# Delivering 100\% throughput in a Buffered Crossbar with Round Robin scheduling 

## Berger, Michael Stübert

Published in:
Proceedings of IEEE Workshop on high performance Switching and Routing

Link to article, DOI:
10.1109/HPSR.2006.1709743

Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Berger, M. S. (2006). Delivering 100\% throughput in a Buffered Crossbar with Round Robin scheduling. In Proceedings of IEEE Workshop on high performance Switching and Routing (pp. 403-407). IEEE. DOI:
10.1109/HPSR.2006.1709743

## DTU Library

Technical Information Center of Denmark

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal


# Delivering 100\% throughput in a Buffered Crossbar with Round Robin scheduling 

Michael. S. Berger<br>COM•DTU<br>Technical University of Denmark<br>DK-2800 Kgs. Lyngby<br>Email: msb@com.dtu.dk


#### Abstract

Buffered crossbars with Virtual Output Queuing are considered an alternative to bufferless crossbars mainly because the latter requires a complex global scheduling algorithm that matches input with output. Buffered crossbars require only simple schedulers that operate independently for each output crosspoint queue column and independently for each port card. In this paper, fluid-model techniques will be utilized to show that the necessary and sufficient speedup for a $N x N$ buffered crossbar with 1 -cell crosspoint buffers is $2-1 / N$ to deliver $100 \%$ throughput. Round robin scheduling is assumed, both among Virtual Output Queues in port cards and among crosspoint buffers in output columns.


Index Terms-Packet
Throughput, Fluid models.

## I. Introduction

CROSSBAR switch fabrics have been studied extensively in the literature. In combination with Virtual Output Queuing (VOQ), the architecture provides a scalable solution with respect to memory access bandwidth. The crossbar can be either unbuffered or contain a small amount of buffering in each crosspoint. A bufferless crossbar requires a complex scheduling mechanism that matches input with output. Due to the complexity of scheduling algorithms for bufferless crossbars, buffered crossbars are considered as an alternative. As originally proposed in [1], by adding a small buffer capacity in each crosspoint, it is possible to perform the scheduling decision independently among the output columns. The crosspoint buffers generate backpressure signals towards the VOQ's in the port card to avoid overflow. The minimum crosspoint buffer size to maintain full throughput is determined by the round trip delay for the backpressure mechanism. As an alternative to small crossbar buffers in combination with VOQ, one may consider pure crosspoint buffering, however, this requires large buffer capacity in each crosspoint to reduce cell loss.

It has recently been shown that a buffered crossbar switch with a speedup of 2 can emulate a pure output queued switch [7]. A similar result is available for bufferless crossbars: Emulation of an output buffered switch of size $N x N$ is obtainable with a speedup of $2-1 / N$ [8]. The emulation algorithm proposed in [8] is, however, much more complex than the one proposed in [7]. This result indicates that QoS is more easily supported in the buffered crossbar architecture.

The performance of buffered crossbars with VOQ has been
studied in various papers. The architecture was originally proposed in [1] where a simple round-robin scheduling scheme was compared to a more advanced scheme taking into account buffer size and cell age. In [5], a stability analysis is performed for a CICQ (Combined Input and Crosspoint Queued) switch with one cell sized crosspoints. The switch uses Longest Queue First VOQ schedulers. Different combinations of scheduling algorithms are compared in [10]. Longest Queue First, Oldest Cell First and round-robin were considered for VOQ scheduling in combination with Oldest Cell First and round-robin for the crossbar. The paper concludes that the performance is quite similar and recommends the round-robin approach due to its simplicity. The combined input one cell crosspoint buffered switch (CIXB) is compared to iSLIP and pure output queuing (OQ) in [10]. The delay performance of CIXB is better than iSLIP and close to that of an OQ switch. For unbalanced traffic, that is traffic with an uneven distribution of destinations, the CIXB will not support $100 \%$ throughput even if the traffic is admissible. The throughput for unbalanced traffic is, however, better for CIXB compared to iSLIP. In [2] the study is extended to cover more than one buffer location in the crosspoints. Due to the round trip time for backpressure signals, a single buffer location in each crosspoint is not feasible. The high memory consumption is the main drawback of this architecture, and the results are mainly interesting from a theoretical point of view.
Another benefit of a buffered crossbar compared to a bufferless crossbar is the less stringent synchronisation requirement between the port cards and the switch cards [9]. Bufferless crossbars require that all port cards are synchronized to the same clock.

