# 多级交换中支持包保序的交换结构及调度算法

胡宇翔 兰巨龙 邬钧霆
(国家数字交换系统工程技术研究中心 郑州 450002)

**摘 要**：现有单级交换结构在其规模的有效扩展方面存在瓶颈。该文提出了一种新的中间级带缓存的高可扩展多级交换结构，并建立了该结构的排队论模型。针对交换网络内部的信元乱序问题，该文基于上述结构提出一种新的包保序算法，该算法通过严格同步输入级和中间级调度指针，能够简单有效地实现信元保序。理论分析结果表明，该结构能够获得100%的吞吐量且实现代价较小。仿真实验表明，该算法不仅能够获得较为理想的高吞吐量，并且在高负载强度下的平均时延性能优势明显。

**关键词**：多级交换网络；交换结构；包保序；调度
**中图分类号**：TP393.08 **文献标识码**：A **文章编号**：1009-5896(2010)02-0272-06
**DOI**: 10.3724/SP.J.1146.2009.00299

# The Switch Structure and Scheduling Algorithm for Maintaining Packet Order in Multistage Switching Fabric

Hu Yu-xiang Lan Ju-long Wu Jun-ting
(*National Digital Switching System Engineering & Technological R&D Center*, *Zhengzhou* 450002, *China*)

**Abstract**: Current single-stage switch structure encounters its bottleneck in scalability. This paper proposes a novel central-stage buffered scalable multistage switch structure, and establishes its mathematical model by queuing theory. For the problem of cell disorder, this paper puts forward a new algorithm to maintaining packet order simply and effectively by controlling the point in input stage and central stage strictly. The results of academic analysis show that this structure could provide 100% throughput which costs less as well. The results of simulation show that this algorithm not only could provide perfect throughput performance, but also take on a better delay performance in heavy load.

**Key words**: Multistage switching fabric; Switch structure; Maintaining packet order; Scheduling

## 1 引言

随着信息化进程的不断深入，因特网用户数量急剧增长，网络业务类型日趋多样化。目前，网络业务量的爆炸式增长带动现代交换技术的迅猛发展：分组交换技术正在朝着更高的性能、灵活的可扩展性、高度的可靠性及良好的经济性的目标演进。传统单级交换结构由于内部加速和实现复杂度等问题，其可扩展性受到了较大限制。于是，国内外的许多学者开始尝试利用多个较小规模的单级交换结构来构建具有更大规模的多级交换网络，其目的是综合多级交换网络的技术优势，设计具有高扩展性、多端口数的交换结构，并取得了较为丰富的研究成果。

Eiji 等人在交换网络的第 1 级和第 3 级分别采用共享缓存方式的输入/输出排队交换单元，中间级采用不带缓存的 Crossbar 交换单元的空分交换方式，提出一种 Memory-Space-Memory(MSM)型 Clos 交换结构[1]。该结构的输入级缓存针对交换网络的不同输出端口构建不同的虚拟输出队列(Virtual Output Queuing，VOQ)，既可以保证到达不同输出端口和不同优先级的分组能够公平利用交换资源，又可以暂存当前时隙无法传送到其目的输出端口的分组。但 MSM 结构的内部链路拥塞较为严重，且未能充分考虑输入业务特性，不管是基于随机匹配策略的 Random Dispatching(RD)算法[2]还是基于轮询匹配的 Concurrent Round-Robin-Based Dispatching Scheme(CRRD)算法[1]，在非均匀流量下均性能不高。针对以上分析，Nikos 等对 MSM 结构进行改进，提出一个 1024×1024 规模、总容量可以达到 10Tbps 的全缓存交换结构[3]，并从理论上推导出了该结构的内部无阻塞特性，然而集中式的调度策略严重限制了该结构的可扩展性。Wang 等人也提出了一种四级对称 Clos 交换结构[4]，并证明了当 $m \geqslant 4n-1$ 时该结构无需加速即可模拟输出排队交

