

# 综合业务网理论与技术 论文选编

(1996)

综合业务网理论及关键技术国家重点实验室



西安电子科技大学

# 目 录

## 第一部分 网络与交换

1. 共享存储器ATM交换机性能分析 ..... 赵群标 刘增基 (3)
2. On Some Key Technologies of ATM Switching ..... Liu Zengji YangXinhui等(7)
3. 基于Banyan网的快速分组交换网络的硬件复杂度分析 ..... 贺飞云 陶锡生 蔡元龙 (12)
4. 采用ATM传输方式的卫星综合宽带通信网 ..... 段晓明 鞠德斌 (19)
5. 一种具有优先级队列的ATM交换在多种类型业务下的性能分析 ..... 贺飞云 陶锡生 蔡元龙 (23)
6. B-ISDN与现有网络互连的研究 ..... 彭雄飞 刘增基 胡 征 (30)
7. Novell网通过ATM微波网的互连及互通 ..... 郑智亮 刘增基等 (34)
8. Comparison of two recycle dilated banyan network with output buffers .....  
..... Ning Zhou and Zeng Ji Liu (41)
9. 分组预约多址 (PRMA) 的性能仿真 ..... 李建东 李明远 李维英 (43)
10. 短波自组网网络技术及其实验系统 ..... 李建东 李维英 陈家祺等(48)

## 第二部分 信源编码及语音图象处理

1. 一种基于双正交小波变换和格型矢量量化的视频编码算法 ..... 薛向阳 樊国信 陈学青 (55)
2. 变速率视频编码的统计特性及其模型研究 ..... 常义林 胡 征 (62)
3. 一种用于序列图象位移估值的匹配函数 — 比特位相关匹配函数 .. 常义林 李兵兵 李飞鹏 (67)
4. 共面五条直线的联合射影不变量 ..... 徐正伟 吴成柯 (72)
5. 空间平面曲线的透视不变性分段 ..... 徐正伟 吴成柯 (76)
6. 由文本至口形的媒体变换技术的研究 ..... 潘丹宁 郭 峰 文成义 (81)
7. Effective Simple Y-C Separator for PAL and NTSC TV Signals .....  
..... Y. Ding , X.Gao and K.Yi (85)
8. 语音压缩编码中的两种基音预测器 ..... 鲍长春 戴逸松 樊国信 (87)
9. 基于有限状态径向基函数网络的汉语语音识别研究 ..... 李革营 易克初等 (95)
10. 伪彩环境下逼真显示彩色图像方法研究 ..... 孙献璞 常义林 (99)
11. 图像分割的遗传算法方法 ..... 吴成柯 刘 靖 徐正伟 周凌云 (103)
12. 视频变速率编码的统计模型研究 ..... 常义林 孙献璞 吴 剑 (110)
13. 16kbps低延迟码激励线性预测语音编码的研究 ..... 王 卿 生 樊国信 (114)

## 第三部分 信道编码与密码

1. 两类新的线性分组码的译码 ..... 马建峰 王育民 (121)
2. Golay 码的快速译码 ..... 马建峰 王育民 (127)
3. 一类三元线性分组码的译码 ..... 马建峰 王育民 (133)
4. 分组码的软判决伪序列译码 ..... 喻建平 王新梅 张艳芝 (138)

5. 一类循环码的神经网络软判决译码算法 .....	曾建平	马建峰	王新梅	(142)
6. 基于矩阵分解的代数几何码的译码 .....	任剑	王新梅	肖国镇	(146)
7. 移动通信网中的认证与密钥分配 .....	徐胜波	武传坤	王新梅	(152)
8. 几类格的快速译码 .....	马建峰	王育民		(156)
9. 基于算法存储中的编码方法 .....	马建峰	赵晓星	王新梅	(160)
10. 二次剩余码的有效译码 .....	马建峰	王育民		(166)
11. 关于二元线性分组码的快速译码 .....	马建峰	王育民		(173)
12. 流密码中非线性组合函数的分析与设计 .....	陈木都	肖国镇		(179)
13. Bent 函数与其变元的相关特性 .....	冯登国	肖国镇		(184)
14. 加强广义 ElGamal 型签名方案的安全性 .....	祁明	肖国镇		(187)
15. 布尔函数的对偶性和线性点 .....	冯登国	肖国镇		(192)
16. 基于 Lam 签名方案的远距离通行字认证方案 .....	祁明	肖国镇		(197)
17. 一类相关免疫函数的非线性性和扩散特性 .....	冯登国	肖国镇		(203)
18. 布尔函数非线性度的谱分析 .....	武传坤			(208)
19. 有限域的伪对偶基 .....	周坤	肖国镇		(217)
20. 满足 $k$ 次扩散准则的布尔函数的谱特征 .....	冯登国	肖国镇		(222)

#### 第四部分 其它

1. 用数字信号处理器实现的短波自适应选频控制器 .....	杜桂义	朱晓明		(231)
2. 频偏对自适应选频系统的影响及估值方法 .....	袁拥军	金力军		(235)
3. 短波自适应选频控制器的 DSP 实现 .....	杜桂义	朱晓明	金力军	(243)
4. 高速数据传输中的自动增益控制 .....	刘乃安	曾兴宽	陈健等	(246)
5. 直接数字式频率合成器的谱展研究 .....	王育红	张殿盛		(252)
6. 相位累加器溢出脉冲的频谱分析 .....	王育红	张殿盛	樊晋信	(260)
7. 高阶滤波器的最佳设计与实现 .....	袁昌幸	刘乃安	刘彦明	(267)
8. 图像空间基于深度元素模型的数字加工几何仿真研究 .....	汤毓宁	蔡长荣		(272)
9. 带限条件下频率闪烁噪声 $N$ 样本抽样方差的精确解 .....		宁德成		(279)