In this chapter, a buffered crossbar with one cell sized crosspoint buffers and round robin arbitration among VOQs and output crosspoint buffers is examined. It is shown, through fluid model techniques that a speedup of $2-1 / N$ is sufficient to deliver $100 \%$ throughput. It is furthermore shown, by example, that this value is also necessary, that is, no speedup values lower than $2-1 / N$ can deliver $100 \%$ throughput.

Here, the switch architecture is described in more detail in section II. A lower bound on speedup is determined in section III based on a specific deterministic arrival pattern. In section IV, it will be shown by fluid model techniques that this speedup is also sufficient for delivering $100 \%$ throughput for a
wide range of admissible traffic patterns. Finally, section V presents proofs and theoretical background for results stated in section IV.

## II. Switch Model

A crossbar buffered switch system of size $N x N$ consists of $N$ Input/Output port cards and a switch card implementing the $N^{2}$ crosspoint buffers as shown in Fig. 1. Each input port card contains VOQs with one buffer for each of the $N$ outputs. The switch model uses round-robin scheduling between VOQs in the port cards and also between crosspoint buffers in an output column. The output port card contains a buffer to store cells in case of speedup. In order to avoid overflow, the crossbar buffers will generate a backpressure signal towards the corresponding VOQ buffer in the port card.


Fig. 1. Buffered crossbar with 1 -cell sized crosspoint buffers. The input buffers are organized as Virtual Output Queues, and the output buffers are required to perform rate adaptation due to speedup.

Each crosspoint buffer has a capacity of one cell. It is assumed that in one switch card timeslot, a cell can be transferred from a VOQ buffer to a crosspoint buffer and further on to the output buffer, if possible. That is, the crosspoint buffers support cut-through. Backpressure from a crosspoint is asserted only if the cell must wait in the crosspoint buffer. These assumptions ensures full throughput for a single traffic stream at $100 \%$ load between input $i$ and output $j$. On the other hand, if the crosspoint buffers would not have cut-through support, then a minimum speedup of 2 would have been required to support $100 \%$ load for this flow, since a cell must wait in the crosspoint buffer for one slot time.

From a performance point of view, the crossbar buffered switch behaves like an output buffered switch for very large crosspoint buffers, and $100 \%$ throughput is achieved for all admissible traffic patterns. This is, however, not the case for limited size crosspoint buffers. In [2], the reduction in switch throughput has been investigated for unbalanced traffic. To
increase throughput, a speedup $S$ can be introduced between the port card and the switch card. The egress port card must then contain buffering to adapt between the different rates, as shown in Fig. 1. The results in [2] were based on a simulation study, however in this paper, fluid model techniques will be utilized to show, that a $100 \%$ throughput can be achieved with $S=2-1 / N$. Fluid models were used in [3] to show that with $S=2,100 \%$ throughput is obtained with a maximal matching algorithm (e.g. iSLIP [4]) in a bufferless crossbar. In [5], a fluid model analysis is performed for a buffered crossbar with one cell sized crosspoints and Longest Queue First (LQF) VOQ schedulers. With all traffic matrix elements less than or equal to $1 / N$, the switch was shown to deliver $100 \%$ throughput. Furthermore, in [6] it was shown that a speedup of 2 is sufficient to deliver $100 \%$ throughput for all types of work-conserving schedulers. Also, algorithms exist that can make buffered and bufferless crossbars emulate an output queued switch, requiring a speedup of 2 and $2-1 / N$, respectively [7][8]. Here, the special case of round robin scheduling [10] is examined. The main contribution of this paper is to examine a specific scheduling implementation (RR) and to derive a specific result which is a lower speedup than the general result in [6]. In the following section, a specific traffic scenario will be utilized to derive a necessary speedup value, $2-1 / N$. Later, it will be shown by fluid model techniques, that this speedup value is also sufficient.