2009-03-09 收到，2009-07-27 改回
国家 973 计划项目(2007CB307102)和国家 863 计划项目(2008AA01A323，2008AA01Z214)资助课题
通信作者：胡宇翔 huyuxiang1982@yahoo.com.cn

换结构，然而该结构的中间级和输出级的缓存都需要 $m$ 倍加速，且缓存需求数量较大。

## 2　CB-3Clos 交换结构

多级交换网络拓扑结构的选取直接决定着交换结构的性能及其可扩展性，为解决内部链路拥塞，现有交换网络多采用复杂而高效的集中式调度策略以严格控制业务流的行为，这在实际的交换系统中是很难做到的。一种新的想法就是在交换网络内部也设置缓存，内部缓存器的设置可以有效缓解多级交换网络的内部链路拥塞，从而获得较高的吞吐量性能，并能够有效减小调度算法紧迫度，虽然实现复杂度较高，但目前电路工艺水平的发展也使得这种想法成为可能。本文基于在中间级设置缓存这一思想，提出一种中间级带缓存的多级交换结构。

### 2.1 CB-3Clos 交换的模型

本节首先给出这种中间级带缓存的多级交换结构—CB-3Clos(Central-stage Buffered Three- stage Clos switch)的拓扑结构模型，如图1所示。CB-3Clos 交换结构采用较小端口规模的单级交换单元(Switch Cell，SC)作为其组成元素，构建成多级交换阵列，从而能够极大提高交换结构的端口数。

CB-3Clos 交换结构可以分为输入级(Input Stage，IS)、中间级(Central Stage，CS)和输出级(Output Stage，OS)，其输入级和输出级采用无缓存的空分交换结构，由 $r$ 个无缓存的交换单元组成，交换单元的端口规模分别为 $n\times m$ 和 $m\times n$；中间级由 $m$ 个输入端口带缓存的 $r\times r$ 规模的交换单元组成，任意 IS 的输出端口和 CS 的输入端口之间有且仅有一条带缓存的链路连接，整个 CB-3Clos 交换结构形成一个 $N\times N$ 规模的交换系统，其中 $N=n\times r$。每条链路上的缓存都由一个分路器(dispatcher)和合路器(synthesizer)来管理，分别完成该链路上不同目的端口的信元的分路和合路功能。

在此，本文假定交换网络仍采用固定长度信元交换方式，数据包的切割与重组在线路接口卡处完成。链路传输或者接收一个定长信元所需的时间就称为一个时隙[5]，目前，多数调度算法要求其调度过程在一个时隙内完成。缓存队列管理策略采用 FIFO(First-In-First-Out)方式或者基于特定的 QoS (Quality of Service)策略的 PIFO(Push-In-First-Out)方式。

### 2.2　CB-3Clos 交换结构性能分析

**2.2.1 CB-3Clos 稳定性分析**　为更直观说明 CB-3Clos 交换结构的特性，本小节首先给出一个简单示例：假定有一业务流从 CB-3Clos 结构的 1 号输入端口进入交换网络，其目的为 1 号输出端口，为方便描述，在此去掉了 CB-3Clos 交换结构中与该业务流不相关的交换单元及内部链路，如图 2 所示。

由图 2 可以看出，CB-3Clos 的信元调度过程就可以抽象为业务流经输入级负载均衡到中间级后，由输出级调度出交换网络的过程。即：在输入级，调度器根据中间级缓存及链路的状态，将该业务流的信元均衡地分配到与之相连的所有中间级交换单元上，然后信元经中间级交换单元及输出级调度路