# 第一部分

## 网络与交换



# 共享存储器 ATM 交换机性能分析\*

## Performance Analysis of Shared Buffer ATM Switching

赵豫彪 刘增基

(西安电子科技大学综合业务网国家重点实验室, 西安 710071)\*\*

**【摘要】** 本文提出了一种分析共享缓冲器 ATM 交换机性能的新方法, 该方法为复杂的排队问题的求解提供了新的途径。该方法是采用 Markov 链一步转移概率矩阵求得信元队长稳态概率分布的精确算法。计算机模拟结果验证了该方法的有效性和正确性。

**关键词:** 性能分析, ATM 交换机, 排队算法

**Abstract:** A novel approach is presented for performance analysis of shared buffer ATM switching. It provides a new means to solve problems in more complex queueing system. It is an accurate algorithm by employing a one-step transition probability matrix of the Markov chain. Exactitude and effectiveness are verified by computer simulation.

**Key words:** Performance analysis, ATM switching, Queueing algorithm

### 一、引言

自从国际标准化组织把异步转移模式(ATM)作为宽带综合业务数字网(BISDN)的统一基础模式以来,在世界范围 ATM 已成为愈来愈引起人们关注的课题<sup>[1]</sup>。其中正在研究的 ATM 交换系统的一个主要方向就是交换结构。ATM 交换机现有多种结构,就 ATM 交换机的主要性能指标交换延迟、信元丢失率和吞吐率来说,共享存储器 ATM 交换结构是实现 ATM 交换机的最佳选择<sup>[2,3]</sup>。

从理论上讲,共享存储器交换方式和输出缓存方式的交换延迟和吞吐率性能是相同的,故本文仅对信元丢失率进行研究。对共享存储器 ATM 交换结构的这一主要性能分析,以往的方法不太精确。例如:

①在分析交换机的某一个输出端口的情况下,给出该端口的信元排队队长分布函数,写出其相应的概率母函数,利用求大量独立随机变量之和的尾

部分布界的 Chernoff 界,给出所有的输出端口信元共享一个缓冲器时低于一定的信元丢失率所需要的存储器容量<sup>[4]</sup>。

②同样是分析一个输出端口的分布情况,认为各输出端口信元分布不相关,然后采用各输出端口分布函数相卷积的方法,或者采用概率母函数相乘,再用其反变换方法求出信元在共享存储器中的排队分布情况,从而确定达到一定的信元丢失率所需共享存储器的容量等指标<sup>[5]</sup>。

③从输出端口信元分布是两两相关分析出发,求出两个队长的均值和协方差,从而导出共享时所有端口总的信元队长的均值和协方差,并假设其总的信元分布服从 Gamma 分布,进而求出相应的性能指标<sup>[6]</sup>。

方法①、②由于假设各输出端口之间信元排队队长分布是相互独立的,在相同情况下达到一定的信元丢失率指标,计算出的存储空间比实际所需的存储空间偏大。方法③在文献[6]中首次指出各输出

端口之间信元排队队长不是不相关而是负相关的。但是该方法一方面从各输出端口信元队长分布是两两相关这一近似出发,另一方面未经推导直接认为总的信元队长分布服从 Gamma 分布。这一方法较以上两种方法更接近实际情况,但是文中采用的这两方面假设使之求出的结果仍不够严谨,计算出的存储空间比实际所需的存储空间仍偏大。

本文充分考虑这样一个事实:在各时隙(slot)到达输入端口的信元数转移到各输出端口并非不相关,而是负相关的,并基于这一事实,首次提出了求共享存储器 ATM 交换机信元排队分布的精确解的方法,并给出在一定的交换端口数和一定的业务量下,交换机的存储容量与信元丢失率指标的准确关系。最后,把该方法与计算机模拟结果作了对比,证明了该方法的正确性。

## 二、问题的提出

对于有  $N$  个输入端口和  $N$  个输出端口的共享存储器 ATM 交换机,在每个时隙同时到达交换机的总信元数不会大于  $N$ 。在某一时刻到达交换机输入端口的所有信元中有一部分到达某一输出端口,那么这部分信元就决不会到达别的输出端口;极端的情况,到达交换机各输入端口的所有信元都被送到某一输出端口,则别的输出端口就不会有信元到达。所以在各时隙到达输入端口的总信元数被转移到各输出端口并非不相关,而是负相关的。这个负相关特性使得实际系统总的队长小于将各输出端口假定为统计独立时的各端口的队长的总和。这就是以往的分析方法所估计的为达到一定的信元丢失率指标所需的存储器容量偏大的原因。

## 三、分析方法

信元在 ATM 交换系统中都是以时隙为时间单位到达交换机或被传送到输出线的,如果在系统中信元刚被送出交换机的时刻观察交换机中的总信元数,则该总信元数仅与上一时隙系统中总信元数有关,而与上一时隙以前的系统中总信元数无关,故把信元刚被送出交换机时刻交换机中的总信元数作为这个排队系统的状态,这样就形成了一个离散状态 Markov 链。若能求出这个 Markov 链的一步转移概率矩阵,就可以求出系统的稳态概率分布,即交换机

存储器中等待的总信元数的稳态概率分布,从而可以得到为达到一定的信元丢失率指标所需的存储器容量的准确关系。

分析的思路是这样的:①首先考虑信元到达情况,写出仅有信元到达而无信元输出的条件下一步转移概率矩阵  $P_{0,i}$ ;②求出仅有信元服务而无信元到达的条件下一步转移概率矩阵  $P_{i,0}$ ;③求出系统状态的一步转移概率矩阵  $P$ ,进而求出平稳状态的系统队长分布。具体求解过程如下:

①共享存储器 ATM 交换机具有  $N$  个输入端口和  $N$  个输出端口。假定在每个时隙,信元到达每个输入端口服从概率为  $P$  ( $P > 1$ ) 的 Bernoulli 过程,即有信元到达的概率为  $p$ ,没有信元到达的概率为  $1-p$ ,每个随机到达的信元被等可能地送往  $N$  个输出端口之一,即在每个输入端口,相继到达的信元被独立地送到相应的输出端口。显然这里  $p$  就是交换机每个输入端口的业务量。

基于以上的假设,在每个时隙有  $i$  个信元到达交换系统的概率为  $a_i, i=1, 2, \dots, N$ , 即

$$a_i = C_N^i p^i (1-p)^{N-i} \quad (1)$$

由此得

$$P_0 = \begin{bmatrix} a_0, a_1, \dots, a_N \\ a_0, a_1, \dots, a_N \\ \vdots \\ a_0, a_1, \dots, a_N \end{bmatrix} \quad (2)$$

矩阵  $P_0$  的行序号和列序号均从 0 开始,其序号分别为不考虑信元的服务时,系统内当前信元数和下一个时隙系统内的信元数。这就是仅有信元到达而无信元输出的条件下系统内信元总数的一步转移概率矩阵。

②共享存储器 ATM 交换机属于输出排队型的交换结构,只是所有的输出信元队列共享一个存储器,对应每个输出端口的队列都是被相应的存储地址链连接构成的逻辑队列,从而提高了存储器的利用率。所有有信元排队的输出端口在每一个时隙都有一个信元被送往输出线路,即被服务。就每个输出端口来看,是定长服务。就所有输出端口来看,服务情况是所有有信元的输出端口数的函数,即有几个输出端口有信元排队则在一个时隙,这些非空队列的排头信元(HOL: Head Of Line)就被服务。一个时隙内被服务的信元数就是系统中的非空队列数。于是就归结为这样一个问题。

有  $n$  个信元排在  $N$  个输出端口中的  $i$  个端口的概率  $b_{ni}$ , 就是此时一个时隙能够传送到输出线路  $i$  个信元的概率, 其中,  $b_{ni}, i=1, 2, \dots, m, m=\min[n, N], n \neq 0$ , 为在系统内共有  $n$  个信元的条件下, 交换机输出端口中  $i$  个端口有信元的概率。

由于到达交换机的所有信元都相互独立地等可能地到达各个输出端口排队, 于是经推导<sup>[1]</sup>可得

$$b_{ni} = \frac{\sum_{j=0}^{i-1} C_{i-1}^{j-1} (-1)^j (i-j)^{n-1}}{(i-1)!} \cdot \frac{N!}{N^i (N-i)!} \quad (3)$$

由此得

$$P_n = \begin{bmatrix} b_{11} \\ b_{21} & b_{22} \\ \dots \\ b_{N1} & b_{N2} & \dots & b_{NN} \\ & b_{N1} & b_{N2} & \dots & b_{NN} \\ & & \dots & \dots & \dots \end{bmatrix} \quad (4)$$

设  $b_{00}=1$ , 显然  $b_{11}=1$ 。同样, 矩阵  $P_n$  的行序号和列序号均从 0 开始, 其序号分别为不考虑信元的到达时, 系统当前信元数和下一个时隙系统的信元数, 这就是仅有信元服务而无信元到达的条件下系统内信元总数的一步转移概率矩阵。

③ 由于交换系统的信元到达和服务是相互独立的随机过程, 故由以上求出的  $P_n$  和  $P_s$  可以求得交换系统内信元总数的一步转移概率矩阵  $P$  为

$$P = P_n P_s \quad (5)$$

$P$  为宽度为  $2N$  的不对称带状矩阵, 其行序号和列序号均从 0 开始, 其序号分别为当前和下一个时隙系统中所有排队的信元数, 即系统中的状态。

每个信元到达交换机的输出端口是均匀等概的, 当系统信元数  $n$  远远大于交换机的端口数  $N$  时, 所有的信元几乎占据所有的输出端口, 系统每个时隙输出  $N$  个信元的概率接近 1, 即当  $n \rightarrow \infty$  时, 则  $b_{nN} \rightarrow 1$ 。此时系统饱和, 其归一化服务率为 1, 而系统的每个输入端口业务量, 即归一化到达率为  $\rho < 1$ , 这样系统的业务强度小于 1, 故该系统具有稳态分布。系统状态数大于某一值后, 随着系统状态数  $n$  的增加, 相应状态的概率变得极小并趋于 0。从而可以选择足够大的系统状态数, 用有限状态 Markov 链来求解其稳态分布问题, 使得由此而引起的误差可以忽略。

任给该系统一个充分大的状态数  $s < \infty$ , 比如,

使该状态的稳态概率  $p_s < 10^{-20}$ , 从而求解这个一步转移概率矩阵相应的稳态概率分布, 采用一步转移概率矩阵  $k$  次幂并使  $k \rightarrow \infty$ , 或用稳态分布方程求解均可。