## III. Minimum Speedup

The maximum number of timeslots required for a Head of Line cell in a specific $V O Q$ buffer to be transmitted out of the switch card is composed of a delay in the VOQ scheduler in the port card and the output column scheduler in the switch card. Assume that a cell is in head of line position in timeslot $T=1$. The VOQ scheduler will transmit the cell no later than in timeslot $N$, that is, $(N-1)$ timeslots later due to Round robin scheduling. The output column scheduler will schedule the packet after ( $\mathrm{N}-1$ ) timeslots at maximum, which is timeslot $T=2 N-1$. The next VOQ cell can be considered for scheduling in timeslot $2 N$.

The worst case departure rate for a specific VOQ is thus $1 / T$ $=1 /(2 N-1)$. It is furthermore possible to find arrival distributions that reach this worst-case throughput. Consider the following $N x N$ traffic matrix that specifies arrival rates to $V O Q_{i j}$ :

$$
\left[\begin{array}{llll}
0 & 0 & 0 & \lambda \\
0 & 0 & 0 & \lambda \\
0 & 0 & 0 & \lambda \\
\lambda & \lambda & \lambda & \mu
\end{array}\right]
$$

$V O Q_{N N}$ is assumed to contain more than one cell, and the arrival rates is given by $\lambda=1 /(2 N-1)$. Now, the matrix below shows the arrival timeslots to the VOQs:

$$
\left[\begin{array}{cccc}
0 & 0 & 0 & N \\
0 & 0 & 0 & N \\
0 & 0 & 0 & N \\
1 & 2 & N-1 & N
\end{array}\right]
$$

The timeslots of departure is given in the matrix below:

$$
\left[\begin{array}{cccc}
0 & 0 & 0 & N \\
0 & 0 & 0 & N+1 \\
0 & 0 & 0 & N+2 \\
1 & 2 & N-1 & 2 N-1
\end{array}\right]
$$

In this case, the cell in crosspoint $(N, N)$ will be scheduled in timeslot $2 N-1$, and in timeslot $2 N$, the above arrival pattern is repeated. This example shows that the departure rate from crosspoint $(N, N)$ is limited to $\mu=1 /(2 N-1)$.
If the traffic matrix elements fulfill the following condition

$$
\lambda_{i j} \leq 1 / N, \quad i, j=1, \ldots, N
$$

then the switch will deliver $100 \%$ throughput with a speedup of:

$$
S=\lambda_{\max } / \mu=2-1 / N
$$

At this point, it has been shown that a speedup of $\mathrm{S}=2-1 / \mathrm{N}$ is necessary. In the following section, it will be shown that this speedup is also sufficient to deliver $100 \%$ throughput for a very wide range of admissible traffic.

## IV. Fluid Model

Fluid models have previous been utilized to prove stability of various switch systems. The basic fluid theory is very comprehensive, and will not be derived here. The contribution of this paper is an application of fluid models to show stability of a specific switch/scheduling architecture. The reader should consult [3] and its references for a full theory of fluid models, however, a very brief description of fluid limits is given in section V.

The number of packets in $V O Q_{i j}$ in the beginning of timeslot $n$ is denoted $Z_{i j}(n)$. The cumulative number of arrivals and departures at the beginning of timeslot n are denoted $A_{i j}(n)$ and $D_{i j}(n)$, respectively. Thus, the following relation holds:

$$
Z_{i j}(n)=Z_{i j}(0)+A_{i j}(n)-D_{i j}(n), \quad n \geq 0, \quad i, j=1, \ldots ., N
$$

It is assumed that the arrival process obeys the strong law of large numbers, that is:

$$
\lim _{n \rightarrow \infty} \frac{A_{i j}(n)}{n}=\lambda_{i j}, \quad i, j=1, \ldots, N
$$

where $\lambda_{i j}$ is the arrival rate to $V O Q_{i j}$. The switch is, by definition, rate stable if:

$$
\lim _{n \rightarrow \infty} \frac{D_{i j}(n)}{n}=\lambda_{i j}, \quad i, j=1, \ldots, N
$$

If the switch is rate stable for an admissible traffic matrix, i.e. traffic that fulfills:

$$
\sum_{i} \lambda_{i j} \leq 1, \quad \sum_{j} \lambda_{i j} \leq 1
$$

then the switch is said to deliver $100 \%$ throughput.
The fluid model of the switch is given by:

$$
\bar{Z}_{i j}(t)=\bar{Z}_{i j}(0)+\lambda_{i j} t-\bar{D}_{i j}(t), \quad t \geq 0
$$

The fluid equation is established by a limiting procedure as illustrated in [3]. This is discussed in sec V.A.