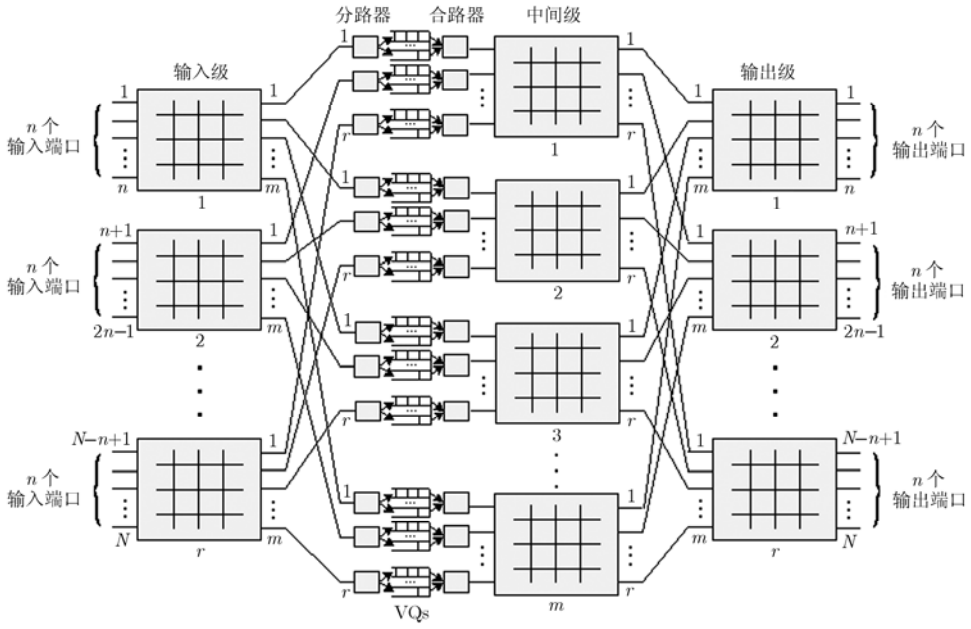


图 1 CB-3Clos 交换结构模型

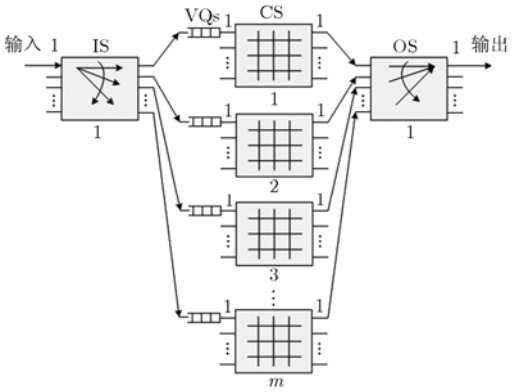


图 2 CB-3Clos 信元调度过程的一个简单示例

由出交换网络。在将业务流均衡到中间级链路过程中，共有 $m$ 条路径可以选择，这种均衡特性可以由以下定义来描述。

**定义 1** CB-3Clos 交换结构的扩展比：设 CB-3Clos 交换结构所支持的最大端口链路速率为 $R$，中间级虚拟队列所支持的信元最大离开速率为 $R'$，则定义 CB-3Clos 交换结构的扩展比为

$$S = \frac{m \times R'}{R} \tag{1}$$

交换结构的扩展比可以刻画交换结构的规模、内部路径数、中间级链路利用率及其可扩展性等。在讨论交换网络内部冲突的概率时，$S$ 越大，交换网络内部冲突的概率越小；而讨论交换结构的中间级链路利用率时，并非 $S$ 越大，利用率越高；同时，考虑交换结构的规模及其可扩展性，为获得交换结构的较好的性能和较小的实现代价，需要综合考虑 $S$ 的取值。

**定理 1** 当 $S>n$ 时，CB-3Clos 交换结构是稳定的。

**证明** 由前文分析可知，考察 CB-3Clos 交换结构的稳定性，只需考察单个业务流通过交换网络内部缓存时所构成的排队系统的稳定性，如果该排队系统是稳定的，那么交换网络内部的缓存就都是稳定的，则 CB-3Clos 交换结构就是稳定的。

不失一般性，在此仍只分析输入端口为 1 号输入端口，其目的端口为 1 号输出端口的业务流，如图 2 所示。不妨定义该业务流通过交换网络内部缓存时所构成的排队系统为 $P$，同时假定 CB-3Clos 结构的中间级缓存容量为无限大。