求出的稳态分布就是在共享存储器中信元总数的稳态分布, 用这个分布可准确地分析交换机的主要指标, 为达到一定的信元丢失率所需的存储容量。

#### 四、结果分析

基于以上介绍的方法, 我们作了大量的计算得到了很多有意义的结果。以往的理论分析方法很难得到信元总数分布情况, 而用计算机模拟的方法得到精确信元队长分布是很费时的。比如, 为了模拟交换机运行  $10^8$  个时隙时, 交换机中的信元分布情况, 在 Intel 486-33 采用 NDP FORTRAN 约需 320 小时; 而使用本文所介绍的方法得到这个结果还不到 1 分钟。以往的分析方法之所以不严格, 存在一些偏差, 是因为不能找出总的存储信元数分布的精确解而引起的。从以下的对比曲线中可以明显地看到这个结果。

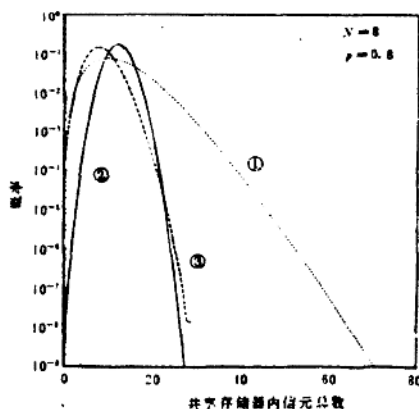


图 1

图 1 中给出了共享存储器 ATM 交换机内信元总数的稳态分布, 这里交换机的输入/输出端口数  $N=8$ , 业务量  $\rho=0.8$ 。图中的三条曲线分别对应于: ① 输出队列不相关的卷积方法, ② 本文提出的解法和 ③ 计算机模拟方法得到的信元分布结果。从曲线中可以明显看到, 计算机模拟结果与本文方法很接



近,而与卷积方法相差很大.这说明用本文的分析方法更接近实际情况,而用输出队列不相关的卷积方法分析引起了很大的偏差,不符合各输出队列不是独立的而是相关的且是负相关的这一实际情况.这一结论与文献[5]中的硬件仿真结果一致.

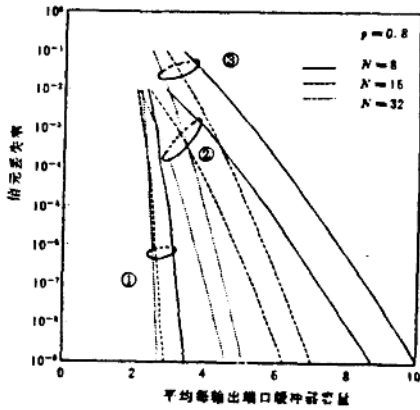


图 2

图 2 给出了共享存储器交换机在一定的业务量和端口数下,平均每线存储信元数与达到相应的信元丢失率之间的关系曲线.这里交换机的  $N=8, 16, 32, \rho=0.8$ . 图中比较了用①本文方法、②卷积方法和③Chernoff 界方法得到的结果.从图中可以明显看出在相同的情况下,达到一定的信元丢失率用卷积方法和用 Chernoff 界方法每线所需存储信元数偏差都很大,后者比前者偏差更大,且都比本文方法每线所需存储信元数偏大.用 Chernoff 界方法比用卷积方法偏差更大,这是因为 Chernoff 界虽然是捕捉小概率事件的一个比较严格的界,但是毕竟不是一个确界,同时 Chernoff 界方法也是基于输出队列独立得到的结果.同文献[6]中的曲线相比,本文方法更接近计算机模拟情况,尤其是在交换机端口数较小时.显然,随着交换机的输入/输出端口数增大,业务量的减小,以及信元丢失率要求的降低,从而引起输出队列的相关性减弱,使得用以往的分析方法引起的偏差降低.

虽然以上假设信元到达 Bernoulli 过程,实际上该方法对于其它的到达模型也是适合的.只要保证到达交换机的所有信元都相互独立地等可能地到达

各个输出端口,就可以用本文中的仅有信元服务而无信元到达的条件下一步转移概率矩阵求解这类问题;否则需要导出其相应的一步转移概率矩阵.

## 五、结 论

本文通过分别考虑共享缓冲器 ATM 交换机的信元到达和信元服务过程,得到系统的状态转移矩阵,从而求解系统中的信元总数的稳态概率分布.该方法得到的信元总数的稳态概率分布与计算机模拟结果较为一致.文中基于此,分析了交换机的主要性能指标.结果表明该方法比以往的分析方法更为准确.本文还为复杂的排队系统的求解提供了新的途径.

## 参 考 文 献

- 1 Draft Revised Recommendation I.432<sup>2</sup>, CCITT Document, Mar., 1993
- 2 Y. Shobatake, et al. A one-chip scalable  $8 \times 8$  ATM switch LSI employing shared buffer architecture. IEEE J-SAC, 1991, 9(8): 1248~1254
- 3 T. Kozaki, et al.  $32 \times 32$  shared buffer type ATM switch VLSI's for B-ISDN's. IEEE J-SAC, 1991, 9(8): 1239~1247
- 4 L. Kleinrock, Queueing System 1, Theory. New York: Wiley, 391~393
- 5 Noboru Endo, et al. Sared buffer memory for an ATM exchange. IEEE Trans. Commun., Jan. 1993, COM-41(1): 237~245
- 6 A E. Eckberg and T C. Hou. Effects of Output Buffer Sharing on Buffer Requirements in an ATDM Packet Switch. Proc. INFOCOM 1986: 459~466
- 7 Zhao Yu-biao and Liu Zeng-ji. A solution for a classical probability problem. submitted to ISIT' 95, Canada

## On Some Key Technologies of ATM Switching\*

Liu Zengji (刘增基), Yang Xinhui, Zhou Daiqi, Qiu Zhiliang

(National Key Laboratory on ISN,

Xi'dian University, Xi'an 710071, P. R. China)

### Abstract

In this paper, the functions and the architecture of an ATM exchange are summarized. The switching fabrics, the line interface module and the control module are discussed in detail. An ATM switching system that adopts the scheme described in this paper has been developed.