Now, the fluid model of a switch is said to be weakly stable, if for every fluid model solution $(\bar{D}, \bar{Z})$, with $\bar{Z}(0)=0$, $\bar{Z}(t)=0$ for $t \geq 0$. The switch is rate stable if the corresponding fluid model is weakly stable. This result is shown in [3]. Note that all quantities will be fluid limits in the remainder of this section, therefore, no specific indication (bar) is given.
As shown in the previous section, if $\lambda_{i j} \leq 1 / N$ and $S \geq 2-1 / N$, then $Z_{i j}(t)=0$ for $t \geq 0$, which implies that $V O Q(i, j)$ is rate stable. This result is actually based on specific knowledge of the scheduling algorithm that is, Round Robin both among Virtual Output Queues in port cards and among crosspoint buffers in output columns.
Now, if on the other hand $\lambda_{i j} \geq 1 / N$, the fluid model must be shown to be weakly stable. To prove this result, the following fact shown in [3] is used: Let $f$ be a non-negative, absolutely continuous function with $f(0)=0$. Assume that for almost every $t$ such that $f(t)>0, f^{\prime}(t) \leq 0$. Then $f(t)=0$ for almost every $t \geq 0$. In order to find the function $f(t)$, a matrix $C_{i j}(t)$ will be introduced:

Assume that $\lambda_{i j} \geq 1 / N$ and $Z_{i j}(t)>0$. Then, it can be derived that (see section V.B):

$$
\sum_{i^{\prime} \neq i} Z_{i^{\prime} j}^{\prime}(t)+\sum_{j^{\prime} \neq j} Z_{i j^{\prime}}^{\prime}(t)+Z_{i j}^{\prime}(t) \leq \sum_{i^{\prime}} \lambda_{i^{\prime} j}+\sum_{j^{\prime}} \lambda_{i j^{\prime}}-\lambda_{i j}-S
$$

Now define:

$$
C_{i j}(t)=\sum_{i^{\prime} \neq i} Z_{i^{\prime} j}(t)+\sum_{j^{\prime} \neq j} Z_{i j^{\prime}}(t)+Z_{i j}(t),
$$

thus

$$
C_{i j}^{\prime}(t) \leq \sum_{i^{\prime}} \lambda_{i^{\prime} j}+\sum_{j^{\prime}} \lambda_{i j^{\prime}}-\lambda_{i j}-S
$$

Since the traffic is assumed to be admissible, then $C_{i j}^{\prime}(t) \leq 0$ for $Z_{i j}(t)>0, \quad \lambda_{i j} \geq 1 / N$ and $S \geq 2-1 / N$. The function $f$ is now defined as:

$$
f(t)=\sum_{i, j} Z_{i j}(t) \cdot C_{i j}(t)
$$

Note that if $f(t)=0$ then also $Z_{i j}(t)=0$ for all $i, j$, and the switch is thus rate stable. The goal is therefore to prove that $f(t)$ equals zero. It is proven in section V.C that:

$$
f^{\prime}(t)=2 \sum_{i, j} Z_{i j}(t) \cdot C_{i j}^{\prime}(t)
$$

Since $Z_{i j}(t) \geq 0$ and $C^{\prime} i j(t) \leq 0$ it follows that $f^{\prime}(t) \leq 0$ for $f(t)>0, f(t)=0$ for almost every $t \geq 0$. As stated previously, this implies that $Z_{i j}(t)=0$ for all $i, j$. Hence, the switch is rate stable (delivers $100 \%$ throughput). It has now been proven that a speedup of $2-1 / N$ is sufficient to deliver $100 \%$ throughput for a wide range of admissible traffic patterns.

At this point it is interesting to note, that by having no assumptions on the scheduling algorithm, except that it is work-conserving, then the switch would be stable for $\lambda_{i j} \geq 0$ and $S \geq 2$. This general result is, as stated previously, derived in [6] by using fluid models.

