# 一种新型的 PPS 交换机

王 斌[①②] 陈 斌[①②] 张小东[②] 丁 炜[①]
[①](北京邮电大学 通信网络综合技术研究所 北京 100876)
[②](中国科学院上海微系统与信息技术研究所 上海无线通信研究中心 上海 200050)

**摘 要** 并行分组交换机(PPS)是一种利用多个低速交换结构实现高速交换的新技术，它是当前高速交换领域的一个研究热点。该文首先分析了 PPS 交换机的研究现状。然后，提出了一种新型的 PPS 交换机结构，运用流指数控制的方法保证了信元通过 PPS 交换机时不会乱序。仿真分析表明该交换结构和算法能提供良好的吞吐率和平均延迟。

**关键词** 并行分组交换机，两阶段交换机，平均时延，流指数
**中图分类号**：TN916.4 **文献标识码**：A **文章编号**：1009-5896(2006)11-2135-05

# A New Type of PPS Switches

Wang Bin[①②] Chen Bin[①②] Zhang Xiao-dong[②] Ding Wei[①]
[①](*Institute of Communication Networks Integrated Technique*, *Beijing University of Posts and Telecommunications*, *Beijing* 100876, *China*)
[②](*Shanghai Institute of Microsystem and Information Technology*, *China Academy of Sciences*, *Shanghai* 200050, *China*)

**Abstract** Parallel Packet Switch (PPS) can realize high speed switching by using low speed switches, which is a hot spot in switching area. So, this paper analyzes the current situation, and presented the PPS_RR scheduling algorithm by using the concept of flow index. The math analyzing validates that the system can guarantee packet sequence in a data flow. Finally, computer simulation shows the type of PPS has good performance on average delay.

**Key words** PPS, Two-stage switch, Average delay, Flow index

## 1 引言

自 90 年代初以来，互联网业务一直在迅猛增长。为了处理由此带来剧增的分组业务，WDM(Wavelength Division Multiplexing)被广泛地应用于传输领域。相应地，高速交换机的线速要求也已经由 10Gb/s(OC192)增长到 40Gb/s(OC768)，甚至达到 160Gb/s(OC3072)。因此，传统的交换结构已经不能适应传输带宽的迅猛增长。

近几年来，并行分组交换结构[1](Parallel Packet Switch，PPS)由多个低速并行的单级交换单元组成。它能够克服单级Crossbar的缺点，因而更适于太比特路由器。文献[1]采用了无缓存的解复用器以及OQ(Output Queuing)交换单元的结构，并证明了该结构的某些优点。文献[2]对其实用化方面提出了一些优化方法，但由于OQ结构的OQ交换机中的内存工作速度为$(N+1)R$，它不能很好地适用于端口数量多的太比特路由器。为了解决此问题， OQ交换机可以用加速比为 2 的CIOQ交换机来模拟[3]，降低所需内存的工作速度，从而实现一个内部缓存工作速度低于端口速率的PPS交换结构，但为了实现用CIOQ交换机模拟OQ交换机，CIOQ交换机调度算法的复杂度是$O(N^2)$，因此文献[2]的PPS结构的实现复杂度是$O(K\times N^2)$，离实际运用还有很长一段距离。

本文提出了一种具有更好的扩展性，易于用硬件实现，适合太比特交换场合新型的 PPS 结构，与以往的结构相比，它具有 4 大优点：(1)它不但能够满足内部线速 $r$ 低于外部线速 $R$，而且能够把分组的乱序控制在一定范围之内。(2)调度算法复杂度极低，仅有 $O(K\times N)$。它的中间交换单元为基于循环阵的两阶段交换机，这种两阶段交换机的算法复杂度是$O(1)$，而解复用器的调度算法 PPS_RR(Round Robin for PRS )是 RR 算法的改进，算法复杂度仅为 $O(K)$，因此，整个 PPS 的算法复杂度为 $O(K\times N)$)，远远低于文献[1,2]中的交换结构。(3)对缓存的要求很低。两阶段交换机中间缓存的工作速率是两阶段交换机线速的两倍(一读一写)，不存在 OQ 交换机缓存带宽的固有缺点。(4)对各种业务流具有良好的适应性。因为基于循环阵的两阶段交换机对各种业务流具有良好的适应性，特别是对非均匀流具有更好的适应能力，所以该 PPS 交换结构将继承此优点。