**Key words:** ATM, Switching fabric, Bus type SF, Output buffer

### 1. Introduction

Asynchronous transfer mode (ATM) is a new system of multiplexing and switching, which is defined as the object transfer mode of the broadband integrated services digital network (B-ISDN). As the way of multiplexing, ATM is a kind of the statistical time division multiplexing and the channels are not distinguished by the location of time slots, but by the label of time slots. The length of a time slot is 53 bytes, the length of a cell. As far as the switching is concerned, ATM is a fast packet switching scheme in which the store-and-forward of packets is realized by hardware and the basic unit of transferred message is a fixed length of minipacket—cell. The way of connection oriented is also used in ATM. This connection is not a physical as in circuit switching system but a virtual like that in the packet switching system.

Being convenient in management, the virtual circuits of ATM are divided into two levels; virtual channel (VC) and virtual path (VP), here a VP is a bundle of VCs. Correspondingly, there are VP and VC connections.

The first step of the B-ISDN is the implementation of the VC switching and the VP cross connection. The principle of VC switching is also suitable for VP cross-connection and VP switching. In the paper, the functions and the architecture of ATM

switching are presented; the key technology of realizing the switching fabrics, the line interface module and the control module of the ATM switches are discussed; and the plan of upgrading the version in future is described.

### 2. Functions and Architecture of the ATM Switches

According to the reference model of the protocol of B-ISDN shown in Fig. 1, the ATM switches as the equipment of network-node should finish the following functions:

(1) functions of physical layer and ATM layer at the network side;

(2) functions of the ATM adaptation layer (AAL) on the control plane;

(3) higher layer functions on the control plane;

(4) traffic controls including connection admission control (CAC), usage parameter control (UPC), flow control (FC), and priority control (PC);

(5) functions of the management plane, including performance management, configuration management, fault management, accounting management and security management.

The basic architecture of ATM switching is shown in Fig. 2. It can be divided into switching fabric (SF), line interface (LI), control module (CM) and operation administration maintenance (OAM) module. The

functions of the modules are described as below:

SF: transferring cells from any input port to the assigned output port;

LI: performing the functions of the physical layer and the ATM layer at the network side;

CM: performing the functions of the control plane at the network side (mainly performing the transmission and processing of the signalling);

OAM: performing the functions of OAM of the network.

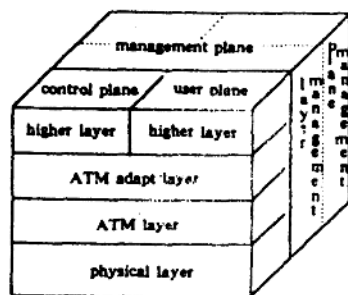


Fig. 1 B-ISDN protocol reference model

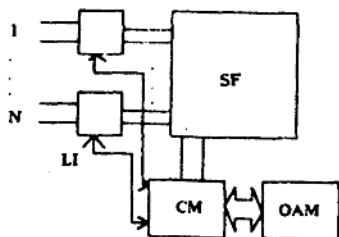


Fig. 2 Structure of ATM switch

### 3. The Switching Fabric

The ATM SF consists of the transmission medium and the cell buffers between inputs and outputs. According to the types of the transmission medium, the SF could be classified into space division type, shared buffer type, and broadcast (bus) type. According to the location of the buffers, the SF could be classified into input buffer

type, output buffer type, and mixed type. The switching fabric implemented and described in this paper is the bus type with output buffer.

#### 3.1 The principle of bus type SF with output buffer

As shown in Fig. 3(a), a  $4 \times 4$  unit of SF has four buses (on each of which 8 bit data in parallel), 4 input ports, 4 output ports, and 4 extended ports. Each output buffer consists of 5 FIFO.

The cells that come from each bus are written into relevant FIFOs respectively according to the routing mark — the value of Tag. The cells belonging to the same output buffer are read out from the relevant FIFOs in proper order by polling. The rules of polling are as follows: If FIFO is idle, jump over and examine the state of the next FIFO, otherwise read one cell from it, then to the next FIFO, and this is carried on cyclically. The polling is implemented by hardware so that the transition time could be ignored.

The  $8 \times 8$  ATM SF can be composed of four  $4 \times 4$  SF units, as shown in Fig. 3(b). In the same way the  $16 \times 16$  ATM SF could also be made up.

#### 3.2 The realization of bus type SF with output buffer

For  $4 \times 4$  fabric, each output buffer is constituted by 5 FIFO chips, and the capacity of each chip is  $2k \times 8$  bits to store 37 cells at most. In order to make the cell loss rate less than  $10^{-9}$  under the conditions of random traffic and utilization of 0.9, the capacity of output buffer at each port must be more than 23 cells (one cell has 53 bytes). Obviously, the buffer capacity is enough. For burst traffic, some measures to control the traffic must be added (see section 5.3).

The input and output data of each FIFO are 8 bits in parallel. If the bit rate on each line is 155.52 Mbps, the rate of accessing FIFO is  $155.52 \times 1/8 = 19.44$  Mb/s, that is, the access period is about 51ns, no difficulty to reach. One FPGA chip has been used to realize the writing/reading control circuit of one output buffer.

