

66-69

ATM 交换机缓冲策略研究

On Study of Buffering Tactics in ATM Switches

7/9/13-24

陈永光

(解放军电子工程学院 合肥230037)

卢锡城

(国防科技大学计算机学院)

摘要 This paper analyzes fabrics of several typical buffering systems, discusses the role of link speedup and emphatically studies HOL blocking in FIFO input buffered systems, as well as the improvement way of HOL blocking. With the help of computer simulation the paper proposes buffering tactics of ATM switches useful for practical systems.

关键词 ATM, Switch, Buffer, HOL blocking

ATM 交换机设计中首先进行最彻底的调查领域是缓冲系统的大小和位置,在均匀业务的假设下,无加速交换结构的输入缓冲型纵横式交换机的吞吐率被限制在58.6%以下^[1];很明显宽带交换系统需要比这好的性能,而纯输出缓冲型 ATM 交换机虽然能实现最佳的吞吐-延迟性能,但其中大部分却深受高度复杂性之害,因此,适当地选择和构造缓冲系统,对于 ATM 交换机的性能是至关重要的。

为了对各种缓冲系统的规模有所了解,我们需要测量缓冲器中队列长度的概率群函数(pmf),pmf与概率密度函数(pdf)相似,不同之处是它用于离散随机变量,而非连续随机变量;输出队列长度的 pmf 是发现队列为特定长度的概率。根据队列长度的 pmf 曲线,我们便能够预测存储器的规模。

1 纯输出缓冲系统

1.1 输出缓冲系统结构

通过对均匀分布业务模式的仔细观察,可知 10^7 次独立试验中一个64口交换机的某个指定输出口上有36.5%次试验中没有业务去往,由于均匀分布对于一指定输出总的负载近似为1,因此就意味着在同样多的试验中该口上有一个以上的信元同时去往,这就造成输出竞争,在竞争中失败的信元如果不加以保留就会丢失。由于存在同一口上有多个信元同时到达的可能性,因此在每个输出口上需要有缓冲器来保留历史。图1是纯输出缓冲型交换机缓冲器配置及带宽要求示意图;交换机必须具有带宽不对称

现象,即输出带宽 N 倍于输入带宽,交换机结构必须是 $N \times N^2$,这一并行性对于处理可能发生的输入向输出的集中是必要的。每个输出缓冲器和输入侧也必须能够 N 倍于输入带宽,假如 BW 是一条链路的带宽,则纯输出缓冲型交换机的总缓冲系统输入带宽是 $N^2 BW$ 。

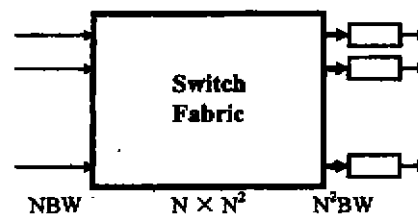


图1 输出缓冲

输出缓冲型交换机以高度复杂性为代价而具有最优吞吐-延迟性能,这类中一个著名的例子是 Knockout 交换机^[2],对于 $N \times N$ 交换机而言,它需有全互连拓扑结构; N^2 个不同的物理通路, N^2 个地址过滤器及 $N(N \times L)$ 个集中器。

1.2 链路加速的效果

当输出缓冲型 ATM 交换机的链路没有加速时,缓冲器队列为任意长度的概率几乎相等,从而使延迟变得无法确定,如果我们想使延迟有界,就必须在链路上有一点加速。加速以后,队列长度的分布就变得很好了,延迟大大减少,任何可能产生的队列都

陈永光 博士后,副教授,主要研究方向为 ATM 交换技术、电子战仿真技术,卢锡城 系主任、教授、博士生导师,主要研究方向为计算机网络与通信。

很快清除。

图2表现的是64口输出缓冲型 ATM 交换机当链路有3%和2倍加速时,在 10^6 次独立试验之后的残余队列情况。仿真结果表明,3%的加速把零 pmf 队列规模减小到大约116个信元,随着输出队列服务速度增加到原先的两倍,对于队列规模大于10个信元的,其 pmf 值已近似为零。

