

Using Constraint Sets to Achieve Delay Bounds in CIOQ Switches

Sundar Iyer, *Associate Member, IEEE*, and Nick McKeown, *Senior Member, IEEE*

Abstract—We recently proposed Constraint Sets as a simple technique to analyze routers with a single stage of buffering. In this letter, we extend the technique to analyze combined input and output (CIOQ) routers with two stages of buffering.

Index Terms—100% throughput, combined input and output (CIOQ) switch, constraint sets, delay guarantees, input queued switch.

I. INTRODUCTION

IN PREVIOUS work [3] we described single buffered (SB) routers: a general class of routers in which packets are buffered exactly once as they pass through. SB routers include some well-known architectures such as input queueing, output queueing and centralized shared memory, as well as some routers with more complex arrangements of buffers, such as the Parallel Packet Switch [2] or the Distributed Shared Memory router [3]. In [2], we showed how the Constraint Set technique (a generalization of the pigeon-hole principle), can be used to determine the number of memory devices needed for a deterministic SB router, and how packets should be allocated to each memory to emulate an ideal output queued (OQ) router. The Constraint Set technique captures the physical constraints in a router, in particular the limitations imposed by access to memory devices. It appears to be a natural technique for analyzing deterministic SB routers.

In this letter, we extend our results in [3] to show how Constraint Sets can be applied to a router with more than one stage of buffering. Specifically, we show how Constraint Sets can be used to analyze a combined input and output queued (CIOQ) router, which has *two* stages of buffering. The analysis of CIOQ routers is usually quite involved, leading to impractical and complex scheduling algorithms [4].

As we will see, applying the Constraint Set technique to the CIOQ router leads to an intuitive understanding of the physical constraints, and a simpler scheduling algorithm. In this case, it simplifies the well-known result by Charny [1] that—with leaky

bucket constrained arrivals—a CIOQ router with a speedup¹ of two can emulate a first in first out (FIFO) OQ router with a bounded delay difference. In other words, when subject to the same leaky bucket constrained arrival patterns, packets depart from the CIOQ router and the FIFO-OQ router at the same time, or at least within a fixed bound of each other.

In what follows, we will use Constraint Sets to analyze the same CIOQ router under the same arrival conditions.

A. Definitions

Shadow FIFO-OQ Switch: Assume that there exists a first in FIFO-OQ switch, called the “shadow FIFO-OQ switch,” with the same number of input and output ports as the CIOQ switch. The ports on the shadow FIFO-OQ switch receive identical input traffic patterns and operate at the same line rate as the CIOQ switch. As the name suggests, the shadow FIFO-OQ switch serves packets destined for each output in FIFO order.

Single Leaky Bucket Constrained Traffic (B): The traffic arriving at a switch is said to be single leaky bucket constrained if for every output j , the number of packets which arrive at the switch destined to j in the time interval (t_1, t_2) is given by $N(t_1, t_2) \leq \lambda_j(t_2 - t_1) + B_j$, where B_j is some constant.

Note that we require, $0 \leq \lambda_j < 1$ for the traffic to be admissible. We define $B = \max\{\forall j, B_j\}$. Also, in this traffic model we shall assume that at most one cell arrives at each input of the CIOQ switch in any given time slot.

FIFO-OQ Departure Time (DT): Consider a cell that arrives to the CIOQ switch. The FIFO-OQ departure time, DT , is the departure time of that cell from the shadow FIFO-OQ switch.

II. ACHIEVING DELAY BOUNDS IN A FIFO-CIOQ SWITCH

A. Background

In [1], Charny proved the following theorem.

Theorem 1—(Sufficiency): Any maximal algorithm with a speedup $S > 2$, which gives preference to cells which arrive earlier,² ensures that any cell arriving at time t will be delivered to its output at a time no greater than $t + [B/(S - 2)]$, if the traffic is single leaky bucket B constrained.

Proof: Proved in [1, Sec. II-C, Th. 5].³ □

¹A CIOQ switch is said to have a speedup of S , for $S \in \{1, 2, 3, \dots, N\}$ if it can remove up to S cells from each input and transfer at most S cells to each output in a time slot. In the rest of this letter we shall assume that all packets are split into cells of fixed size. We take the arrival time between cells as the basic time and refer to it as a time slot.

²This is defined in [1, Sec. II-C].

³Charny uses a dual leaky bucket traffic model. The result in [1] has been restated here for the single leaky bucket model to facilitate comparison between Theorems 2 and 1.

Manuscript received October 16, 2002. The associate editor coordinating the review of this letter and approving it for publication was Prof. C. Douligeris. This work was supported by the National Science Foundation under NGI Contract ANI-9872761, by the Stanford Networking Research Center, and by Cisco Systems.

