

# 一种综合型宽带接入系统的研究与实现<sup>\*</sup>

Some Research for a Synthesized Broad Accessing System

刘琼 江浩 吴志美

(中国科学院软件研究所 北京100080)

**Abstract** The virtues and defects for Crossbar and Sharing Memory technologies are analyzed firstly in this paper. Then the improvements that Crossbar and Sharing Memory fabric are integrated in a broad accessing system are indicated. Finally the synthesized specialties for some application are discussed.

**Keywords** Crossbar, Sharing memory, Ethernet switch

## 一、引言

随着多媒体信息的数字化表示与处理,电信网、电话网和数据网的融合已成为通信网的发展趋势,骨干网正逐步统一采用IP网络,不久,将实现光纤铺设到大楼或住宅小区,宽带接入已成为当前研究与发展的热点。宽带接入的主要技术包括xDSL、Cable Modem(HFC)和局域网。住宅小区建设宽带局域网,意味着采用交换技术,把网络连接到千家万户,使每个用户独享10M/100Mbps带宽,是一种经济实用、符合中国国情的理想方案。

本文以局域网宽带接入设备的研究开发为背景,从应用角度分析了交换矩阵和共享存储器技术的优缺点,研究了综合这两种技术实现一种宽带接入系统时做出的改进,并讨论了它们在实际应用中的特点。

## 二、交换矩阵与共享存储器技术

### 2.1 交换矩阵

交换矩阵(如图1)和共享存储器是交换机常用的两种实现技术,交换矩阵以点对点方式连接,同时提供多条总线交换,大大提高了系统交换容量,传统的电信设备(如程控交换机等)常采用这类结构,但是交换矩阵结构存在以下主要缺点:

·中心调度模块控制交叉接点状态,调度速度可能形成交换速度的瓶颈。

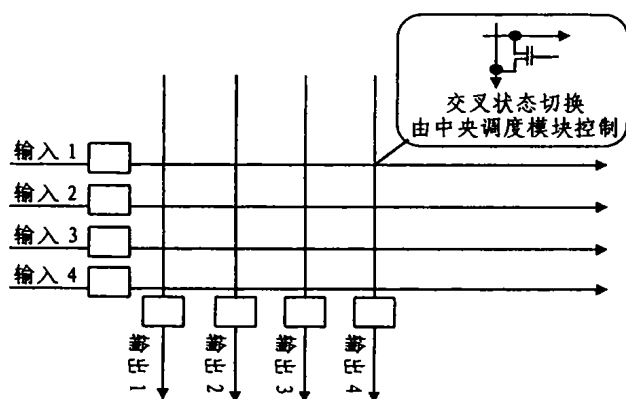


图1 交换矩阵原理图

·交换矩阵的输入/输出端口只采用简单的FIFO队列机制,某输入缓冲区队首信元因目标端口繁忙一时未发出时,阻止着后续其它目的端口信元的及时发送,形成“线头(Head-of-Line)阻塞”。

·输入阻塞:采用VOQ(Virtual Output Queue)技术的交换矩阵,同一输入端口不同VOQ中的信元竞争出线而产生的阻塞。

·输出阻塞:不同输入端首信元指向同一目标输出端口。

### 2.2 共享存储器

共享存储器(sharing memory)结构与交换矩阵相比可算是后起之秀,原理上可以克服交换矩阵的不足。所有的输入数据存入同一存储器,动态分配与端口关联的存储器空间,数据报文交换,只需移动报文的地址指针,交换简单灵活,转发效率高,存储空间利用充分。存储器价格的不断降低推动了共享存储器技术的应用。但是在共享存储器结构中,每个接收包都要通过中央交换引擎检测,扩展端口比较困难,访问速度和总线带宽制约着交换速度的进一步提高,当存储容量要求过大时,共享存储器交换结构技术的实现十分复杂,限制了它的应用范围。