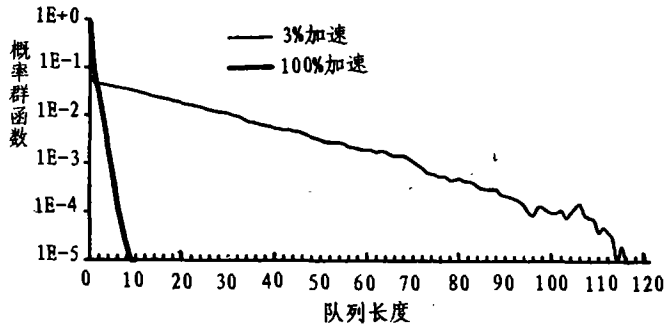


图2 链路加速因子与输出缓冲器队列长度的关系

需要指出的是,仿真中如果把输入载荷限制在全负载之下(如70%、80%或者90%),则可以得到与链路加速相同的效果。

2 输入缓冲系统

2.1 FIFO 输入缓冲系统的 HOL 阻塞

输入缓冲型交换机将所有进入的信元置于输入口处的缓冲器里,并在交换之前即解决竞争问题。对于简单的 FIFO 输入缓冲器,均匀业务下不管在竞争队列头(HOL)信元过程中所用的选择策略如何,对非阻塞交换机吞吐量的最大限制是58.6%,这么差的性能是所谓 HOL 阻塞的结果。我们定义 HOL

阻塞成立时当且仅当:□队列的第一个信元在本次周期中被阻止去往其目的地;□队列中第二个信元的 VC 与第一个信元的 VC 不同;□本周期中第二个信元的目的端口上无业务去往。

为了表现 HOL 阻塞的影响,以64口交换机为例做仿真,用 10^6 次独立试验中去向指定的输入口经历了 HOL 阻塞的输入队列个数 pmf 值绘制了图3。从

图3可见,如果所有的阻塞信元都能够取得进展,而不被 HOL 阻塞所挡住的话,交换机可以增加19%的利用率。

从一个周期到下一周期存储的信元引起的阻塞可以起到阻止下面的信元到达输出而加剧交换机利用率不足的作用,但是由竞争暂时阻塞的缓冲信元却也能在本次周期输入原本无载荷时通过向输出提供存储信元而增加利用率。由于 HOL 阻塞对交换机利用率存在正反两方面的影响,所以有必要对其整体影响作出判断。

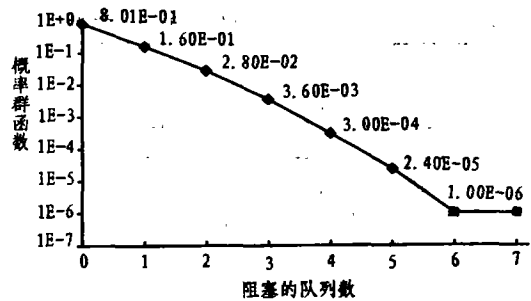


图3 FIFO 输入缓冲系统的 HOL 阻塞

表1 HOL 阻塞对交换机利用率的影响

输入数(N)	0	1	2	3	4	5	6	7	8	9	10
无缓冲(%)	36.49	37.01	18.55	6.09	1.47	0.28	0.04	0.01	0	0	0
FIFO 缓冲(%)	41.33	33.17	16	6.23	2.18	0.72	0.24	0.09	0.02	0.01	0

表1是同时有 N 个输入去往某指定输出口的次数在总试验次数中所占比例的统计情况,仿真结果表明:在 10^6 次独立试验中,一个无加速结构的 ATM 交换机输入加缓冲的净结果是将其利用率从无存储系统时的63.5%降低到有存储时的58.6%。

图4所示为 FIFO 输入缓冲系统的基本结构示意图,对于纯输入缓冲型的 ATM 交换机,每个输入上的缓冲器只需用链路速度运行。如果所有链路速度相同,缓冲器系统总输入带宽为 $N \times BW$,交换结构是 $N \times N$,结构的输入和输出带宽是 $N \times BW$,输出

链路带宽是 $1 \times BW$ 。由此可见,FIFO 输入缓冲型与输出缓冲型交换机相比具有结构简单,带宽要求不

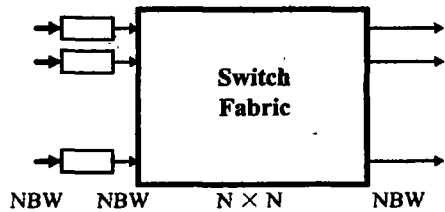


图4 输入缓冲

高等优点,但是必须克服 HOL 阻塞的影响才能使其得到有效的利用。

2.2 HOL 阻塞的改善

根据 HOL 阻塞成立的条件,为减少 HOL 阻塞的影响,我们可以在输入缓冲器中以一个固定大小的窗口观察输入 FIFO,并用交互配对的方法找出可以进展的信元。为了保证两个来自同一输入口并去向同一输出口的连续信元的顺序完整性,在观察窗口时必须以流水线方式运作,即以先进先观察的次序在输入缓冲器的窗口中找出可以进展的信元;在缓冲器输出链路带宽不变的情况下,每个窗口中每个周期只能有一个信元进展,该信元取走后,后面的顺序递补;如果窗口中没有信元能够找到空闲输出口,则只有下一周期再重新开始这一过程。