对于排队系统 $P$，假定信元的到达过程服从参数为$\lambda$的泊松过程，易知$\lambda \leqslant R$。考虑到输入级交换单元的流量均衡效应，那么任意第 $x$ 个中间级链路上 VQ 的信元到达过程就服从参数为 $\lambda_x$($1 \leqslant x \leqslant m$, $0 \leqslant \lambda_x \leqslant \lambda$)的泊松过程，且满足条件

$$\sum_{x=1}^{m} \lambda_x = \lambda, \quad 1 \le x \le m \tag{2}$$

考虑到各中间级 VQ 的地位完全相同，且假定输入级交换单元能够获得严格的负载均衡，那么就有

$$\lambda_x = \lambda_y, \quad x \ne y,\ 1 \le x, y \le m \tag{3}$$

由式(2)和式(3)可得

$$\lambda_x = \lambda_y = \frac{\lambda}{m}, \quad x \ne y,\ 1 \le x, y \le m \tag{4}$$

对于中间级上的各 VQ，其调度过程可以视为该 VQ 为信元提供服务的过程。由于每个 VQ 最多可由 $n$ 个输入端口的业务流共享，因此，每一个 VQ 的信元到达和离开过程就可以抽象为输入过程服从参数为 $(n \times \lambda)/m$ 的泊松过程，服务时间服从参数为 $R'$ 的指数分布的 $M/M/1$ 排队系统[6]，在此定义该 VQ 所构成的 $M/M/1$ 系统为系统 $Q$，那么排队系统 $P$ 可抽象为由 $m$ 个排队系统 $Q$ 组成。

对于单个 VQ 所组成的 $M/M/1$ 系统 $Q$，其出生率和死亡率分别为

$$\lambda_i = \frac{n \times \lambda}{m}, \quad i \ge 0 \tag{5}$$

$$\mu_i = R', \quad i \ge 0 \tag{6}$$

其中 $i$ 为排队系统 $Q$ 中的顾客数。

下面考察排队系统 $Q$ 的稳定性。

$M/M/1$ 系统的稳定性可以由其平稳分布刻画。由生灭过程的相关结论知，当

$$\sum_{j=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j} < \infty \tag{7}$$

时，该过程的平稳分布存在，即该系统是稳定的。

把式(5)和式(6)代入式(7)可得

$$\sum_{j=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j} = \sum_{j=1}^{\infty} \frac{\left(\frac{n \times \lambda}{m}\right)^j}{(R')^j} = \sum_{j=1}^{\infty} \left(\frac{n \times \lambda}{m \times R'}\right)^j < \infty \tag{8}$$

令 $\rho = \dfrac{n \times \lambda}{m \times R'}$，由于 $\lambda \le R$，则当 $\rho \le \dfrac{n \times \lambda}{m \times R'} < 1$时，该系统的平稳分布是存在的，即该系统是稳定的。

由于 $\rho_{\max} = \dfrac{n \times R}{m \times R'} = \dfrac{n}{s}$，那么就有：当 $S = \dfrac{m \times R'}{R} > n$ 时，任意 VQ 组成的排队系统 $Q$ 是稳定的。由于各 VQ 并行工作且是彼此独立的，则对于 CB-3Clos 交换结构，其中间级缓存组成的排队系统 $P$ 的平稳分布是存在的，即当 $S > n$ 时，CB-3Clos 交换结构是稳定的。 证毕

**2.2.2 CB-3Clos 吞吐量性能分析** 稳定性是交换结

构获得高性能的必要条件，通过在 CB-3Clos 交换结构内部引入缓存，可以有效缓解由于交换网络内部竞争而引起的拥塞，其性能的提升直接反映在 CB-3Clos 交换结构吞吐量上，下面以推论的形式给出 CB-3Clos 交换结构的吞吐量性能分析结果。

**推论 1**　当 $S>n$ 时，CB-3Clos 交换结构能够达到 100%吞吐量。

**证明**　由定理 1 可知，当 $S>n$ 时，CB-3Clos 交换结构是稳定的，即其中间级任意缓存的状态也是稳定，所以就能够存储暂时因为竞争而拥塞的信元从而确保不丢信元[4,7]，由吞吐量的定义可知，此时 CB-3Clos 多级交换结构就能够获得 100%的吞吐量。　　证毕

