

文章编号:1004-1478(2011)03-0088-04

# 基于输入排队的最大匹配 调度算法仿真分析

景志勇, 方娜, 王珏

(郑州轻工业学院 计算机与通信工程学院, 河南 郑州 450002)

**摘要:**针对路由器输入队列最大匹配调度算法中存在每个时隙内输入端口和输出端口匹配率低的问题,提出:在算法设计中,可通过改变更新轮询指针的方法提高端口匹配数目,解决输出端口同步问题,从而有效提高系统的吞吐率.仿真实验表明,采用轮询指针的调度算法可以有效地提高交换机的性能.

**关键词:**调度算法;输入队列;轮询指针;仿真

**中图分类号:**TP393

**文献标志码:**A

## An simulated analysis of maximal matching based on input-queuing

JING Zhi-yong, FANG Na, WANG Jue

(College of Comp. and Com. Eng., Zhengzhou Univ. of Light Ind., Zhengzhou 450002, China)

**Abstract:** Focused on the lower matching scores of input and output port in every time slot based on input-queuing the maximal matching of router, the proposal is that by changing the methods to update polling pointer so as to improve the number of port matching, solve the synchronization problems on output port, and effectively optimize throughput rate of system. The simulation experiment showed that using matching of polling pointer was able to effectively optimize the performance of switches.

**Key words:** scheduling algorithm; input-queuing; polling pointer; simulation

## 0 引言

互联网是一个庞大的分组交换网络,随着互联网用户爆炸式的增长和网络多媒体业务的不断涌现,设计高性能大容量的路由器变得越来越迫切.其中调度算法的设计是路由器设计中一个重要部分,在很大程度上决定了路由器的性能.传统的设计方法有实验和分析2种方法:第1种方法的实现成本过高;第2种方法虽然实现成本低,但对复杂情况下一些细节把握不准,容易造成设计风险.于是,

仿真技术应运而生.当前对交换结构及其调度算法的设计研究主要通过计算机仿真来实现<sup>[1]</sup>.

交换网调度算法按照交换网缓存结构可划分为输入排队调度算法和输出排队调度算法2种.输出排队调度算法简单、直观且时延性能良好,但对加速比要求较高,当交换网的容量增大时,现有的硬件工艺水平很难实现,所以交换网调度算法的研究主要集中在输入排队算法.输入排队算法在加速比为1的情况下可以实现100%的吞吐率<sup>[2]</sup>.现有输入排队调度算法可分为最大(无权重)匹配、最大

收稿日期:2011-03-01

作者简介:景志勇(1983—),男,河南省泌阳县人,郑州轻工业学院助教,硕士,主要研究方向为计算机网络.

权重匹配、稳定结合匹配3类,其中最大匹配算法具有实现简单且运算速度快等特点.本文将主要针对最大匹配调度算法进行仿真比较分析,提出算法设计时可以通过设置轮询指针来增加输入和输出端口的匹配数目,以期有效地提高系统吞吐率.

## 1 几种典型最大匹配调度算法

### 1.1 PIM 调度算法

PIM (parallel iterative matching) 是较早提出的一种调度算法. PIM 算法的基本思想是在每个信元发送的时隙,通过多步迭代,在输入端口和输出端口之间匹配尽可能多的发送通道,利用随机选择来解决输入端口与输出端口的竞争问题.在信元开始发送时,所有的输入端口和输出端口状态都标为空闲.在每一步迭代中,都只考虑未匹配的端口,在每一步迭代之后,都有一部分输入端口和输出端口被随机匹配上. PIM 每次迭代包含3个步骤,即请求、授权、接收,由于其随机选择的局限性,算法性能和公平性并不好,上限为63%左右<sup>[3]</sup>.

### 1.2 RRM 调度算法

RRM (round-robin matching) 算法是最简单、最直观的迭代循环算法,它形成了一个循环判定的二维数组,在每个输入与每个输出之间循环调度信元. RRM 算法克服了 PIM 算法中不公平性与复杂性的缺点,并使用更快的循环仲裁器替代了 PIM 中的随机仲裁器.优先级的轮循调度算法能够使请求连接更公平,也可以均匀地分配带宽. RRM 的问题主要在于其输出端轮转指针的更新规则带来的输出端指针同步问题,因而算法不稳定,无法达到100%吞吐率<sup>[4]</sup>.