For large switch systems, the term $1 / N$ is insignificant, and the required speedup for the 1 -cell crosspoint buffered switch approaches 2 , which is also the speedup required for bufferless crossbars with maximal matching schedulers, e.g. iSLIP. The performance analysis in [2] shows by simulation that buffered crossbars with round robin scheduling performs significantly better than bufferless crossbars with iSLIP. However, in the general case, the worst case performance is quite similar, as indicated by the results in this paper.

## V. Proofs and theory

This section presents proofs of results from previous sections and sketches theoretical background for fluid models. [3] and its references contains a full theory of fluid models. Here, the fluid models and the fluid limits are only briefly covered in section A . The following section B derives an inequality which is a central part of the stability proof given in the previous section. Finally, section C shows

## A. Fluid Limits

The fluid model of a switch is determined by a limiting procedure, as illustrated in the following. First, continuous functions for the cumulative number of arrivals and departures $(A(n)$ and $D(n))$ and are defined for $t \in[n, n+1)$ :

$$
\begin{aligned}
& A_{i j}(t)=A_{i j}(n) \\
& D_{i j}(t)=D_{i j}(n)+(t-n)\left(D_{i j}(n+1)-D_{i j}(n)\right)
\end{aligned}
$$

Based on the continuous functions $(Z, A, D)$, three new functions are defined:

$$
\begin{aligned}
& \bar{A}_{i j}^{r}(t, \omega)=r^{-1} A_{i j}(r t, \omega) \\
& \bar{D}_{i j}^{r}(t, \omega)=r^{-1} D_{i j}(r t, \omega) \\
& \bar{Z}_{i j}^{r}(t, \omega)=r^{-1} Z_{i j}(r t, \omega)
\end{aligned}
$$

Note the explicit indication of the sample path $\omega$. It can be shown [3] that for a fixed $\omega$ and any sequence $\left\{r_{n}\right\}$ with $r_{n} \rightarrow \infty$ for $n \rightarrow \infty$, there exists a subsequence $\left\{r_{n_{k}}\right\}$ and continuous functions $(\bar{A}, \bar{D}, \bar{Z})$ such that for any $t \geq 0$ as $k \rightarrow \infty$ :

$$
\left(\bar{A}^{r_{n k}}(t, \omega), \bar{D}^{r_{n k}}(t, \omega), \bar{Z}^{r_{n k}}(t, \omega)\right) \rightarrow(\bar{A}(t), \bar{D}(t), \bar{Z}(t)) .
$$

Any functions ( $\bar{A}, \bar{D}, \bar{Z}$ ) obtained through this limiting procedure is said to be a fluid limit, and it is easy to see, that fluid limits are solutions to the fluid equation of the switch.

## B. Inequality

Consider crosspoint $(i, j)$. Assume that $\bar{Z}_{i j}(t)>0$, then by continuity of $Z, \exists \delta$ such that $\bar{Z}_{i j}\left(t^{\prime}\right)>0$ for $t^{\prime} \in[t, t+\delta]$. Set $a=\min _{t \in[t, t+\delta]} \bar{Z}_{i j}\left(t^{\prime}\right)$. Thus, for large enough $k$, $\bar{Z}_{i j}^{r_{n k}}\left(t^{\prime}\right) \geq a / 2$ for $t^{\prime} \in[t, t+\delta]$ and for large enough $k$, $r_{n_{k}} a / 2 \geq 1$. Thus, $Z\left(t^{\prime}\right) \geq 1$ for $t^{\prime} \in\left[r_{n_{k}} t, r_{n_{k}}(t+\delta)\right]$, which
 $\left[r_{n_{k}} t, r_{n_{k}}(t+\delta)\right]$.
Now, the number of departures from $\operatorname{VOQ}(i, j)$ in the long time interval $\left[r_{n_{k}} t, r_{n_{k}} t^{\prime}\right], t^{\prime} \in[t, t+\delta]$ fulfills:
$D_{i j}\left(r_{n_{k}} t^{\prime}\right)-D_{i j}\left(r_{n_{k}} t\right) \geq S r_{n_{k}}\left(t^{\prime}-t\right)-\sum_{i^{\prime} \neq i}\left(D_{i^{\prime} j}\left(r_{n_{k}} t^{\prime}\right)-D_{i^{\prime} j}\left(r_{n_{k}} t\right)\right)$
$-\sum_{j^{\prime} \neq j}\left(D_{i j^{\prime}}\left(r_{n_{k}} t^{\prime}\right)-D_{i j^{\prime}}\left(r_{n_{k}} t\right)\right)-K$
The equation states that there will be a departure from $V O Q(i, j)$ in each timeslot except for the number of timeslots corresponding to the number of departures from VOQs belonging to the same row or column as $\operatorname{VOQ}(i, j)$ (worstcase). The equation above is valid for all types of workconserving scheduling disciplines, and is not restricted to round robin. Note that a small constant $K$ is subtracted to compensate for e.g. truncations; however, it is not significant in the fluid limit.