## 3　基于令牌环的包保序算法

目前，公开的文献中至少有两种办法来防止信元乱序：一种办法是在输入端线卡中建立整形机制，但是这种办法不能解决信元在输出端口失序的问题[8]；另一种是针对不同的交换结构建立相应的包保序技术，例如全缓存交换结构的 LDVSA 调度算法[9]和并行交换中的 VIQ&SKRR 技术[10]，但他们都不适用于 CB-3Clos 交换结构。

解决 CB-3Clos 交换结构乱序问题的关键在于 CB-3Clos 的中间级缓存，而中间级缓存引起信元乱序的根本问题就是调度过程中只根据信元的目的端口进行调度，却忽略了信元的输入端口信息，导致拥有相同输入端口的信元不能严格按照输入顺序进行调度、输出交换网络。

为实现保序，本节对CB-3Clos交换结构的中间级缓存进行扩展：中间级交换单元的每个输入端口都与一个拥有 $r$ 个队列的缓存调度器相连，分别对应 $r$ 个输出级交换单元，且每个队列都可以虚拟成 $n$ 个虚拟队列，分别对应输入交换单元的 $n$ 个输入端口。因此，中间级的每一个虚拟队列都可以用[输入端口、中间级链路、输出端口]唯一标识。在调度过程中只需添加一次对信元输入端口，对应到缓存上就是对虚拟队列序号进行判决的过程，就可以有效保证信元的顺序。在此，虚拟队列定义为$\mathrm{VQ}^{(t)}(i,j,h)$，$1\leqslant i\leqslant m, 1\leqslant j\leqslant r, 1\leqslant h\leqslant r, 1\leqslant t\leqslant n$。为实现保序，同一业务流的信元以轮询方式均匀分配到中间级链路上，中间级调度仍以相应的轮询的方式选择信元。

基于以上分析，本节介绍一种基于令牌环的支持包保序的算法 —— RCLDS(Reordering-supported Central-stage Load-balanced-based Distributed Scheduling)。对调度过程进行功能分解，可以将RCLDS的调度过程分为两部分：对于输入级来说，调度过程就是实现信元负载均衡的过程；对于中间级和输出级来说，调度过程就是实现中间级缓存和输出端口之间匹配的过程。RCLDS算法实现信元保序的关键所在就是实现输入级负载均衡过程和输出级调度过程的轮询指针的完全同步。

### 3.1　输入级负载均衡过程

为较好地实现信元负载均衡，输入级各调度器需要维护一个 1 维的状态矩阵来记录当前时隙与之相连的队列的长度，各调度器的状态矩阵彼此独立。

那么，输入级的负载均衡过程就可以描述为：输入级各调度器根据到达信元的数目取得队长较短的队列进行匹配。假定当前时隙有 $x$ 个输入端口有信元到达，则各调度器取得队长较短的 $x$ 个队列，并基于特定的 QoS 策略与输入端口进行匹配。默认情况下，优先级最高的信元与队列最短的缓存进行匹配，其余匹配依次类推。同时，为实现信元保序，每一个输入端口设置一个同步轮询指针 $A_I(j,t)$(模 $n$ 轮询)，用于记录同一业务流的信元的分配状态，即以轮询的方式将信元按时间先后顺序均匀地分配到相应的虚拟队列中。

### 3.2　中间级和输出级匹配过程

考虑到输入级上的信元分配方式，为保证信元的同步调度，逻辑上可以通过设置令牌的方式来标识同一业务流中到达虚拟队列的信元的先后顺序，因此，用于存储同一业务流的不同信元的相应虚拟队列就可以组成一个令牌环。为实现信元保序，同一业务流中只有那些获得令牌的虚拟队列才可以向其目的端口发送匹配请求。