The authors are with the Computer Systems Laboratory, Stanford University, Stanford, CA 94301-9030 USA (e-mail: sundaes@cs.stanford.edu; nickm@stanford.edu).

Digital Object Identifier 10.1109/LCOMM.2003.812712

In [1], Charny uses a maximal matching algorithm (called oldest cell first) which gives priority to cells which arrive earlier to the CIOQ switch, and uses the fact that for any maximal algorithm, if there is a cell waiting at input i destined to output j , then either input i is matched or output j is matched (or both).⁴ The proof counts all cells (called competing cells) that can prevent a particular cell from being transferred and classifies the competing cells into two types—cells at input i , or cells destined to output j . It is shown that, after it arrives, a cell cannot be prevented from being transferred to output j for more than $\lceil B/(S-2) \rceil$ time slots. The argument is somewhat complex for two reasons. First, a cell can repeatedly be prevented—by competing cells—from being transferred over multiple time slots. Second, it is possible that a cell is overtaken by cells that arrive later at different inputs, and then need to be resequenced at the output. While this can't be prevented, in what follows we'll fix the transfer time of a cell as soon as it arrives. This way, a cell's transfer time can't be affected by cells arriving later.

B. Alternative Approach Based on Constraint Sets for FIFO-CIOQ Router

First consider the physical structure of a CIOQ router. If a cell arrives at input i destined for output j , the CIOQ router is constrained to transfer the cell only when input i and output j are both free. Constraint Sets are a convenient accounting method to maintain and update information about when the inputs and outputs are free, and to analyze the conditions under which the router will emulate a FIFO OQ router.

We will use the following algorithm:

When a cell arrives at input i destined to output j with FIFO-OQ departure time DT , the cell is scheduled to depart at the first time in the future (larger than DT) when both the input i (output j) are free to send (receive) a cell.

More formally, the algorithm is as follows. We start by describing the algorithm when speedup $S = 1$, before generalizing to larger speedup values:

- 1) *Maintaining Constraint Sets*: All inputs and outputs maintain a constraint set. Each entry in the constraint set represents an opportunity to transmit a cell in the future; one entry for each future time slot. For each future time slot that an input is busy, the corresponding entry in its constraint set represents a cell that it will transmit across the switch fabric to an output. Similarly, for each future time slot that an output is busy, the entry represents a cell that it will receive from one of the inputs. If, at some time in the future, there is no cell to be transferred from an input (or to an output), then the corresponding entry is free and may be used to schedule newly arriving cells.
- 2) *Negotiating a Constraint-Free Time to Transfer*: When a cell arrives at input i destined to output j , input i communicates its input constraint set to output j and requests a time in the future for it to transmit that cell. Output j then

picks the first time in the future t_f in the interval $(DT, DT+k)$ (where k is a constant which we will determine shortly) for which both input i and output j are free to transmit and receive a cell, i.e., time index t_f is free in the constraint sets of input i and output j . Output j grants input i the time slot t_f in future for transmitting the cell.

- 3) *Updating Constraint Sets*: Both input i and output j update their respective constraint sets to note the fact that time t_f in the future is reserved for transmitting the cell from input i to output j in the CIOQ switch.

When the speedup $S > 1$, an entry in the input constraint set is said to be free in a particular time slot if the input is scheduled to send fewer than S cells. Likewise, an entry in the output constraint set is said to be free if the output is scheduled to receive fewer than S cells (from any input) in the corresponding time slot.

We now find the value of k for which every packet in the CIOQ switch is transferred from its input to its output within k time slots of its FIFO-OQ departure time, i.e., $t_f \leq DT+k$ or $t_f \in (DT, DT+k)$ (where t is the arrival time of a cell and k is a constant). The larger the speedup the smaller the value of k .

Lemma 1: The number of time slots available in the input constraint set (ICS) for any input i at any given time is greater than $\lfloor k - \lfloor (k+B)/S \rfloor \rfloor$.

Proof: Consider a cell that arrives to input i at time t , destined for output j with FIFO-OQ departure time DT . The cell is scheduled to be transferred from input i to output j in the CIOQ switch in the interval $(DT, DT+k)$. Since the traffic is single leaky bucket B constrained, no cell which arrived before time $DT-B$ at input i has a FIFO-OQ departure time in the interval $(DT, DT+k)$. Hence, no cell which arrived before time $DT-(B+k)$ at input i , is allocated to be transferred from input i in the CIOQ switch in the interval $(DT, DT+k)$. If the speedup is S , then the number of time slots available in the input constraint set for the newly arriving cell is at least $\lfloor k - \lfloor (k+B)/S \rfloor \rfloor$. \square

Lemma 2: The number of time slots available in the output constraint set (OCS) for any output j at any given time is greater than $\lfloor k - \lfloor k/S \rfloor \rfloor$.

