

基于输入队列的 ATM 信元交换调度策略定性分析研究

刘宴兵¹, 李秉智¹, 幸云辉²

(1. 重庆邮电学院计算机系, 重庆 400065; 2. 北京邮电大学计算机学院, 北京 100876)

摘要: 目前基于输入队列技术的 ATM 交换机的研究日益活跃. 使用 FCFS 调度策略其吞吐率只有约 59%, 为此提出了 3 种改进的调度策略, 即输入扩展、窗口选择和信元丢弃. 本文用排队理论对改进策略进行定性分析研究, 并通过仿真和数值计算验证, 得到更加吻合实际的解析式.

关键词: 输入队列; ATM 交换机; 调度策略; 排队系统

中图分类号: TN913.24 **文献标识码:** A **文章编号:** 0372-2112(2001)11-1546-03

Performance Study of Scheduling Policies for ATM Cell Switches with Input Queue

LIU Yan-bing¹, LI Bing-zhi¹, XING Yun-hui²

(1. The Department of Computer Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. The Department of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Currently the research on input queued ATM switches is one of the most active research fields. The limiting throughput of it with input queue under the FCFS scheme is 0.59. Three policies are proposed to improve the throughput of switch: input expansion policy, window selection policy and cell discarding policy. In this paper, we study performance of these scheduling algorithms, and using simulation and numerical results, we get analytic formulas that are better than available formulas in practice.

Key words: input queued; ATM switch; scheduling policies; queueing system

1 引言

交换节点的资源调度策略是保证实时业务服务质量和提高网络资源利用率的关键^[1], 输入缓冲无阻塞交换结构中存储器和交换结构速率与端口线速相等, 非常适用于高速和大容量交换. 因此, 近年来许多高速交换机和路由器均倾向于采用输入缓冲无阻塞交换结构.

输入队列技术的优点是其交换结构简单, 内部交换速度要求较低, 缺点是其调度比较困难, 因为一个信元不但必须和其它信元竞争输出端口, 而且还要与同一输入端口的信元竞争. 输入队列技术一般使用 FCFS 队列, 在这种条件下, 研究表明输入队列 ATM 交换机在单播和均匀业务流量条件下其吞吐量只能达到 59%^[2]. 为了提高输入缓冲方式的吞吐量, 可以对输入缓冲器的队列设置和排队规则加以改进. 由于信元到达的相关性使得排队理论(话务理论)复杂化, 目前尚未找到一个与实践比较吻合的解析式, 这限制了理论对实践的指导. 本文用排队理论对几种改进策略进行定性分析、比较研究, 得到比参考文献更加吻合实际仿真的解析式.

2 交换调度模型

图 1 是交换调度模型: $N \times N$ 的内部无阻塞、带有输入队列的 ATM 交换结构. 每个时隙的宽度相当 1 个信元位置, 每

个输入端有一个输入缓冲. 在 1 个时隙内如果有个信元要传输到输出端口 j , 只有 1 个信元交换到目的端口 j ($1 \leq j \leq N$), 其余 $k-1$ 个信元随策略的不同作丢弃或存储处理.

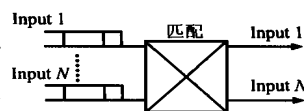


图 1 交换调度模型

3 改进调度策略及其定性分析

3.1 输入扩展策略

所谓输入扩展, 就是将每条输入线在进入交换结构之前扩展为 s 个端口, 使得原来的交换网络成为 $Ns \times N$ 的交换网络. 这样, 每个输入队列可有 s 个信元参与出线竞争.