令牌的管理和分配方式如下：令牌的分发控制由同步调度器 $A_C(h,l)$执行。最先进入交换网络的信元获得令牌，一旦该信元被调度出交换网络，令牌即以轮询方式在令牌环内向下传递，如下一虚拟队列为空，则由 $A_C(h,l)$收回令牌。图 3 给出一个简单示例，在此仍假设有一业务流，其输入端口为 1 号端口，目的为 1 号输出端口。

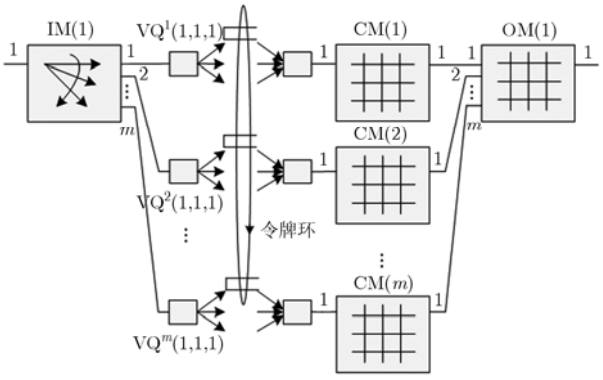


图 3 RCLDS 算法中缓存的令牌调度过程

下面给出 RCLDS 算法的中间级和输出级匹配的详细描述。

阶段 1：CM 内的迭代匹配

步骤 1　所有获得令牌的未匹配的非空 $VQ^{(t)}(i,j,h)$向链路 $L_j(j,h)$的仲裁器 $A_L(j,h)$发送请求；

步骤 2　接收到请求且尚未匹配的仲裁器 $A_L(j,h)$通过查询其轮询指针 $P_L(j,h)$的位置，以轮询方式选择一个非空 VQ，然后向该 VQ 回复应答；

步骤 3　$A_L(i,j,h)$查询其轮询指针 $P_L(i,j,h)$的位置，在接收到的所有应答信号中，以轮询方式选择一个并给予确认。

阶段 2：CM 和 OM 之间的迭代匹配

步骤 1　阶段 1 完成后，被 $A_V(i,j,h)$选中的链路 $L_j(j,h)$向 OM($h$)发送请求，$A_C(h,l)$根据其轮询指针 $P_C(h,l)$的位置，选择一个请求，向 $L_j(j,h)$和 $VQ^{(t)}(i,j,h)$回复应答，同时向 $VQ^{(t+1)}(i,j,h)$发放令牌，并更新其指针 $P_C(h,l)$的值；

步骤 2　如果 $VQ^{(t+1)}(i,j,h)$和 $L_j(j,h)$接收到 $A_C(h,l)$的应答，则发送相应 $VQ$ 中的一个信元，同时取消令牌，并更新指针 $P_V(i,j,h)$和 $P_L(j,h)$的值。

## 4　仿真分析

本文通过仿真实验对 RCLDS 算法的性能进行评估，并与 RD 算法和 CRRD 算法进行比较。实验环境如下：业务源模型分别采用贝努利均匀业务源、突发均匀业务源(突发强度为 10)及 diagonal 业务源；缓存容量为 200 个信元；$n$=$m$=$r$=8，仿真时间为 $10^6$ 个时隙。

图 4 给出了各算法在不同均衡系数下的吞吐量变化曲线。当 $w$=0 时，RD 算法的吞吐量取最小值 0.667，之后开始增长，这是由于 RD 算法采用随机匹配的策略：当均衡系数较小时，随机匹配的算法匹配成功的概率较小，从而导致吞吐量较低。而 CRRD 和 RCLDS 由于采用轮询匹配的策略，能够