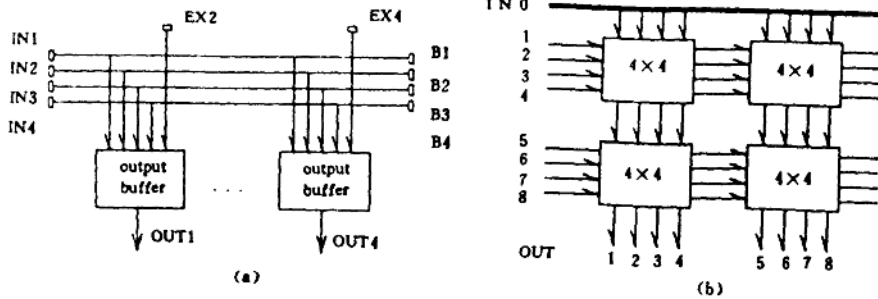


Fig. 3 Bus type ATM SF  
(a) 4x4 ATM unit (b) 8x8 ATM SF

#### 4. Line Interface

The functions at the network side that would be finished by the line interface (LI) of the ATM switches are as follows;

(1) functions of physical layer, including O/E and E/O converting, assembly and disassembly of STM-1 frame, cell delimiting, error control of the cell header (CH), idle cell deleting and inserting, etc.;

(2) functions of ATM layer, such as the processing of CH, identifying of VPI and VCI, registering of CLP, forming of Tag and translating the values of VPI and VCI, etc.

For realizing the VC switching, the first thing is to identify the value of (VPI+VCI) in the input cell header, then to take out the corresponding new value of (VPI+VCI) to be transformed and the value of Tag from the routing table in every input port, and then to change the value of (VPI+VCI) in the cell header.

The format of routing table is designed as in Fig. 4. In the table, the first column is the values of (VPI+VCI) in input cells, the second column is the corresponding values of (VPI+VCI) in output cells, and the third is the value of Tag. The table is realized by using a double-port RAM as the hardware. The input (VPI+VCI) is the address of the RAM. The transformed value

of (VPI+VCI) and Tag are stored contents that have been written in the control module by the processor. For permanent and semi-permanent connection, these stored contents are determined by the management entity. For switching connection, they are decided by the result of signalling processing in the period of establishing a call connection.

(VPI+VCI) input	(VPI+VCI) output	Tag	
0			permanent connection
005H	005H		semi-permanent connection
...	...		dynamic connection

Fig. 4 Routing Table of VC Switching

#### 5. Control Module

A main job of the control module (CM) is to perform the functions of control plane of the protocol of B-ISDN. At present, it should implement at least the call/connection control, the management and assignment of VP/VC and the resource of bandwidth.

The functions of control are realized by the signalling system involving UNI signalling and NNI signalling.

### 5.1 The ATM layer of signalling system

As a signalling point, the ATM switch will complete the functions of physical layer, ATM layer, AAL layer and higher layers of the signalling system. The physical layer and the ATM layer are common with respect to the signalling and user information, but the signalling and the user information are transferred by different VCs. The VC for transfer signalling is called SVC. When the signalling end points are in point-to-point configuration, the VC of VCI = 5 is fixed as the special SVC, but in point-to-multipoint configuration, the SVC consists of point-to-point signalling VC (PVC) and selective broadcast signalling VC (SVC). They are established by the Metasignalling that is transferred in a special VC (MSVC). In every VP, the VC of VCI = 1 is arranged as MSVC. The Metasignalling is used to establish, maintain, and release the connection of SVC. Its end point is the management entity of ATM layer. The related protocol may refer to the recommendation Q. 1420.

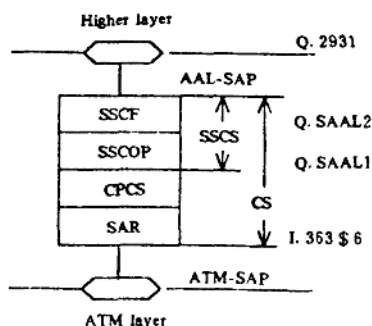


Fig. 5 SAAL

### 5.2 Signalling AAL

For signalling ATM adaptation layer (SAAL), the AAL type 5 (AAL5) is adopted. As in Fig. 5, this layer is subdivid-

ed into segmentation and reassemble (SAR) sublayer and convergence sublayer (CS). The CS is further subdivided into common part CS (CPCS) and services specified CS (SSCS). The SSCS could be divided into the service specific connection oriented protocol (SSCOP) that will support the signalling transmission and the service specific coordination function (SSCF). The specifications of the SSCOP and the SSCF conform to recommendation Q. 2110 and Q. 2130, and the specifications of CPCS + SAR comply with recommendation I. 363 § 6.

### 5.3 Signalling higher layer

The following services are supported by the signalling higher layer:

(1) The call connection control; to establish, maintain and release the VCC which is used to transfer user's information, selection or assignment of VPI/VCI during establishing the connection;

(2) Management of resources; to ensure the rational use and assignment about the resources of VP/VC and resources of bandwidth.

#### 5.3.1 Signalling message and information element

To implement the basic call connection control of the B-ISDN, the following message types have been chosen:

(1) Call establishing messages; SET-UP, CALL PROceeding, ALERTing, CONNect, CONNect ACKnowledge;

(2) Call release messages; REL ease, REL ease COMPLETE;

(3) Other messages; STATUS, STATUS ENquiry.

The common information elements contained in every message are protocol identifier, call reference, message type and message length.

Some variable-length information elements could be contained in a message such as: parameter of AAL, peak rate of ATM user cells, connection identifier (VPCI, VCCI), called party number, calling party subaddress, calling party number, calling party subaddress, broadband bearer capability, lowerlayer information of broad-

band, higher layer information of broadband, causes, etc. Some of above information elements are mandatory, and some of them are optional. The format and code of each message and information element should comply with the recommendation of Q. 2931.