## 三、综合型宽带接入系统的交换结构

### 3.1 系统结构

本文研究的综合型宽带接入交换机采用分布式系统结构,将交换矩阵和共享存储器技术合为一体,如图2。所示系统分两级,核心级交换背板采用交换矩阵,吞吐量可达15Gbps以上,每个交换模块总线端口全双工交换带宽达2.88Gbps。交换矩阵以直通方式中继交换模块间的交互信息,交换模块之间或交换模块与管理模块之间的互访利用交换矩阵提供中继桥路,建立直接存储器访问(DMA)通道,利用共享存储器技

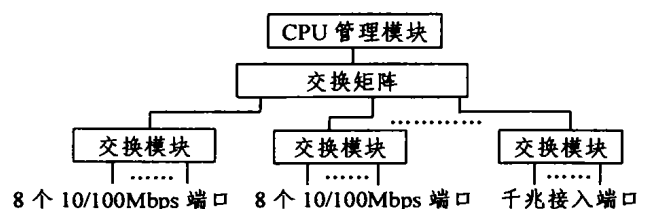


图2 综合型宽带接入系统结构框图

<sup>\*</sup>本文得到国家重点基础研究发展规划973项目的支持,刘琼 博士后,副教授。

术和存储-转发方式(store-and-forward)实现。交换模块以智能型交换处理芯片,向上提供与交换矩阵连接的总线接口,向下提供8个10/100Mbps网络端口,内部除了实现IEEE802.3协议外,可以实现地址学习、组播包提取、VLAN协议头添加,以及数据流分类等功能,数据收发采用共享存储器技术和存储-转发方式。

### 3.2 交换背板的实现

交换背板用 Crossbar 技术实现,因为:(1)交换模块与核心交换矩阵间采用了简单的点到点链接,有利提高线路速率;(2)收发通道分开,可以有效控制信号反射并使信号在短时间内得到稳定;(3)多路总线传输,便于扩展系统吞吐量。通常认为 Crossbar 内部无阻塞,主要指它允许所有无冲突配对的输入和输出同时传送数据包。但是交换矩阵操作中经常出现三种阻塞:线头(HOL,Head of Line)阻塞;输入阻塞和输出阻塞。线头阻塞严重影响系统性能,本系统已采取有效手段消除这种阻塞。输入、输出阻塞也采取了相应的抑制措施。

在简单交换矩阵中,接收端口的所有数据包都存入单一的输入 FIFO(先进先出)队列,到达 FIFO 头部的数据包才可能得到中央调度单元的处理。如果多个端口的队首数据包同时竞争同一目标端口,中央调度单元一次只能允许一个数据包通过 Crossbar,其余队列头部的数据包只能按预定调度规则等候(如轮循、优先级等)处理。这对于同一端口队列而言,可能因为头部的数据包得不到即时处理,而延时后续目标端口空闲数据包的发送,形成所谓“线头(HOL)阻塞”问题。在通信量较大系统中,好不容易到达队列之首的数据包,又可能因为目标端口竞争,再次形成线头阻塞,导致输入缓冲队列很快灌满或包丢失。

本系统核心级交换矩阵采用虚拟输出队列解决线头阻塞问题。图3所示的 Crossbar 结构,交换矩阵的4个 GLINK 端口和一个 PCI 接口分别挂4个子交换模块和一个管理模块。每个端口都为其余端口维护一个“输出”队列 FIFO,这样不同源端口数据包要同时发往同一目标端口时,先存入虚拟目标端口 FIFO 中,再经过调度发往物理目标端口,从而解决了线头阻塞问题,代价是每个端口添加了多个 FIFO 和端口调度单元。当发生多个端口向一个目标端口发送数据的极端情况时,如果该目标端口 FIFO 不空,调度算法按规则继续向前轮询,发往其余目标的数据可相对及时到达对端,使系统输入、输出阻塞得到有效改善。