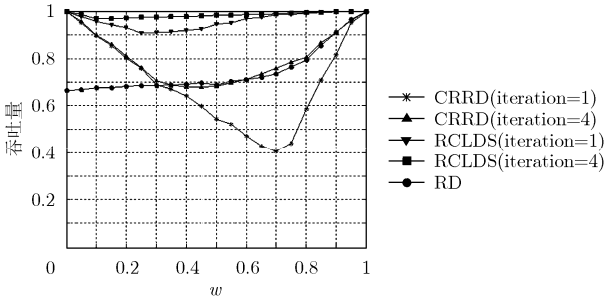


图 4 各算法在不同均衡系数 $w$ 下的吞吐量

在均衡系数较小时获得高匹配数，从而获得较高吞吐量。但随着均衡系数的增大，CRRD 和 RCLDS 算法的吞吐量均有所下降，到达某最小值之后开始上升，其中 CRRD 算法下降迅速，在负载为 0.5 左右时达到最小值 57%。但值得注意的是，RCLDS 算法的系统吞吐量始终保持在 90% 以上，特别是经过 4 次迭代的 RCLDS 算法，几乎可以获得 100%的吞吐量，这说明 RCLDS 算法具有较高的吞吐量特性。

图 5 和图 6 分别给出了不同业务源模型各算法的平均时延性能。可以看出：在贝努利业务源模型下，当负载强度小于 0.55 时，经过 4 次迭代的 RCLDS 算法的平均时延要比 4 次迭代的 CRRD 算法的平均时延要稍微大一些，当负载强度大于 0.55 时，4 次迭代的 RCLDS 算法的时延性能明显优于 4 次迭代的 CRRD 算法。在非均匀 diagonal 业务源模型下，随着负载强度的不断增大，RCLDS 算法的平均时延始终低于 CRRD 算法。这在一定程度上也反映出调度时延与网络端口数成正比。在 RCLDS 算法中网络端口数为 $n$，而 CRRD 算法的网络端口数为 $N$，因此 RCLDS 算法时延的增长速率明显要比 CRRD 慢得多。另一方面，在非均匀 diagonal 业务源模型下，RCLDS 算法由于在输入级采用负载均

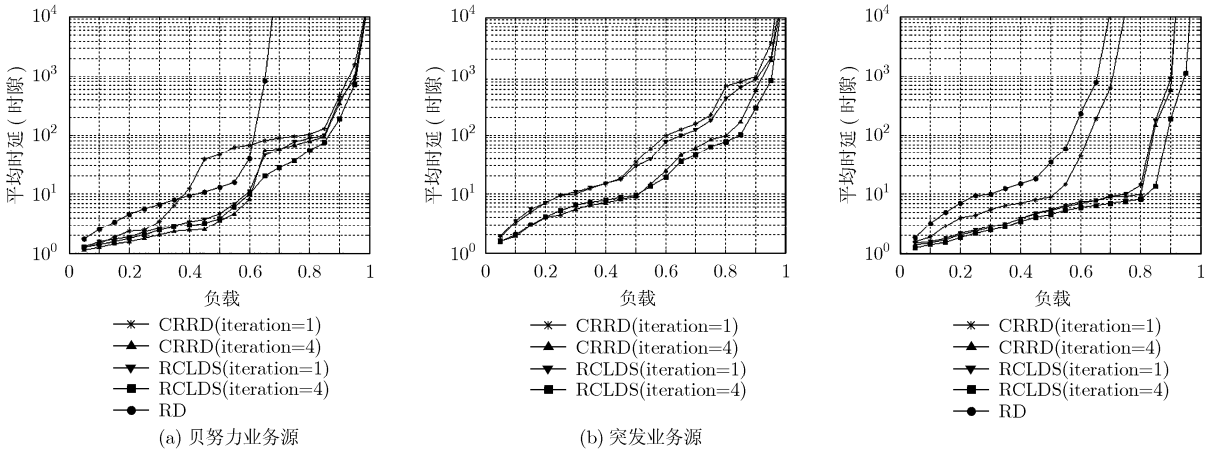


图 5 均匀业务源下各算法的平均时延性能

图 6 非均匀 diagonal 业务源下各算法的平均时延性能

衡，其输入级和中间级的轮询指针能够获得很好的去同步效应，因此 RCLDS 算法能够获得较好的平均时延性能。

从前面的分析和仿真结果可以看出，RCLDS 算法是一种建立在分布式交换网络结构和负载均衡调度思想基础之上的调度算法，与传统的集中式调度算法相比具有以下优点：可扩展性好；调度算法实现简单；在非均匀流量下能够获得较高的吞吐量等。同时，该算法很好地解决了信元乱序问题，不需要在网络输出端口设置大容量缓存来进行信元重排序，降低了交换结构的实现成本，具有较高的理论意义和实际应用价值。