## 2 PPS 的结构

如图 1 所示，PPS 的基本结构包括解复用器、层平面交换机和复用器。其中，解复用器有 $N$ 个，任意解复用器 D($i$)与所有层平面交换机的第 $i$ 个端口都相连。D($i$)中有 $N$ 个 VOQ(虚拟输出队列)，它们分别对应于一个复用器。

该 PPS 结构的中间部分包含 $K$ 个层平面交换机，其中每个层平面交换机采用两阶段交换机(two-stage switch)结构，其

2005-03-09 收到，2006-02-27 改回

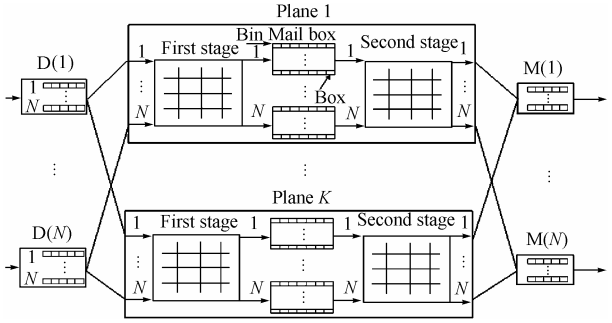


图 1 PPS 交换机的结构

Fig.1 Fabric of PPS

输入端口的线速是 PPS 输入线线速的 1/$K$。层平面交换机的两个阶段的交换机都采用循环阵交换机，两个循环阵交换机的中部有 $N$ 个中间缓存块，第 $i$ 个中间缓存块与第 1 阶段的循环阵交换机的第 $i$ 个输出端口相连的同时，也与第 2 阶段的循环阵交换机的第 $i$ 个输入端口相连。第 $i$ 个中间缓存块有 $N$ 个 VOQ 队列，它们分别对应于 1 个复用器，每个 VOQ 队列的最大长度为 $F$。

PPS 交换机有 $N$ 个复用器，$M(j)$与每个层平面交换机的第 2 阶段交换机的第 $j$ 个输出端口相连。每个复用器中都有重排序缓存，用来恢复数据流中的信元顺序。

## 3 PPS 的调度算法

为了下文的叙述方便，我们在详细介绍调度算法之前，首先阐明文中用到的几个概念、符号。

**帧和超帧** 帧的长度为连续 $K$ 个时隙，在这 $K$ 个时隙中最多可以传送 $K$ 个信元，它是 PPS 交换机的层平面交换机操作的最小时间单元。超帧的长度为连续 $N$ 个帧，它有 $KN$ 个时隙。在文中以后的部分中，我们将第 $p$ 个超帧的第 $m$ 个帧记为［$p$,$m$］，第 $p$ 个超帧的第 $m$ 个帧的第 $s$ 个时隙记为 ($p$,$m$,$s$) ，其中 $p$=0,1,⋯； $m$=1,2,⋯,$N$； $s$=1,2,⋯,$K$。

**输入线群** 输入线群Wr($i$)是解复用器D($i$)与所有层平面交换机的$i^{\text{th}}$输入端口相连的输入线的集合。输入线群有$N$个。

**数据流** 由D($i$)到M($j$)的有序信元组成的集合，记为$f_{ij}$。

**邮箱**MB 邮箱是层平面交换机的中间缓存块，$\text{MB}_{ij}$是指第$i$层平面交换机的第$j$个中间缓存块。

**箱柜** Bin 箱柜是层平面交换机中间缓存块的虚队列(VOQ)，每一个邮箱有 $N$ 个箱柜，分别对应于一个复用器，记为 Bin。

**邮箱组** 邮箱组$\text{GMB}_j$是指所有层平面交换机的第$j$个邮箱的集合。

**盒子** 每一个箱柜有 $F$ 个盒子，盒子是层平面交换机存放分组的缓存单元。

**信元指数** 信元指数是信元在邮箱中占据的盒子编号，例如：$\varepsilon_{ij}$是数据流$f_{ij}$的一个信元在某个邮箱的$j^{\text{th}}$个Bin中占据盒子的编号。由于所有邮箱中的虚队列在一帧内将被层平面交换机中的第二阶段交换机取出，所以信元指数在每帧(层平面交换机操作的最小时间单位)的开始都减少 1。

