表完成后查表停止，进行维护，维护后再开始正常的查表。而本文实现方式有两路维护需求，如果分别响应，逻辑则较复杂、效率较低。本文将更新指令与处理机维护指令合路到一个指令队列，简化了查表逻辑、提高了效率。

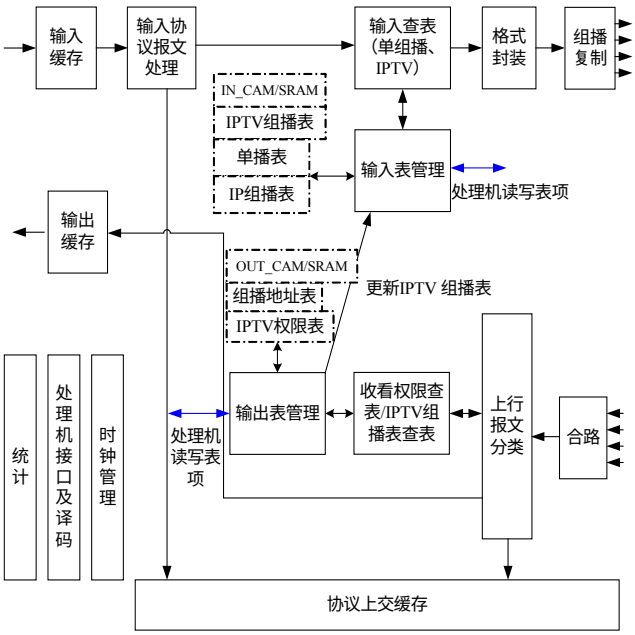

图 4 FPGA 实现

使用 Xilinx 开发工具 ISE8.1 仿真 IPTV 组播部分 FPGA 程序的波形如图 5 所示。

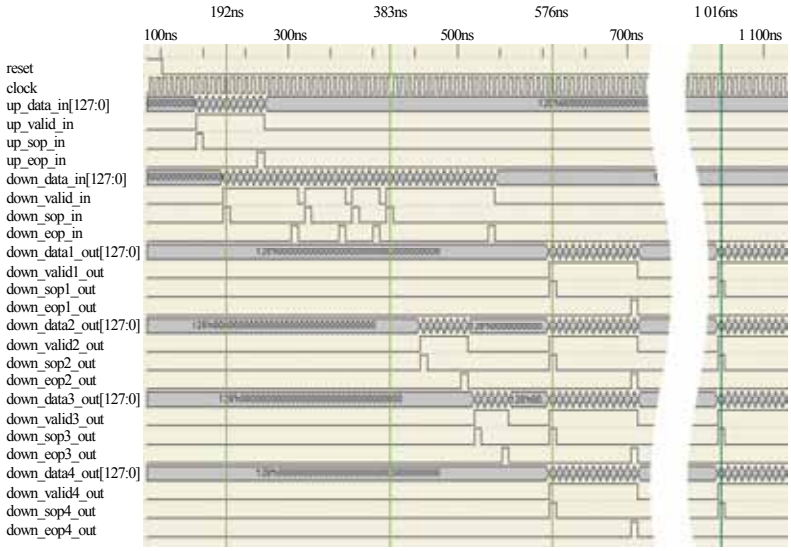

图 5 IPTV 模块仿真波形

图 5 所示信号命名规则为 up/down_XXn_in/out。其中，up 表示上行数据流；down 表示下行数据流；XX 表示该信号功能；*n* 表示第 *n* 路，in 表示模块输入；out 表示模块输出。

rval、sop、eop 分别为包有效信号、包头指示信号、包尾指示信号。reset、clk 为复位信号、时钟信号，这些信号都是高有效。data 为 128 位宽的内部数据总线。为了测试模块功能，设计了特殊的激励信号，上行数据流输入的为用户 IPTV 换台请求(Req)，假设所换台号为 X，下行数据流中输入的依次为 IPTV 组播包(MultiA)、单播包(SingleA)、单播包(SingleB)、IPTV 组播包(MultiB)，两个组播包都是 IPTV 组播台 X 的包。从模块仿真后产生的输出可以看到，由于 MultiA 等待查表结果，因此，SingleA、SingleB 先于 MultiA 输出；MultiA 在经过约 300ns 后组播复制输出；而由于 Req 产生的 IPTV 组播表修改，使得 MultiB 从输入到输出约 330ns 的时间，完成 1 次由于用户换台引起的表项修改只需要不到 30ns，即使 960 个用户在 1s 内同时换台，由于换台引起的时延也只有 30*960≈30ms，因此，用户完全可以接受的。在图 5 中，单播包多了 1 个周期的数据，组播包多了 2 个周期的数据，这是为了方便后续设备处理封装的内部包头。

5 结束语

本文根据 ACR 路由器对 EMD IPTV 组播的要求，设计了 EMD 10Gb 的查表结构、查表电路，编写了 IPTV 组播管理的 FPGA 程序并进行了仿真。仿真结果显示该实现能够满足人们对 IPTV 组播的性能要求。

参考文献

1 张兴明. 大规模接入汇聚路由器总体技术规范[Z]. 郑州：国家数字交换系统工程技术研究中心, 2005-02.

2 赵永忠，兰巨龙，刘勤让. 基于 FPGA 实现的 10Gbbps 线路接口设计分析与实现[J]. 微电子学与计算机, 2005, 22(2).

3 李　锋. 一种 IPTV 承载网的解决方案——大规模接入汇聚路由器[J]. 计算机工程与应用, 2006, 42(1).

4 Montrose M I. 电磁兼容和印刷电路板：理论、设计和布线[M]. 北京：人民邮电出版社, 2002-12.

（上接第 106 页）

3 结论

本文针对无线分组网络数据传输及具体应用的特点，设计实现了基于 TDMA 的无线网络应用系统。采用时分多址接入机制将无线信道资源分配给网络中各节点，避免信道竞争，保证各节点的数据信息在其专有时隙及时、可靠发送。并可通过预留时隙适应节点的移动切换、突发通信以及复杂的通信流程变化需求，同时避免了由于传输延迟导致的传输率下降、传输层通信中断以及慢恢复等问题，更适合无线分组网络战术信息传输的实时可靠性要求。

参考文献

1 Jackson W K W, Victor C M L. Improving End-to-End Performance of TCP Using Link-layer Retransmissions over Mobile Internet-works[C]//Proceedings of IEEE ICC'99, NY, USA. 1999-06: 324.

2 潘　路，赵文娟，欧阳宁. 分组无线网在侦察系统中的应用[J]. 桂林电子工业学院学报, 2003, 23(2): 43-46.

3 程伟明. 无线移动自组网及其关键技术[J]. 数据通信, 2002, 23(3): 56-58.

4 魏云升，郭　治，王校会. 火力与指挥控制[M]. 北京：北京理工大学出版社, 2003.