### 5.3.2 The procedure of the call/connection control

The procedure that the messages are transferred between higher layer entities in every period of call connection control under normal conditions is shown in Fig. 6.

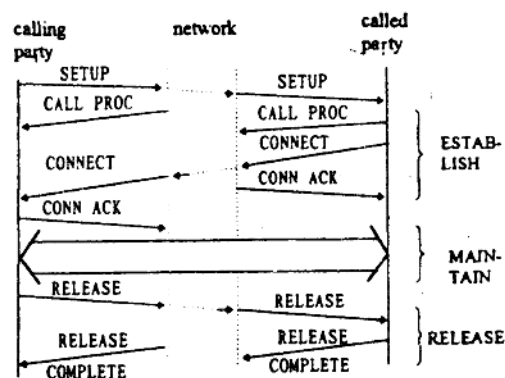


Fig. 6 Call/connection control procedure

## 6. Conclusion

The first phase (the implementation of the B-ISDN) is successful. The basic work

in this period is to realize the VC switching, the VP cross connection, the bandwidth assignment according to peak cell rate, the point to point connection, and the QOS without different classes, and to support the transparent transmission of circuit emotion digital information with constant bit rate and the data services with connection oriented and connectionless, and a few supplement services.

The next phase will realize the assignment of the bandwidth according to peak cell rate and average cell rate, the point-to-multipoint connection, the QOS with multiple classes, the function of OAM, and to provide the video service with variable bit rate and the multimedia service.

The hardware of the ATM switch could be suitable for the needs in the next phase. The enhancement of network functions and services will be supported by upgrading the software version continuously.

## References

- [1] Prycker M De et al. B-ISDN and The OSI Protocol Reference Model IEEE Network, March 1993
- [2] Takahiko Kozaki et al.  $32 \times 32$  Shared Buffer Type ATM Switch VLSI's for B-ISDN, *IEEE Journal SAC*, 1991, 9(8)
- [3] ITU-T Rec. Q. 2931, 1993

## 基于 Banyan 网的快速分组交换网络的硬件复杂度分析

贺飞云<sup>\*\*</sup> 闻懋生<sup>\*\*\*</sup> 蔡元龙<sup>\*\*\*\*</sup>

(西安交通大学 西安 710049)

**摘要** 为提高交换网络的性能,人们已提出了许多不同的网络结构技术,如并联 Banyan、串联 Banyan、扩张 Banyan,并对这些结构的性能进行了分析,但关于这些结构的硬件复杂性的研究却很少。本文通过建立几种结构技术的性能与网络设计参数的直接表达式,研究了给定性能下它们的硬件复杂度随网络规模的变化关系,并给出了它们复杂度的量级。研究表明,扩张 Banyan 是一种更好的网络结构技术。

**关键词** 分组交换 网络实现 宽带 ISDN

## Hardware Complexity Analysis of Fast Packet Switch Networks Based on Banyan Network

He Feiyun Wen Maosheng Cai Yuanlong

(Xi'an Jiaotong University, Xi'an 710049)

**Abstract** In order to improve network's performance, many architecture technologies, such as parallel Banyan, tandem Banyan and dilated Banyan, have been reported. Their performance has also been well-understood. But little work has been done in analysing and comparing their complexity. This paper presents a theoretical analysis to compare the complexity of these architectures under given performance. The complexity orders are obtained. The results show that dilated Banyan is a better method.

**Key words** packet switch, network implementation, broadband-ISDN

### 1 引言

快速分组交换网络是构成能支持多媒体业务的宽带综合业务数字网的基本要素,在文献中提出了许多快速分组交换网络的设计方案<sup>[1],[2]</sup>。为了适应网络高的传输速率,减小网络协议处理的开销,这些方案采用硬件交换网络实现分组交换,利用 VLSI 技术以硬件方式实现低层交换协议,因而在交换网络的设计中,硬件的复杂性和网络的性能一样是衡量网络设计的重

本文选自《通信学报》1996年第17卷第3期

• 西安电子科技大学综合业务网理论及关键技术国家实验室开放项目资助

要指标。

近年来,有关硬件复杂性的研究获得人们极大关注。如文献[3]研究了 Clos 网的复杂性、互连代价和所需的 VLSI 部件数;文献[4]比较了管脚受限时几种结构所需的芯片数随网络规模的变化关系等。

Banyan 网是由一些基本交换单元以一定的互连规则连接起来的多级网络。在 Banyan 网中,任意输入线到任意输出线间仅有一条通路,由于级间链路的共享,它的内部阻塞使交换网络的通过率比较低。提高 Banyan 网性能的主要方法有:交换单元内部设缓存的 Banyan 网;排序 Banyan 网和无缓存多通路 Banyan 网。本文集中在无缓存多通路 Banyan 结构上。这种结构的基本思想是通过提供入出线间多条物理通路,从而减小内部阻塞,提高网络性能。显然,性能的提高是以硬件的增加为代价的。

典型的多通路 Banyan 网结构方法是并联 Banyan、串联 Banyan 和扩张 Banyan。关于它们的性能,文献[5]给出了 Banyan 网通过率的递推公式。文献[6]给出了并联和扩张 Banyan 网的通过率的计算公式,并给出了一种近似表达式。本文遵循文献[6]的近似思想,用统一的方法建立了网络性能与网络参数间的直接表达式,在此基础上研究了不同的结构方法的复杂性随网络规模的变化关系,给出了复杂度的量级。

在本文中,硬件复杂度指达到一定的性能所需的交换单元的数目。由于网络性能与网络业务有关,讨论假设为均匀业务,即各输入线的业务独立、同分布;各到达分组以相等的概率均匀选择输出口。