**流指数** 为解决乱序问题，系统为每个数据流分配一个流指数，它是建立在信元指数基础上的概念，与信元指数一样，流指数在每帧开始都减少 1。另外，当解复用器向层平面交换机的中间缓存块发送一个数据流 $f$ 的信元 $B$ 时，若 $B$ 获得的信元指数比数据流 $f$ 的流指数要大，那么 $f$ 的流指数更新为 $B$ 的信元指数，否则数据流 $f$ 的流指数不变。

**最大偏移量** 最大偏移量是指信元指数$\varepsilon_{ij}(t)$与流指数$\mu_{ij}(t)$的最大差值的绝对值。在本文中我们规定最大偏移量为$\delta$。它是控制乱序的重要参数。

另外，对于该PPS系统我们有以下几个约束条件。首先，整个交换系统在开始时刻，所有缓存都是空的。显然，这是容易满足的。其次，对于每个流$f_{ij}$的速率$r_{ij}$必须满足“可容许”的条件，即

$$\sum_{i=0}^{N-1} r_{ij} \le 1, \quad \sum_{j=0}^{N-1} r_{ij} \le 1, \quad i,j = 0,\cdots,(N-1)$$

### 3.1 层平面交换机的操作

在本文中，层平面交换机工作的最小时间单位是帧，层平面交换机的结构采用两阶段交换机，第 1 阶段和第 2 阶段交换机均采用循环阵交换机，循环阵交换机的定义如下：

**循环阵交换机** 循环阵交换机是指输入端口和输出端口满足下面式子：

$$j = f(i,T) = (T - i - 1) \bmod N \tag{1}$$

式(1)告诉我们，在 $pN$+$i$+1 帧内，第 $i$ 个输入端口与第 0 个输出端口相连;⋯; 在 $pN$+$i$+$N$ 帧内，第 $i$ 个输入端口与第 $N-1$ 个输出端口相连，其中 $p$ 为正整数，$i,j$=0,1,⋯,$(N-1)$。$T$ 是指第 $T$ 帧。

### 3.2 解复用器的操作过程

解复用器的操作过程是一个解复用器向层平面交换机的中间缓存发送信元的过程。由式(1)，可以得出解复用器与邮箱组的连接关系，即

$$s = f(i,T) = (T - i - 1) \bmod N \tag{2}$$

其中$s$为邮箱组号，$i$为解复用器号。式(2)告诉我们，在$pN$+$i$+1帧内，解复用器D($i$)与邮箱组$\text{GMB}_0$相连;⋯; 在$pN$+$i$+$N$帧内，解复用器D($i$)与邮箱组$\text{GMB}_{(N-1)}$相连，其中$p$为正整数，$i$,$j$=0,1,2,⋯,$(N-1)$。$T$是指第$T$帧。

在本文中，我们提出了 PPS_RR(Round Robin for PPS)算法。在阐述 PPS_RR 调度算法之前，我们先讨论一下流指数，流指数主要是为了解决乱序问题而提出的概念，它的更新方法是：

(1)在每一帧的开始，所有不为 0 的流指数都减少 1。

(2)当数据流$f_{ij}$的一个信元被发送到PPS的中间缓存块时，若该信元的信元指数$\varepsilon_{ij}>\mu_{ij}$，那么$f_{ij}$的流指数更新为$\mu_{ij}=\varepsilon_{ij}$，否则保持不变。

PPS_RR算法实质上是一种改进的RR算法，不失一般性，考察在$pN$+$i$+1 帧内，解复用器D($i$)发送信元的情况。由式(2)

可知，在$pN+i+1$帧内，D($i$)与邮箱组$GMB_0$相连。D($i$)发送信元的要点如下：

(1)在D($i$)中，系统维护一个轮询指针Pr，来指示下一次匹配考虑哪一个$VOQ_{ij}$。

(2)对于$GMB_0$中的某个$MB_{s0}$邮箱而言，D($i$)用轮询指针Pr来轮询所有在D($i$)中不为空的$VOQ_{ij}$。

(3)如果选出了合格的$VOQ_{ik}$ (该VOQ将向$MB_{s0}$邮箱的第$j$个箱柜发送信元)，那么轮询指针Pr更新为Pr=($k$+1) mod $N$(该Pr指针为下一个轮询周期的起始位置)。若$\varepsilon$为第 0 个邮箱的第$k$个箱柜的第 1 个空闲的盒子，为了使$VOQ_{ik}$合格，$f_{ik}$的流指数$\tau_{ij}(t)$和$\varepsilon$必须满足：

$$\max\{\tau_{ij}(t)-\delta,1\}\le\varepsilon\le\min\{\tau_{ij}(t)+\delta,F\} \tag{3}$$