以64口 ATM 交换机为例,用仿真手段可以得到输入缓冲器开窗对吞吐率的影响。表2关于均匀分布模式下的仿真结果表明(10⁵次独立试验),大小为2的窗口已经消除了 HOL 阻塞的大部分影响;而大小超过4的窗口,其效果变得不明显。

表2 输入缓冲器开窗尺寸对吞吐率的改善效果

窗口尺寸	2	3	4	5	6	7	8
吞吐率 (%)	69.19	76.9	81.4	84.87	85.11	86.23	88.15

2.3 输出链路组织输入缓冲系统

处理并行到达数据和 HOL 阻塞以提高交换机吞吐量的另一个方法是用每个输出链路组织输入缓冲,而不是一个简单的 FIFO。如果交换机拥有 VC 缓冲器,则通过把去向同一输出口的 VC 编组即可实现这一缓冲结构。由于 VC 相同的信元被放置在同一缓冲器里并且每次以随机方式只取走一个信元,自然就消除了信元的竞争同一出口问题;这样以圆形方式服务 VC 的方法在输入缓冲器系统的输出侧实现了平行性,而不用增加缓冲器系统或交换机结构的带宽。阻塞影响被来自缓冲器系统的并行通路所克服。用控制信息做随机配对,每个输出口都能与有数据并以它为终点的输入匹配上^[3]。用此方法增加了控制机构而非交换机结构和缓冲器的带宽;在均匀分布业务模式下,这种方式可以与输出缓冲系统的效果相媲美,但却没有增加交换机的结构,这是用于高速交换机一个很吸引人的技术,因为在高速交换机中,技术的限制,不允许任何给交换机结构或缓冲器系统大量加速的办法。

图5以输入载荷与最大队列长度的关系曲线表

现了输出链路组织输入缓冲系统可以与输出缓冲系统相媲美的效果。仿真试验仍然以64口 ATM 交换机为例,以0.0~1.0(每次递增0.05)的负载率分别加入三种典型缓冲结构到交换机中,经过10⁵次试验后,观察记录缓冲队列的最大历史长度。

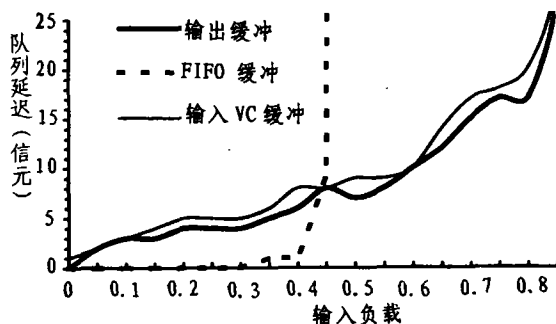


图5 最大队列长度与载荷的关系

3 组合型缓冲系统

3.1 基于输入的组合型缓冲系统

如果在 ATM 交换机的输入输出都加上缓冲器,则可减少交换机结构和输出缓冲器输入的加速。在输出缓冲一些数据可以用捕获多路输入上数据的方法增加链路利用率。交换机结构不过分的加速可以让比链路运送的更多数据进入输出缓冲器,这样就可输出缓冲器里建立起储备,也可防止 HOL 阻塞变成严重的制约。输入上的缓冲器提供了以链路速度同时从所有输入接收数据所需的带宽,输入缓冲器的输出侧和输出缓冲器的输入侧必须以交换机结构速度(S)运行。

放置缓冲器的大部分有两种选择。若缓冲器的大部分放在输出,则每一输出必须有足够的缓冲器来处理从所有输入链路积累的数据。若缓冲器的大部分放在输入,则每个输入只需有足够的缓冲器处理进到该单个输入链路的数据即可。

任一给定输出可用的总缓冲量等于输入缓冲器规模乘上向输出集中的输入数,后者可以高达 N。如果缓冲移到输出侧,则总的交换机缓冲器要求为输入缓冲器数的 N²倍才能对所有输出产生同样的效果。