Proof: Consider a cell that arrives at input i at time t destined for output j with FIFO-OQ departure time DT . The cell is scheduled to be transferred from input i to output j in the CIOQ switch in the interval $(DT, DT+k)$. Since all cells are scheduled to be transferred in the CIOQ switch within k time slots of their FIFO-OQ departure time, no more than k cells which have FIFO-OQ departure times in the interval $(DT-k, DT-1)$ can already have been allocated to be transferred to output j in the CIOQ switch in the interval $(DT, DT+k)$. Thus if the speedup is S , then the number of time slots available in the output constraint set for the newly arriving cell is at least $\lfloor k - \lfloor k/S \rfloor \rfloor$.⁵ \square

Theorem 2: (Sufficiency) With a speedup $S > 2$, the algorithm ensures that each cell in the CIOQ switch is delivered to its output within $\lceil B/(S-2) \rceil$ time slots of its FIFO-OQ departure time, if the traffic is single leaky bucket B constrained.

⁵We do not consider cells which have a FIFO-OQ departure time in the interval $(DT+1, DT+K)$ since the output policy is FIFO and these cells will be considered only after cell C is allocated a time $T_f \in (DT, DT+k)$ for it to be transferred from input i to output j in the CIOQ switch.

⁴A similar analysis was used in [5].

Proof: (Using Constraint Sets). Consider a cell which arrives at time t . It should be allocated a time slot t_f for departure such that, $t_f \in ICS \cap OCS$. A sufficient condition to satisfy this is that $[k - \lfloor (k + B)/S \rfloor] > 0$, $[k - \lfloor k/S \rfloor] > 0$ and $[k - \lfloor (k + B)/S \rfloor] + [k - \lfloor k/S \rfloor] > k$. This is always true if we choose, $k > B/(S - 2)$. \square

III. OBSERVATIONS

In [1] it was shown that a *maximal matching* algorithm would lead to the main result (Theorem 1). The result relied on a scheduler that examines the contents of the input queues during each time slot to determine which cells to schedule. In contrast, the Constraint Set technique leads to an almost identical result (Theorem 2), using a simpler algorithm that schedules cells as soon as they arrive. While algorithms for IQ and CIOQ switches that schedule cells immediately upon arrival have been proposed before [6]–[9], we are not aware of any previous work which shows when such algorithms can achieve 100% throughput or give bounded delay.

By showing that the FIFO-CIOQ router emulates a FIFO-OQ router, it immediately follows from Theorem 2 that the router has bounded delay, and 100% throughput.

REFERENCES

- [1] A. Charny, "Providing QoS guarantees in input-buffered crossbars with speedup," Ph.D. dissertation, M.I.T., Cambridge, Sept. 1998.
- [2] S. Iyer and N. McKeown, "On the speedup required for a parallel packet switch," *IEEE/ACM Trans. Networking*, Apr. 2003, to be published.
- [3] S. Iyer, R. Zhang, and N. McKeown, "Routers with a single stage of buffering," in *Proc. ACM SIGCOMM*, Pittsburgh, PA, Sept. 2002.
- [4] S. T. Chuang, A. Goel, N. McKeown, and B. Prabhakar, "Matching output queueing with a combined input output queued switch," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1030–1039, Dec. 1999.
- [5] B. Prabhakar and N. McKeown, "On the speedup required for combined input and output queued switching," *Automatica*, vol. 35, no. 12, Dec. 1999.
- [6] M. Akata, S. Karube, T. Sakamoto, T. Saito, S. Yoshida, and T. Maeda, "A 250 Mb/s 32×32 CMOS crosspoint LSI for ATM switching systems," *IEEE J. Solid-State Circuits*, vol. 25, pp. 1433–1439, Dec. 1990.
- [7] M. Karol, K. Eng, and H. Obara, "Improving the performance of input-queued ATM packet switches," in *INFOCOM '92*, pp. 110–115.
- [8] H. Matsunaga and H. Uematsu, "1.5Gb/s 8×8 cross-connect switch using a time reservation algorithm," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1308–1317, Oct. 1991.
- [9] H. Obara, S. Okamoto, and Y. Hamazumi, "Input and output queueing ATM switch architecture with spatial and temporal slot reservation control," *Electron. Lett.*, pp. 22–24, Jan. 1992.