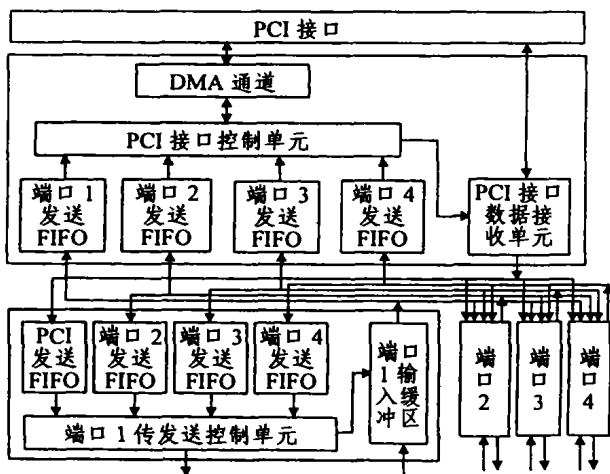


图3 并行“输出”队列 Crossbar 结构示意图

### 3.3 子模块交换机制

以8端口10/100Mbps 交换模块为例,在 DRAM 中开辟一个共享缓冲区,将其划分为390或1024个分区(取决于 DRAM 容量),每分区容量占1.5kB,缓冲分区“空”否,由专用状态标志表(390或1024位构成)标识。每端口占用分区数按端口当前接收数据量状况动态分配,但是每端口至多只能占用512个缓冲分区。

进入缓冲分区的数据帧由9个发送队列(对应模块内部的8个下行交互端口和1个上行交互口或网络管理口)输出,每个队列拥有一个描述符环,每个描述符环含512个描述符,每个描述符占32位(存放分区地址、帧字节数、帧类型:M/U),描述符环也存放 DRAM。图4表示描述符环指向缓冲分区,形成发送队列的情况。

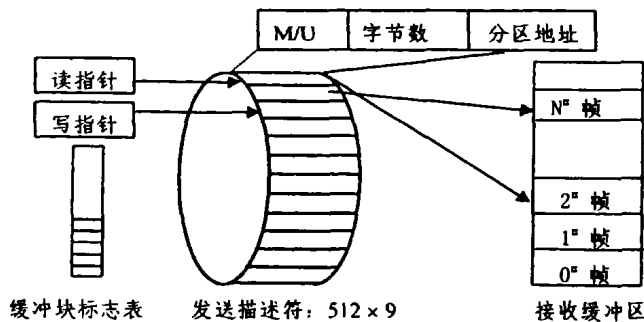


图4 子模块的交换机制示意图

子交换模块内部帧的“存储转发”操作步骤如下:

1. 接收帧进入事先分配缓冲分区的同时,完成帧地址(SA 和 DA)识别和帧校验。
2. 如果接收帧完好,将帧类型 M/U(多播/单播),字节数,分区地址分别写入“写指针”指向的发送描述符(图4),至此该分区列入发送队列。
3. 写指针加一,继续指向“待写”描述符。只要读/写指针所指向的描述符编号不等,队列中便存有待发数据帧。
4. 每结束一次帧发送,清除相应状态标志位,释放该缓冲分区,读指针加一,指向下一待读帧描述符。

子交换模块之间或子交换模块与管理模块的交互,需要建立 DMA 通道,并配合相应的操作规则。

## 四、系统的实现与讨论

交换矩阵和共享存储器技术共存同一系统有利于扬长避短,充分发挥交换矩阵连接简单,技术实现可靠,借总线结构吞吐量大的优势而共享存储器技术存储资源利用率高、实现简单灵活的特点也得到发扬,提高了交换机整体综合性能。交换矩阵的线头阻塞问题在本系统中已基本得到解决,唯一可能出现的短暂阻塞仅发生在多对一发送的极端情况。综观本系统实现,还可以看到以交换矩阵为核心的分布式结构,形成弹性吞吐量,系统扩展方便。以交换矩阵为核心的分布系统可以灵活集中多个交换模块,并可随意选配10/100Mbps 或1000Mbps 交换模块。交换矩阵带 PCI 接口,便于系统实现网络管理或添加高层应用。

从硬件实现来看,目前可以做到将介质访问控制内核、足够容量的存储器、以及相关的物理层收发器统一做在较小体积的单一芯片上,大大简化共享存储器交换模块的设计实现。

(下转第61页)