### 1.3 iSLIP 调度算法

N. McKeown<sup>[5]</sup>为了解决 RRM 算法中输出端口仲裁同步的问题,提出了 iSLIP (iterative round robin matching with serial line interface protocol) 算法,可以在一个时隙内采用多次迭代来选择交叉开关的配置,使输入端口和输出端口尽量达到最大匹配. iSLIP 算法采用优先级轮循指针方式来轮流调度每个有效输入和输出,通过更新轮循指针来解决输出端同步问题,有效地提高了系统的吞吐率,甚至可以达到100%. iSLIP 只进行1次迭代称为1-SLIP.

### 1.4 DRRM 调度算法

DRRM (the dual round-robin matching) 算法具有

低复杂度和良好可扩展性的特点. DRRM 算法在每个输入端放置  $N$  个虚拟输出队列,依照轮询规则更新每个端口的调度指针. 输入调度器选择非空虚拟输出队列向输出端发送请求,输出端最多接收  $N$  个请求,从轮转顺序中选择1个应答.

与 iSLIP 算法相比, DRRM 算法的输入端向输出端只发送  $N$  个请求,而 iSLIP 算法发送  $N \times N$  个请求;另外 DRRM 算法最后一步不需要输入端向输出端发送确认信号,而 iSLIP 算法需要发送. 因此 DRRM 更加简单. 采用轮询方式后调度算法是公平的,并且不会出现饿死现象<sup>[6]</sup>.

## 2 流量模型和仿真工具

### 2.1 Bernoulli 和 Bursty 流量模型

流量建模是宽带网络设计和性能分析的一个核心问题. 目前描述到达输入端口的业务性能,一般使用信元到达输入端口的随机过程和信元目的端口地址分布2个性能指标,本文采用随机和突发2种业务模型对交换结构进行性能分析. 随机业务模型是假定信元到达输入端口是一独立同分布 Bernoulli 过程,即在任意给定的时隙内,信元到达某一输入端口的概率为  $P(0 < -P < -1)$ , 无信元概率为  $(1 - P)$ ,  $P$  也称为输入负荷. 突发传输常用 ON/OFF 模型描述,由交替的活动期 (ON) 和沉默期 (OFF) 组成,是一种两状态马尔可夫过程,其中 ON 和 OFF 相互独立并服从几何分布,活动期内产生信元,沉默期不产生信元.

### 2.2 仿真工具

通过计算机仿真的交换结构是  $16 \times 16$  Crossbar 结构,所有链路信元传输速率相同. 对输入队列结构采用虚拟输出队列缓存方式,通过2种流量模型对不同算法进行仿真. 仿真过程中模拟的时间区间为  $[1, 500\ 000]$ , 单位是信元时隙. 主要考察各调度算法的时延性能,绘制不同环境下的仿真结果曲线并且进行比较分析. 为了精确绘制负载-平均时延曲线,取样点数目选为100点/曲线,负载增量为0.01.

## 3 仿真结果和性能比较

### 3.1 2种流量模型下 PIM, RRM 和 iSLIP 算法性能比较

PIM, RRM 及 iSLIP 算法在 Bernoulli 流量模型

下的仿真性能如图 1 所示.