Dividing this equation with $r_{n_{k}}$ and letting $k \rightarrow \infty$, fluid limits are obtained. Furthermore, by dividing with $t^{\prime}-t$ and letting $t^{\prime} \rightarrow t$, the inequality is expressed by the derivative of the fluid limit:

$$
\bar{D}_{i j}^{\prime}(t) \geq S-\sum_{i^{\prime} \neq i} \bar{D}_{i^{\prime} j}^{\prime}(t)-\sum_{j^{\prime} \neq j} \bar{D}_{i j^{\prime}}^{\prime}(t)
$$

Now since $\bar{D}_{i j}^{\prime}(t)=\lambda_{i j}-\bar{Z}^{\prime}{ }_{i j}(t)$ according to the switch fluid model, the following inequality is finally obtained:
$\sum_{i \neq i} \bar{Z}_{i^{\prime} j}^{\prime}(t)+\sum_{j^{\prime} \neq j} \bar{Z}_{i j^{\prime}}^{\prime}(t)+\bar{Z}_{i j}^{\prime}(t) \leq \sum_{i^{\prime}} \lambda_{i^{\prime} j}+\sum_{j^{\prime}} \lambda_{i j^{\prime}}-\lambda_{i j}-S$,
when $\bar{Z}_{i j}(t)>0$.

## C. Derivative of $f(t)$

The function $f(t)$ defined in sec. IV is given by:

$$
\begin{aligned}
& f(t)=\sum_{i, j} Z_{i j}(t) \cdot C_{i j}(t)= \\
& \sum_{i, j} Z_{i j}(t)\left(\sum_{i \neq i} Z_{i^{\prime} j}(t)+\sum_{j^{\prime} \neq j} Z_{i j^{\prime}}(t)+Z_{i j}(t)\right)= \\
& \sum_{i, j} Z_{i j}(t)\left(\sum_{i^{\prime}} Z_{i^{\prime} j}(t)+\sum_{j^{\prime}} Z_{i j^{\prime}}(t)-Z_{i j}(t)\right)= \\
& \sum_{i, j} Z_{i j}(t)\left(\sum_{i} Z_{i j}(t)\right)+\sum_{i, j} Z_{i j}(t)\left(\sum_{j^{\prime}} Z_{i j^{\prime}}(t)\right) \\
& +\sum_{i, j} Z_{i j}(t)\left(Z_{i j}(t)\right)
\end{aligned}
$$

Consider the first part of the function $f$ given by:

$$
f_{1}(t)=\sum_{i, j} Z_{i j}(t) \sum_{i^{\prime}} Z_{i^{\prime} j}(t)=\sum_{i, j, i^{\prime}} Z_{i j}(t) Z_{i^{\prime} j}(t) .
$$

Thus,

$$
f_{1}^{\prime}(t)=\sum_{i, j, i^{\prime}} Z_{i j}^{\prime}(t) Z_{i^{\prime} j}(t)+\sum_{i, j, i^{\prime}} Z_{i j}(t) Z_{i^{\prime} j}^{\prime}(t) .
$$