## 5　结论

多级分组交换技术是实现高可扩展大容量路由交换设备的关键技术之一。文本提出一种中间级带缓存的多级交换结构——CB-3Clos 并建立其排队论模型，从而证明该结构的稳定性并推导出其吞吐量性能。此外，针对多级交换技术中的信元保序这一未被很好解决的公开难题，本文基于 CB-3Clos 提出一种新的包保序方案，包括 $n$ 倍扩展虚拟队列结构和包保序调度算法 RCLDS，分析及仿真结果表明该方案实现简单，易于扩展且公平高效，因此，综合考虑当前电路工艺的水平及其实现代价，有理由认为这种中间级带缓存的 CB-3Clos 交换结构及基于中间级负载均衡的调度思想将是一个很好的多级交换网络解决方案。

## 参 考 文 献


[1] Oki E, Jing Zhi-gang, and Rojas-Cessa R, *et al*.. Concurrent round-robin-based dispatching schemes for clos-network switches[J]. *IEEE/ACM Transactions on Networking*, 2002, 10(6): 830–844.

[2] Chiussi F M, Kneuer J G, and Kumar V P. Low-cost scalable switching solutions for broadband networking: The ATLANTA architecture and chipset[J]. *IEEE Communications Magazine*, 1997, 35(12): 44–53.

[3] Chrysos N and Katevenisz M. Scheduling in non-blocking buffered three-stage switching fabrics[C]. Proc. IEEE Globecom2006, Francisco, IEEE Computer Society, 2006: 6–13.

[4] Wang Feng, Zhu Wen-qi, and Hamdi M. The central-stage buffered clos-network to emulate an OQ switch[C]. Proc. IEEE Globecom2006, Francisco, IEEE Computer Society, 2006: 1–5.

[5] Li X and Elhanany I. A scalable frame-based multi-crosspoint packet switching architecture[C]. Proc. HPSR, Brooklyn, USA, 2007: 61–65.

[6] 盛友招. 排队论及其在现代通信中的应用[M]. 北京: 人民邮电出版社, 2007: 50–55.

[7] Shen Yanming, Panwar S S, and Chao H J. Providing 100% throughput in a buffered crossbar switch[C]. Proc. HPSR, Brooklyn, USA. 2007: 1–8.

[8] Iyer S and Mckeown N. Making parallel switches practical[C]. Proc. INFOCOM2001, Alaska, IEEE Computer Society, 2001: 1680–1687.

[9] 杨君刚, 鲍民权, 刘增基等. 一种具有信元保序能力的Clos网络分布式调度算法[J]. 计算机学报, 2008, 31(3): 467–475.
Yang Jun-gang, Bao Min-quan, and Liu Zeng-ji, *et al*.. A distributed scheduling algorithm maintaining cells order for three-stage clos networks[J]. *Chinese Journal of Computers*, 2008, 31(3): 467–475.

[10] 兰巨龙, 董雨果, 陈越, 温建华. 并行交换中支持包保序的缓存结构及调度算法[J]. 电子学报, 2004, 32(12): 35–38.
Lan Ju-long, Dong Yu-guo, Chen Yue, and Wen Jian-hua. The buffer structure and scheduling algorithm for maintaining packet order in the parallel switch[J]. *Acta Electronica Sinica*, 2004, 32(12): 35–38.



胡宇翔：　男，1982 年生，博士生，从事高速交换及调度、路由器体系结构等方面的研究.

兰巨龙：　男，1962 年生，副总工程师，教授，博士生导师，从事高性能宽带信息网络、路由与交换技术等方面的研究.

邬钧霆：　男，1979 年生，博士生，从事路由器体系结构及网络体系架构等方面的研究.