并行输入缓冲器向总缓冲器系统提供了 N 个带宽,因为它们是分开的。通过割裂缓冲器系统,输入缓冲器的输出侧、输出缓冲器的输入侧以及交换机结构只需不过分的加速因子(1~2X);并且把缓冲器的大部分放在交换机的输入侧可使整个交换机的缓

语流中录制合成音节单元的策略,对系统的合成质量有较明显的改进。

3.4 抽取有效语音单元

以音节为合成单位进行波形连接的文语转换方法,音节单元录制的质量对系统的总体效果有重大的影响,本系统采用了以下策略以保证语音库音节单元的质量,即重复录音,概率择优^[6]。对于每个音节单元,录制8遍,在建立合成音节库时,选取其中一个作为音节单元。选取的策略如下:

a) 舍弃时长小于所有该类语音单元平均时长80%的语音单元。这样处理的目的是,一方面,这些被删除的语音单元的音质往往不太好。另一方面,在

合成时如果使用这些被删除的语音单元,在调整时长时,时长的增加有可能超过限定的时长增长因子(目标语音单元时长/原始语音单元时长)的阈值。造成语音信号的失真或不自然。在本系统中,时长增长因子设为1.22。

b) 计算所有语音单元平均采样能量,舍弃平均采样能量大于该值的语音单元。这样处理可以减小在合成时声音的剧烈起伏振荡。

c) 对于进行以上两个步骤后剩下的语音单元,通过试听和合成测试,最后,在每一类型语音单元选定一个作为系统使用的语音单元。

(下转第89页)

(上接第4页)

主要参考文献

- [1] H. T. Kung et al., Systolic arrays (for VLSI), Sparse Matrix Proc. Society for Industrial and Applied Mathematics, 1978
- [2] H. T. Kung, Why systolic arrays, IEEE Computer, vol. 15 Jan. 1982
- [3] S. Hauck, Multi-FPGA Systems, PH. D Thesis, Dept. CS&E, University of Washington, 1995
- [4] B. K. Fawcett et al., Reconfigurable Processing with Field Programmable Gate Arrays, in Inter. Conf. Application Specific Array Processors, 1996
- [5] M. Annaratone, et al., The Warp Computer: Architecture, Implementation and Performance, IEEE Trans. Computer, 36(12)1987
- [6] 杨超峰, 颜玲, 傅宇卓, 胡铭曾, 支持 MPEG-2 标准的 ME 芯片设计, 体系结构年会'97, 1997. 10
- [7] W. Shang et al., On uniformation of affine dependence algorithms, IEEE Trans. Computer, 45 (7) 1996
- [8] Y. Wong et al., Transformation of broadcast into propagation in systolic arrays, J. of Parallel and Distributed Computing, 14(2)1992
- [9] J.-C. Tsay et al., Design of efficient regular arrays for matrix multiplication by two step regularization, IEEE Trans. Parallel Distrib. System, 6(2)1995
- [10] W. Li et al., A singular loop transformation framework based on non-singular matrices, Int. J. Parallel Programing, 22(2)1994
- [11] H. J. Lee et al., Automatic generation of modular mapping, ASAP'96, 1996
- [12] D. I. Moldovan et al., Partitioning and mapping algorithms in fixed size systolic arrays, IEEE Trans. Comput, vol. C-35, 1986
- [13] W. Shang et al., On time mapping of uniform dependence algorithms into lower dimensional processor arrays, IEEE Trans. Parallel and Distributed Systems, 3(3)1992
- [14] S. Y. Kung, VLSI Array Processors, Englewood Cliffs, NJ: Prentice Hall, 1988
- [15] G. -J. Li et al., The design of optimal systolic arrays, IEEE Trans. Comput. vol. C-34, 1985
- [16] K. Ganapathy et al., Optimal Synthesis of Algorithm-specific Lower-Dimensional Processor Arrays, IEEE Trans. on Parallel and Distributed Systems, 7 (4)1996
- [17] V. K. Prasanna Kumar et al., Designing linear systolic array, J. Parallel and Distrib. Computing, vol. 7, 1989
- [18] A. Darte, Regular partitioning for synthesizing fixed-size systolic array, J. VLSI Integration, vol. 12, 1991
- [19] K. Ganapathy, et al., Design a scalable processor array for recurrent computations, IEEE Trans. on Parallel and Distributed Systems, 8(8)1997
- [20] A. L. Fisher et al., Synchronizing large VLSI processor arrays, IEEE Trans. Computers, C-34(8) 1985
- [21] J. Teich, et al., Scheduling of partitioned regular algorithms on processor arrays with constrained resources, ASAP'96, 1996