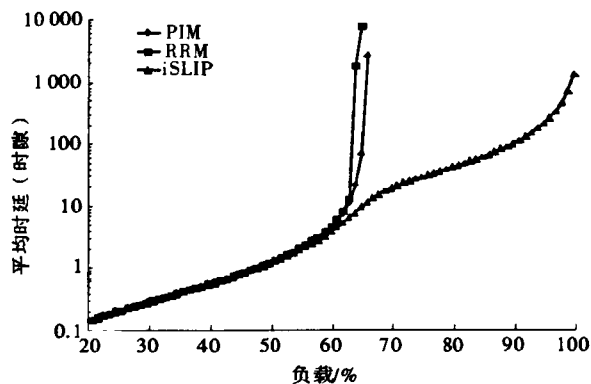


图 1 Bernoulli 流量模型下 PIM, RRM 和 iSLIP 算法性能比较

由图 1 可知,在负载较低时,到达交换网的信元数量较少,3 种算法的平均时延比较接近,iSLIP 只是略好一点.但随着负载的增加,PIM 及 RRM 分别因为其随机仲裁的局限性和输出端指针的同步问题导致时延增大,吞吐量下降.而 iSLIP 算法的平均时延增长相对缓慢,指针更新具有规律性,吞吐率可以稳定在 100%.这是因为 iSLIP 算法有效地去除了输出端口同步问题,提高了每时隙成功匹配的数量.故 iSLIP 算法相比于 PIM 和 RRM 这 2 种算法性能更加优越,且其循环调度器比随机调度器在硬件方面更容易实现.

PIM, RRM 及 iSLIP 算法在 Bursty 流量模型下的仿真性能如图 2 所示.

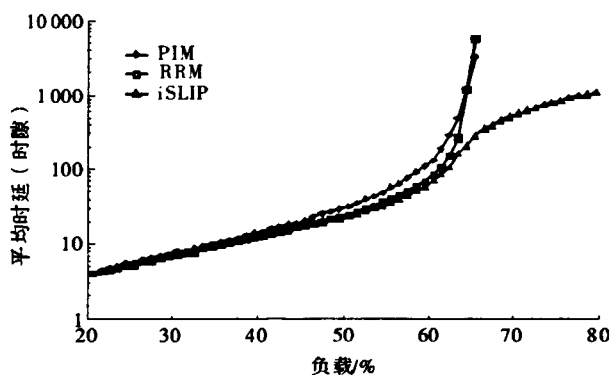


图 2 Bursty 模型下 PIM, RRM 和 iSLIP 算法性能比较

由图 2 可知,在低负载情况下 3 种算法性能相似,但对比图 1 可以发现,三者的平均时延要比 Bernoulli 流量模型下高出很多倍;随着负载的增加,PIM 和 RRM 算法性能急剧恶化,而高负载情况下 iSLIP 算法的性能下降趋势比较缓慢,并能够达到

100% 的吞吐量.故 iSLIP 算法相比于 PIM 和 RRM 2 种算法性能更为优越.

### 3.2 2 种流量模型下 1-SLIP 和 DRRM 算法性能比较

1-SLIP 和 DRRM 算法在 Bernoulli 流量模型下的仿真结果如图 3 所示.

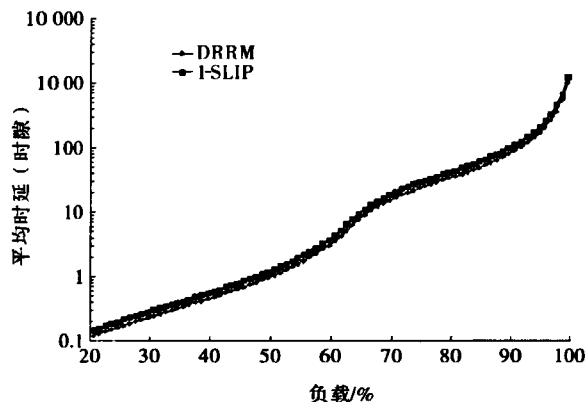


图 3 Bernoulli 流量模型下 1-SLIP 算法和 DRRM 算法的性能比较

DRRM 在每次调度开始时向输出端口发送  $N$  个端口数目的请求,而 iSLIP 算法发送  $N^2$  个请求,DRRM 算法节省了系统带宽;另外 DRRM 仲裁过程只有请求和授权 2 个步骤,有效地节省了系统内的通信量,而 iSLIP 算法需要请求、响应和确认 3 个步骤,所以 DRRM 算法具有更好的性能.从图 3 可以看出,由于两者都可以达到 100% 的吞吐量,但在流量负载较低时 DRRM 平均时延比 iSLIP 略高,负载较高时 DRRM 平均时延明显小于 iSLIP 算法.

1-SLIP 和 DRRM 算法在 Bursty 流量模型下的仿真结果如图 4 所示.