其中 $\delta$ 为偏移量，$F$ 为箱柜的长度(含有 $F$ 个盒子)。

在 $pN+i+1$ 帧的开始，D($i$)依据上面过程最多经过 $KN$ 次匹配，能够选出要发送的信元，D($i$)依据判决的结果将选取的信元发送到相应的盒子中。

### 3.3　复用器的操作过程

由式(1)，可以得出复用器与邮箱组的连接关系，即

$$j=f(s,t)=\{(T-s-1)\ \bmod\ N\} \tag{4}$$

其中 $s$ 为邮箱组号，$j$ 为复用器号，$T$ 是指第 $T$ 帧。

复用器读取邮箱中信元的过程是一个回收盒子的过程，在第$T$帧内，第$j$个复用器与邮箱组$GMB_s$相连，$GMB_s$中所有的第$j$个Bin的第 1 个盒子的分组将在该帧内被传送到复用器M($j$)，在每个第$j$个Bin的 2nd，3rd，$\cdots$，$F$th个盒子中的分组将被向前移到 1st，2nd，$\cdots$，$(F-1)$th个盒子中。换句话说，所有不为 0 的信元指数在每帧开始都要减少 1。但是在工程上，我们在每个帧的开始，只需要更新每个箱柜的最后一个信元的信元指数，因为流指数的更新只参考箱柜的最后一个信元的信元指数。

## 4　理论分析

本节主要研究PPS的层平面交换机对信元造成的乱序问题。首先，我们分析层平面交换机对信元造成的延时问题，然后进一步证明，每流中信元的乱序是有界的。下面首先导出引理 1。

**引理 1**　若$f_{i,h}$流中有一信元$X$在[$M$, $i+j+1$]帧离开解复用器D($i$)，若它被放入某一层平面交换机的某个邮箱的第$h$个箱柜的第$q$个盒子，则

(1)它到达复用器 M($h$)的帧是[$M+q$, $j+h+1$]。其中 $i, j, h=0,1,\cdots,(N-1)$；$M=0,1,2,\cdots$；$q=1,2,\cdots,F$。

(2) PPS 的层平面交换机对信元 $X$ 的延迟为

$d=qN+(h-i)$，($d$ 的单位是以帧为单位)

**证明**　(1)由式(1)和已知条件可知，信元 $X$ 必然在［$M$,$i+j+1$］帧到达某一层平面交换机的第 $j$ 个邮箱的第 $h$ 个箱柜的第 $q$ 个盒子中。又由层平面交换机对输入缓存的管理方式、式(2)和复用器的工作方式可知，信元 $X$ 必然在［$M+q$, $j+h+1$］帧内到达复用器 M($h$)。

(2)由上面的结论，我们可以直接得到：

$$d=[(M+q)N+(j+h+1)]-[(MN+(i+j+1)]=qN+(h-i)$$

证毕

根据引理 1，可以推出下面的引理，从而进一步得到层平面交换机造成分组乱序是有上界的。

**引理 2**　如果 $\delta(\delta<F)$是 PPS 交换机的最大偏移量，若 PPS 采用上面调度算法，那么在乱序的情况下，同一数据流中的两个相邻信元的信元指数的最大差值是 $\delta$。

**证明**　不失一般性地假设$X$，$Y$是$f_{i,h}$流的两个相邻的信元，$X$和$Y$被解复用器D($i$)发送的时间分别是$t_1$和$t_2$，其中$t_2\ge t_1$，$t_2-t_1=\omega$，$t_2$和$t_1$的单位是帧，$f_{i,h}$流发送$X$和$Y$的信元指数分别是$\alpha(t)$和$\beta(t)$。在$t_1$和$t_2$时，$f_{i,h}$的流指数分别为$\tau(t_1)$和$\tau(t_2)$。下面分两种情况进行讨论：