设 B_m^i 为输入队列 i 在第 m 时隙被阻塞的信元数目; 同时定义 A_m^i 为第 i 时隙以输出端口 i 为目的端口的出现在队头的新信元数目. 则:

$$B_m^i = \max(B_{m-1}^i + A_m^i - 1, 0) \quad (1)$$

当 $N \rightarrow \infty$ 且 $s=1$, A_m^i 转化为一个带参数 ρ_0 的泊松分布^[3]. 等式(1)表示一个 $M/D/1$ 排队过程, 由排队理论可推导出其平均队列长度:

$$\bar{B}^i = \rho_0^2 / [2(1 - \rho_0)] \quad (2)$$

在稳定状态下, 第 $m-1$ 个时隙到达以输出端口 i 为目的端口的出现在队头得新信元数目即为第 $m-1$ 个时隙通过交换

已被传输的信元数目 F_{m-1} , F_{m-1} 可由以下式子得出:

$$F_{m-1} = sN - \sum_{i=1}^N B_{m-1}^i \quad (3)$$

那么输出端口 i 的吞吐量为:

$$\rho_0 = \bar{F}/N \quad (4)$$

这里 $\bar{F}$ 表示每个时隙平均交换信元的个数. 由等式(2), (3)和(4)可解得:

$$s - \rho_0 = \rho_0^2 / [2(1 - \rho_0)] \quad (5)$$

因此由上式可推导出吞吐量 ρ_0 :

$$\rho_0 = s + 1 - \sqrt{s^2 + 1} \quad (6)$$

表 1 给出几种不同扩展因子下由等式(5)计算出的吞吐量. 注意, 当 $s = 1$ 时此策略与带 FCFS 的缓冲器输入队列情况一致, 进一步验证了公式正确. 参考表 1, 可以看到: 该策略的吞吐量随着扩展因子 s 的增加而增加.

表 1 交换吞吐量与扩展因子 s 大小的对照分析

扩展因子	1	2	3	4	5	6
公式结果	0.59	0.76	0.84	0.88	0.90	0.92
仿真结果	0.59	0.75	0.83	0.87	0.89	0.92

3.2 窗口选择策略

窗口策略不再局限只有输入队列的队首信元才能参与输出竞争, 而是输入队列的前 w 个信元都可以参与输出竞争.

注意到, 通过与上面求输入扩展策略的交换吞吐量公式和参考文献[3]对窗口选择策略推导出的一个类似的结论比较, 以及大量实验仿真表明窗口选择策略获得的吞吐量通常比输入扩展策略获得的吞吐量低 0.07. 由此, 计算窗口选择策略的吞吐量公式如下, 这里 $s = w$.

$$\rho_0 = s + 1 - \sqrt{s^2 + 1} - 0.07 \quad (7)$$

表 2 给出几种不同窗口大小下由公式(7)以及计算机仿真得到的吞吐量. 由表 2 看出公式计算结果和仿真结果非常接近, 而且这比参考文献[4]中得到的解析公式结果更接近仿真结果.

表 2 由公式和仿真获得的吞吐量和窗口大小 w 的分析

窗口大小	1	2	3	4	5	6
仿真结果	0.59	0.70	0.76	0.80	0.83	0.85
公式结果	0.59	0.69	0.77	0.81	0.83	0.85

2.3 信元丢弃策略——中断的贝努里过程

在这种策略中, 如 k 个信元在 1 个时隙内要同时到同一目的输出端口时, 只有其中 1 个信元交换到输出端口, 其余的 $k-1$ 个信元作简单的丢弃处理.

现实网络的流量往往具有突发性. 假定 1 个突发中的所有信元要传输到同一输出端口地址 i , 设 k^i 代表 1 个时隙内要被交换到输出端口的信元个数, 并设与此相联系的转移矩阵为 W^* . W^* 的平稳状态分布可由以下线性方程组解得:

$$\prod_{W^*}(k^i) = \prod_{W^*}(k^i) W^* \quad (8)$$

则输出吞吐量为:

$$\rho = 1 - \pi_W^*(0) \quad (9)$$

设突发信元从输入端口 i 交换到输出端口 j 的概率为 r_{ij} ,

$\sum_{i=1}^N r_{ij} = 1$. 以下考虑输入业务量的聚合(由转移矩阵 W^* 构成): 对于 $1 \leq j, k \leq N$, 定义源 i 的转移矩阵 Q_i 如下:

$$Q_i[(1, j), (1, j)] = 1 - \alpha_i$$

$$Q_i[(1, j), (0, j)] = \alpha_i$$

$$Q_i[(1, j), (1, y)] = 0, y \neq j$$

$$Q_i[(0, j), (0, j)] = 1 - \beta_i$$

$$Q_i[(0, j), (1, k)] = \beta_{ijk}$$

注意到等式 $Q_i[(1, j), (1, y)] = 0, y \neq j$ 促使 1 个突发中的所有信元被交换到同一输出端口. 再定义 $\pi_i(x_i, y_i)$ 为第 i 个源状态 (x_i, y_i) 的平稳分布概率, 那么 $\pi_i(0, j) = (1 - \beta_i) r_{ij}$, $\pi_i(1, j) = \rho_{ij}$, $i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, N\}$.

聚合转移矩阵 W^* 可如下推导出. 首先, 聚合的源的平稳分布概率 $\pi_W^*(j), j = 0, 1, \dots, N$ 可由以下等式递归地乘 $2N^2$ 次计算出:

$$\begin{aligned} \sum_{i=0}^N \pi_W^*(i) s^i &= \prod_{i=1}^N [\sum_y \pi_i(0, y) + \sum_{y \neq j} \pi_i(1, y) + \pi_i(1, j) s] \\ &= \prod_{i=1}^N [(1 - \beta_i) + \rho_i(1 - r_{ij}) + \rho_{ij} s] \end{aligned} \quad (10)$$

上面等式需要得知概率生成函数. 那么下一步求出平稳状态概率的概率生成函数 $W^*(i, l)$, 由以下等式给出:

$$\sum_{i=0}^N \sum_{l=0}^N s^i t^l \pi_W^*(i) W^*(i, i+l) = \prod_{i=1}^N (A_i + B_i t + C_i s + D_i s t^{-1})$$

其中:

$$\begin{aligned} A_i &= \sum_{(x', y') \in F} \pi_i(x, y) Q_i[(x, y), (x', y')] \\ B_i &= \sum_{(x, y) \in F} \pi_i(x, y) Q_i[(x, y), (1, j)] \\ C_i &= \pi_i(1, j) Q_i[(x, y), (1, j)] \\ D_i &= \sum_{(x', y') \in F} \pi_i(1, j) Q_i[(1, j), (x', y')] \end{aligned} \quad (11)$$

这里 $F = \{(x, y) \neq (1, j)\}$ 和 $F' = \{(x', y') \neq (1, j)\}$. 为了获得因子 $s^i t^l$ 需要大略 $4N^2$ 次计算. 把等式(10)中的 $\pi_W^*(i)$ 除以这因子, 就可以得到矩阵 W^* 的转移概率 $W^*(i, i+l)$.

下面给出一些数值计算结果, 这里假定交换为正常均匀路由(即路由概率 $r_{ij} = 1/N$)和交换尺寸 N 为 8, 由公式迭代得到的吞吐量称为正常吞吐量. 表 3 表明了突发业务长度对吞吐量的影响. 从表 3 清楚地看到增加突发业务的长度对吞吐量没有影响, 这也是对此策略不需输入信元排队的解释(丢弃未被交换的信元).

表 3 突发业务长度对吞吐量影响的对照分析

策略 3 (交换尺寸: $N = 8$ 输入业务量 $\rho = 0.1$)					
突发业务长度(单位: 信元)	5	10	20	30	100
正常吞吐量	0.957	0.957	0.957	0.957	0.957
仿真吞吐量	0.948	0.948	0.948	0.948	0.948

从表 4 看到正常吞吐量随着业务量的增加而降低, 当业务量 $\rho \rightarrow 1$ 时, 吞吐量大约为 0.656. 这个结果告诉我们: 交换机在此策略工作下, 较好的情况是在业务量等于或小于 0.5 时, 此时的吞吐量可获得 0.8 以上. 表 4 还表明此策略优越于

FCFS 策略, FCFS 的峰值吞吐量为 0.59.