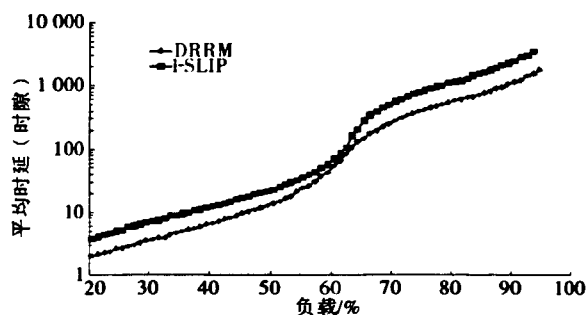


图 4 Bursty 流量模型下 1-SLIP 算法和 DRRM 算法性能比较

由于 DRRM 算法减少了系统内的控制信元流量,突发流量下性能优于 1-SLIP 算法.

## 4 结语

通过计算机仿真,本文分别得出几种典型的最大匹配调度算法在不同的流量模型下的性能:PIM和RRM类算法由于算法设计上的局限性而性能较差;iSLIP算法在均匀流量下能很好地达到100%的吞吐量;DRRM算法在性能负载较低时表现优于iSLIP算法.仿真结果表明,通过针对输入输出端口设置轮询指针并适当地变换更新轮询指针的方法,能够增加每个时隙内端口匹配的数目,有效地避免输出端同步问题发生,降低信元的平均时延.因此,在以后设计改进调度算法时,可以通过增加每个时隙中的端口匹配数目,使极大匹配更加接近最大数量匹配,来改善交换机的性能.

(上接第87页)

出了一种新的有效的网页净化算法.利用网页的DOM树,避开对半结构化数据的直接处理,使用合理的启发性规则,提高了算法的效率,通过对节点内容的比较,保证了算法的准确率.面对日益复杂的网页结构,如何提出更加合理的启发性规则,进一步提高网页主题内容的获取率和噪声内容的消除率是今后需要努力的方向.

### 参考文献:

- [1] 陈治昂,周知予,李大学.一种基于模板的快速网页文本自动抽取算法[J].计算机研究应用,2009,26(7):2646.
- [2] Yi Lan, Liu Bing. Eliminating noisy information in Web

### 参考文献:

- [1] Nong G, Hamdi M. On the provision of quality-of-service guarantees for input queued switches[J]. IEEE Com Magazine, 2000, 38(12):62.
- [2] 庞斌,贺思敏,高文.高速IP路由器中输入排队调度算法综述[J].软件学报,2003,14(5):1011.
- [3] Anderson T, Owicki S, Saxes J, et al. High speed Switch Scheduling for local area networks[J]. ACM Trans on Comp Syst, 1993, 11(4):319.
- [4] Marsan M A, Bianco A, Leonardi E, et al. RPA: A flexible scheduling algorithm for input buffered switches[J]. IEEE Trans on Com, 1999, 47(12):1921.
- [5] McKeown N. The iSLIP scheduling algorithm for input-queued switches[J]. IEEE/ACM Trans on Networking, 1999, 7(2):188.
- [6] Chao H J. Saturn: A terabit packet switch using dual round robin[J]. IEEE Com Magazine, 2000, 38(12):78.

pages for data mining[C]//Proce of Int Conf on Knowledge Discovery and Data Mining, New York: ACM Press, 2003:296-305.

- [3] 张恒,屈景辉,张亮.网页文本信息提取及结果评价[J].微计算机应用,2007,28(9):921.
- [4] 邱江涛,唐常杰,李川,等.基于块分布的新闻网页内容提取[J].吉林大学学报:工学版,2009,39(5):1326.
- [5] Lin S H, Ho J M. Discovering informative content blocks from Web documents[C]//Proc of the 8th ACM SIGKDD Int Conf, New York: ACM Press, 2002:588-593.
- [6] 王琦,唐世渭,杨冬青,等.基于DOM的网页主题信息自动提取[J].计算机研究与发展,2004,41(10):1787.
- [7] 王实,高文,李锦涛. Web数据挖掘[J]. 计算机科学, 2000, 27(4):28.