(1)当$t_2>t_1$时　(a)当$f_{i,h}$的流指数在$t_1$时刻发生了更新。此时必然有$\alpha(t_1)>\tau(t_1)$，流指数更新为$\alpha(t_1)$。又由于信元指数和流指数在每帧的开始都减少 1，因而，在$t_2$帧的起始时刻，$X$信元的信元指数和流指数分别更新为

$$\begin{aligned}&\alpha(t_2)=\alpha(t_1)-\omega\\&\tau(t_2)=\tau(t_1)-\omega=\alpha(t_1)-\omega=\alpha(t_2)\end{aligned} \tag{5}$$

由式(3)和式(5)可得

$$\left|\beta(t_2)-\alpha(t_2)\right|=\left|\beta(t_2)-\tau(t_2)\right|\le\delta \tag{6}$$

(b)当$f_{i,h}$的流指数在$t_1$时刻不发生更新。则在$t_2$帧的起始时刻，有

$$\alpha(t_2)=\alpha(t_1)-\omega \tag{7}$$

$$\tau(t_2)=\tau(t_1)-\omega \tag{8}$$

由式(7)和式(8)可得

$$\left|\alpha(t_2)-\tau(t_2)\right|=\left|\alpha(t_1)-\tau(t_1)\right|\le\delta \tag{9}$$

而

$$\left|\beta(t_2)-\tau(t_2)\right|\le\delta \tag{10}$$

进一步，由式(9)和式(10)可得，在乱序的情况下：

$$\left|\beta(t_2)-\alpha(t_2)\right|\le\delta \tag{11}$$

(2)当$t_1=t_2$时，$X$和$Y$将在同一帧内被发送　(a)当$f_{i,h}$的流指数在$t_1$时刻发生了更新。此时必然有$\alpha(t_1)>\tau(t_1)$，流指数更新为$\alpha(t_1)$。那么根据式(3)，当$f_{i,h}$发送$Y$时，必然有

$$\left|\beta(t_1)-\alpha(t_1)\right|=\left|\beta(t_1)-\tau(t_1)\right|\le\delta \tag{12}$$

(b)当$f_{i,h}$的流指数在$t_1$时刻不发生更新。此时必然有$\alpha(t_1)<\tau(t_1)$。

由引理 1 可知，当$\beta(t_1)>\tau(t_1)$，$X$和$Y$不会发生乱序，则当$X$和$Y$发生乱序时，必有

$$\left|\beta(t_1)-\alpha(t_1)\right|\le\delta \tag{13}$$

综上所述：$\left|\beta(t_1)-\alpha(t_1)\right|\le\delta$，引理 2 成立。　　证毕

由引理 2 和引理 1，可以直接得到下面的定理。

**定理**　如果 $\delta(\delta<F)$是 PPS 交换机的最大偏移量，若 PPS 采用上面调度算法，那么在 PPS 系统中，同一数据流中最大

信元乱序是 $\delta$。

**证明** 不失一般性地假设$X$，$Y$是$f_{i,h}$流的两个相邻的信元，$X$和$Y$被解复用器D($i$)发送的时间分别是$t_1$和$t_2$，其中$t_2>t_1$。若在时间$t$时，$X$和$Y$都在层平面交换机的中间缓存中，它们的信元指数分别是$\alpha(t)$和$\beta(t)$。由引理2可知：$|\beta(t)-\alpha(t)|\le\delta$。证毕

下面，我们分 3 种情况进行讨论：

(1)当 $\alpha(t)=\beta(t)$时，由复用器的工作方式容易知道，$X$ 和 $Y$ 到达复用器时不会乱序。

(2)当 $\beta(t)>\alpha(t)$时，由引理 1 可知，$X$ 和 $Y$ 不会乱序。

(3)当 $\beta(t)<\alpha(t)$时，由引理 1 可知，$X$ 和 $Y$ 发生了乱序，乱序的最大数量是 $\delta$。

定理表明同一数据流中最大信元乱序是 $\delta$，因此在每个复用器当中必须要有 $KN\delta$ 个缓存来对信元进行重新排序，信元在复用器当中被延迟的最大上界是 $KN\delta$ 个时隙。