表 4 正常吞吐量和业务量的对照关系

策略 3 (交换尺寸: $N=8$; 突发业务长度=30)						
业务量 ρ	0.1	0.3	0.5	0.7	0.9	0.999
正常吞吐量	0.957	0.877	0.8066	0.7418	0.6834	0.6566

4 结束语

本文用排队理论对输入队列 ATM 交换机的几种调度策略作了定性分析进行研究, 并给出相应的一些改进的解析式. 通过理论分析和仿真模拟表明这些解析式和实际吻合, 为选择不同的调度策略提供了参考依据, 也对改善吞吐量、保证 QoS 和公平服务等具有现实指导意义. 研究输入队列调度策略不仅是设计高速 ATM 交换机的需要, 而且对未来高速路由器的设计也有借鉴作用.

参考文献:

- [1] 黄立群, 黄载禄. FQLP: ATM 网中一种新的实时业务调度算法 [J]. 电子学报, 2000, 28(4): 20-23.
- [2] 戴礼森, 洪佩琳. 高速信元交换调度算法研究 [J]. 电子学报, 2000, 28(5): 96-98.
- [3] Karol M J, et al. Input versus output queueing on a space division packet switch [J]. IEEE Trans. on Commun, 1988, 35(12): 1347-1356.

- [4] HLUCHJ M G, KAROL M J. Queueing in high performance packet switching [J]. IEEE J. Sel. Areas Commun, 1998, 6(9): 1587-1597.
- [5] LEE T T. A modular architecture for very large packet switches [J]. IEEE Trans. on Commun, 1990, 38(7): 1097-1106.

作者简介:



刘宴兵 男, 1971 年生于四川省遂宁市. 讲师, 取得 Cisco 网络技术教员资格. 北京邮电大学硕士研究生. 主要研究方向为宽带网络性能分析, 发表论文近 20 篇.



李秉智 男, 1946 年生于河北省唐山市. 教授, 博士生导师, 重庆邮电学院计算机系主任. 主要研究方向为宽带网络技术. 曾获国家科技成果进步二等奖等多项奖励.

幸云辉 女, 1938 年 5 月出生于湖南省长沙市. 教授, 主要从事计算机网络及应用研究.

(上接第 1545 页)

- [37] Lin W, Zheng R, Hou J. How to make assured services more assured [A]. Proceedings of ICNP [C], Toronto, Canada, Oct. 1999: 182-191.
- [38] I Yeom, A Reddy. Impact of marking strategy on aggregated flows in a differentiated services network [A]. IWQoS'99 [C], May 1999.
- [39] M May, J Bolot, A Jear Marie, C Diot. Simple performance models of differentiated services schemes for the internet [A]. Proc. INFOCOM'99, New York City, NY, March 1999: 1385-1394.
- [40] S Sahu, P Nain, D Towsley, C Diot, V Firoiu. On achievable service differentiation with token bucket marking for TCP [R]. UMASS CMPSCI technical report, pp. 99-72.
- [41] Yeom I, Reddy A. Modeling TCP behavior in a differentiated services network [J]. To appear in IEEE/ACM Transactions on Networking, Feb. 2001.
- [42] R Gibbens, et al. An approach to service level agreements for IP net-

works with differentiated services [DB/OL], <http://www.statslab.cam.ac.uk/~richard/research/papers/sla/>, January 2000.

- [43] Haitao Wu, Keping Long, Shiduan Cheng, Jian Ma. A direct congestion control scheme for non responsive flow control in diff serv IP networks. Internet Draft <draft-wuht-diffserv-dcos-00.txt> [S]. IETF, Aug. 2000.
- [44] Haitao Wu, Keping Long, Shiduan Cheng, Jian Ma. Direct congestion control scheme for non responsive flow control in diff serv IP networks [P]. Invention Reports Applying for Patent, March, 2000.
- [45] Haitao Wu, Keping Long, Shiduan Cheng, Jian Ma. A self configuring TCP friendly marker based on random token bucket [P]. Invention Reports Applying for Patent, March, 2000.
- [46] Haitao Wu, Keping Long, Shiduan Cheng, Jian Ma. A general three color marker [P]. Invention Reports Applying for Patent, July, 2000.