查询的完整性和准确性以及时效性。一个理想的移动环境下的查询策略应该保证符合要求的查询的完整性、准确性和时效性,并且要求消耗较小的资源同时缩短查询的响应时间。这几个方面的目标有些是相互矛盾的,在某些极端情况下(移动设备长时间处于盲区)需要制定特定的查询策略来满足实际的需求。

**结论** 移动环境下的查询策略主要是受到了移动计算环境下的移动设备资源相对比较紧张的限制。在保证查询正常完成的情况下,减少对瓶颈资源使用和保证查询事务正确性是一个值得进一步研究的课题。

### 参考文献

1 Bowman I T. Hybrid Shipping Architectures: A Survey. ACM SIGMOD Conf. Montreal, Quebec, Canada, ACM Press, 2001

- 2 Franklin M J, et al. Performance tradeoffs for client-server query processing. In: H. V. Jagadish and Inderpal Singh Mumick, eds, Proc. ACM SIGMOD Conference, Montreal, Quebec, Canada, ACM Press, 1996
- 3 林怀忠, 陈纯, 吴震华. 数据复制与一致性. 计算机工程与应用, 2001, 27
- 4 Franklin M J, Carey M J. Client-server caching revisited. In: M. Tamer Ozsu, U. Dayal, and Patrick Valduriez, eds, Distributed Object Management. Morgan-Kaufmann Publ. Co., San Mateo, CA, USA, May 1994. International Workshop on Distributed Object Management
- 5 Kossmann D, Franklin M J. A study of query execution strategies for client-server database systems. In: Proc. ACM SIGMOD Conf. 1995. Also available at: University of Maryland Technical: [Report CS-TR-3512]

(上接第42页)

本文已在技术研究的基础上,研制出一种综合型宽带接入交换机,初步测试表明综合交换矩阵与共享存储器的体系结构能更好地满足宽带接入系统的要求,便于实现低、中、高端网络交换设备。

### 参考文献

1 Galilio Inc. GT-48303A, 6-Port Crossbar switch datasheet, Pre-

liminary Revision 1.0, April 17, 2000

- 2 GT-48310A, GT-48320A Switched Ethernet Controller datasheet, Preliminary Revision 1.2, April 03, 2000.
- 3 刘琼. 高速以太网交换技术的研究及千兆位交换机的研制. [博士后出站报告]. 北京邮电大学信号与信息处理专业, 2000
- 4 McKeown N. Fast Switched Backplane for a Gigabit Switched Router, Department of Electrical Engineering Stanford University, CA 94305-9030

## 中国计算机学会全国第四次程序设计语言发展与教学学术会议

### 征文通知

全国第四次程序设计语言发展与教学学术会议定于2003年春季在江苏扬州召开。本次会议由东南大学承办,扬州大学、南京大学、武汉大学等院校协办,现将有关事项通知如下:

#### 一、征文范围

A) 程序设计语言历史、现状与发展; B) 面向对象语言及相关技术; C) 各类建模语言及其设计、实现与应用; D) 面向网络应用的程序设计语言(XML、HTML、PERL等); E) 其它各种新型程序设计语言(包括逻辑型语言、函数型语言等); F) 程序设计语言分析、评价与比较; G) 程序设计语言语法、语义与语用以及形式化描述技术与方法; H) 并发、并行与实时程序设计语言; I) 软件开发过程中各类描述语言(包括软件体系结构描述语言等); J) 第四代语言与数据库语言; K) 程序设计语言教学、教材与课件; L) 各类写作语言与工具; M) 其它。

#### 二、征文要求

来稿一般不得超过6000字,并且未被其它会议、期刊录用或发表。为了便于正式出版论文集,来稿必须附中英文摘要、关键词与主要参考文献,注明作者姓名、工作单位、详细通讯地址(包括电子邮件地址与电话)与作者简介。

欢迎电子投稿,来稿不退,请自留底稿。

**三、来稿地址** 南京 东南大学 计算机科学与工程系 徐宝文; 邮编: 210096;

电话: (025) 3793977; Email: bwxu@seu.edu.cn

**四、重要日期** 征文截止日期: 2002年10月15日 录用通知发出日期: 2002年11月15日

修改稿截止日期: 2002年12月15日