## 5 计算机仿真分析

本节采用计算机仿真的方法对PPS交换机的性能进行分析。计算机仿真主要集中在信元平均时延和稳定性两个方面的性能上。仿真选择 PPS 交换结构的规模为 16×16，单端口的输入线速是 80Gbit/s，层平面交换机共有 4 个，层平面交换机的单端口速率是 20Gbit/s，信元的长度为 53byte。各解复用器中的 VOQ 长度不做限制。仿真的时长是 500000 个时隙(时隙为发送一个信元需要的时间)，样本取样从第 50000 时隙才开始，平均时延是在[50000,500000]期间计算出来的。符合可容许条件的业务流采用均匀分布的贝努力业务流(UBP)和突发业务流(UIBP)两种。

稳定性的研究采用稳定参数进行评估。根据文献[4]，当稳定参数 $L(n)$满足 $E(L(n))<\infty$ 时，调度算法是稳定的。$L(n)$的表达式是：

$$L(n)=\Big[\mathrm{VOQ}_{00}(n)^2+\cdots+\mathrm{VOQ}_{1N}(n)^2+\cdots+\mathrm{VOQ}_{N1}(n)^2+\cdots+\mathrm{VOQ}_{(N-1)(N-1)}(n)^2\Big]^{1/2}$$

$$L=E(L(n))$$

在本节中，我们不打算从理论上证明 $L<\infty$，仅仅依靠仿真来分析这种 PPS 的时延特性和稳定性能。下面，我们分两种情况分别进行讨论。

### 5.1 UBP 业务流

如图 2 所示，PPS 交换机的最大偏移量为 3。当输入线路利用率 $\rho$ 低于 0.9 时，PPS 交换机的时延增长速度非常缓慢，交换机的平均时延低于 150 时隙。当输入线路利用率 $\rho$ 高于 0.9，交换机的平均时延急剧增长；输入线路利用率 $\rho$ 为 0.94 时，交换机的平均时延是 1068.415 时隙；输入线路利用率 $\rho$ 为 0.98 时，交换机的平均时延是 12678.4515 时隙。总之，在输入流为 UBP 和 $\rho$ 高于 0.9 的情况下，PPS 交换机的平均时延特性才开始急剧恶化。

如图 3 所示，当输入线路利用率 $\rho$ 低于 0.9 时，PPS 交换机的稳定参数低于 10。当 $\rho$ 在[0.9,0.94]时，稳定参数快速增长。当 $\rho$ 大于 0.94 时，交换机的稳定参数 $L$ 开始急剧增长。因此，我们可以得出这样一个结论：当输入业务流是 UBP 且 $\rho$ 小于 0.94 时，PPS 交换机工作在稳定状态。

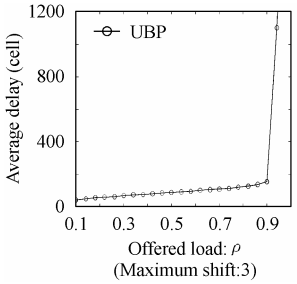


图 2 UBP 业务流下的时延分析

Fig.2 Average delay performance under UBP traffic

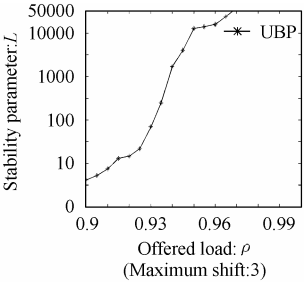


图 3 UBP 业务流下的稳定性分析

Fig.3 Stability performance under UBP traffic

### 5.2 UIBP 业务流

如图 4 所示，PPS 交换机的最大偏移量为 3。当输入线路利用率 $\rho$ 低于 0.5 时，PPS 交换机的时延增长速度非常缓慢。当输入线路利用率 $\rho$ 高于 0.5 且低于 0.82 时，交换机的平均时延增长速度开始加快。当 $\rho$ 高于 0.82 时，交换机的平均时延急剧增长；输入线路利用率 $\rho$ 为 0.9 时，交换机的平均时延是 1468.2817 时隙；输入线路利用率 $\rho$ 为 0.98 时，交换机的平均时延是14642.02019时隙。总之，在输入流为UIBP和 $\rho$ 高于 0.82 的情况下，PPS 交换机的平均时延特性才开始恶化，当 $\rho$ 高于 0.9 时，交换机的平均时延急剧恶化。

如图 5 所示，当输入线路利用率 $\rho$ 低于 0.6 时，PPS 交换机的稳定参数低于 70。当 $\rho$ 在[0.6,0.9]时，稳定参数快速增长。当 $\rho$ 高于 0.9 时，稳定参数急剧增长。因此，我们可以得出结论：当输入业务流是 UIBP 且 $\rho$ 小于 0.90 时，PPS 交换机工作在稳定状态。