## 2 网络模型

在多通路 Banyan 网中,通过丢失某些冲突的分组和让其它分组无延时地通过交换网络来解决网络固有的冲突;要求的网络性能通过增加更多的通路来达到。下面给出本文讨论的几种网络结构方法的模型:

(1) 并联 Banyan 网:一个并联度为  $d$  的网络定义为由  $d$  个相同的 Banyan 网并联组成。它们对应的输入来自同一条外部输入线,输出统计复用到同一条输出线上。来自外部输入线的分组在输入端随机地选择  $d$  个 Banyan 网中的一个通过交换网络,其结构如图 1(a)所示。

(2) 串联 Banyan 网:一个串联度为  $d$  的网络定义为  $d$  个相同的 Banyan 网的级联,各 Banyan 网的对应输出统计复用在一条输出线输出。在 Banyan 网中使用改向选路方法<sup>[2]</sup>,当发生冲突时,一个分组顺利输出,另一个改向在其它出线输出。在一个 Banyan 网的输出端,正常输出的分组到输出端,改向输出的分组进入下一个 Banyan 网再竞争,直到最后一个 Banyan,不能正常输出的分组被丢失。其结构如图 1(b)所示。

(3) 扩张 Banyan 网:扩张度为  $d$  的网络定义为将一个 Banyan 网的交换单元的每条出线用  $d$  条替代形成的网。一个进入网络的分组可在一个交换单元的  $d$  条出线的任一条进入下一级交换单元。在输出端, $d$  条输出统计复用到输出线上。一个  $8 \times 8$  的结构如图 1(c)所示。

在文献[6]中给出了计算并联和扩张 Banyan 网各级交换单元输入线上有分组的概率的叠代公式,用它可以计算网络的丢失率  $P_{\text{loss}}$ ,但是它是一个叠代公式, $P_{\text{loss}}$ 不是明显地用网络参数表示的,不适于复杂度的分析。为了获得  $P_{\text{loss}}$ 与网络参数的直接表达关系,需要某种近似方法。因此,本节先建立网络的基本关系,为以后的分析作好准备。

假定 Banyan 网由  $2 \times 2$  的交换单元组成,则一个  $N \times N$  的 Banyan 网有  $n = \log_2 N$  级,每级有  $N/2$  个交换单元。对一个扩张度为  $d$  的 Banyan 网,每一交换单元有 2 组输入和 2 组输



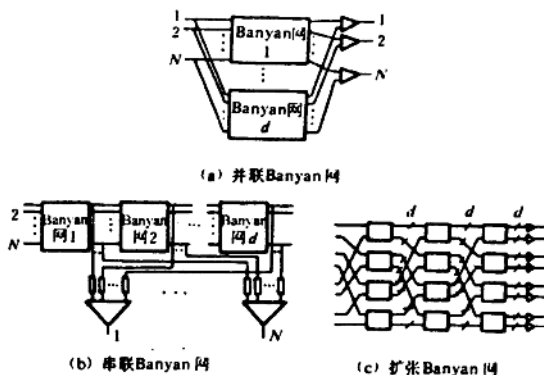


图1 几种基于Banyan网的交换结构

出,每组都有  $d$  条线。假设网络业务为均匀业务,业务到达率为  $\rho$ 。如果第  $m$  级有  $i$  个分组在一个交换单元中竞争同一组输出的  $d$  条出线,则  $i > d$  时随机地选择  $d$  个通过; $i \leq d$  时全部通过。令  $P_m$  为第  $m$  级交换单元的一条输入线上有分组到达的概率,则第  $m$  级该交换单元的一条输出线上有分组的概率,也即对应输出连向的第  $m+1$  级交换单元的一条输入线上有分组的概率为

$$P_{m+1} = \frac{1}{d} \left\{ \sum_{i=1}^d i \binom{2d}{i} \left(\frac{P_m}{2}\right)^i \left(1 - \frac{P_m}{2}\right)^{2d-i} + \sum_{i=d+1}^{2d} d \binom{2d}{i} \left(\frac{P_m}{2}\right)^i \left(1 - \frac{P_m}{2}\right)^{2d-i} \right\} = P_m - \frac{1}{d} \sum_{i=d+1}^{2d} (i-d) \binom{2d}{i} \left(\frac{P_m}{2}\right)^i \left(1 - \frac{P_m}{2}\right)^{2d-i} \quad (1)$$

其中  $P_1 = \rho$ 。

由  $P_m$  的定义可知扩张度为  $d$  的 Banyan 网的丢失率可由下式求得:

$$P_{\text{loss}} = 1 - \frac{P_{m+1}}{P_1} \quad (2)$$

普通 Banyan 网即是  $d=1$  时的扩张 Banyan 网,所以其  $P_m$  满足

$$P_{m+1} = P_m - \frac{1}{4} P_m^2 \quad (3)$$

### 3 并联 Banyan 网

并联 Banyan 网是几个 Banyan 网并联而成的。假设输入的分组随机地进入一个并联度为  $d$  的 Banyan 网的任意一个而通过交换网络,则分组的丢失率等于以交换网到达率  $\rho$  的  $1/d$  为输入业务时一个普通 Banyan 网的丢失率。先分析 Banyan 网的丢失率的直接近似表达式。

对于表达式(3),如果把变量  $m$  看作是连续变量,将  $P_{m+1}$  展开成泰勒级数并忽略高次项,则可用下列方程近似地表示  $P_m$ :

$$\begin{cases} \frac{dP_m}{dm} = -P_m^2/4 \\ P_1 = \rho \end{cases}$$

解方程可得:

$$P_m = \frac{4\rho}{m\rho + 4 - \rho} \quad (4)$$