The two sums above yield the same result because the index ranges and products are identical. Hence,

$$
f_{1}^{\prime}(t)=2 \sum_{i, j, i^{\prime}} Z_{i j}(t) Z_{i^{\prime} j}^{\prime}(t)=2 \sum_{i, j} Z_{i j}(t) \sum_{i^{\prime}} Z_{i^{\prime} j}^{\prime}(t) .
$$

Identical results hold for the remaining parts of $f$, thus

$$
f^{\prime}(t)=2 \sum_{i, j} Z_{i j}(t)\left(\sum_{i^{\prime}} Z_{i^{\prime} j}^{\prime}(t)+\sum_{j^{\prime}} Z_{i j^{\prime}}^{\prime}(t)-Z_{i j}^{\prime}(t)\right)
$$

and it is finally concluded that the derivative of $f$ is given by:

$$
f^{\prime}(t)=2 \sum_{i, j} Z_{i j}(t) \cdot C_{i j}^{\prime}(t)
$$

## VI. Conclusions

Buffered crossbars have several advantages compared to non-buffered crossbars including simpler arbitration, synchronisation relaxation and better performance. In this paper, it has been shown through fluid model techniques that a buffered crossbar with 1-cell crosspoint buffers delivers $100 \%$ throughput with a speedup of $2-1 / N$ when Round Robin scheduling is assumed. The only assumption on the arrival process is that it obeys the strong law of large numbers. The contribution of this paper is an application of fluid models to show stability of a specific scheduling mechanism. It has previously been shown that a speedup of 2 is sufficient to deliver $100 \%$ throughput for any work conserving scheduler, however, by utilizing specific properties of the scheduling algorithm, a lower bound of $2-1 / N$ could be obtained. In a practical switch system, a larger crosspoint buffer size is typically required, due to backpressure latency. Therefore, it will be interesting to extend this work to cover more than one buffer position in the crosspoints.

It was previously believed that the performance of a 1 -cell crosspoint buffered switch would be significantly better than the performance of a buffless crossbar with iSLIP arbitration. However, since the $-1 / N$ is insignificant for large switch sizes, it is concluded that the worst case performance of the two
switch systems in terms of throughput is quite similar,

## References

[1] M. Nabeshima, "Performance evaluation of a combined input- and crosspoint-queued switch", IEICE Transactions on Communications, Vol.E83-B Issue.3, p737-41, 2000
[2] R. Rojas-Cessa, E. Oki, H.J. Chao, "CIXOB-k: combined input-crosspoint-output buffered packet switch", Proceedings of IEEE GLOBECOM 2001, p 2654-2660, Vol. 4
[3] J.G. Dai, B.Prabhakar, "The throughput of data switches with and without speedup", Proceedings IEEE INFOCOM 2000, p556-64 vol. 2 2000
[4] N. McKeown, "The iSLIP scheduling algorithm for input-queued switches", IEEE/ACM Transactions on Networking, p 188 -201, Vol. 7 Issue.2, 1999.
[5] T. Javidi, R. Magill, T. Hrabik, "A high-throughput scheduling algorithm for a buffered crossbar switch fabric", Proceedings of IEEE ICC 2001, p 1586-1591, vol.5.
[6] S. Chuang, S. Iyer, N. McKeown, "Practical Algorithms for Performance Guarantees in Buffered Crossbars". Tech. Report, Stanford univ.
[7] R. B. Magill, C. E. Rohrs, R. L. Stevenson, "Output-queued switch emulation by fabrics with limited memory", IEEE Journal on Selected Areas in Communications, Volume: 21, Issue: 4, May 2003.
[8] S. T. Chuang, A. Goel, N. McKeown, B. Prabhakar, "Matching Output Queuing with a Combined Input Output Queued Switch", IEEE Journal on Selected Areas in Communications, vol.17, n.6, Dec.1999, pp. 10301039.
[9] F. Abel, C. Minkenberg, R. P. Luijten, M. Gusat, I. Iliadis, "A FourTerabit Single-Stage Packet Switch with Large Round-Trip Time Support", 10th Symposium on High Performance Interconnects HOT Interconnects, p.5, 2002
[10] I. Radusinovic, M. Pejanovic, Z. Petrovic, "Impact of scheduling algorithms on performances of buffered crossbar switch fabrics", Proceedings of IEEE ICC 2002, p 2416-2420, Vol. 4