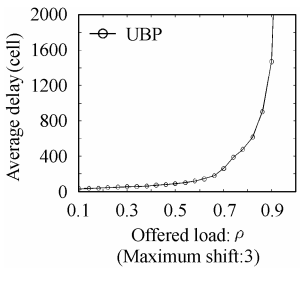


图 4 UIBP 业务流下的时延分析

Fig.4 Average delay performance under UIBP traffic

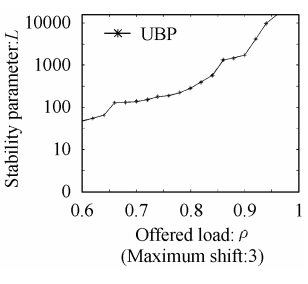


图 5 UIBP 业务流下的时延分析

Fig.5 Stability performance under UIBP traffic

### 5.3 最大偏移量 $\delta$ 对交换机的性能影响

如图 6 所示，进入 PPS 交换机的业务流是 UBP 流。当输入线路利用率 $\rho$ 低于 0.9 时，$\delta$ 的值越小，交换机的平均延时越大。当 $\rho$ 高于 0.9 时，$\delta$=1 的情况的平均延时高于其它两种情况，但是，随着 $\delta$ 值的增加($\delta$>3)，$\delta$ 对平均时延的

影响已经非常小。

如图7所示，进入PPS交换机的业务流是UIBP流。由图7中的曲线可以看出，3条曲线几乎重合。因此，我们可以得出结论，$\delta$值对稳定参数$L$的影响很小。

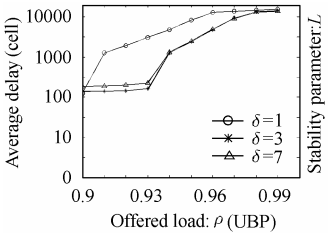


图6　$\delta$对平均时延的影响

Fig.6 The influence of $\delta$ on average delay performance

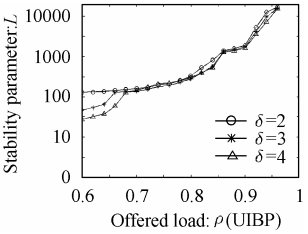


图7　$\delta$对稳定性的影响

Fig.7 The influence of $\delta$ on stability performance

## 6　结束语

在本文中，我们提出了一种基于流指数概念的PPS交换机。理论分析表明，该PPS能够将信元的失序程度控制在$\delta$(最大偏移量)范围内。仿真分析也表明，这种PPS结构在UBP和UIBP类型的数据流作用下，具有良好的时延特性和稳定性。调度算法PPS_RR是RR算法的改进，算法复杂度极低，仅有$O(KN)$，容易用硬件实现，非常适合于高速交换。

## 参 考 文 献

[1] Iyer. S, Awadallah. A, *et al*.. Analysis of packet switch with memories munning mlower than the mine mate[A]. IEEE INFOCOM'00, Tel Aviv, Israel, April 2000, vol.2: 529–537.

[2] Iyer S, McKeown N. Making parallel packet switches practical. IEEE INFOCOM, May 2001, vol.3: 1680–1687.

[3] S-T Chuang, Goel A, McKeown N, Prabhakar B. Matching output queueing with a combined input output queued switch. *IEEE J. on Select. Areas Commun*, 1999,17: 1030–1039.

[4] Chang C S, Lee D S, Jou Y S. Load balanced Birkhoff-von Neumann switches Part I: one-stage buffering. *Computer Communications*, 2002, 25(6): 611–622.

王　斌：男，1970年生，博士后，研究方向为交换技术和调度算法、MPLS、GMPLS、3GPP LTE 和 WCDMA.

陈　斌：男，1973年生，博士，研究方向为IP网与多媒体通信、人工智能、3GPP LTE 和 WCDMA.

张小东：男，1971年生，博士，研究员，研究方向为无线网络、接入网和小波技术.

丁　炜：男, 1935年生, 教授，博士生导师，主要研究方向为高速交换和路由技术、MPLS 和流量工程、网络安全、以及下一代网(NGN)的关键技术等.